Publication status: This preprint has not been published elsewhere.

Environmental Drivers of Skipjack tuna (Katsuwonus pelamis) Abundance in the Indian Ocean: Insights from Satellite Remote Sensing and Catch Data

Teshan Ishara, Kasun Anuruddha Bandara, Sandamali Sakunthala Herath

https://doi.org/10.1590/2675-2824074.2512

Submitted on: 2026-08-06
Posted on: 2026-08-12 (version 1)
(YYYY-MM-DD)

Ocean and Coastal Research

http://dx.doi.org//10.1590/2675-2824074.25127

ISSN 2675-2824

Environmental Drivers of Skipjack tuna (*Katsuwonus pelamis*) Abundance in the Indian Ocean: Insights from Satellite Remote Sensing and Catch Data

Teshan Ishara[1,*], Kasun Anuruddha Bandara[1], Sandamali Sakunthala Herath[1]

T.I.: https://orcid.org/0009-0009-8818-2972

K.A.B.: https://orcid.org/0000-0003-1351-3959

S.S.H.: https://orcid.org/0000-0001-9331-6845

1 Department of Fisheries and Aquaculture, Faculty of Fisheries and Marine Sciences & Technology, University of Ruhuna, Wellamadama, Matara, Sri Lanka.

* Corresponding author: teshan.ishara@gmail.com

Abstract

Skipjack tuna (*Katsuwonus pelamis*) is an economically and ecologically important fish species with a wide global distribution, particularly abundant in tropical oceans influenced by strong climatic variability. In the Indian Ocean, monsoonal cycle and mesoscale oceanographic processes produce rich upwelling areas that serve as productive fishing grounds. . Four-year satellite-derived environmental parameters, Sea Surface Temperature (SST), Sea Surface Chlorophyll (SSC), Sea Surface Height (SSH), Mixed Layer Depth (MLD), Sea Surface Salinity (SSS), and Eddy Kinetic Energy (EKE), data were obtained. These satellite data was merged with the catch data obtained from the Department of Fisheries and Aquatic Resources (DFAR) of Sri Lanka. The effect of the above oceanographic factors on skipjack tuna abundance and distribution was studied using the generalized additive model (GAM), and the empirical cumulative distribution function (ECDF). Findings showed clear seasonal and spatial variability with elevated CPUE consistently associated with moderate SST (26.5°C to 29.0°C), 0.00 mgm^{-3} to 0.50 mgm^{-3} of SSC, 0.35 m to 0.55 m of SSH, 0.00 m^2s^{-2} to 0.90 m^2s^{-2} of EKE, 31.5 PSU to 35.25 PSU of SSS, and 10.00 m to 27.50 m of MLD. The catch per unit effort (CPUE) shows variability in the main monsoon seasons of the Indian Ocean. These results showed SST, EKE, and SSC are important environmental parameters affecting the distribution and abundance of the skipjack tuna in the Indian Ocean. The EKE showed a strong association with the skipjack tuna catch rates when paired with the SST, suggesting that ocean eddies and sea surface temperature patterns play an important role in skipjack tuna distribution and abundance.

Keywords: Monsoonal cycle, abundance, remote sensing, oceanographic factors

INTRODUCTION

Skipjack tuna is an economically and ecologically important fish species widely distributed throughout the Indian Ocean (Stamouli et al., 2017). They are highly migratory and one of epipelagic species typically inhabits environmental conditions such as sea surface temperatures ranging from 25 ° C to 30 ° C and they are known to be aggregated in areas with high primary productivity (Grande et al., 2014). Oceanographic factors such as sea surface temperature (SST), sea surface chlorophyll (SSC), sea surface height (SSH), and mixed layer depth (MLD) have been strongly influencing the distribution and abundance of skipjack tuna species. In addition to distribution and abundance, those oceanographic factors affect the thermal habitat and the prey availability (Zainuddin et al., 2006).

Remote sensing (RS) systems play a very important role in the generation of essential information concerning oceanographic parameters for mapping and monitoring purposes of the oceans and hence become ideal tools for ocean observation due to the ability to cover huge and distant areas (Moradi et al., 2016). Previous studies have demonstrated that the integration of remotely sensed environmental data with fishery data can effectively assess the spatial variation of pelagic species such as tuna (Su et al., 2008; Chang et al., 2013). Statistical modeling approaches such as Generalized Additive Model (GAM) are widely used to assess the complex non-linear relationships between oceanographic variables and the species abundance. Additionally, Empirical cumulative distribution function (ECDF) has been used to characterize environmental variables and quantify the probability of species presence throughout the different oceanographic conditions (Druon et al., 2011; Hidalgo et al., 2016).

Indian Ocean is seasonally influenced by the monsoonal cycle and large-scale Indian Ocean dipole, because of these incidents the oceanographic parameters are varying temporally and spatially (Schott, Friedrich A. et al., 2009). But there is a gap of comprehensive studies focused on skipjack tuna in the Indian Ocean considering the longline catch data and effort data from Sri Lankan fisheries. Understanding environmental dynamics affecting the distribution and abundance of skipjack tuna is necessary to improve the fisheries management and to ensure the sustainable exploitation of these resources.
Globally, skipjack tuna is the most commercially important tuna species, with global catches exceeding 3 million metric tonnes annually, representing approximately 60% of all tuna landings (FAO, 2022). The Indian Ocean accounts for roughly 20–25% of global skipjack catches, with key fishing nations including Sri Lanka, Maldives, Indonesia, and distant-water fleets from Japan and South Korea (IOTC, 2022). In Sri Lanka, skipjack tuna is a commercially significant species captured primarily using pole-and-line and purse seine methods in coastal and offshore waters, while longline vessels targeting yellowfin and bigeye tuna also record incidental skipjack catches in offshore and high-seas operations (Amarasiri and Joseph, 1986). The present study utilizes longline catch data from the Department of Fisheries and Aquatic Resources (DFAR) of Sri Lanka, in which skipjack tuna CPUE is used as a relative abundance index reflecting spatial and temporal distribution patterns rather than absolute fishing efficiency.

The objective of this study is to identify the fishing grounds of skipjack tuna (*Katsuwonus pelamis*) and to model their spatiotemporal distribution in the Indian Ocean using remotely sensed data and fishery dependent data from longline catch data from Sri Lanka. We hypothesize that the catch distribution of skipjack tuna is associated with the oceanographic factors. The Generalized Additive

Model was employed to identify the key predictors of catch rates and, while the Empirical Cumulative Distribution Function approach was used to define the environmental preferences and the potential fishing grounds. Furthermore, this study supports data driven strategies for sustainable fisheries management and improves the understanding of tuna environment preferences in the Indian Ocean.

METHODS

STUDY AREA

The geographical area extending from the latitudes 30°S- 20°N and the longitudes 50°E– 90°E, was selected as the study area (Figure 1).

The study area encompasses the waters surrounding Sri Lanka and extends into the central and northern Indian Ocean. Sri Lanka is situated at the heart of this region, approximately 50 to 300 nautical miles from the nearest offshore fishing grounds. The Indian Ocean is strongly influenced by the monsoon cycle, with two dominant seasons: the Southwest (Summer) Monsoon (May–October) and the Northeast (Winter) Monsoon (November–April). These seasonal wind reversals drive major oceanographic processes including upwelling along the southwestern and northeastern coasts of Sri Lanka, mesoscale eddy formation, and significant variability in sea surface temperature, chlorophyll concentration, and mixed layer depth across the study area. These monsoon-driven dynamics play a critical role in shaping the availability of prey and the spatial distribution of skipjack tuna within the defined study domain.

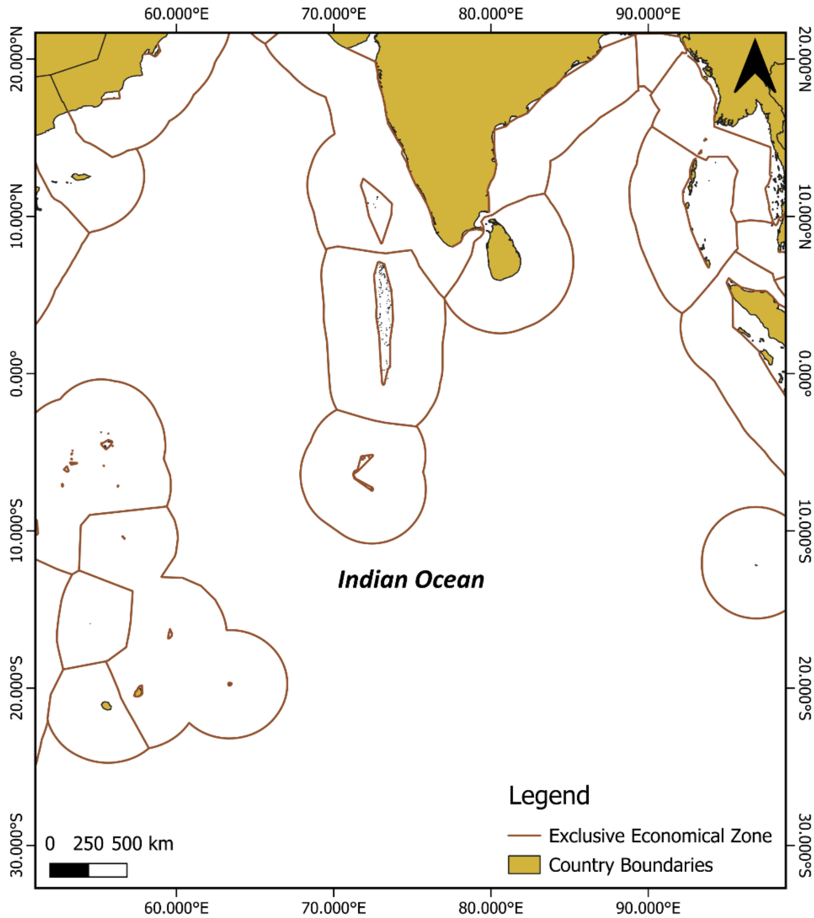

Figure 1: The study area of the research (latitudes 30 ° S- 20 °N and the longitudes 50°E - 90°E)

Fishery data processing

Longline catch data and effort data of skipjack tuna from offshore and high seas boats were collected from the Department of Fisheries and Aquatic Resources (DFAR) In Sri Lanka for the study period of January 2019 to December 2023. The dataset was derived from official logbook records submitted by offshore and high-seas vessels, which are routinely collected and validated by DFAR as part of the national fisheries monitoring program. Each record included information on date of operation, fishing location, catch quantity, fishing effort. These catch and effort data were standardized by calculating catch per unit effort (CPUE – Catch per 1000 hooks were considered). Then, the average CPUE for each month for the study area was calculated.

It is acknowledged that skipjack tuna is primarily targeted commercially using purse seine and pole-and-line methods globally. However, in Sri Lanka's offshore and high-seas longline fishery, skipjack tuna is regularly recorded as bycatch alongside primary target species such as yellowfin tuna (*Thunnus albacares*) and bigeye tuna (*Thunnus obesus*). The CPUE derived from longline logbook records is therefore interpreted as a relative abundance index, reflecting spatial and temporal patterns in skipjack tuna occurrence rather than absolute catch efficiency. This approach is consistent with previous studies that have used longline CPUE as a proxy for skipjack tuna abundance in the Indian Ocean (Rajapaksha et al., 2013; Sisira et al., 2023).

Prior to analysis, the fishery dataset was screened to remove duplicates, records with missing environmental matches, and outliers. Zero catch records were retained to avoid bias in CPUE standardization.

Environmental data processing

Monthly average SST, SSC, SSH, SSS, MLD, and EKE data for the study period were extracted from Copernicus Marine Service (**Product URL**) for the study area with ¼° spatial resolution and one-month temporal resolution.

Environmental data were spatially and temporally matched to the fishery records by extracting monthly mean values corresponding to the fishing locations and time.

Statistical analysis

Satellite data and filtered fisheries catch data were combined into a csv file compatible with MS Excel. Python was used for all statistical data analysis following the methodology described by Rajapaksha et al. (2013). The Generalized Additive Model (GAM) was utilized to obtain smooth functions of linear models between catch rates and oceanographic characteristics, as well as to demonstrate the nonlinear relationship between covariates (independent variables/ oceanographic parameters) and target variables (CPUE). The Empirical Cumulative Distribution Function (ECDF) model was used to estimate the probabilistic distribution of data. These functions determine the relationship between ocean factors and catch rates and optimize them.

In Table 1, $P(X^2)$ refers to the chi-square p-value derived from the significance test of each GAM smooth term. For each smooth function s(x) in the GAM, an approximate chi-square statistic is computed to test the null hypothesis that the smooth term has no significant effect on CPUE. A p-value of less than 0.05 indicates that the smooth term is statistically significant, confirming a meaningful non-linear relationship between the respective oceanographic variable and the catch rates of skipjack tuna.

Histogram analysis was used to identify the frequency distribution. The ECDF was used to calculate the percentage of sample observations that are less than or equal to a specific value, given a set number of observations (n). ECDF is a step function that increases $\frac{1}{n}$ at each data point. The ECDF estimates the cumulative distribution function of the generated sample points (Pernot and Savin, 2018).

$$(t) = \frac{1}{n}\sum_{i=1}^{n} l(x_i) \qquad \text{(1)}$$

With the indication function,

$$l(x_i) = \begin{cases} 1, & If\ x_i \le t; \\ 0, & Otherwise \end{cases}$$

$$f(t) = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\bar{y}}l(x_i) \quad \text{———} \quad \textbf{(2)}$$

$$D(t) = max|f(t) - g(t)| \quad \text{———} \quad \textbf{(3)}$$

In the equation, *n* denotes the number of fishing days, *xi* represents the first measurement of an ocean parameter, and *t* indicates the sequence of observations from lowest to highest value of the oceanographic variable. *Y* represents the catch weight on the fishing day *i*, while $\bar{y}$ shows the estimated mean catch weight across all fishing days.

The link between ocean variables and catch rates is strongest when capture weights exceed the predicted mean catchweight, and vice versa. The Kolmogorov-Smirnov tests can measure the strength of a relationship by comparing the difference between two curves, $\boldsymbol{f(t)}$ and $\boldsymbol{g(t)}$, at any point. The largest value of $\boldsymbol{D(t)}$ indicates the oceanographic variables that yield the highest catch weight value (Andrade and Garcia, 1999).

The Generalized Additive model was used to analyze the association between the skipjack tuna's CPUE and six oceanographic factors. Developing a model for the association between CPUE and ocean factors is challenging due to noisy, distributed data from fishing sources (Hanselman et al., 2015). The GAM model offers greater flexibility in determining the optimal fit between catch rates and ocean factors. The relationship was predicted to be non-linear. Following the shape of relationships. The GAM model structure can be expressed as;

$$g(E(y)) = a + s_1(x_1) + \cdots + \cdots + s_p(x_p) \quad \text{———} \quad \textbf{(4)}$$

In this equation, *Y* represents the dependent variable, *E(Y)* is the expected value, and *g(Y)* is the link function that connects the expected value to the predictor variables (*x1,… xp*). $\boldsymbol{s1\ (x1)}$ represent smooth, non-parametric functions. Using the GAM function, the shapes derived from the additive model between the predictor variable (ocean variable) and the responder variable (CPUE of fish species) were parameterized in the Generalized Linear Model. The piecewise GLM function was used to accurately replicate the forms of the GAM model's functions.

$$ln(CPUE) = a + s(SST) + s(SSH) + s(SSC) + s(SSS) + s(MLD) + s(EKE) + e \quad \text{———} \quad \textbf{(5)}$$

a represents a constant, *s()* smoothed the variables (SST, SSH, SSC, EKE, SSS, MLD), and e is a random error. The CPUE's asymmetric frequency distribution was normalized using a logarithmic technique. The GAM model, a non-parametric alternative to multiple linear regressions, is less reliant on data distribution assumptions (Hastie and Tibshirani, 1986).

The GAM function identified a nonlinear link between predictor variables and CPUE. Three sorts of scores (derived coefficients) are utilized to explain the link between oceanographic variables.
The following metrics were used: GCV score, percentage of explained deviation, and Akaike's information criteria. The GAM model created best-fitting models for individual parameters in the study, as well as a combined model to enhance fitting percentage (deviance explained). A model's fit to the original data improves as the explained deviation increases (Gomez-Rubio, 2018).
The Generalized Cross-Validation score (GCV) evaluates the fit of the GAM model. The GCV score indicates the degree of smoothness of the function. Smoothing functions minimize the model's prediction error (Wood, 2017). The best-fitted model is the one with the lowest GCV score and the maximum explained variation (Gomez-Rubio, 2018). Akaike's Information Criteria (AIC) evaluates a model's fit to the original data set. Using AIC values, we may choose the best-fitting model based on parameter combinations and CPUE for the chosen fish species. Lower AIC values indicate greater model fit (Gomez-Rubio, 2018).

RESULTS

The histogram analysis of the skipjack tuna fishing operations considering the satellite derived environmental parameters showed that there were specific ranges of oceanographic conditions where they tend to be aggregated (Figure 2).

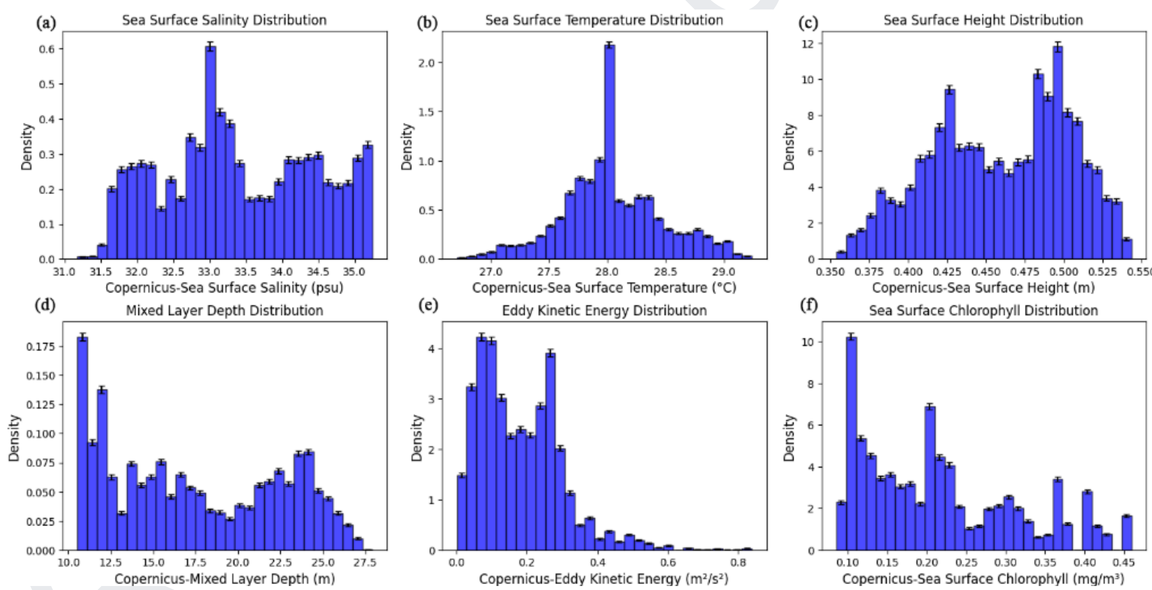

Figure 2: Histograms showing the frequencies of skipjack tuna caught in (a) Sea Surface salinity (SSS), (b) Sea Surface temperature (SST), (c) Sea Surface Height (SSH), (d) Mixed Layer Depth (MLD), (e) Eddy Kinetic Energy (EKE), (f) Sea Surface chlorophyll (SSC)

According to the histogram analysis results (Figure 2), the sea surface salinity varied from 31.0 – 35.5 PSU. Most of the catches were concentrated within the 32.5 - 35 PSU of SSS. The fishing activities were done in areas of SST ranging from 26.5°C – 29.5°C. However, the highest frequency of fishing activities was observed between 27.5°C –28.5 °C. SSH showed variations from 0.35 m – 0.55 m. However, the most catches were recorded in the range of 0.400 m– 0.525 m. The variation of MLD was from 10 m to 27.5m, while most of the catches were recorded between 10 m– 12.5 m of MLD. The EKE showed variations similar to SSC and SSH. However, the highest frequency of catches was recorded between 0.0 m^2s^{-2} – 0.4 m^2s^{-2} of EKE. The SSC varied from 0.0 mgm^{-3}– 0.50 mgm^{-3}. Most catches were concentrated between 0.1 mgm^{-3}– 0.25 mgm^{-3}.

The logline fleet operate throughout the Indian Ocean, with high fishing effort in tropical areas of the central Indian Ocean between 15 °N and 15 °S. High catch rates of skipjack tuna occurred in the equatorial region for all seasons, whereas the catch rates were lower in North and South Indian Ocean. For the GAM analysis, the residuals from the lognormal distribution confirmed that their assumptions were fulfilled based on the quantile-quantile plots: the log-spaced residuals appeared normally distributed except lower and upper extremes (Figure 3) and independent of the values for all the environmental factors used in the study (Figure 3).

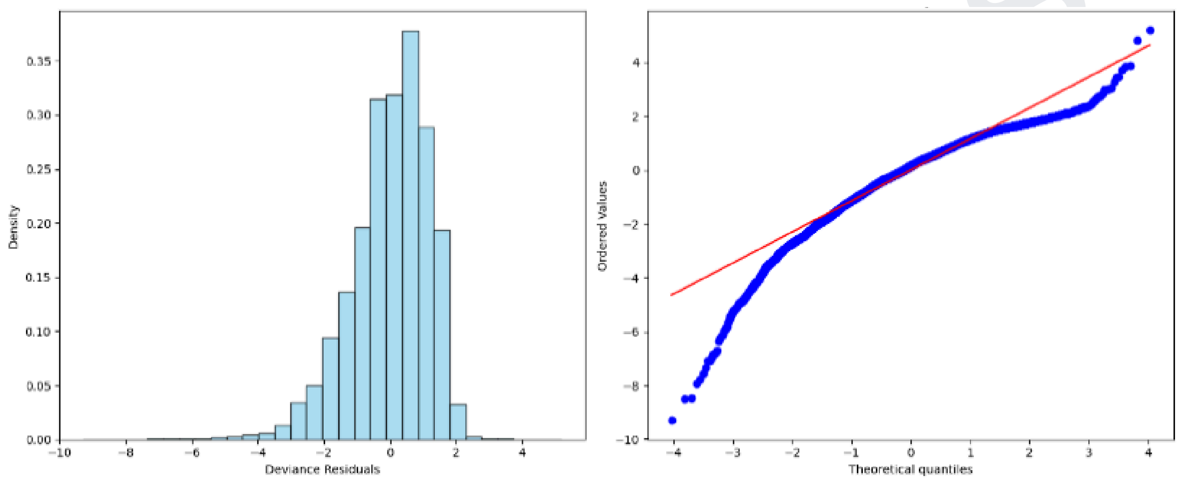

Figure 3: Residual distributions and quantile–quantile plots for diagnostic analysis of the final generalized additive model (GAM) with prediction variables (presented in Table 1) based on remote sensing and catch data.

Residual plots were examined to study the assumptions of homoscedasticity and independence. In several predictor variables, non-random patterns were observed. The residual plot of EKE clearly shows a heteroscedasticity. Similarly, SSC and MLD exhibit patterns that deviate from the expected homoscedasticity, with noticeable clustering and spread inconsistencies across their value ranges (Figure 4).

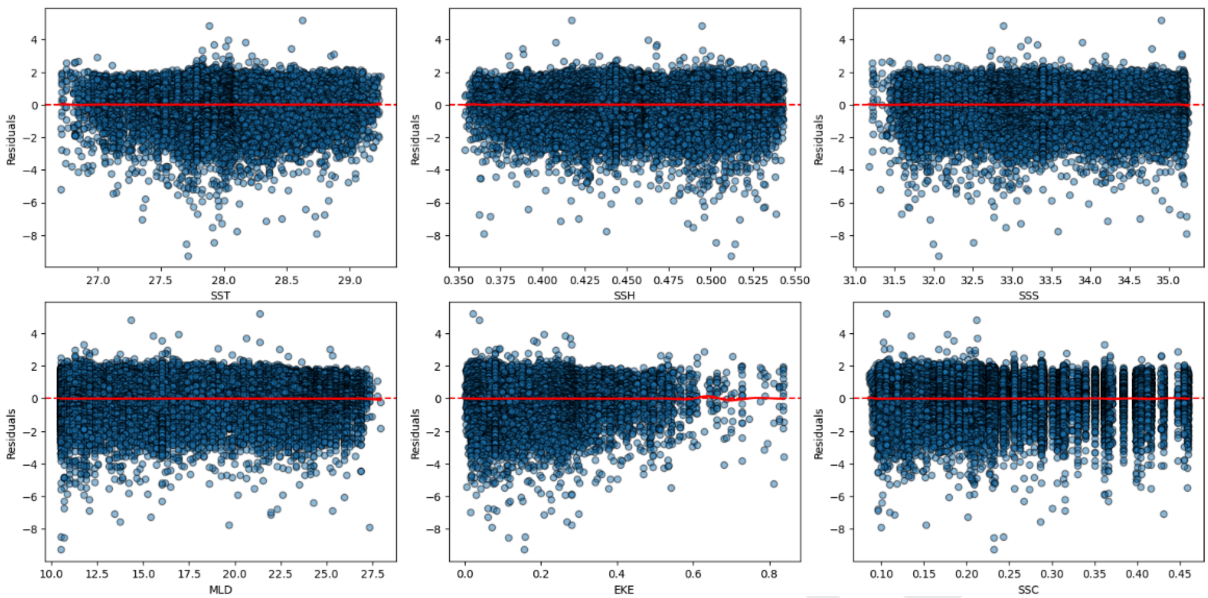

Figure 4: Scatterplots of residuals for environmental factors in the final GAM with variables listed in Table 1 based on fishery and remote sensing data. The red solid lines in each panel represent the pattern of residuals fitted using a spline function for each environmental variable.

Table 1: The Applied Generalized Additive Model results for 12 Models of individual and combined ocean variables with the Catch Per Unit Effort (CPUE) of skipjack tuna in Indian Ocean. Deviance explained%, generalized cross validation score (GCV), Akaike's Information Criteria (AIC) value, and the $P(X^2)$ for the GAMs selected to analyze the fishery data and remote sensing oceanographic variables.

Parameter (Model combination)	**Deviance explained %**	**GCV**	**AIC**	**$P(X^2)$**
SST	0.42%	1.43	80687.38	< 0.05
SSS	0.30%	1.43	80729.41	< 0.05
SSC	0.55%	1.43	80647.00	< 0.05
SSH	0.35%	1.43	80711.76	< 0.05
EKE	1.31%	1.42	80395.15	< 0.05
MLD	0.48%	1.43	80670.58	< 0.05
SST+SSC	1.46%	1.42	80393.38	< 0.05
SST+EKE	1.77%	1.41	80288.96	< 0.05
SST+SSH	1.03%	1.42	80533.61	< 0.05
SST+MLD	1.15%	1.42	80495.50	< 0.05
SST+SSS	1.05%	1.42	80527.97	< 0.05
SST+SSS+SSC+SSH+EKE+MLD	3.49%	1.40	79896.02	< 0.05

The GAM analysis results were obtained for the model combinations of the ocean variables and the results in presented in Figure 5. The individual models were derived for each parameter and then the combinations with more than one parameter. The graphical representation by the following GAM plots with individual predictor variables conveys the most favorable condition of each parameter for the skipjack tuna.

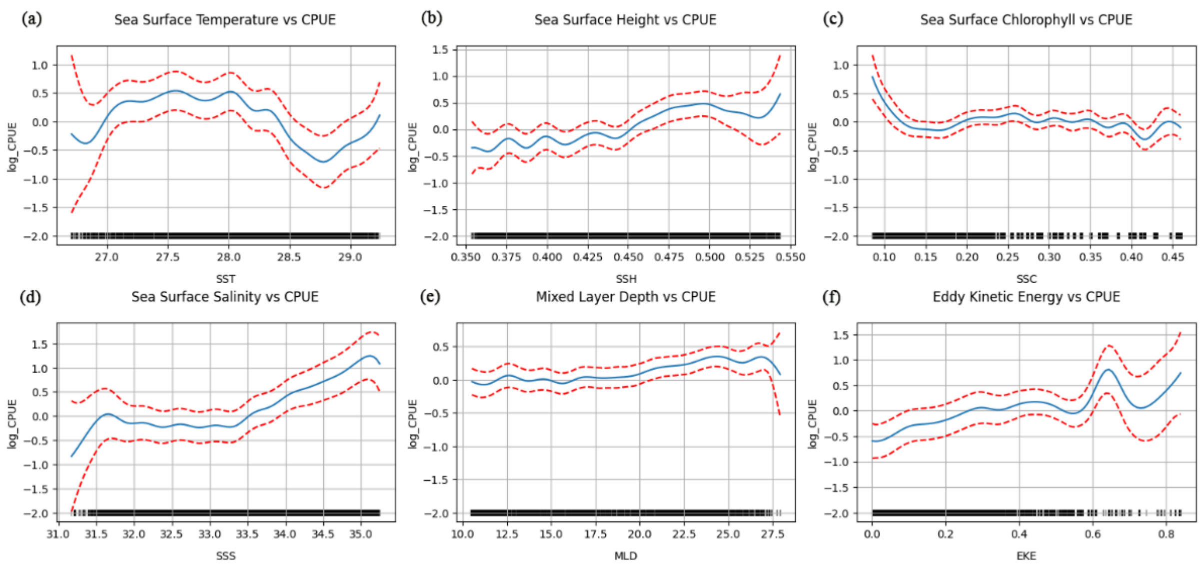

Figure 5: Generalized Additive Model (GAM) derived effect of oceanographic variables (a) Sea surface temperature (SST), (b) Sea surface height (SSH), (c) Sea surface chlorophyll (SSC), (d) Sea surface salinity (SSS), (e) Mixed layer depth (MLD) (f) Eddy kinetic energy (EKE) on skipjack tuna catch rates. The rug line/marks at the abscissa axis represents the observed data points. The solid line describes the function, and the dashed line indicates 95% confidence interval, equivalent to two standard deviations (± 2S.D.)

The GAM-fitted model with SST and the CPUE (Figure 5 a) showed that the most favorable SST range was 27.0°C – 29.0°C for the skipjack tuna. The SSC and CPUE model (Figure 5 c) indicated that most of the chlorophyll values were concentrated within the 0.10 mgm^{-3} – 0.25 mgm^{-3} range. The SSH and the CPUE model (Figure 5 b) showed that most of the sea surface height values were concentrated within the 0.35 m– 0.55 m range. The EKE and the CPUE model (Figure 5 f) showed that most of the eddy kinetic energy values were concentrated within the 0.0 m^2s^{-2} – 0.6 m^2s^{-2} range. The SSS and the CPUE model (Figure 5 d) showed that most of the salinity values were within the 31.5 – 35.25 PSU range. The MLD and the CPUE model (Figure 5 e) showed that most of the mixed layer depth values were within the 10.2 m– 27.5 m range. Table 1 shows the relevant Deviance explained (%), GCV score, AIC score, and the p values derived from the GAM models which were applied for CPUE with the ocean variables for skipjack tuna.

The derivation of the GAM model with the combination of the CPUE and the individual ocean variables showed that the lowest GCV score, and the AIC values resulted from the GAM-derived EKE model with the highest percentage deviation (deviation explained) of 1.31%. The next highest percentage deviations were shown by the GAM-derived SSC model (0.55%) and the GAM-derived MLD model (0.48%) in order. Then the GAM model results were obtained for combined variables. The SST was kept as the main variable, and the other variables were coupled with SST. The lowest GCV score, and the AIC value were obtained from the SST-EKE model which showed the highest percentage deviation of 1.77%. Finally, the GAM model results were obtained for all six parameters, and this showed the lowest GCV score and the AIC value with the highest percentage deviation of 3.49%. The six parameters were combined for GAM derivation to increase the accuracy and form the best-fit model for the original data. The model derivation with all six parameters was the best-fit model for the skipjack tuna according to Table 1.

The relationship between the oceanographic variables and the catch rates (CPUE) of skipjack tuna was derived using the ECDF models (Figure 6). The cumulative distribution curves of the six variables were different and the degrees of differences (D) between the two curves (catch weighted (g(t)) and the ocean variable/reference variable (f(t)). The optimum values were obtained from the highest region of significant difference (degree of difference) between the curves.

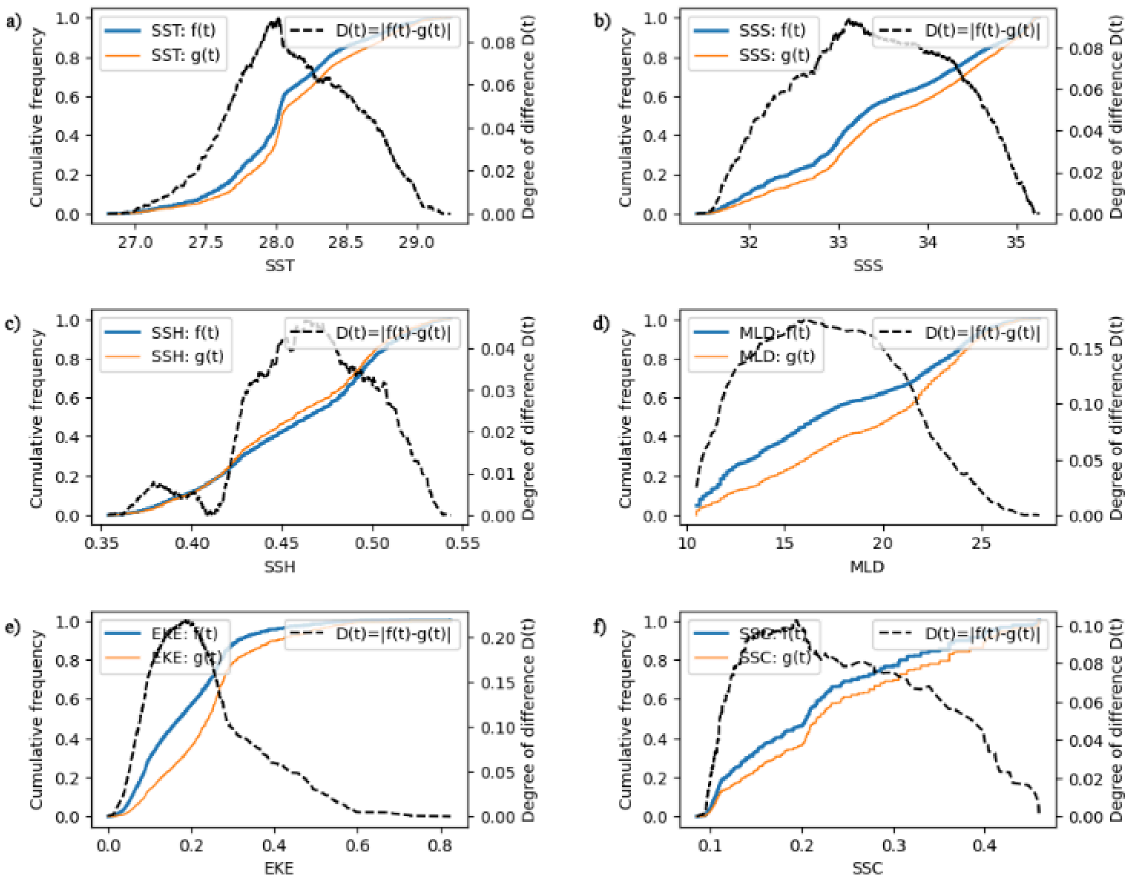

Figure 6: Empirical Cumulative Distribution Function (ECDF) frequencies for the (a) Sea surface temperature (SST), (b) Sea surface salinity (SSS), (c) Sea surface height (SSH), (d) Mixed layer depth (MLD), (e) Eddy kinetic energy (EKE), (f) Sea surface chlorophyll (SSC) as weighted by skipjack tuna caught during 2019 to 2023.

According to the Figure 6 and Table 2, SST values were within the 27.78°C - 28.29 °C and SSC values were within 0.13 mgm^{-3} - 0.29 mgm^{-3} range. The SSH values were within the 0.42 m- 0.50 m range. The EKE values were within 0.09 m^2s^{-2} - 0.26 m^2s^{-2} range. SSS variables showed that salinity values were between 32.62 PSU and 34.27 PSU. The MLD values were between 12.36 m and 22.28 m. Table 2 shows the summary of the ECDF model results.

Table 2: The model results of the Empirical cumulative distribution curve (ECDF) for the six oceanographic parameters including the optimum range and the most favorable value of the oceanographic parameter. (Maximum degree of difference value considered, significant at $p<0.05$).

Parameter	Optimum range	The most favorable value of the ocean parameter (Highest degree of difference)
SST	27.78 - 28.29 °C	28.22 °C
SSC	0.13 - 0.29 mgm^{-3}	0.24 mgm^{-3}
SSH	0.42 - 0.50 m	0.47 m
EKE	0.09 - 0.26 m^2s^{-2}	0.21 m^2s^{-2}
SSS	32.62 - 34.27 PSU	33.31 PSU
MLD	12.36 - 22.28 m	21.83 m

Figure 7 shows the spatial distribution of both the GAM predicted and nominal CPUE values pf skipjack tuna in the Indian Ocean using the four quarters of the year. Overall, both the model and observed data indicate higher CPUE values concentrated around the equatorial region (approximately 15 °N and 15 °S).

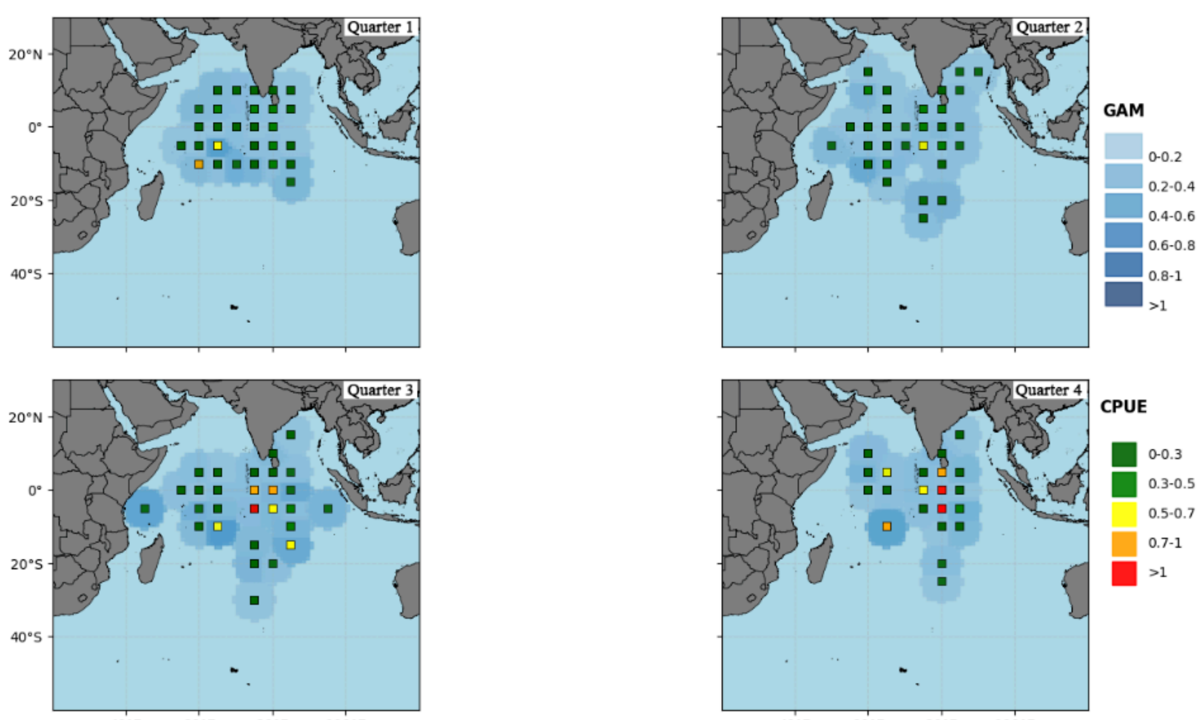

Figure 7: Distributions of observed nominal catch rates of skipjack tuna for 2019–2023 overlapping on

the maps of predicted relative densities using the final GAM with variables listed in Table 1 based on fishery and satellite-based remotely sensed environmental data.

DISCUSSION

The oceanographic variables of Sea Surface Temperature (SST), Sea Surface Chlorophyll (SSC), and Eddy Kinetic Energy (EKE) have statistically significant influence on skipjack tuna distribution and abundance in the Indian Ocean. According to the GAMs results the combined effect of all six environmental factors providing the most appropriate explanation of spatial distribution pattern of skipjack tuna.

Direct comparison of fishery and oceanographic parameters

Our findings on the optimal temperature range from 26.5 °C to 29.5 °C for skipjack tuna distribution and abundance in the Indian Ocean reveals the important biogeographical patterns that align with the previous studies in other oceanic regions. In the western Pacific Ocean, the reported optimal temperature was 29 °C to 30 °C for skipjack tuna distribution and abundance (Hsu et al., 2021). This SST range corresponds to the optimum SST conditions for skipjack tuna in tropical waters. These values align with the previous studies highlighting the species' preferences for warmer temperatures (Dueri and Maury, 2013). The findings showed that the SSC ranged between 0.00 mgm^{-3} and 0.50 mgm^{-3}, and the fishing activities were concentrated in that range. The highest fishing activities were observed in the locations where SSC was recorded between 0.10 mgm^{-3} to 0.25 mgm^{-3}. These findings suggest that skipjack tuna prefers waters with moderate levels of chlorophyll (Kumar et al., 2014). Fishing activities were observed in locations with SSH ranging from 0.35 m to 0.55 m, with the highest concentrations of fishing activities occurring between 0.425 m and 0.525 m. With the southwest monsoon season (May to October), these conditions frequently coincide. The highest SSH caused by stronger ocean currents and wind patterns (Amarasiri and Joseph, 1986). The high SSH during this period promotes skipjack tuna aggregation, increased nutrient mixing, and influx of cooler deeper water to the top, which increases the prey availability of skipjack tuna. The data suggest that the fishing activities for skipjack tuna occur when high SSH, moderate SST, and moderate levels of chlorophyll concentrations. These values and conditions align with the previous studies highlighting the species' preferences for higher SSH (Schott, Friedrich A. et al., 2009) The association between EKE and the skipjack tuna abundance and distribution found that the fishing activities were highly focused where the EKE values ranged between 0.00 m^2s^{-2} to 0.9 m^2s^{-2}. EKE is a critical component in nutrient mixing, providing favorable conditions for primary productivity and thereby prey availability of skipjack tuna. EKE is highest during the southwest monsoon period and it drives vertical nutrient transport, and increases the food availability for pelagic species such as skipjack tuna (Kumar, 2011). Maximum fishing activities were recorded with moderate EKE values aligning with optimum SST and SSC ranges. These findings highlighted the importance of observing EKE to predict the skipjack tuna fishing zones and to optimize fishing efforts.

The SSS ranged from 31.0 PSU to 35.5 PSU, and the skipjack tuna fishing activities were recorded. The most fishing activities were concentrated between the SSS values of 32.5 PSU and 33.5 PSU. The SSS has a significant impact on the vertical stratification of the water column. It has a greater impact on nutrient distribution. High salinity levels can affect nutrients and prey availability (Sisira et al., 2023). The seasonal variabilities of skipjack tuna against SSS were observed during the southwest monsoon. Because of the monsoonal rain and freshwater input can lead to a temporary decrease in SSS, which can influence tuna distribution (Sisira et al., 2023).

The MLD range varied between 10.0 m– 27.5 m in the study area with a higher concentration of fishing activities observed the MLD values between 10.0 m– 12.5 m. This MLD is important in the vertical distribution of nutrients and the temperature in the ocean, and it affects the availability of skipjack tuna's prey (Borja et al., 2019). It is a well-known phenomenon, that supports to increase in the nutrients in upwelling and primary productivity (Yapa, 2009).

The Indian Ocean undergoes the strongest monsoon during the southwest monsoon season (Summer monsoon). The Summer Monsoon Current (SMC) shows the movement from the western Indian Ocean to the Eastern Indian Ocean and this ocean current transports high-saline water from the Arabian Sea to the eastern Indian Ocean including the Bay of Bengal (BOB) (De Vos et al., 2014). During the Northeast monsoon (winter monsoon), the North East Monsoon Current (NMC) shows a flow direction reverse to the SMC. NMC transports low-saline water from the Bay of Bengal to the western Indian Ocean (Schott, Friedrich A. et al., 2009). During South west monsoon period the SMC directly hits the western coast of Sri Lanka and passes to the east through the southern coastal belt. After passing the Southern coastal belt, the current shows an upward motion and divides into two main current branches as the East Indian Coastal Current (EICC) and the Sri Lankan Dome (SLD). During this period the catch rates of relatively high in the southern region below to Sri Lanka skipjack tuna are relatively high in the southern region below to Sri Lanka (Upeksha Swarnamalie et al., 2021).

Integrating histogram, ECDF, and GAM results

The results from histogram analysis, Generalized Additive Model (GAM), and Empirical Cumulative Distribution Function (ECDF) models summarized the most favorable conditions for skipjack tuna. The histogram analysis was done to find the parameter range in the highest catch frequency of the fish species observed. The GAM model analysis was done to investigate the best-fit ocean parameter model that fits the catch rates of the fish. The ECDF model analysis was done to investigate the most optimum condition of oceanographic parameters that have the highest probability of the catches of skipjack tuna.

For the skipjack tuna, the GAM functional analysis showed that Eddy Kinetic Energy (EKE) was the major parameter that affects the distribution of the skipjack tuna. The Optimum range was observed as 0.0 m^2s^{-2} – 0.6 m^2s^{-2} (Figure 5). The combined model analysis showed that the Sea Surface Temperature (SST) showed a better fit than the individual EKE model. The EKE/SST model showed the highest deviance explained by the percentage of 1.77% with the lowest Generalized Cross-Validation (GCV) score and the Akaike's Information Criteria (AIC). The six-parameter combined GAM model showed the highest deviance, explaining the percentage of 3.49%, expressing that all the parameters together make the best-fit model for the survival and distribution of skipjack tuna. All the models were significant at the 0.05 confidence level, depicting that there is a significant relationship between the catch rates and the ocean variables (Table 1).

The ECDF model results proved the GAM model results of the optimum parameter ranges for the skipjack tuna. The cumulative distribution curves of the six ocean variables were different, and the degrees of difference between the two curves were highly significant ($p < 0.05$) (Figure 6). The highest association was shown by the EKE and SST with the CPUE. The other parameters also showed a statistically significant association with the CPUE by expressing the preferable or optimum conditions for the survival of skipjack tuna, i.e., 0.09 m^2s^{-2} - 0.26 m^2s^{-2} (EKE), 27.78°C - 28.29°C (SST), 0.13 mgm^{-3} - 0.29 mgm^{-3} (SSC), 0.42 m- 0.50 m (SSH), 32.62 PSU- 34.27 PSU (SSS), and 12.36 m- 22.28 m (MLD). The highest association of the six ocean variables occurred at 0.21 m^2s^{-2} of EKE, 28.22 °C of SST, 0.24 mgm^{-3} of SSC, 0.47 m of SSH, 33.31 PSU of SSS, and 21.83 m of MLD, depicting that most prominent and favorable oceanographic conditions for the skipjack tuna (Table 2). The catch rates show a

SciELO Preprints - This document is a preprint and its current status is available at: https://doi.org/10.1590/2675-2824074.2512

decreasing trend outside these center values. This decreasing trend in catch rates outside the optimal ranges of the environmental variables suggests that the skipjack tuna is highly sensitive to specific environmental conditions. The identified most favorable values represent the most favorable habitat conditions for their feeding, migration, and spawning behavior.

CONCLUSION

The outcomes and predictions of the models, such as the spatial distribution of skipjack tuna and catch rates are strongly related by the environmental variables, particularly Sea Surface Temperature (SST), Eddy Kinetic Energy (EKE), and Sea Surface Chlorophyll (SSC). The GAM and ECDF models consistently identified the specific optimal ranges of these environmental parameters, indicating that skipjack tuna aggregate in highly defined oceanic conditions.

This study indicates that skipjack tuna exhibit a strong preference for moderate SST and SSC levels coupled with low to moderate EKE, suggesting that areas characterized by balanced thermal conditions and enhanced but stable productivity provide favorable environments. The strong relationship between EKE and CPUE emphasizes the role of mesoscale eddies in conditioning pelagic habitats through nutrient redistribution and prey aggregation.

Although the GAM model explained a modest proportion of CPUE variability, this result is consistent with the inherently complex and multi factorial nature of fisheries catch data. Where fishing behavior, gear characteristics, and unobserved ecological mechanisms also influence catch rates.

The observed temporal variability in skipjack tuna CPUE, particularly the concentration of higher catch rates in equatorial waters during specific quarters, depicts the strong influence of monsoonal circulation on oceanographic conditions in the Indian Ocean. These outcomes underscore the importance of considering large scale climatic variabilities when interpreting tuna distribution patterns.

Future research should aim to incorporate additional drivers such as prey dynamics, fishing effort behavior, and large scale climate indices (e.g., Indian Ocean Dipole and ENSO) to further improve the model performance.

DATA AVAILABILITY STATEMENT

The data used in our manuscript is available online. The compilation of bibliographic data carried out in this article is available at the Supplementary Material.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGMENTS

We thank the valuable comments and suggestions from the three anonymous reviewers and the editor, as well as the team members of the Ocean and Coastal Research (OCR), the Department of Fisheries and Aquatic Resources (DFAR) of Sri Lanka, and the Copernicus marine services for their assistance in data preparation.

AUTHOR CONTRIBUTION

All authors contributed to this study;
T.I.: Conceptualization; Writing – original draft; Methodology; Software; Formal Analysis;
K.A.B.: Supervision; Writing – review & editing;
S.S.H.: Supervision; Writing – review & editing.
All authors have read and agreed to the published version of the manuscript.

AI USE DISCLOSURE

The authors declare that no generative artificial intelligence (AI) tools were used in the preparation, writing, or editing of this manuscript.

REFERENCES

Amarasiri, C. & Joseph, L. 1986. Skipjack tuna (K. pelamis) - Aspects of the biology and fishery from the western and southern coastal waters of Sri Lanka. *In*: *FAO Indo-Pacific Tuna Development and Management Programme. Collective volume of working documents presented at the Expert Consultation on Stock Assessment of Tunas in the Indian Ocean. 4-8 December 1986*.

Andrade, H. A. & Garcia, C. A. E. 1999. Skipjack tuna fishery in relation to sea surface temperature off the southern Brazilian coast. *Fisheries Oceanography*, *8*(4). DOI: https://doi.org/10.1046/j.1365-2419.1999.00107.x

Borja, A., Amouroux, D., Anschutz, P., Gómez-Gesteira, M., Uyarra, M. C. & Valdés, L. 2019. The Bay of Biscay. *World Seas: An Environmental Evaluation Volume I: Europe, the Americas and West Africa*, 113–152. DOI: https://doi.org/10.1016/B978-0-12-805068-2.00006-1

Chang, Y. J., Sun, C. L., Chen, Y., Yeh, S. Z., Dinardo, G. & Su, N. J. 2013. Modelling the impacts of environmental variation on the habitat suitability of swordfish, Xiphias gladius, in the equatorial Atlantic Ocean. *ICES Journal of Marine Science*, *70*(5). DOI: https://doi.org/10.1093/icesjms/fss190

Druon, J. N., Fromentin, J. M., Aulanier, F. & Heikkonen, J. 2011. Potential feeding and spawning habitats of Atlantic bluefin tuna in the Mediterranean Sea. *Marine Ecology Progress Series*, *439*. DOI: https://doi.org/10.3354/meps09321

Dueri, S. & Maury, O. 2013. Modelling the effect of marine protected areas on the population of Skipjack tuna in the Indian Ocean. *Aquatic Living Resources*, *26*(2). DOI: https://doi.org/10.1051/alr/2012032

Gomez-Rubio, V. 2018. Generalized Additive Models: An Introduction with R (2nd Edition) . *Journal of Statistical Software*, *86*(Book Review 1). DOI: https://doi.org/10.18637/jss.v086.b01

Grande, M., Murua, H., Zudaire, I., Goñi, N. & Bodin, N. 2014. Reproductive timing and reproductive capacity of the Skipjack tuna (Katsuwonus pelamis) in the western Indian Ocean. *Fisheries Research*, *156*. DOI: https://doi.org/10.1016/j.fishres.2014.04.011

Hanselman, D. H., Lunsford, C. R. & Rodgveller, C. J. 2015. Assessment of the sablefish stock in Alaska. *Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska*, *2014*(December 2015).

Hastie, T. & Tibshirani, R. 1986. Generalized additive models. *Statistical Science*, *1*(3). DOI: https://doi.org/10.1214/ss/1177013604

Hidalgo, M., Secor, D. H. & Browman, H. I. 2016. Observing and managing seascapes: Linking synoptic oceanography, ecological processes, and geospatial modelling. *ICES Journal of Marine Science*, *73*(7). DOI: https://doi.org/10.1093/icesjms/fsw079

Hsu, T. Y., Chang, Y., Lee, M. A., Wu, R. F. & Hsiao, S. C. 2021. Predicting Skipjack tuna fishing grounds in the western and central pacific ocean based on high‐spatial–temporal‐resolution satellite data. *Remote Sensing*, *13*(5). DOI: https://doi.org/10.3390/rs13050861

Rajapaksha, J., Surasinghe, L. & Gunathilaka, A. A. J. K. (2013).. Environmental preferences of yellowfin tuna in the north east Indian Ocean: an application of satellite data to longlien catches. *International Journal of Fisheries and Aquatic Sciences*, *2*(4).

Kumar, P. 2011. *Sea Level and Eddy Kinetic Energy Variability in the Bay of Bengal Using Satellite Altimetry*. Retrieved from www.jason.noaa.gov.org

Kumar, P. S., Pillai, G. N. & Manjusha, U. 2014. El Nino Southern Oscillation (ENSO) impact on tuna fisheries in Indian Ocean. *SpringerPlus*, *3*(1). DOI: https://doi.org/10.1186/2193-1801-3-591

Moradi, N., Hasanlou, M. & Saadatseresht, M. 2016. OCEAN COLOR RETRIEVAL USING LANDSAT-8 IMAGERY IN COASTAL CASE 2 WATERS (CASE STUDY PERSIAN AND OMAN GULF). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLI-B8*. DOI: https://doi.org/10.5194/isprs-archives-xli-b8-1161-2016

Pernot, P. & Savin, A. 2018. Probabilistic performance estimators for computational chemistry methods: The empirical cumulative distribution function of absolute errors. *Journal of Chemical Physics*, *148*(24). DOI: https://doi.org/10.1063/1.5016248

Schott, Friedrich A., Xie, Shang Ping & McCreary, Julian P. 2009. Indian ocean circulation and climate variability. *Reviews of Geophysics*. DOI: https://doi.org/10.1029/2007RG000245

Sisira, S., Haputhantri, K. & Gunasekara, S. S. 2023. *Accounting for spatial, temporal and operational effects in the Catch Per Unit Effort standardization of Skipjack tuna in tuna drift gillnet fishery in Sri Lanka*. Retrieved from https://www.researchgate.net/publication/376615148

Stamouli, C., Akel, E. H. K., Azzurro, E., Bakiu, R., Bas, A. A., Bitar, G., Boyaci, Y., Cakalli, M., Corsini-Foka, M., Crocetta, F., Dragičević, B., Dulčić, J., Durucan, F., El Zrelli, R., Erguden, D., Filiz, H., Giardina, F., Giovos, I., Gönülal, O., Hemida, F., Kassar, A., Kondylatos, G., Macali, A., Mancini, E., Ovalis, P., De Mendoza, F. P., Pavičić, M., Rabaoui, L., Rizkalla, S. I., Tiralongo, F., Turan, C., Vrdoljak, D., Yapici, S. & Zenetos, A. 2017. New Mediterranean Biodiversity Records (December 2017). *Mediterranean Marine Science*, *18*(3), 534–556. DOI: https://doi.org/10.12681/mms.15823

Su, N. J., Yeh, S. Z., Sun, C. L., Punt, A. E., Chen, Y. & Wang, S. P. 2008. Standardizing catch and effort data of the Taiwanese distant-water longline fishery in the western and central Pacific Ocean for bigeye tuna, Thunnus obesus. *Fisheries Research*, *90*(1–3). DOI: https://doi.org/10.1016/j.fishres.2007.10.024

Upeksha Swarnamalie, M., Rajapaksha, J. & Gunatilake, J. 2021. Effect of Ocean Currents and Temperature Variability on Yellowfin Tuna Catch Rates of Sri Lankan Longline Fishery. *International Journal of Scientific and Research Publications (IJSRP)*, *11*(8). DOI: https://doi.org/10.29322/ijsrp.11.08.2021.p11662

De Vos, A., Pattiaratchi, C. B. & Wijeratne, E. M. S. 2014. Surface circulation and upwelling patterns around Sri Lanka. *Biogeosciences*, *11*(20). DOI: https://doi.org/10.5194/bg-11-5909-2014

Wood, S. N. 2017. *Generalized additive models: An introduction with R, second edition*. *Generalized Additive Models: An Introduction with R, Second Edition*. DOI: https://doi.org/10.1201/9781315370279

Yapa, K. K. A. S. 2009. Upwelling phenomena in the southern coastal waters of Sri Lanka during southwest monsoon period as seen from MODIS. *Sri Lankan Journal of Physics*, *10*(0). DOI: https://doi.org/10.4038/sljp.v10i0.3831

Zainuddin, M., Kiyofuji, H., Saitoh, K. & Saitoh, S. I. 2006. Using multi-sensor satellite remote sensing and catch data to detect ocean hot spots for albacore (Thunnus alalunga) in the northwestern North Pacific. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *53*(3–4). DOI: https://doi.org/10.1016/j.dsr2.2006.01.007