Modelling the distribution of sea turtles in the Western Indian Ocean based on bycatch data from the French pelagic longline and purse seine fisheries

Jonathan R. Monsinjon[1]*, Philippe S. Sabarros[2,3], Pascal Bach[2], Sylvain Bonhommeau[1], Jérôme Bourjea[2]

[1] IFREMER - Indian Ocean Delegation (DOI), Le Port, La Reunion, France

[2] MARBEC, Univ Montpellier, CNRS, Ifremer, IRD, Sète, France

[3] IRD, Ob7, Sète, France

*Corresponding author: jonathan.monsinjon@gmail.com

Abstract

Species Distribution Models (SDMs) are valuable predictive tools to anticipate bycatch risk in fisheries. Bycatch of sea turtles, which are of conservation concern worldwide, could negatively affect populations through direct mortality or decreased post-release fitness. With a better understanding of the environmental variables driving their distribution, one could provide successful bycatch mitigation strategies. However, this remains an important knowledge gap for sea turtles in the Western Indian Ocean. To address this, we used two modelling approaches, namely logistic regression and Random Forest, to identify and quantify the importance of 15 candidate environmental predictors for loggerhead (TTL), olive ridley (LKV), and green (TUG) turtles. Using on-board observer data from the French pelagic longline and purse seine fisheries, we show that sea surface height and the Dipole Mode Index could be important predictors of bycatch events for the three turtle species. Our results should prove useful to select appropriate environmental variables depending on the focal species to fit SDMs from bycatch data. Nevertheless, the modelling approaches used here have limitations that warrant consideration. We discuss those and provide recommendations for further improvement.

Keywords

Marine turtles | Habitat modelling | Bycatch | Pelagic longline | Purse seine | Floating objects | Western Indian Ocean

Introduction

Species Distribution Models (SDMs) are useful tools for Dynamic Ocean Managements (Abrahms et al., 2019; Hazen et al., 2017; Maxwell et al., 2015) and are promising to reduce bycatch:target species ratios (Hazen et al., 2018; Howell et al., 2008; Stock et al., 2020). For large marine species, such models are commonly fitted using presence data from tracking studies but true absence data are lacking in this case, which requires generating pseudo-absences using a variety of techniques (Hazen et al., 2021; Iturbide et al., 2015; O'Toole et al., 2021; Raymond et al., 2015). On-board observer data from fisheries are extremely valuable since they provide true absences. Although they might not cover uniformly the full distributional range of focal bycatch species, they can feed SDMs to produce bycatch risk maps and thus inform mitigation strategies.

Many types of SDMs exist and their application has become easier for non-statisticians via R packages. In addition, most of the environmental variables potentially driving the distribution of marine species are now available at more ecologically relevant spatiotemporal scales thanks to advances in remote sensing techniques and global climate reanalyses. This allowed recent studies to fill substantial knowledge gaps regarding the habitat characteristics of many marine species, including sea turtles (Chambault et al., 2020; Chambault et al., 2021b; Chambault et al., 2021c; Scales et al., 2015) that are negatively affected by fisheries worldwide (Wallace et al., 2013). The main challenge now lies in the choice of the most suited SDMs and predictive variables, especially for sea turtles in the Western Indian Ocean for which we know very little about their pelagic habitat.

Our ultimate goal is to use SDMs to predict sea turtle distribution in the Western Indian Ocean and provide recommendations to reduce bycatch risk in pelagic longline and purse seine fisheries. Here we present preliminary results on the influence of a series of environmental variables on bycatch risk of three sea turtle species using observer data from the French pelagic longline and purse seine fisheries.

Material and methods

1. Case studies

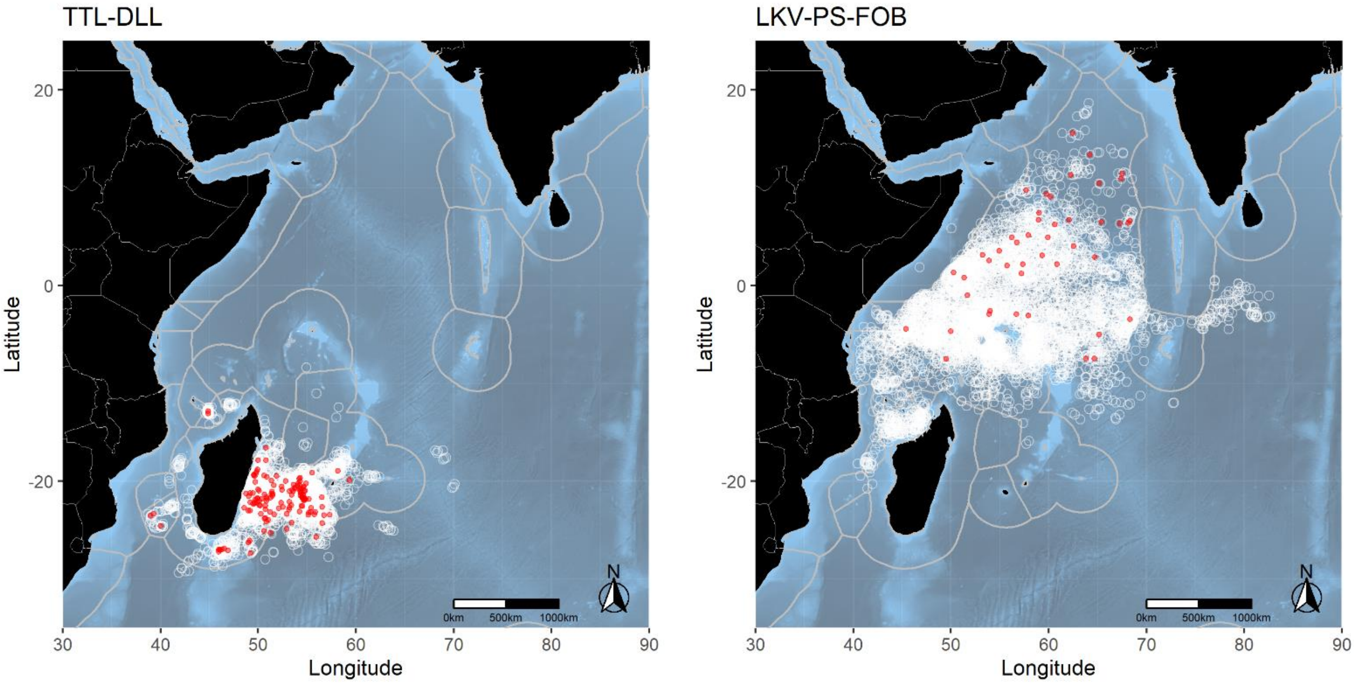
The loggerhead turtle (FAO code: TTL) is the most common turtle bycatch in drifting longlines (2007-2021), representing 50.35% of the individuals incidentally caught (Table 1). Olive ridley turtles (FAO code: LKV) and green turtles (FAO code: TUG) are the most commonly caught turtles in purse seines on floating object-associated tuna schools (2005-2021), representing together more than 50% of the individuals (Table 1). As case studies, we focused on these three situations (Fig. 1). We did not consider bycatch events in purse seines on free-swimming tuna school because the number of individuals caught remains very small for every turtle species (Table 1).

Table 1: Number of turtles caught (and relative proportions) in drifting longlines (DLL) and purse seines on floating object-associated tuna school (PS-FOB) or free-swimming tuna school (PS-FSC). Species: loggerhead turtles (TTL), olive ridley turtles (LKV), green turtles (TUG), leatherback turtles (DKK), hawksbill turtles (TTH), and unidentified species (TTX).

Species	DLL	PS-FOB	PS-FSC
TTL	145 (50.35%)*	16 (10.53%)	1 (7.14%)
LKV	22 (7.63%)	52 (34.21%)*	2 (14.29%)
TUG	45 (15.62%)	40 (26.32%)*	3 (21.43%)
DKK	38 (13.19%)	1 (0.66%)	2 (14.29%)

TTH	15 (5.2%)	24 (15.79%)	6 (42.86%)
TTX	23 (7.99%)	19 (12.5%)	0 (0%)
Total caught	288	152	14

* Case studies presented here.



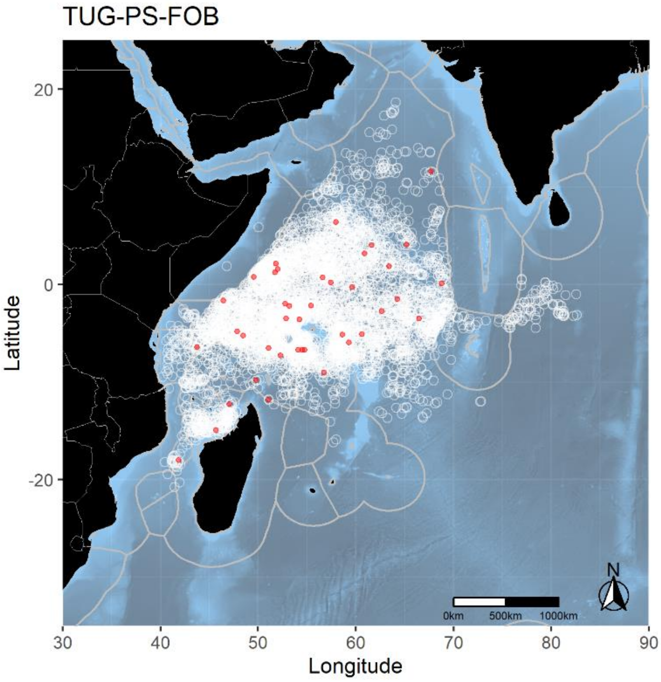


Figure 1: Fishing events (white circles) and bycatch locations (red dots) of loggerhead turtles in drifting longlines (TTL-DLL) and olive ridley and green turtles in purse seines on floating object-associated tuna school (LKV-PS-FOB and TUG-PS-FOB, respectively). Grey polygons delineate the boundaries of all exclusive economic zones in the study area.

2. Modelling approach

To identify the environmental drivers potentially involved in bycatch events, we used two modelling approaches: the logistic regression (i.e., parametric approach via a generalized linear model) and the Random Forest (i.e., non-parametric approach via a machine-learning algorithm for classification). In both cases, we treated each bycatch event with at least one individual as a presence location. We chose the logistic regression for its interpretability regarding the relationship between predictors and presence/absence probabilities and the Random Forest for its ability to cope with complex interactions among the predictor variables (Cutler et al., 2007) and its good predictive performance in species distribution modelling (Chambault et al., 2021a; Scales et al., 2016; Siders et al., 2020; Stock et al., 2020). We performed our analyses using R version 4.2.1 (R Core Team, 2022).

2.1. Logistic regression

For each species, we built generalized linear models with a binomial distribution and a logit function to predict presence/absence probabilities from a series of environmental predictors (see Table 2) using functions from the 'stats' R package. To avoid multicollinearity issues (Zuur et al., 2010), we first removed highly correlated variables one by one until no absolute correlation higher than 0.6 remained in our datasets. We then performed a model selection based on the lowest Akaike Information Criterion corrected for finite sample size (AICc) by comparing all possible combinations of environmental predictors (without interaction terms) and we considered all models with a difference in AICc values below (or equal to) two as best candidates (Burnham & Anderson, 2002). To evaluate the influence of each environmental variable on the likelihood to catch at least one turtle, we compare the odds ratios from the best candidate models. In short, odds ratios greater than one suggest that a bycatch event is more likely to occur as the predictor increases while odds ratios below one indicate that the event is less likely to occur. Finally, we computed bycatch probabilities within the range of each predictor according to the best model while holding the others constant at their respective means.

2.2. Random Forest

As the Random Forest approach is less sensitive to multicollinearity (Cutler et al., 2007), we considered all environmental variables as potential predictors and we subsequently evaluated their importance in the classification procedure (i.e., their predictive power). Using functions from the 'caret' R package, we trained a series of random forests via a *k*-fold cross-validation procedure repeated 5 times and varying the following parameters: *k* (number of test/training partitions of the dataset: 10, 5, and 3), *mtry* (number of randomly selected predictors: number of predictors minus one), *ntree* (number of trees in the forest: 50, 250, 500, 1000, and 2500), *nodesize* (minimum size of terminal nodes: from 5 to 50 every 5), and *maxnodes* (maximum amount of terminal nodes: from 5 to 50 every 5). We conducted this exploratory grid search to identify the set of parameters that were more likely to yield good performances based on the area under the receiver operating characteristic curve (AUC). The AUC is a measure of a classifier's ability to discriminate between true and false positives, where values close to 1 indicate very good classification performances (typically between 0.9 and 1) and values close to 0.5 near-complete failures (typically between 0.5 and 0.6). During this first step, we retained the set of parameters associated with AUC values greater than (or equal to) the 97.5th percentile (i.e., the best performing ones among all trained random forests) and we trained again the random forests with these best candidate parameters. From this second series of random forests, we retained the ones with AUC values greater than (or equal to) 0.6 to identify the environmental variables that have the most predictive power.

3. Environmental variables

We extracted the following variables within a 25-km radius ellipsoidal buffer around presence/absence locations and we computed the inverse-distance weighted average when relevant (see "_idw" in the variables' short name). We chose those variables for their potential role in the ecology of highly mobile marine species (Abrahms et al., 2019; Becker et al., 2020; Brodie et al., 2018; Chambault et al., 2021b; Hazen et al., 2018; Scales et al., 2017; Virgili et al., 2019).

Table 2: Environmental variables considered in this study.

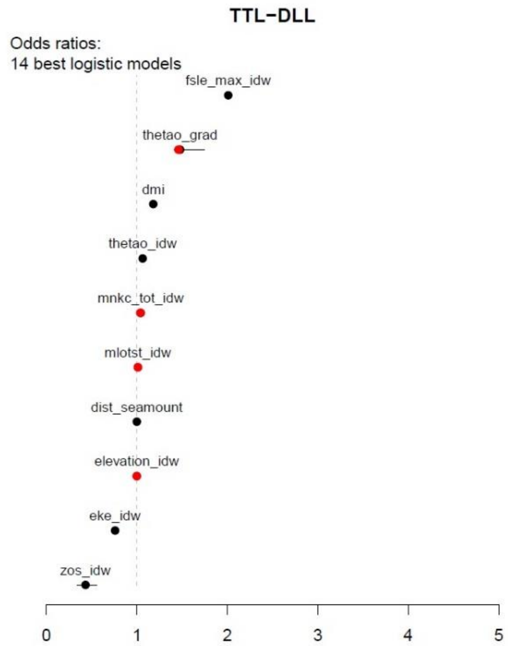
Variable	Short name	Source
Bathymetry (m) - 15 arc-second spatial resolution	elevation_idw	https://www.gebco.net/data_and_products/gridded_bathymetry_data
Distance to the coast (m)	dist_coast	Calculated from bathymetry data
Distance to the 200-m isobath (m)	dist_isobath200m	Calculated from bathymetry data
Distance to seamounts (m)	dist_seamount	Calculated from the location of seamounts (http://data.unep-wcmc.org/datasets/41)
Sea surface temperature (°C) - daily; 0.083° spatial resolution	thetao_idw	https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_PHY_001_030/INFORMATION
Sea surface temperature gradient (°C) - daily	thetao_grad	Calculated from sea surface temperature data (maximum-minimum within the ellipsoidal buffer)
Mixed layer thickness (m) - daily; 0.083° spatial resolution	mlotst_idw	https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_PHY_001_030/INFORMATION
Sea surface height above geoid (m) - daily; 0.083° spatial resolution	zos_idw	https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_PHY_001_030/INFORMATION
Finite size Lyapunov exponent (days^{-1}) - daily; 0.04° spatial resolution	fsle_max_idw	https://www.aviso.altimetry.fr/en/data/products/value-added-products/fsle-finite-size-lyapunov-exponents.html
Eddy kinetic energy (m^2/s^2) - daily	eke_idw	Calculated from surface zonal and meridional velocities (https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_PHY_001_030/INFORMATION) using the following formula: $(U^2+V^2)/2$
Net primary production (mg C m^{-2} day^{-1}) - daily; 0.083° spatial resolution	npp_idw	https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_BGC_001_033/INFORMATION
Zooplankton biomass (g C m^{-2}) - daily; 0.083° spatial resolution	zooc_idw	https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_BGC_001_033/INFORMATION
Epipelagic micronekton biomass (g C m^{-2}) - daily; 0.083° spatial resolution	mnkc_epi_idw	https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_BGC_001_033/INFORMATION
Total micronekton biomass in the epipelagic zone* (g C m^{-2}) - daily; 0.083° spatial resolution	mnkc_tot_idw	Calculated from micronekton biomass data (https://resources.marine.copernicus.eu/product-detail/GLOBAL_MULTIYEAR_BGC_001_033/INFORMATION) using the following formula: epipelagic micronekton + (24 - day length) x (migrant upper mesopelagic micronekton + highly migrant lower mesopelagic micronekton)
Dipole mode index (°C) - monthly	dmi	https://psl.noaa.gov/gcos_wgsp/Timeseries/DMI/

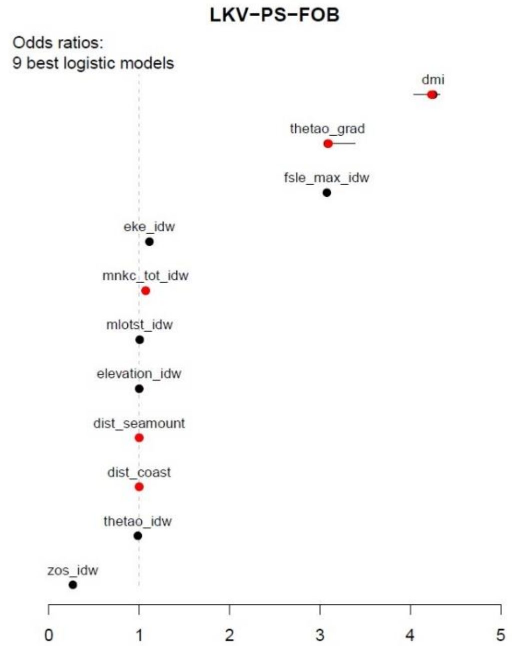
Results

We retained 14 logistic models for loggerhead turtles (TTL) in drifting longlines (DLL), 9 for olive ridley turtles (LKV) and 26 for green turtles (TUG) in purse seines on floating object-associated tuna school (PS-FOB), with odds ratios ranging between 0 and 5 (Fig. 2). For every species, all best models estimated very similar odds ratios. According to the best model for TTL, the probability to catch at least one turtle increases as elevation_idw decreases, and increases with mlotst_idw, mnkc_tot_idw, and thetao_grad (Supp. Info Fig. S1). For LKV, bycatch probabilities increase with dist_coast, dist_seamount, dmi, mnkc_tot_idw, and thetao_grad (Supp. Info Fig. S2). For TUG, bycatch probabilities decrease with dist_coast and increase with dmi (Supp. Info Fig. S3).

To evaluate the importance of our candidate environmental predictors, we retained 284 random forests for TTL, 386 for LKV, and 36 for TUG (Fig. 3). For TTL, there is no clear consensus in the contribution of each predictor, except for dist_coast, dist_isobath200m, and zooc_idw for which at least 97.5% of the forests agreed on a contribution below 50%. For LKV, at least 97.5% of the forests agreed on a contribution above 50% for dmi and zos_idw, and below 50% for thetao_grad, dist_coast, and dist_isobath200m. For TUG, at least 97.5% of the forests agreed on a contribution above 50% for zos_idw and npp_idw, and below 50% for dist_isobath200m, dist_coast, dist_seamount, and eke_idw.

Overall, a consistent pattern emerged for dmi with a median importance above 50% and a clear positive effect on bycatch of the three turtle species (Table 3). For zos_idw, which has a similarly high importance across the three species, we found an opposite incidence on bycatch probabilities when comparing TTL and LKV with TUG. We observed the same situation for fsle_max_idw, although this predictor is potentially worthless for TUG and not powerful for the two other species.





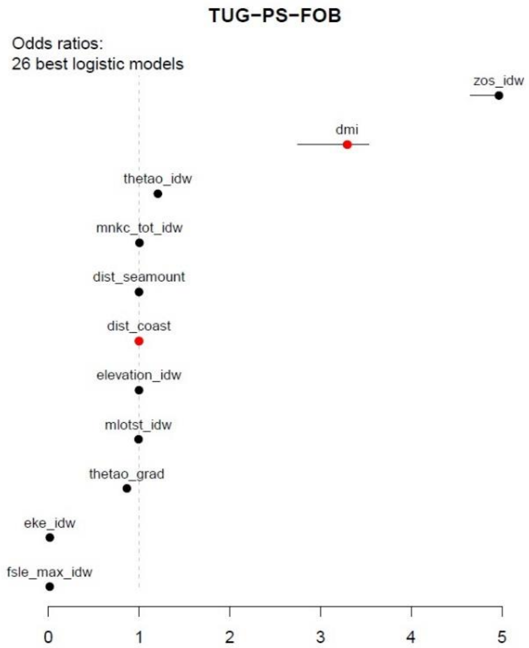
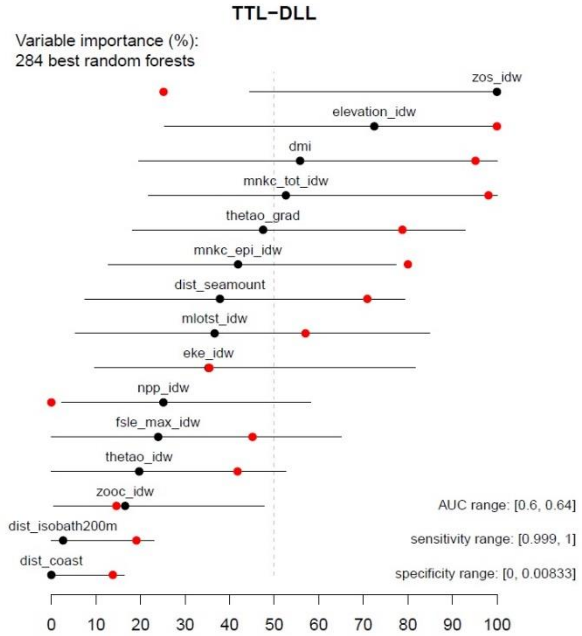
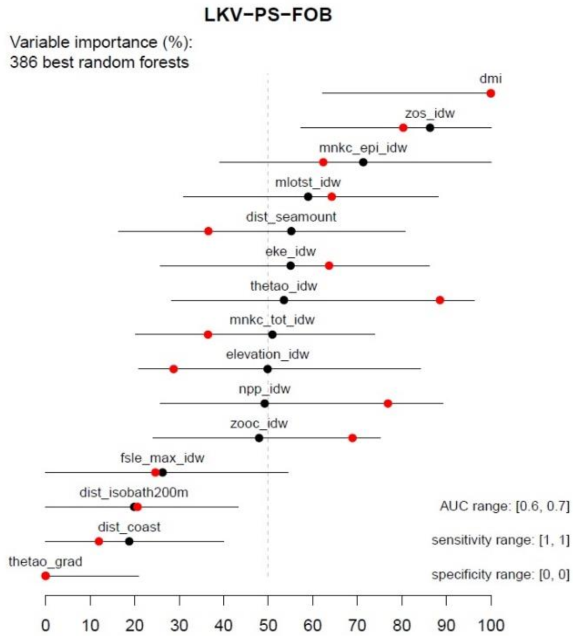


Figure 2: Odds ratios of the best logistic models for loggerhead turtles in drifting longlines (TTL-DLL) and olive ridley and green turtles in purse seines on floating object-associated tuna school (LKV-PS-FOB and TUG-PS-FOB, respectively). Black dots represent the median and horizontal lines the 2.5th and 97.5th percentiles. Red dots indicate odds ratios of the best model (i.e., with the lowest AICc). Vertical dashed lines indicate the situation where a predictor would have no effect on the probability to catch at least one turtle.





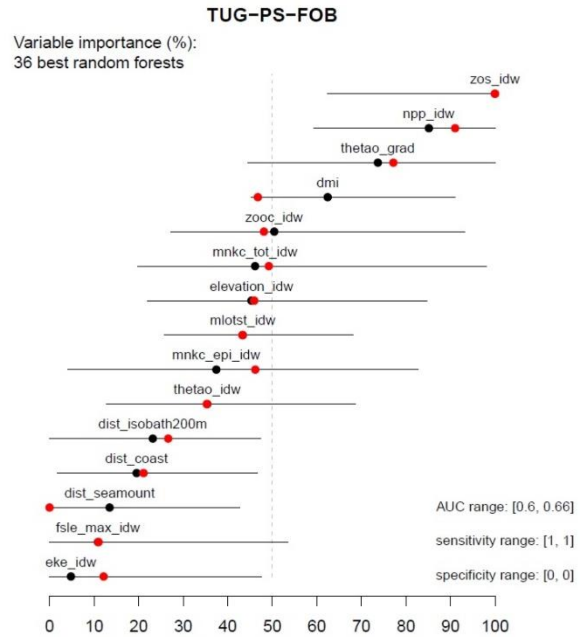


Figure 3: Variable importance (i.e., AUC of each variable scaled to 100%) of the best random forests for loggerhead turtles in drifting longlines (TTL-DLL) and olive ridley and green turtles in purse seines on floating object-associated tuna school (LKV-PS-FOB and TUG-PS-FOB, respectively). Black dots represent the median and horizontal lines the 2.5th and 97.5th percentiles. Red dots indicate variable importance of the best forest (i.e., with the highest AUC). Vertical dashed lines indicate the situation where a variable would have a predictive power half as high as the best one.

Table 3: Environmental variables with a clear positive (odds ratio >1.15) or negative (odds ratio <0.85) effect on the odds to catch at least one turtle. The variables with a median importance over 50% likely have good predictive power while the ones below 20% are probably worthless. Letter codes indicate whether loggerhead turtles in drifting longlines (TTL) and olive ridley and green turtles in purse seines on floating object-associated tuna school (LKV and TUG, respectively) met these conditions.

Variable	**Odds ratio >1.15**	**Odds ratio <0.85**	**Importance >50%**	**Importance <20%**
elevation_idw			TTL	
dist_coast				TTL, LKV, TUG
dist_isobath200m				TTL, LKV
dist_seamount			LKV	TUG
thetao_idw	TUG		LKV	TTL
thetao_grad	TTL, LKV		TUG	LKV

mlotst_idw			LKV	
zos_idw	TUG	TTL, LKV	TTL, LKV, TUG	
fsle_max_idw	TTL, LKV	TUG		TUG
eke_idw		TTL, TUG	LKV	TUG
npp_idw			TUG	
zooc_idw			TUG	TTL
mnkc_epi_idw			LKV	
mnkc_tot_idw			TTL, LKV	
dmi	TTL, LKV, TUG		TTL, LKV, TUG	

Discussion

Our study sheds light on the environmental variables that potentially relate to bycatch events of loggerhead turtles (TTL) in drifting longlines (DLL) and olive ridley (LKV) and green turtles (TUG) in purse seine on on floating object-associated tuna school (PS-FOB). We found that sea surface height above geoid (zos_idw) and the Indian Ocean Dipole Mode Index (dmi) might be important predictors of bycatch for these three species.

Interestingly, our results suggest that sea surface height above geoid influences bycatch events in opposite directions for loggerhead and olive ridley turtles in one hand and green turtles in the other hand, with increasing bycatch risk of loggerhead and olive ridley turtles at higher surface heights. Because high surface heights occur in upwelling regions, this reinforces a previous finding that loggerhead turtles likely forage in highly productive waters around mesoscale eddies and thermal fronts (Scales et al., 2015) and this could be the case for olive ridley turtles too as shown here. The reason why bycatch events of green turtles more likely occur at lower surface heights remains unclear, but could well be due to very different foraging strategies as adult green turtles are mostly herbivorous (Esteban et al., 2020) and the two other species carnivorous (Colman et al., 2014; Revelles et al., 2007).

We found that the Dipole Mode Index likely relates to bycatch events of the three species in the same direction, where it is more likely to catch at least one turtle with higher values of this climate index. Positive values of the Dipole Mode Index indicate warmer than average waters in the tropical western Indian Ocean and cooler than average waters in the tropical eastern Indian Ocean (Saji et al., 1999). In this situation, anomalous winds affect a series of oceanographic processes through feedbacks that influence evaporation, upwelling, and the thermocline depth (An et al., 2022). Such modifications in abiotic conditions could have large-scale ecosystem impacts causing regional depletion or augmentation of food resources for the three turtles species studied here.

Overall, we highlighted clear positive and negative effects of a series of oceanographic variables on bycatch of our three focal turtle species. These preliminary results can serve as a baseline for further development of Species Distribution Models (SDMs) using the most important environmental predictors identified in Table 3. However, both modelling approaches used here (namely, logistic regression and Random Forest) have limitations that must be addressed. For instance, one could use generalized additive models instead of logistic regressions to handle better non-linear and non-monotonic relationships via smooth functions, albeit at the expense of interpretability and with an increased risk of overfitting. In addition, we treated each bycatch event as a presence location – hence predicting the probability to catch at least one individual – but one could use delta-models to jointly model bycatch probability and the number of individuals likely to be caught while accounting for fishing effort and location (e.g., Stock et al., 2020; Stock et al., 2019). Provided that fishing events often catch multiple individuals at once, this approach can be advantageous because the environmental variables driving the occurrence and the density of individuals could be different. In practice, however, delta-models are difficult to apply to sea turtles because they typically represent rare bycatch events with too few multiple-individual catches to fit meaningfully the positive component of delta-models (i.e., number of individuals). In our case, bycatch of loggerhead turtles in drifting longlines consisted in 3.22% of non-zero observations (with only 6/145 events with multiple individuals), and, in purse seines on floating object-associated tuna school, 0.66% for olive ridely turtles and 0.6% for green turtles (with respectively 7/52 and 0/40 events with multiple individuals). Random Forest is a good alternative to overcome issues related with such severely imbalanced data. However, this might result in good predictive performance for absences (i.e., the majority class) but not for presences (He & Garcia, 2009). This might explain why the random forests retained in our study yielded fairly good, but not ideal, overall performances (i.e., with high sensitivity but low specificity and maximum AUC values of 0.64, 0.7, and 0.66, respectively for loggerhead, olive ridley and green turtles; Fig. 3). With a similar percentage of non-zero observations, Stock et al. (2020) predicted bycatch of loggerhead and leatherback turtles in the Pacific Ocean with AUC values over 0.8 by implementing two alternative sampling techniques to the Random Forest algorithm: (1) setting the sample size to a proportion of the number of presences (down-sampling), and (2) combining down-sampling of the absences with over-sampling of the presences (SMOTE; Chawla et al., 2002). Using a third alternative sampling technique, an iterative procedure using bootstrapping and balancing over an ensemble of random forests (ERFs), Siders et al. (2020) reached even greater predictive performances for other rarely caught species. This recent improvement of the Random Forest approach will prove useful to refine the results presented here, and model sea turtle distribution and bycatch risk in the Western Indian Ocean.

References

Abrahms, B., Welch, H., Brodie, S., Jacox, M. G., Becker, E. A., Bograd, S. J., . . . Hazen, E. L. (2019). Dynamic ensemble models to predict distributions and anthropogenic risk exposure for highly mobile species. *Divers. Distrib., 25*(8), 1182-1193. doi:https://doi.org/10.1111/ddi.12940

An, S.-I., Park, H.-J., Kim, S.-K., Shin, J., Yeh, S.-W., & Kug, J.-S. (2022). Intensity changes of Indian Ocean dipole mode in a carbon dioxide removal scenario. *npj Climate and Atmospheric Science, 5*(20). doi:https://doi.org/10.1038/s41612-022-00246-6

Becker, E. A., Carretta, J. V., Forney, K. A., Barlow, J., Brodie, S., Hoopes, R., . . . Hazen, E. L. (2020). Performance evaluation of cetacean species distribution models developed using generalized additive models and boosted regression trees. *Ecology and Evolution, 10*, 5759–5784. doi:https://doi.org/10.1002/ece3.6316

Brodie, S., Jacox, M. G., Bograd, S. J., Welch, H., Dewar, H., Scales, K. L., . . . Hazen, E. L. (2018). Integrating Dynamic Subsurface Habitat Metrics Into Species Distribution Models. *Frontiers in Marine Science, 5*(219).

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. New York: Springer-Verlag.

Chambault, P., Dalleau, M., Nicet, J.-B., Mouquet, P., Ballorain, K., Jean, C., . . . Bourjea, J. (2020). Contrasted habitats and individual plasticity drive the fine scale movements of juvenile green turtles in coastal ecosystems. *Mov. Ecol., 8*(1). doi:https://doi.org/10.1186/s40462-019-0184-2

Chambault, P., Fossette, S., Heide-Jørgensen, M. P., Jouannet, D., & Vély, M. (2021a). Predicting seasonal movements and distribution of the sperm whale using machine learning algorithms. *Ecology and Evolution, 11*(3), 1432-1445. doi:https://doi.org/10.1002/ece3.7154

Chambault, P., Gaspar, P., & Dell'Amico, F. (2021b). Ecological Trap or Favorable Habitat? First Evidence That Immature Sea Turtles May Survive at Their Range-Limits in the North-East Atlantic. *Frontiers in Marine Science, 8*, 736604. doi:https://doi.org/10.3389/fmars.2021.736604

Chambault, P., Hattab, T., Mouquet, P., Bajjouk, T., Jean, C., Ballorain, K., . . . Bourjea, J. (2021c). A methodological framework to predict the individual and population-level distributions from tracking data. *Ecography, 44*(5), 766-777. doi:https://doi.org/10.1111/ecog.05436

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research, 16*. doi:https://doi.org/10.1613/jair.953

Colman, L. P., Sampaio, C. L. S., Weber, M. I., & Castilhos, J. C. d. (2014). Diet of Olive Ridley Sea Turtles, *Lepidochelys olivacea*, in the Waters of Sergipe, Brazil. *Chelonian Conservation and Biology, 13*(2), 266-271. doi:https://doi.org/10.2744/CCB-1061.1

Cutler, D. R., Jr., T. C. E., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology, 88*(11), 2783-2792. doi:https://doi.org/10.1890/07-0539.1

Esteban, N., Mortimer, J. A., Stokes, H. J., Laloë, J. O., Unsworth, R. K. F., & Hays, G. C. (2020). A global review of green turtle diet: sea surface temperature as a potential driver of omnivory levels. *Marine Biology, 167*, 183. doi:https://doi.org/10.1007/s00227-020-03786-8

Hazen, E. L., Abrahms, B., Brodie, S., Carroll, G., Welch, H., & Bograd, S. J. (2021). Where did they not go? Considerations for generating pseudo-absences for telemetry-based habitat models. *Movement Ecology, 9*, 5. doi:https://doi.org/10.1186/s40462-021-00240-2

Hazen, E. L., Palacios, D. M., Forney, K. A., Howell, E. A., Becker, E., Hoover, A. L., . . . Bailey, H. (2017). WhaleWatch: a dynamic management tool for predicting blue whale density in the California Current. *Journal of Applied Ecology, 54*, 1415-1428.

Hazen, E. L., Scales, K. L., Maxwell, S. M., Briscoe, D. K., Welch, H., Bograd, S. J., . . . Lewison, R. L. (2018). A dynamic ocean management tool to reduce bycatch and support sustainable fisheries. *Sci. Adv., 4*(5), eaar3001. doi:https://doi.org/10.1126/sciadv.aar3001

He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Trans Knowl Data, 21*(9), 1263−1284. doi:https://doi.org/10.1109/TKDE.2008.239

Howell, E. A., Kobayashi, D. R., Parker, D. M., Balazs, G. H., & Polovina, J. J. (2008). TurtleWatch: a tool to aid in the bycatch reduction of loggerhead turtles Caretta caretta in the Hawaii-based pelagic longline fishery. *Endanger. Species Res., 5*(2-3), 267-278. doi:https://doi.org/10.3354/esr00096

Iturbide, M., Bedia, J., Herrera, S., Hierro, O. d., Pinto, M., & Gutiérrez, J. M. (2015). A framework for species distribution modelling with improved pseudo-absence generation. *Ecological Modelling, 312*, 166-174. doi:https://doi.org/10.1016/j.ecolmodel.2015.05.018

Maxwell, S. M., Hazen, E. L., Lewison, R. L., Dunn, D. C., Bailey, H., Bograd, S. J., . . . Crowder, L. B. (2015). Dynamic ocean management: Defining and conceptualizing real-time management of the ocean. *Mar. Policy, 58*, 42-50. doi:https://doi.org/10.1016/j.marpol.2015.03.014

O'Toole, M., Queiroz, N., Humphries, N. E., Sims, D. W., & Sequeira, A. M. M. (2021). Quantifying effects of tracking data bias on species distribution models. *Methods in Ecology & Evolution, 12*(1), 170-181. doi:https://doi.org/10.1111/2041-210X.13507

R Core Team. (2022). R: A language and environment for statistical computing (Version R version 4.2.1). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Raymond, B., Lea, M.-A., Patterson, T., Andrews-Goff, V., Sharples, R., Charrassin, J.-B., . . . Hindell, M. A. (2015). Important marine habitat off east Antarctica revealed by two decades of multi-species predator tracking. *Ecography, 38*, 121–129. doi:https://doi.org/10.1111/ecog.01021

Revelles, M., Cardona, L., Aguilar, A., & Fernández, G. (2007). The diet of pelagic loggerhead sea turtles (*Caretta caretta*) off the Balearic archipelago (western Mediterranean): relevance of long-line baits. *Journal of the Marine Biological Association of the United Kingdom, 87*(3), 805-813. doi:https://doi.org/10.1017/S0025315407054707

Saji, N. H., Goswami, B. N., Vinayachandran, & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature, 401*, 360–363. doi:https://doi.org/10.1038/43854

Scales, K. L., Miller, P. I., Ingram, S. N., Hazen, E. L., Bograd, S. J., & Phillips, R. A. (2016). Identifying predictable foraging habitats for a wide-ranging marine predator using ensemble ecological niche models. *Diversity and Distributions, 22*(2), 212-224. doi:https://doi.org/10.1111/ddi.12389

Scales, K. L., Miller, P. I., Varo-Cruz, N., Hodgson, D. J., Hawkes, L. A., & Godley, B. J. (2015). Oceanic loggerhead turtles *Caretta caretta* associate with thermal fronts: Evidence from the Canary Current Large Marine Ecosystem. *Marine Ecology Progress Series, 519*, 195-207. doi:https://doi.org/10.3354/meps11075

Scales, K. L., Schorr, G. S., Hazen, E. L., Bograd, S. J., Miller, P. I., Andrews, R. D., . . . Falcone, E. A. (2017). Should I stay or should I go? Modelling year-round habitat suitability and drivers of residency for fin whales in the California Current. *Diversity and Distributions, 23*(10), 1204-1215. doi:https://doi.org/10.1111/ddi.12611

Siders, Z. A., Ducharme-Barth, N. D., Carvalho, F., Kobayashi, D., Martin, S., Raynor, J., . . . Ahrens, R. N. M. (2020). Ensemble Random Forests as a tool for modeling rare occurrences. *Endang. Species. Res., 43*, 183–197. doi:https://doi.org/10.3354/esr01060

Stock, B. C., Ward, E. J., Eguchi, T., Jannot, J. E., Thorson, J. T., Feist, B. E., & Semmens, B. X. (2020). Comparing predictions of fisheries bycatch using multiple spatiotemporal species distribution model frameworks. *Can. J. Fish. Aquat. Sci., 77*, 146–163. doi:https://doi.org/10.1139/cjfas-2018-028

Stock, B. C., Ward, E. J., Thorson, J. T., Jannot, J. E., & Semmens, B. X. (2019). The utility of spatial model-based estimators of unobserved bycatch. *ICES Journal of Marine Science, 76*(1), 255–267. doi:https://doi.org/10.1093/icesjms/fsy153

Virgili, A., Authier, M., Boisseau, O., Cañadas, A., Claridge, D., Cole, T., . . . Ridoux, V. (2019). Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale - Large-scale modelling of deep-diving cetacean habitats. *Global Ecology and Biogeography, 28*, 300–314.

Wallace, B. P., Kot, C. Y., Dimatteo, A. D., Lee, T., Crowder, L. B., & Lewison, R. L. (2013). Impacts of fisheries bycatch on marine turtle populations worldwide: toward conservation and research priorities. *Ecosphere, 4*(3), 1-49. doi:https://doi.org/10.1890/ES12-00388.1

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution, 1*, 3–14. doi:https://doi.org/10.1111/j.2041-210X.2009.00001.x

Supplementary information

Supplementary information includes Figs S1-3.

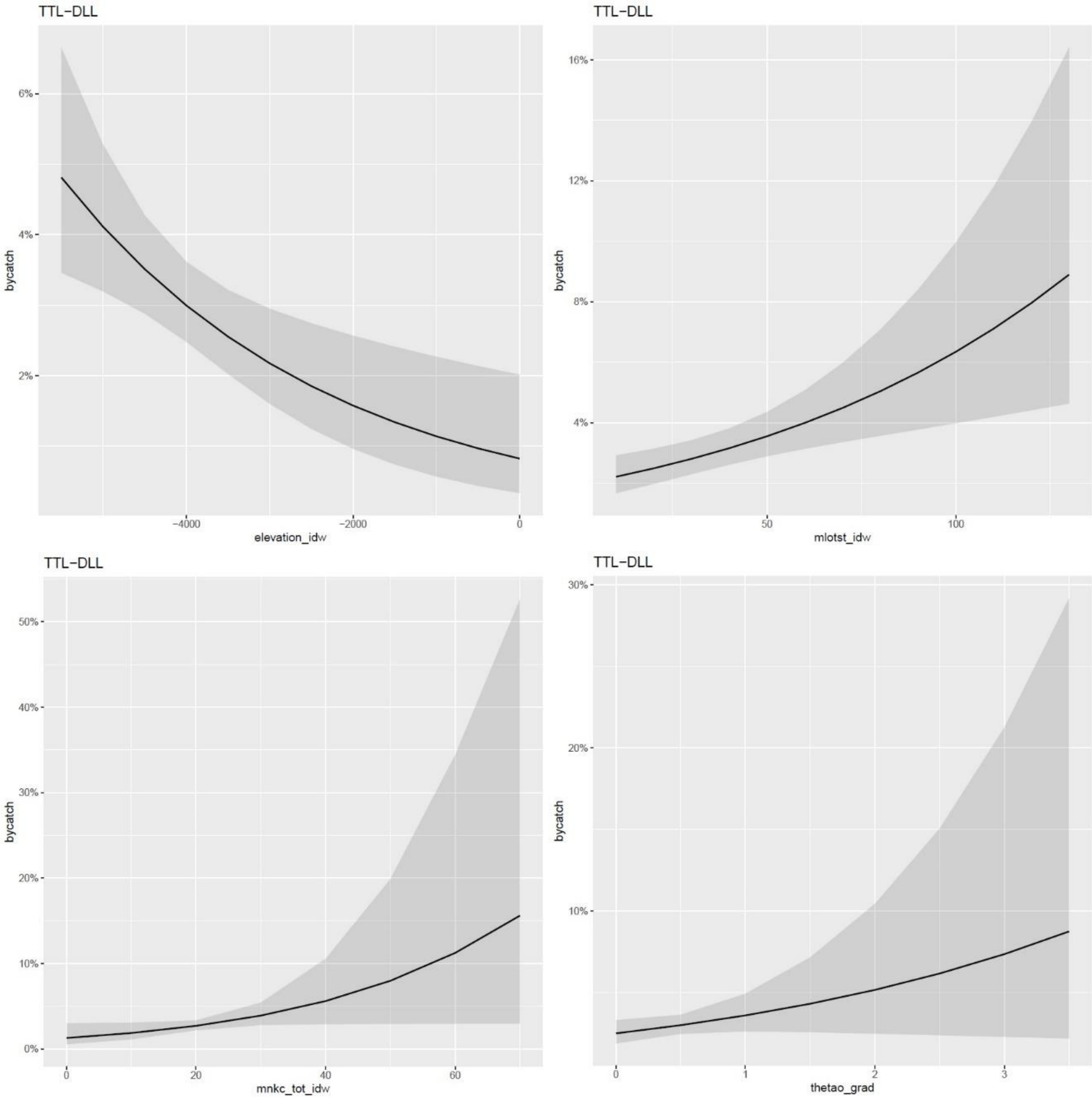
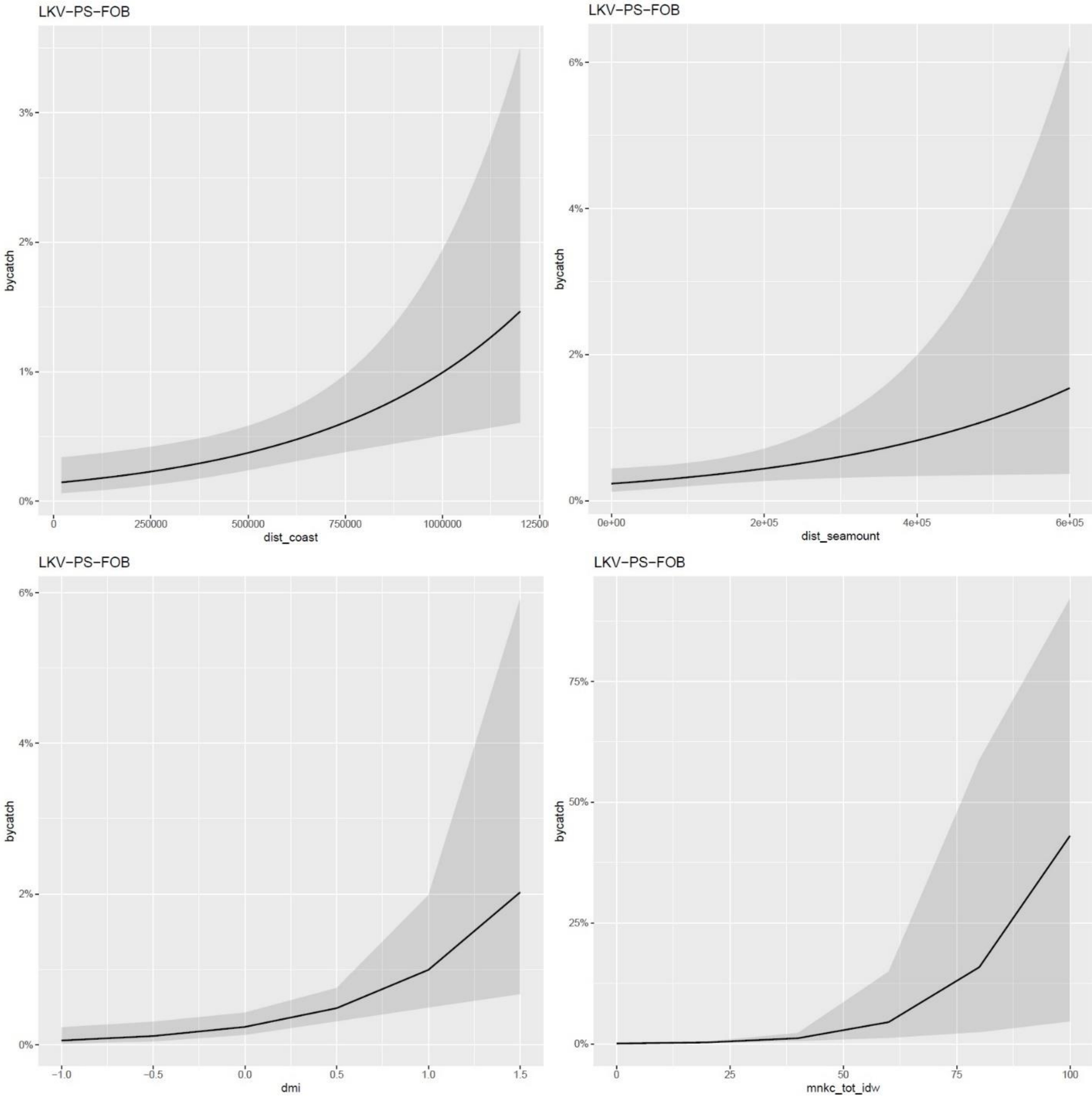


Figure S1: Response curve of the environmental variables from the best logistic model for loggerhead turtles in drifting longlines (TTL-DLL).



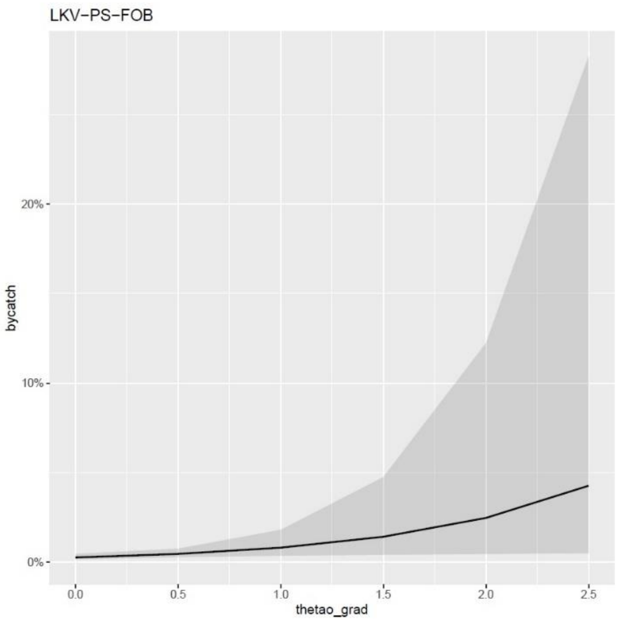


Figure S2: Response curve of the environmental variables from the best logistic model for olive ridley turtles in purse seines on floating object-associated tuna school (LKV-PS-FOB).

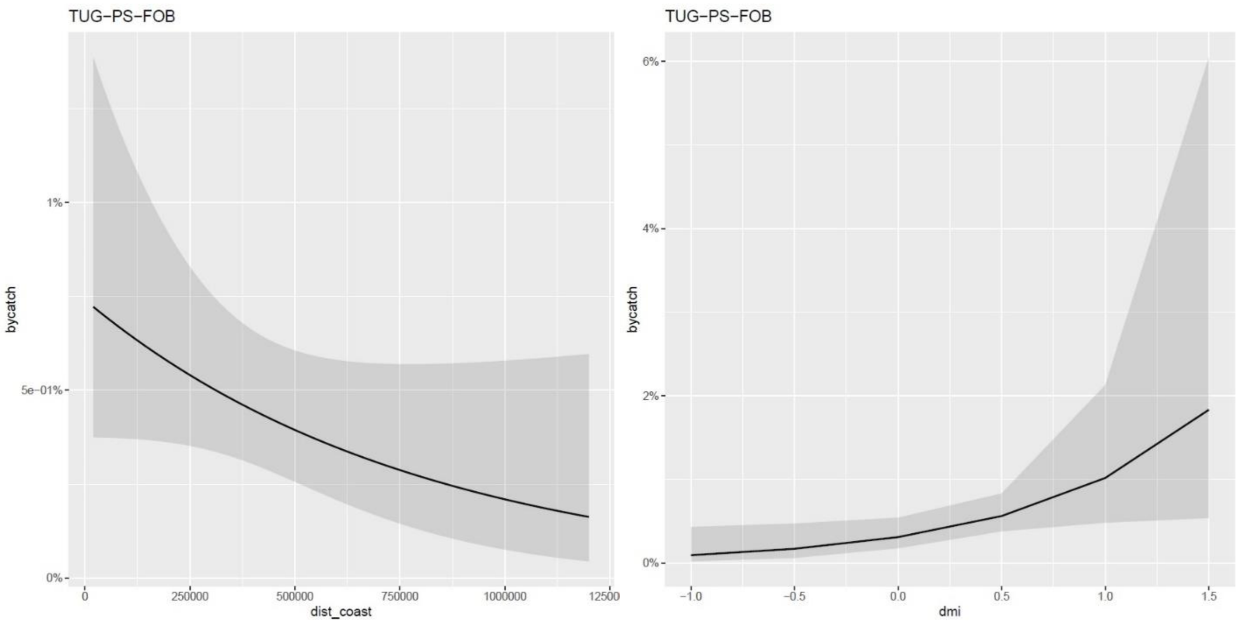


Figure S3: Response curve of the environmental variables from the best logistic model for green turtles in purse seines on floating object-associated tuna school (TUG-PS-FOB).