Indian Journal of Geo Marine Sciences
Vol. 53 (05), May 2024, pp. 353-361
DOI: 10.56042/ijms.v53i05.21663

Unravelling the dynamics of Indian mackerel abundance in the Malabar upwelling region: A comprehensive analysis of oceanographic influences using GLM, GAM, and BRT models

S Jha*, N Jeevanand, S Joseph & T M B Nair

Indian National Centre for Ocean Information Services (INCOIS), Ministry of Earth Sciences, Government of India, Hyderabad, Telangana – 500 090, India

*[E-mail: s.jha-p@incois.gov.in]

Received 30 March 2024; revised 24 May 2024

The Indian mackerel stands as a pivotal small pelagic fish in the southeastern Arabian Sea, particularly along the coast of Kerala. Despite its significant commercial value, endeavours to forecast the abundance and availability of this species in the Indian waters have been notably limited. This study delves into the intricate relationship between diverse oceanographic parameters in the Malabar Upwelling Region (MUR) and the abundance of Indian mackerel. It employs non-parametric statistical models such as the Generalised Linear Model (GLM) and the Generalised Additive Model (GAM), alongside the machine learning methodology using the Boosted Regression Tree (BRT) model, to delineate the relationship between species landings. The response variable and satellite-derived parameters served as predictors. This study utilised 18 years (1995 – 2012) of fish landing data and environmental variables, encompassing rainfall, mixed layer depth, seawater temperature and salinity at 50 m depth; dissolved oxygen; chlorophyll-a and net primary productivity. Smooth terms were used to capture non-linear relationships between fish landing and the predictors. Rainfall, chlorophyll-a, net primary productivity, and dissolved oxygen have statistically significant effects (p-value < 0.05) on fish landing. Seawater temperature at 50 m depth showed a marginally significant effect (0.05 < p-value < 0.1). Mixed Layer Depth (MLD) and salinity at 50 m depth did not statistically affect fish landing (p-value > 0.1).

[**Keywords:** Abundance prediction, Correlation analysis, Indian mackerel, Machine learning model, Malabar upwelling region, Statistical models]

Introduction

India has an Exclusive Economic Zone (EEZ) of more than 2 million square kilometres rich in marine biodiversity that plays a crucial role in providing food security and employment opportunities leading to poverty alleviation for millions of people who rely on ocean resources for their livelihood. Of the 33,059 total fish species in the world, Indian EEZ houses a diversity of about 2,492 marine fish belonging to 941 genera under 240 families of 40 orders, owing to 7.4 % of the total marine fish resources[1]. According to CMFRI[2], the marine fish landings along the coast of mainland India during 2022 were 3.49 million tonnes, and Indian mackerel contributed 3.28 lakh tonnes (9.39 % of the national total).

The Indian mackerel (*Rastrelliger kanagurta* (Cuvier, 1816)), a scombrid fish in the Indo-West Pacific, thrives in the epipelagic zone from the surface to 150 m, often near thermoclines. Adults favour sheltered areas like coastal bays, harbours, and lagoons, particularly with plankton-rich turbid waters with a diet of larger plankton, including larval shrimp and fish. At the same time, juveniles survive on phytoplankton and smaller zooplankton. These fast swimmers form large schools and exhibit seasonal migrations linked to ocean productivity cycles, breeding seasons, currents, and circulation[3,4]. The Indian mackerel plays a crucial role in marine ecosystems by connecting lower trophic level production to higher level consumers[5,6]. However, the complex distribution patterns of this pelagic species make it challenging to pinpoint specific fishing grounds[7].

The Malabar coast, India's southwest coastline, is a global hotspot for upwelling[8]. The area selected for the study (73 – 77° E and 07 – 13° N), spanning from Ratnagiri in the north to Kanyakumari in the south, is a highly productive region, resulting in abundant marine fisheries landing in the nearby states. The upwelling process, characterised by offshore and alongshore currents, displaces warm surface water, allowing cold, nutrient-rich water to rise. This nutrient influx fosters plankton blooms, creating optimal feeding grounds for small pelagic fishes, including Indian mackerel, oil sardine, and anchovies[9]. Upwelling during the summer monsoon (May – September) brings nutrient-rich deepwater to the surface, fuelling a surge in primary and secondary productivity[10,11]. This rich ecosystem contributes to nearly half of India's total marine fish landings, with a peak in landings (dominated by oil sardine and mackerel) occurring from September to December, coinciding with or following the upwelling season[12,13]. The Malabar coast fishery relies heavily on mechanised gear, with trawl nets, purse seines, ring seines, and gill nets dominating landings. These mechanised units contribute to 95 % of the annual haul, leaving only a small portion for traditional, artisanal gears used primarily during the southwest monsoon[14]. For decades, oil sardines and mackerel have been the undisputed champions of Kerala's marine fisheries, although their dominance has fluctuated over time[15].

While physical processes in the upper ocean, like upwelling and currents, are known to influence biological interactions that determine pelagic fish stocks[8], relating fish landing to these factors using traditional methods had limited success in the Indian waters[16-19]; however, remote sensing offers a powerful alternative. Researchers can unravel critical relationships between fish populations and their environment by analysing remotely sensed data alongside fisheries data[20]. Satellites provide estimates of significant environmental parameters at broader spatial and temporal scales than traditional methods like buoys and moorings, ship-base observations, underwater gliders, UAVs, etc., capturing both short-term fluctuations and long-term trends. Integrating these data streams with GIS, statistical analysis, and in-situ fishery abundance data offers the most effective approach for understanding fish stock variability and managing fisheries for long-term sustainability.

Statistical modelling is a cornerstone of ecological research, allowing scientists to delve into the intricate relationships between fish and their environment. Generalised Linear Models (GLM), Generalised Additive Models (GAM), and Boosted Regression Trees (BRT) are adaptable statistical models that are crucial for understanding and predicting complex ecological systems like fisheries. These models effectually analyse disparate datasets to reveal intricate relationships between fish populations and environmental factors. Researchers, policymakers, and fishers can leverage these tools for tasks such as population dynamics modelling, habitat suitability assessment, impact evaluation of climate change, fisheries management, and risk assessment. By providing insights into fish abundance, distribution, growth, and interactions, these models support data-driven decisions for sustainable fisheries management.

GLMs are popular for their flexibility in handling linear, logistic and other data types and offering interpretable results[21]. They are robust to outliers, scalable for complex models, and relatively easy to use. However, there are limitations, too, as choosing the right statistical distribution is crucial and incorrect assumptions can lead to biases. Additionally, they may struggle with highly complex relationships or limited data. GAMs offer an alternative by using smooth functions to capture non-linear relationships without prior assumptions about the functional form[22]. This makes them ideal for analysing complex data, like time series or spatial data, where straight lines may not easily capture patterns. Unlike traditional linear models that assume a straight-line fit, GAMs use smoothing functions to model these non-linear connections. These functions act like flexible curves, detecting patterns that linear models might overlook. The key components of GAM are linear predictor - similar to linear regression, that combines the effects of different variables; smooth functions - which capture the non-linear relationships between each predictor and the response variable; link function - to connect the predicted outcome to the linear predictor, and additive structure - adding the effects of each smooth function together, allowing GAMs to model intricate relationships as a sum of simpler components[23]. In essence, GAMs provide a versatile and robust way to analyse complex data by offering a more flexible framework for capturing real-world phenomena.

Boosted Regression Tree (BRT) is a powerful statistical modelling technique that combines both statistics and machine learning fundamentals[24-26]. Unlike traditional regression methods that aim to produce a single "best" model, BRT utilises the concept of boosting to create a more robust model by combining a large number of relatively simple regression trees in a flexible way[27]. This approach has optimised predictive performance for various ecological applications[24-26]. Despite its advantages, including strong predictive power and reliable identification of important variables and interactions, the use of BRT has been limited in ecology, although some noteworthy works exist[28,29].

Boosted Regression Tree is one of several ensemble learning techniques for improving model performance by fitting numerous models and subsequently combining them for prediction. It takes advantage of two key algorithms: (i) regression trees, models of Classification and Regression Tree (CART) family are popular due to their intuitive and easily visualised information[30], and (ii) boosting, for building and combining a collection of models to enhance overall accuracy. Boosting builds upon the ease of finding and averaging numerous heuristic methods compared to finding a single, highly accurate prediction rule[31]. A unique aspect of boosting is its sequential nature, which follows a forward, stage-wise procedure where regression trees are fitted iteratively to the training data. The process emphasises observations that are poorly modelled by the existing collection of trees, gradually improving the model's ability to handle these challenging cases[32]. Various boosting algorithms exist depending on the quantification of lack of fit and selected settings for subsequent iterations. The original boosting algorithm, AdaBoost[33], was developed for two-class classification problems. Similar to GLMs, BRT models can be fitted to various response types - gaussian, poisson, binomial, etc., by specifying the error distribution and the link function. However, BRT introduces an element of stochasticity into the process. This stochasticity improves predictive performance by reducing the variance of the final model. It achieves this by using only a random subset of data to fit each new tree[32]. The sequential model-fitting process builds upon previously fitted trees, gradually focusing on predicting the most challenging observations. While prediction from a BRT model is straightforward, interpreting the model's inner working requires specific tools. These tools help identify the model's most important variables and interactions and visualise the fitted functions[34].

This study investigates the influence of primary environmental parameters on the abundance of Indian mackerel in the Malabar upwelling region. Given the observed high abundance of the Indian mackerel landings in Kerala, it is hypothesised that these environmental factors include rainfall, mixed layer depth, seawater temperature and salinity at 50 m depth, dissolved oxygen concentration, chlorophyll-a concentration, and net primary productivity. However, for this study, we focused on seawater temperature and salinity at 50 m depth along with mixed layer depth rather than sea surface temperature and sea surface salinity. We posit that these variables are crucial determinants of mackerel abundance as this species often inhabits and reproduces near the mixed layer interface, and their physiological processes are inherently linked to water temperature and salinity conditions at these depths.

Materials and Methods

State-wise fisheries quarterly landing data (1995 – 2012) was acquired from previous endeavours of INCOIS projects. Month-wise annual rainfall data (in mm) for coastal Kerala in the study period was downloaded from the Open Government Data (OGD) Platform, India (https://data.gov.in). Mixed layer depth, seawater temperature and salinity at 50 m depth for the Malabar upwelling region (MUR) were extracted from GOFS 3.0: HYCOM + NCODA Global 1/12° Reanalysis (https://www.hycom.org/data/glbu0pt08/expt-19pt1), available from August 1995 onwards. Data validation confirmed a good correlation with National Centers for Environmental Information (NCEI) OISST data. Chlorophyll-a and net primary productivity volume data were extracted from GLOBAL_ANALYSISFORECAST_BGC_001_028 E.U[ref. 35]. These environmental data were converted into quarterly time series data. The statistical analyses were performed in R (version 4.3.1; R Core Team 2023)[36]. GLM, GAM and BRT were performed using MASS, *mgcv* and *gbm* packages.

Results and Discussion

The results of the trends of mackerel landings and the environmental parameters in the study period are shown in Figure 1. GLM plots of the linear relationship of each variable with the mackerel landing are shown in Figure 2, which show a positive linear relationship between mackerel landing and chlorophyll-a, primary productivity, rainfall and salinity at 50 m depth and a negative relationship with mixed layer depth and seawater temperature at 50 m depth. Dissolved oxygen doesn't exhibit any clear relationship with mackerel landings.

Generalised additive model plots showing the partial effects of variables on the landing of mackerel in MUR are given in Figure 3. The final result of the model is shown in Table 1. These results suggest a

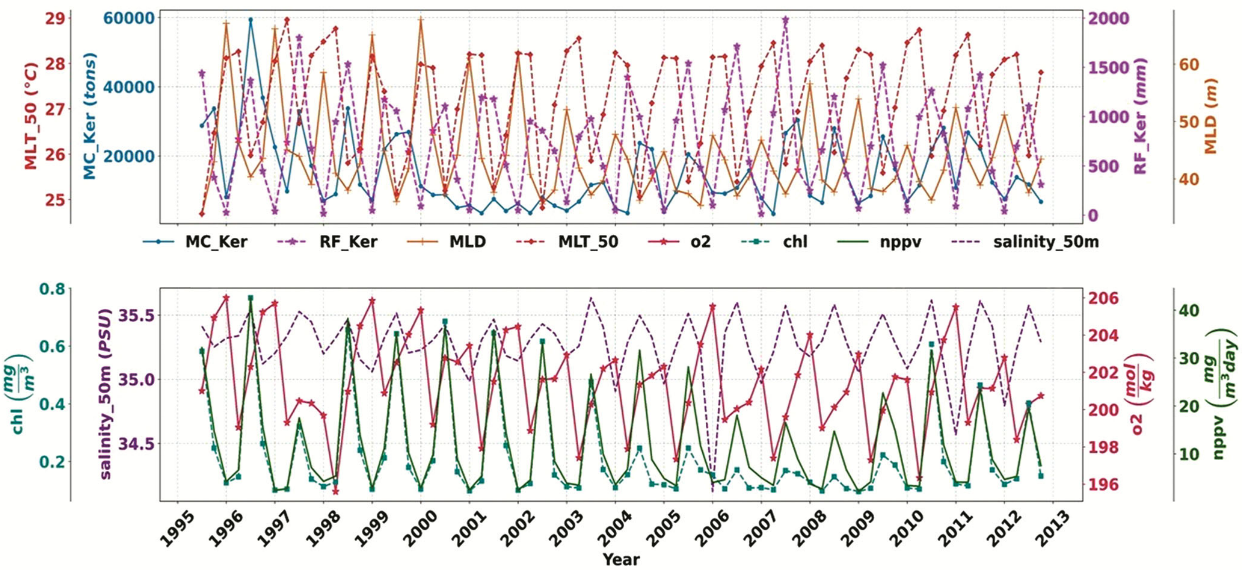

Fig. 1 — Plot showing the trend of mackerel landings in Kerala coast and the environmental variables during the study period. MC_Ker - Mackerel landings, RF_Ker - Rainfall in Kerala coast, MLD - Mixed Layer Depth, MLT_50 - Seawater temperature at 50 m depth, O_2 - Dissolved oxygen concentration, chl - Chlorophyll-a concentration, nppv - Net primary productivity volume, salinity_50 m - Salinity at 50 m depth

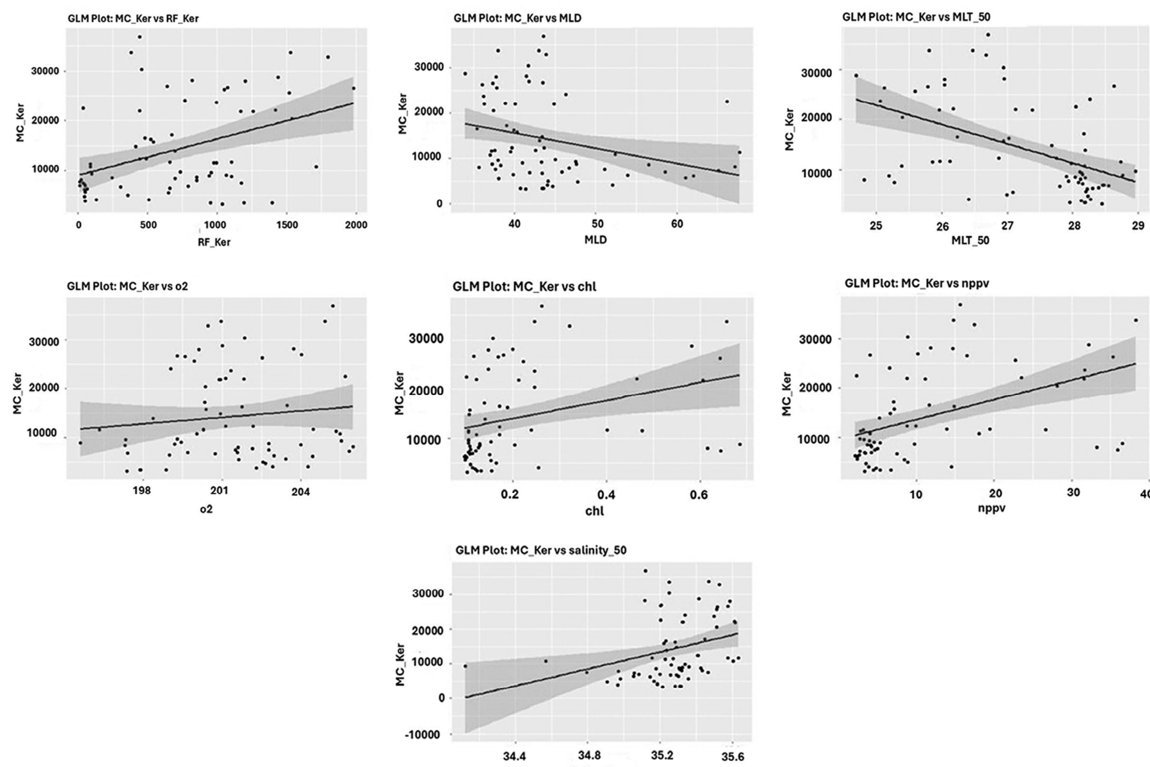

Fig. 2 — GLM plots of the linear relationship of each variable with the mackerel landings during the study period

significant relationship between mackerel landing and these environmental parameters in the region. The highly significant *p*-value (< 2e-16) for the intercept indicates a strong relationship between the predictors and the fish landing. Smooth terms were used to capture non-linear relationships between the fish landing and the predictors. Rainfall, chlorophyll-a, net primary productivity volume and dissolved oxygen

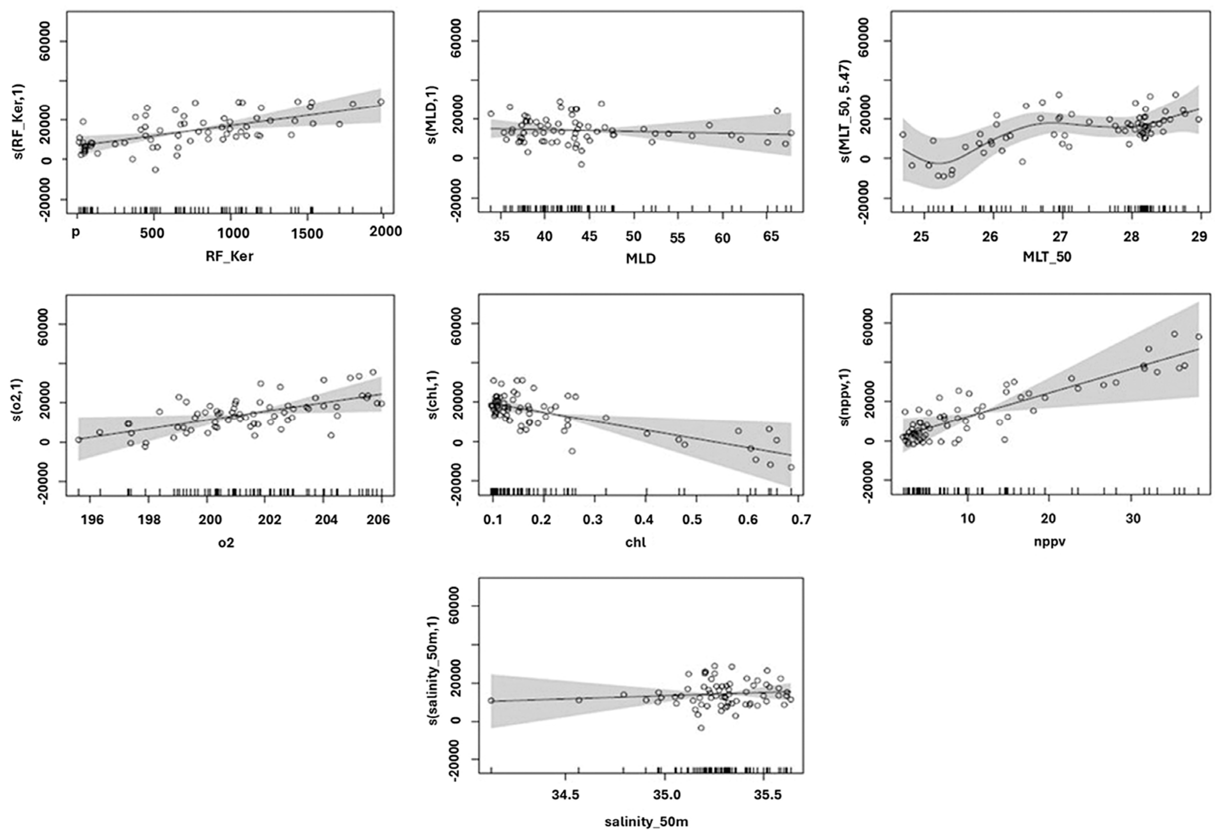

Fig. 3 — GAM plots showing the partial effects of variables on mackerel landing. The tick marks on the x-axis are observed data points. The y-axis represents the partial effect of each variable. The shaded areas indicate the 95 % confidence intervals

Table 1 — Results of the GAM model

	Estimate	Std. error	t value	Pr(>|t|)	
(Intercept)	14255.8	842.7	16.92	<2e-16	***---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:

	edf	Ref.df	F	*p*-value	
s(RF_Ker)	1	1	9.155	0.00372	**
s(MLD)	1	1	0.144	0.70579	
s(MLT_50)	5.471	6.582	2.142	0.05683	.
s(o2)	1	1	5.331	0.02459	*
s(chl)	1	1	6.67	0.0124	*
s(nppv)	1	1	7.123	0.00989	**
s(salinity_50m)	1	1	0.298	0.5874	---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) = 0.429 Deviance explained = 52.5%

GCV = 5.9804e+07Scale est. = 4.8995e+07 n = 69

Table 2 — Result of cross-validation of the GAM model

Method: GCV Optimizer: magic

Smoothing parameter selection converged after 19 iterations

The RMS GCV score gradient at convergence was 1.099299

The Hessian was positive definite

Model rank = 64 / 64

Basis dimension (k) checking results. Low p-value (k-index<1) may indicate that k is too low, especially if edf is close to k'

	k'	edf	k-index	*p*-value
s(RF_Ker)	9	1	1	0.45
s(MLD)	9	1	1.07	0.72
s(MLT_50)	9	5.47	1.1	0.68
s(o2)	9	1	1.09	0.7
s(chl)	9	1	1.31	1
s(nppv)	9	1	1.2	0.98
s(salinity_50m)	9	1	0.88	0.17

statistically affect the fish landing ($p < 0.5$). Seawater temperature at 50 m depth shows a marginally significant relationship ($0.05 < p$-value < 0.1). Mixed layer depth and salinity at 50 m depth do not statistically affect fish landing (p-value > 0.1). After accounting for model complexity, the adjusted R square value indicates 42.9 % of the variance in fish landing. The model’s ability to capture a substantial portion of variation is reinforced by the deviance of 52.5 %.

The smoothing parameter selection, model convergence, and the basis dimension (*k*) adequacy for each smooth term results information is given in Table 2. Generalised Cross-Validation (GCV) and "magic" optimiser are used for smoothing parameter selection. The smoothing parameter selection process convergence occurred after 19 iterations, suggesting the model reached a stable state at that stage. The RMS GCV score gradient at convergence was 1.099299, which, along with the positive values of the Hessian indicates stability in the optimisation process. The k-checking results assess whether the chosen basis dimension is adequate for capturing the relationships in the data. All smooth terms have a '*k*' value of 9. The effective degrees of freedom '*edf*' values vary for each term, indicating the flexibility of the model in capturing non-linear relationships. The '*k-index*' values are mostly close to 1, suggesting that the chosen basis dimension is adequate for most variables. The '*p*-values' mostly > 0.05, indicate that the '*edf*' are not significantly different from '*k*' for most variables. The model rank (64/ 64) meant that it had successfully utilised all available basic functions without overfitting. GAM plots show a positive relationship between mackerel landing and net primary productivity, rainfall and oxygen, while mixed layer depth and chlorophyll show a negative relation. Seawater temperature at 50 m depth shows an interesting non-linear relationship with mackerel landing. The residual analysis plots, using GAM check-plots (Fig. 4), show the presence of non-linear relationships between the predictor and response variables.

The relative influence plot (Fig. 5) in Boosted Regression Trees (BRT), generated as part of the model diagnostics and interpretation process, shows the relative importance of predictor variables in explaining the variation in the fish landing within the

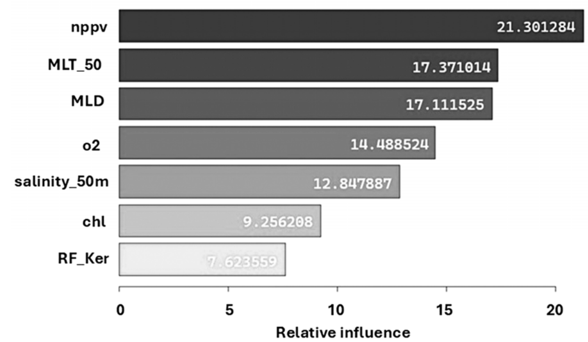

Fig. 5 — Relative influence plots of BRT showing percentage of influence of each variable on the mackerel landing

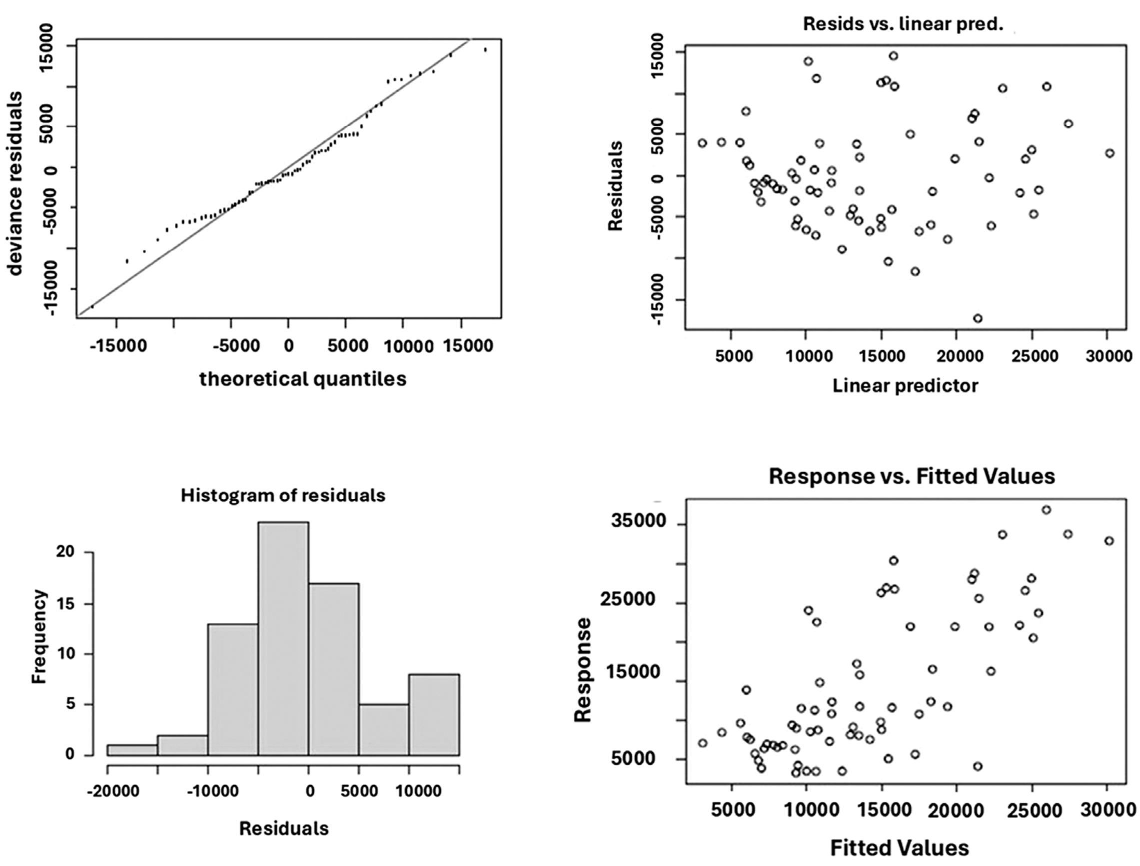

Fig. 4 — GAM check plots showing model performance

model (Table 3). The relative influence is determined considering (i) the contribution of each variable in the improvement model's fit or reducing error when splitting the data for each tree in the BRT ensemble, and (ii) the normalisation of cumulative contributions to represent the relative importance of each variable. The relative importance scores visually show the influence of each variable in the model. In this study, net primary productivity showed the highest importance value (21.30), suggesting that it has the strongest association with fish landing among all these variables. Seawater temperature at 50 m depth (17.37) and mixed layer depth (17.11) has a significant value, indicating a notable influence on fish landing. However, dissolved oxygen (14.49) appears to be slightly less influential than the above variables. Salinity at 50 m depth has a moderate importance value (12.85), while chlorophyll-a (9.26) and rainfall (7.62) have the lowest importance value, indicating a lesser impact on fish landing compared to the other variables. After relative influence analysis, the model was refined for prediction using the three most influential variables. The Stochastic Gradient Boosting (SGB) model was trained and evaluated using 5-fold cross-validation. A cross-validation study (Fig. 6a, b) found that the model was able to capture the underlying trends in the landing as reflected in the R-square value of 0.6. The Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) values were found to be 6504.5 and 5280, respectively, while a bigger dataset may help improve the model performance.

Table 3 — Result of the BRT model performed with the three most influential predictors

Stochastic Gradient Boosting
69 samples
3 predictor
No pre-processing
Resampling: Cross-Validated (5 fold)
Summary of sample sizes: 56, 56, 55, 55, 54
Tuning parameter 'shrinkage' was held constant at a value of 0.1
Tuning parameter 'n.minobsinnode' was held constant at a value of 10
RMSE was used to select the optimal model using the smallest value
The final values used for the model were n.trees = 50, interaction.depth = 2, shrinkage = 0.1 and n.minobsinnode = 10
"RMSE: 6504.48277754415"
"R-squared: 0.595748063299951"
"MAE: 5279.89373200466"

This study highlights the intricate relationship between environmental factors and Indian mackerel landing within the Malabar upwelling zone. Key environmental variables, such as chlorophyll-a, primary productivity, and rainfall, positively influenced mackerel abundance, while mixed layer depth exhibited a negative correlation. The unexpected lack of a clear relationship between dissolved oxygen and mackerel landing highlights the complexity of these marine ecosystems. A generalised additive model effectively captured the non-linear nature of these relationships, with net primary

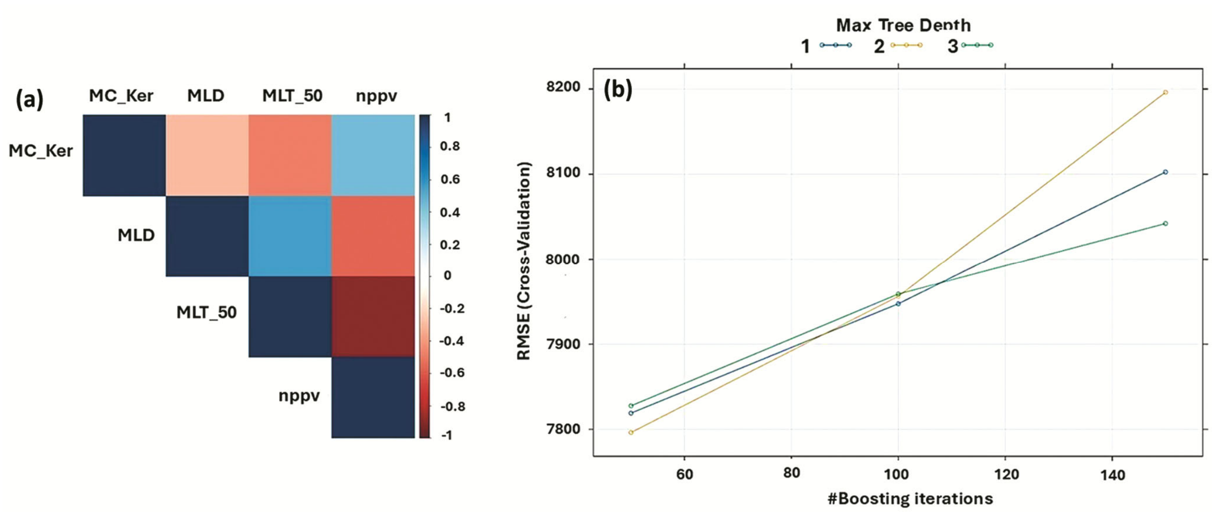

Fig. 6 — BRT performance metrics plots: a) Corplot showing the correlation of the three most influential variables on mackerel landing and their interrelationship; and b) Cross-validation of the number of boosting iterations with RMSE. RMSE decreases as the number of boosting iterations increases, showing that the model's accuracy improves as more trees are added to the ensemble. The rate of improvement slows down from the elbow in the curve at 100 iterations

productivity emerging as the primary driver of mackerel abundance. Although the model demonstrated predictive capability, further research incorporating additional variables and advanced modelling techniques is essential to enhance its accuracy. These findings may provide valuable insights for sustainable fisheries management, enabling informed decision-making regarding landing limits, fishing efforts, and conservation strategies.

Conclusion

The study provides new insights into the complex relationship between environmental factors and Indian mackerel abundance in the Malabar upwelling zone, south-east Arabian Sea. Findings demonstrate that environmental variables significantly influence mackerel landing, with primary productivity emerging as a key factor. While the models employed here offer a robust foundation for understanding these interactions, further research is imperative to refine the predictions and simultaneously design sustainable management strategies. By elucidating these ecological connections, the study contributes to the broader knowledge base on marine ecosystems and commercially important small pelagic fisheries.

Acknowledgements

The authors thank the Director of the Indian National Centre for Ocean Information Services (INCOIS), Hyderabad, India, for the encouragement. We sincerely thank the organizers of the "World Ocean Science Congress-2024" for providing a valuable platform to present this study. We would also like to acknowledge the editors and reviewers for their collaborative efforts in enhancing the clarity and quality of this manuscript. This study has been conducted as a part of "Species-specific potential fishing zone advisory research". This is INCOIS contribution no. 531.

Conflict of Interest

The authors declare that they have no known competing or conflict of interest.

Ethical Statement

This study has used standardised modelling procedures. The remote sensing data used is freely available. The fish catch data used was obtained from the authors' institute repository and permitted for use in internal research and external publications.

Author Contributions

SJ: Conceptualization, formal analysis, investigation, writing-original draft, and writing-reviewing and editing; NJ: Resources and software, SJo & TMBN: Supervision and writing-reviewing and editing.

References

1 Joshi K K, Sethulakshmi M & Varsha M S, Fish biodiversity of Indian Exclusive Economic Zone, In: *Winter School on Structure and Function of the Marine Ecosystem: Fisheries*, 1st to 21st Dec 2017, (CMFRI, Kochi), 2017, pp. 15. http://eprints.cmfri.org.in/id/eprint/12748

2 CMFRI, *Marine Fish Landings in India - 2022*, Technical Report, CMFRI Booklet Series No. 31/2023, (ICAR-Central Marine Fisheries Research Institute, Kochi), 2023, pp. 16. http://eprints.cmfri.org.in/id/eprint/16883

3 Collette B B & Russo J L, Morphology, systematics and biology of the Spanish mackerels (Scomberomous, Scombridae), *Fish Bull*, 82 (1984) 545-692. https://biostor.org/reference/59651

4 Skjoldal H R, Fish stocks and fisheries in relation to climate variability and exploitation, Ocean and aquatic ecosystems, In: *Encyclopedia of life supporting system (EOLSS)*, (UNESCO, Oxford, UK), 2004, pp. 19.

5 Sherman K, The large marine ecosystem approach to marine resources assessment and management, In: *The ecosystem approach to fisheries*, edited by G Bianchi & H R Skjoldal, (CABI Publishing & FAO, Rome), 2008, pp. 47–75. https://doi.org/10.1079/9781845934149.0047

6 Skjoldal H R & Misund O A, Ecosystem approach to management: definitions, principles and experiences from implementation in the Sea, In: *The ecosystem approach to fisheries*, edited by G Bianchi & H R Skjoldal, (CABI Publishing & FAO, Rome), 2008, pp. 209–227. https://doi.org/10.1079/9781845934149.0209

7 Nadira Y K, Mustapha M A & Ghaffar M A, The Indian mackerel aggregation areas in relation to their oceanographic conditions, *Sains Malays*, 48 (11) (2019) 2575-2581. http://dx.doi.org/10.17576/jsm-2019-4811-27

8 Bakun A, Roy C & Lluch-Cota S, Coastal Upwelling and other Processes Regulating Ecosystem Productivity and Fish Production in the Western Indian Ocean, In: *Large Marine Ecosystems of the Indian Ocean-Assessment, Sustainability and Management*, edited by K Sherman, E N Okemwa & M J Nuba, (Malden, Blackwell Science, London), 1998, pp. 103–141.

9 Krishnakumar P K, Mohamed K S, Asokan P K, Sathianandan T V, Zacharia P U, *et al.*, How environmental parameters influenced fluctuations in oil sardine and mackerel fishery during 1926-2005 along the southwest coast of India? *Mar Fish Infor Serv T&E Ser*, 198 (2008) 1-5. http://eprints.cmfri.org.in/id/eprint/3750

10 Vivekanandan E, Srinath M, Pillai V N, Immanuel S & Kurup K N, Trophic Model of the Coastal Fisheries Ecosystem of the Southwest Coast of India, In: *Assessment, Management and Future Directions for Coastal Fisheries in Asian Countries (1705)*, edited by Silvestre G, Garces L, Stobutzki I, Ahmed M, Valmonte-Santos R A, *et al.*, (WorldFish centre, Penang, Malaysia and The Asian

Development Bank, Philippines), 2003, pp. 281-298. http://eprints.cmfri.org.in/id/eprint/8510

11 Smith S & Madhupratap M, Mesozooplankton of the Arabian Sea: Patterns influenced by seasons, upwelling, and oxygen concentrations, *Prog Oceanogr*, 65 (2-4) (2005) 214-239. https://doi.org/10.1016/j.pocean.2005.03.007

12 Manjusha U, Jayasankar J, Remya R, Ambrose T V & Vivekanandan E, Influence of coastal upwelling on the fishery of small pelagics off Kerala, southwest coast of India, *Indian J Fish*, 60 (2) (2013) 37-42. http://eprints.cmfri.org.in/id/eprint/9507

13 Krishnakumar P K & Bhat G S, Seasonal and interannual variations of oceanographic conditions off Mangalore coast (Karnataka, India) in the Malabar upwelling system during 1995–2004 and their influences on the pelagic fishery, *Fish Oceanogr*, 17 (1) (2013) 45–60. https://doi.org/10.1111/j.1365-2419.2007.00455.x

14 Ramamirtham C P & Patil M R, Hydrography of the west coast India during the premonsoon period of the year 1962, Part 11, *J Mar Biol Assoc India*, 7 (1964) 150-168.

15 Mohamed K S, Muthiah C, Zacharia P U, Sukumaran K K, Rohit P, *et al.*, Marine fisheries of Karnataka state, *Naga*, 21 (2) (1998) 10-15. http://eprints.cmfri.org.in/id/eprint/6181

16 Longhurst A R & Wooster W S, Abundance of oil sardine (*Sardinella longiceps*) and upwelling on the southwest coast of India, *Can J Fish Aquat Sci*, 47 (1990) 2407–2419. https://doi.org/10.1139/f90-268

17 Madhupratap M, Shetye S R, Nair K N V & Nair S R, Oil sardine and Indian mackerel: Their fishery, problems, and coastal oceanography, *Curr Sci*, 66 (5) (1994) 340-348. http://drs.nio.org/drs/handle/2264/7405

18 Yohannan T M &Abdurahiman U, Environmental influence on the behaviour of Indian mackerel and their availability to fishing gear along the Malabar coast, *Indian J Fish*, 45 (1998) 239-247. http://eprints.cmfri.org.in/id/eprint/175

19 Jayaprakash A, Long term trends in rainfall, sea level and solar periodicity: A case study for forecast of Malabar sole and Oil sardine fishery, *J Mar Biol Assoc India*, 44 (2002) 163-175. http://eprints.cmfri.org.in/id/eprint/1992

20 Jha S, Sudhakar S D, Majumder S, Joseph S & Nair T M B, A multidecadal study of the Malabar upwelling system influencing Indian Mackerel landings along the coasts of Karnataka and Kerala, south-east Arabian Sea, *J Mar Syst*, 246 (2024) p. 104012. https://doi.org/10.1016/j.jmarsys.2024.104012

21 Müller M, Generalized Linear Models, In: *Handbook of Computational Statistics*, edited by J Gentle, W Härdle & Y Mori, Springer Handbooks of Computational Statistics, (Springer, Berlin, Heidelberg), 2012, pp. 681-709. https://doi.org/10.1007/978-3-642-21551-3_24

22 Hastie T J & Tibshirani R J, Generalized additive models, *Stat Sci*, 1 (2) (1986) 297-318. https://api.semanticscholar.org/CorpusID:270228534

23 Wood S N, *Generalized Additive Models: An Introduction with R*, 2nd edn, (Chapman and Hall/CRC Press, USA), 2017, pp. 496. https://doi.org/10.1201/9781315370279

24 Elith J, Graham C H, Anderson R P, Dudík M, Ferrier S, *et al.*, Novel methods improve prediction of species' distributions from occurrence data, *Ecography*, 29 (2) (2006) 129-151. https://doi.org/10.1111/j.2006.0906-7590.04596.x

25 Elith J & Leathwick F R, Species Distribution Models: Ecological Explanation and Prediction Across Space and Time, *Annu Rev Ecol Evol Syst*, 40 (2009) 677-697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

26 Leathwick J R, Elith J, Chadderton W L, Rowe D K & Hastie T J, Dispersal, disturbance and the contrasting biogeographies of New Zealand's diadromous and non-diadromous fish species, *J Biogeogr*, 35 (8) (2008) 1481-1497. https://doi.org/10.1111/j.1365-2699.2008.01887.x

27 Breiman L, Friedman J, Olshen R & Stone C, *Classification and regression trees*, (Chapman and Hall/CRC Press, New York, US), 1984, pp. 368. https:// doi.org/10.1201/9781315139470

28 James G, Witten D, Hastie T & Tibshirani R, *An introduction to statistical learning with Applications in R*, (Springer, New York), 2021, pp. 607. https://doi.org/10.1007/978-1-0716-1418-1

29 Moisen G G, Freeman E A, Blackard J A, Frescino T S, Zimmermann N E, *et al.*, Predicting tree species presence and basal area in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods, *Ecol Modell*, 199 (2) (2006) 176–187. https://doi.org/10.1016/j.ecolmodel.2006.05.021

30 De'ath G, Boosted regression trees for ecological modelling and prediction, *Ecology*, 88 (1) (2007) 243-251. https://doi.org/10.1890/0012-9658(2007)88[243:BTFEMA]2.0. CO;2

31 Schapire R E, The Boosting Approach to Machine Learning: An Overview, In: *Non-linear Estimation and Classification, Lecture Notes in Statistics*, Vol 171, edited by D D Denison, M H Hansen, C C Holmes, B Mallick & B Yu, (Springer, New York), 2003, pp. 149-171. https://doi.org/10.1007/978-0-387-21579-2_9

32 Friedman J H, Stochastic gradient boosting, *Comput Stat Data Anal*, 38 (4) (2002) 367-378. https://doi.org/10.1016/S0167-9473(01)00065-2

33 Freund Y & Schapire R E, Experiments with a new boosting algorithm, *Proc 13th International Conference on Machine Learning (ICML'96)*, (Morgan Kaufmann Publishers Inc, San Francisco, CA, USA), 1996, pp. 148–156.

34 Elith J, Leathwick J R & Hastie T, A working guide to boosted regression trees, *J Anim Ecol*, 77 (2008) 802-813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

35 GLOBAL_ANALYSISFORECAST_BGC_001_028. E.U. Copernicus Marine Service Information (CMEMS), Marine Data Store (MDS). Available online at: http://doi.org/10.48670/moi-00015; Accessed on September 2023.

36 R Core Team, *R: A Language and Environment for Statistical Computing*, (R Foundation for Statistical Computing, Vienna, Austria), 2023. Available online at: https://www.R-project.org/