



Vulnerability of the Critically Endangered leatherback turtle to fisheries bycatch in the eastern Pacific Ocean. I. A machine-learning species distribution model

Jon Lopez^{1,*}, Shane Griffiths¹, Bryan P. Wallace^{2,3,4,*}, Verónica Cáceres⁴, Luz Helena Rodríguez⁴, Marino Abrego⁵, Joanna Alfaro-Shigueto^{6,7,8}, Sandra Andraka⁹, María José Brito¹⁰, Leslie Camila Bustos¹¹, Iliá Cari¹², José Miguel Carvajal¹³, Ljubitzá Clavijo¹², Luis Cocas¹¹, Nelly de Paz¹⁴, Marco Herrera¹⁰, Jeffrey C. Mangel^{7,8}, Miguel Pérez-Huaripata¹⁵, Rotney Piedra¹⁶, Javier Antonio Quiñones Dávila¹⁵, Liliana Rendón⁹, Juan M. Rguez-Baron^{17,18}, Heriberto Santana¹⁹, Jenifer Suárez²⁰, Callie Veelenturf²¹, Rodrigo Vega¹², Patricia Zárate¹²

¹Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Drive, La Jolla, CA 92037, USA

Full author addresses are given in the Appendix

ABSTRACT: The Eastern Pacific population of leatherback turtles *Dermochelys coriacea* is Critically Endangered, with incidental capture in coastal and pelagic fisheries as one of the major causes. Given the population's broad geographic range, status, and extensive overlap with fisheries throughout the region, identifying areas of high importance is essential for effective conservation and management. In this study, we created a machine-learning species distribution model trained with remotely sensed environmental data and fishery-dependent leatherback presence ($n = 1088$) and absence data ($> 500\,000$ fishing sets with no turtle observations) from industrial and small-scale fisheries that operated in the eastern Pacific Ocean between 1995 and 2020. The data were obtained through a participatory collaboration between the Inter-American Convention for the Protection and Conservation of Sea Turtles and the Inter-American Tropical Tuna Commission as well as non-governmental organizations to support the quantification of leatherback vulnerability to fisheries bycatch. A daily process was applied to predict the probability of leatherback occurrence as a function of dynamic and static environmental covariates. Coastal areas throughout the region were highlighted as important habitats, particularly highly productive feeding areas over the continental shelf of Ecuador, Peru, and offshore from Chile, and breeding areas off Mexico and Central America. Our model served as the basis to quantify leatherback vulnerability to fisheries bycatch and the potential efficacy of conservation and management measures (Griffiths & Wallace et al. 2024; Endang Species Res 53:295–326). In addition, this approach can provide a modeling framework for other data-limited vulnerable populations and species.

KEY WORDS: *Dermochelys coriacea* · Species distribution model · Probability of occurrence · Boosted regression trees · Conservation priority-setting

1. INTRODUCTION

Marine turtles spend most of their lives at sea but are tied to terrestrial habitats (i.e. nesting beaches) for re-

production. This life history strategy exposes these species to indirect (e.g. climate change and habitat loss) and direct (e.g. fishing, pollution and marine debris, egg consumption, coastal development) anthro-

*Corresponding authors: Jlopez@iattc.org

pogenic threats that have caused declines in many populations (Wallace et al. 2011, Northwest Atlantic Leatherback Working Group 2019, Wibbels & Bevan 2019). Consequently, conservation efforts have increased in many regions of the world to address these threats, with varying degrees of success (e.g. Chaloupka et al. 2008, Casale & Tucker 2017, Mazaris et al. 2017).

Of particular conservation concern is the leatherback turtle *Dermochelys coriacea*, the largest (typically > 1.5 m curved carapace length) and most geographically widespread of all marine turtle species (Eckert et al. 2012). Distributed circumglobally from tropical to temperate regions and regularly occurring in coastal as well as high-seas areas, the species comprises 7 regional management units (RMUs) or sub-populations (Wallace et al. 2023). Two RMUs—the East Pacific (EP) and West Pacific—exist in the Pacific Ocean, both of which are currently classified as Critically Endangered on the IUCN Red List (Tiwari et al. 2013, Wallace et al. 2013). The EP leatherback RMU has declined by over 90% since the 1980s, largely due to unsustainable levels of incidental mortality by industrialized and artisanal fisheries, which mainly affect sub-adults and adults, as well as egg consumption by humans (Laúd OPO Network 2020).

Given that industrial and artisanal tuna (e.g. yellowfin *Thunnus albacares*, albacore *T. alalunga*, big-eye *T. obesus*, skipjack *Katsuwonus pelamis*) purse-seine and longline fisheries cover a high proportion of the species' distribution in the east Pacific Ocean (EPO) (Wallace et al. 2023), they unavoidably interact with leatherback turtles and other sea turtle species during their normal fishing operations as they target tunas, billfish, and other species that share similar epipelagic habitats (IATTC 2020). The Inter-American Tropical Tuna Commission (IATTC) is the Regional Fisheries Management Organization that is responsible for the long-term conservation and sustainable management of fisheries that target tuna and tuna-like species in the EPO. Since the Antigua Convention entered into force in 2010, which required a significantly broader ecosystem-based approach to management by the IATTC, several resolutions pertaining to the conservation and management of various sensitive bycatch species have been implemented, including sea turtles (IATTC Resolution C-07-03). However, in light of the declining population of leatherback turtles in the EPO, the IATTC implemented more stringent conservation measures in 2021 to mitigate fishery impacts on sea turtles (IATTC Resolution C-19-04).

Further, the Inter-American Convention for the Protection and Conservation of Sea Turtles (IAC) is a

binding intergovernmental treaty that provides the legal framework for states of North, Central, and South America and the Caribbean Sea to take actions to benefit sea turtles, in both nesting beaches and the Parties' territorial waters. Concerned with the critical status of leatherback turtles in the EPO, the IAC adopted Resolution CIT-COP7-2015-R2 in 2015, updated in 2022 to Resolution CIT-COP10-2020-R6, which requests IAC Parties to implement or improve measures to reduce bycatch of leatherback sea turtles in the eastern Pacific fisheries based on the best scientific information available and using recommendations from IAC Resolution CIT-COP10-2022-R7 to exercise FAO guidelines to reduce sea turtle mortality in fishing operations (FAO Fisheries Department 2009).

Assessing fisheries impacts on bycatch species is challenging due to the frequent lack of reliable biological and catch information, especially for species of little or no commercial value or in data-limited settings. Therefore, assessing bycatch species using traditional stock assessment approaches is often both cost-prohibitive and impractical, thus requiring alternative approaches. To address this issue, Griffiths et al. (2019) developed a flexible, spatially explicit, quantitative ecological risk assessment approach—Ecological Assessment of Sustainable Impacts of Fisheries (EASI-Fish)—to quantify the cumulative impacts of multiple fisheries on data-limited bycatch species, such as sea turtles.

Because the extent of areal overlap between the species and each interacting fishery is of critical importance in the EASI-Fish approach, a reliable species distribution model (SDM) is required. SDMs are built to describe the relationship between a species and environmental conditions and can predict how environmental variability may affect their distribution and habitat choice (Elith & Leathwick 2009). Although SDMs have been widely applied in predicting suitable habitats for marine species over the past few decades (Melo-Merino et al. 2020), only a few studies have focused on bycatch species that are data-poor and are of particularly high conservation importance (e.g. Sequeira et al. 2014, Abrahms et al. 2019, Lezama-Ochoa et al. 2020, Lopez et al. 2020). For example, a simple environmental envelope model was developed for NOAA's TurtleWatch tool (Howell et al. 2008, 2015) to mitigate loggerhead *Caretta caretta* and leatherback turtle bycatch by longline fisheries in the central north Pacific Ocean.

For rarely encountered bycatch species, whether their rarity is due to fishing gear selectivity issues, non-reporting, or declining population size, there is often a small number of observations from which to develop

an SDM. Therefore, this limits the types of models available to develop an SDM that can make use of scant presence records. Machine-learning algorithms, including boosted regression trees (BRT) (Elith et al. 2006), are a relatively new suite of powerful tools that can accommodate non-linear relationships, high-dimensional large data sets, imbalanced classes, and limited species occurrences (Elith et al. 2008, Mi et al. 2017).

The goal of this study was to generate a reliable high-resolution SDM that would support the evaluation of EP leatherback vulnerability to the impacts of fisheries as well as the potential efficacy of conservation management measures (e.g. IATTC C-19-04) using the EASI-Fish approach. This study was part of a multi-year, collaborative effort between IATTC and IAC representatives, under a 2011 Memorandum of Understanding (MoU) between the 2 conventions to inform conservation and management of EP leatherback turtles in the EPO. Previous research using satellite telemetry has estimated overlaps between Atlantic leatherbacks and the potential risk of accidental interactions with fisheries (e.g. James et al. 2005, Fossette et al. 2014). Similarly, for EP leatherbacks, previous studies have developed potential methods for producing SDMs using either individual-based satellite telemetry data (Hoover et al. 2019), presence-only observation data (Degenford et al. 2021), or both (Liang et al. 2023). Furthermore, there is clear value in a detailed investigation of inter- and intra-annual leatherback distributions in relation to large-scale environmental cycles (e.g. El Niño–Southern Oscillation [ENSO] regimes, climate change) to inform adaptive management options that reflect these dynamics (e.g. Hazen et al. 2018, Willis-Norton et al. 2015, Pons et al. 2022). Leatherback movements, habitat use, and life history are known to be strongly influenced by environmental conditions, specifically how those conditions affect resource availability (Saba et al. 2007, 2008, Shillinger et al. 2008, 2010, Wallace & Saba 2009, Bailey et al. 2012, Hoover et al. 2019). Thus, predictions of leatherback occurrence in time and space at different scales would be highly informative and useful for fine-tuning conservation strategies in the EPO.

In this study, we developed a hierarchical machine-learning modeling approach that used a region-wide presence–absence data set for EP leatherback turtles and incorporated different predictive variables and modeling configurations to (1) understand the potential distribution of the species at different spatio-temporal scales, (2) identify the environmental preferences of the species, and (3) develop a final prediction map describing the most plausible distribution for the

species that will be used in a concurrent study to assess the species' vulnerability to fishery interactions. This paper describes the development of a novel SDM that was then used to quantify and mitigate the impacts of EPO fisheries on EP leatherback turtles, helps inform their vulnerability status, and guide the development of appropriate conservation and management decisions (Griffiths & Wallace et al. 2024, this volume). It was also envisaged that if the model was successful for leatherback turtles, it could also be applied to other data-limited vulnerable taxa.

2. METHODS

All data processing and analytical work was carried out using R version 3.4.3 (R Core Team 2017)).

2.1. Fisheries observer data

We used 26 years (1995–2020) of observer and log-book data collected from a variety of industrial and small-scale coastal (artisanal) fisheries (Table 1) operating in 6 countries and the high seas within the IATTC convention area—defined as the region from the Pacific coast of the Americas to 150° W between 50° S and 50° N. Our data set included observations of leatherback turtle presence as well as absence during fishing operations compiled by a participatory collaboration coordinated by the IATTC and IAC to undertake the EASI-Fish leatherback vulnerability assessment.

Nearly 90% of high-seas presence–absence data were derived from large-scale tuna longline fishing vessels (≥ 24 m, hereafter called the 'industrial longline fishery'; 484 active vessels in 2022; <https://iattc.org/en-US/management>) and the industrial tropical tuna large purse-seine fishing fleet (Class 6 with a carrying capacity of > 363 mt, 180 active vessels in 2022; <https://iattc.org/en-US/management>). The data include set-level information on leatherback turtle interactions along with location, date, and time of the observation. The distribution of industrial tuna fishing effort was mostly concentrated ($> 75\%$ of sets) between 20° N and 20° S during the study period. The data for these fleets were collected by IATTC on-board scientific observers or submitted to the IATTC by its members under Resolution C-19-08. The observer coverage rate was close to 100% for purse-seine vessels of Class 6 and approximately 5% for the industrial longline fishery.

In contrast to the industrial fisheries in the EPO, catch and effort by the numerous artisanal fleets that

Table 1. Data sources and period of coverage of data used to build the leatherback species distribution model for the East Pacific Ocean

Country	Gear	First year	Last year	Presence only	Abundance	Effort	No. of presences	No. of ind.	No. total sets	% of presences	Source
Chile	Purse-seine	2015	2019	No	No	—	3	3	4396	0.07	Observers
Chile	Industrial longline	2001	2018	No	Yes	Yes (No. of hooks)	327	365	13828	2.36	Observers
Chile	Artisanal longline	2002	2018	No	Yes	Yes (No. of hooks)	59	62	1831	3.22	Observers
Chile	Artisanal longline (espinel)	2010	2019	No	No	Yes (No. of hooks)	2	2	564	0.35	Observers
Chile	Artisanal gillnet	2007	2019	No	Yes	No	22	24	1399	1.57	Observers
Colombia	Gillnet	2017	2018	Yes	No	No	3	3	3	—	Observers
Colombia	Longline	2018	2018	Yes	No	No	2	2	2	—	Observers
IATTC	Purse-seine	1995	2020	No	Yes	No	272	274	532857	0.05	Observers
IATTC	Longline	2013	2020	No	Yes	No	67	67	24005	0.28	Observers
Panama	Purse seine / longline / gillnet	2018	2020	Yes	No	No	10	10	10	—	Observers
Peru	Longline (surface), gillnet	2001	2019	Yes	No	—	186	186	186	—	Pro-Delphinus
Ecuador	Purse-seine	2019	2020	No	No	—	3	3	2746	0.11	Observers
Ecuador	Longline (bottom)	2017	2020	No	No	No	0	0	766	0.00	Observers
Ecuador	Longline (surface)	2019	2020	No	No	No	2	2	1667	0.12	Observers
Peru	Driftnet / gillnet	1997	2015	Yes	No	No	141	141	141	—	IMARPE/ACOREMA
Peru	Driftnet / gillnet	2013	2020	Yes	Yes	No	21	21	21	—	IMARPE (LAMBAYEQUE)
WWF (various)	Longline	2004	2009	No	Yes	Yes (Various)	20	20	7539	0.27	WWF-IATTC
Costa Rica All	Longline	2005	2012	No	Yes	Yes (No. of hooks)	5	5	2602	0.19	Observers
		1995	2020				1145	1190	594563	0.19	

operate within the exclusive economic zones (EEZs) of countries in the EPO generally have very low (if any) observer coverage and are poorly documented in general. However, based on available data (i.e. collected by opportunistic or sporadic observations, and port-based interviews with fishermen), leatherback turtles have been shown to be heavily impacted by coastal artisanal gillnet and longline fisheries, particularly in foraging areas but also in migratory and reproduction areas (Frazier & Brito 1990, Alfaro-Shigueto et al. 2011, 2018, Quiñones et al. 2021). Reasonably detailed effort data for artisanal longline vessels throughout Central America was available from

IATTC's long-term research program that examined the effects of different hook types on marine turtle bycatch rates, reported in part by Andracka et al. (2013). In addition, unpublished data pertaining to leatherback turtle interaction and fishing effort information for several artisanal fisheries operating in territorial waters of 6 countries in the EPO were opportunistically compiled (Table 1).

Duplicated records, data outside the EPO, and observations without reliable date and location information were removed from the data set ($n = 57$ records, $\sim 5\%$ of the initial presences). The final data set included 1088 leatherback records from nearly

575 000 fishing sets (i.e. 0.19% of sets containing at least one leatherback observation) from all years with available data (Table 1, Fig. 1).

2.2. Predictive variables

A total of 23 variables were included in SDMs, which included 3 spatio-temporal variables, 11 surface variables, 2 subsurface variables, and 2 static variables (Table 2). The 3 spatio-temporal variables included location and date of set, as seasonality can affect catches. Spatio-temporal variables can be confounded with environmental factors and reflect certain natural processes not captured by the environmental (i.e. surface and subsurface) variables. The majority of environmental data was sourced from daily or weekly fields of global data assimilative models (i.e. assimilated data derived from satellites and *in situ* platforms) for the IATTC convention area at 0.25° (~25 km²) resolution (available at <http://marine.copernicus.eu/> and <https://www.aviso.altimetry.fr/>). The 0.25° spatial resolution, combined with a fine temporal scale, is considered adequate for habitat modeling (Scales et al. 2017a).

The 11 surface variables chosen included sea surface temperature (SST) and its gradient (SST_grad; calculated as the change in temperature at the same pixel over a period of 7 d), salinity (Sal), sea surface height (SSH), current speed (Vel), current direction (Dir), eddy kinetic energy (Eke), finite size Lyapunov exponents (FSLE), front index (FrontIndex; estimated as a count of the front pixels in the grid cell for the 7 d window), chlorophyll *a* (chl), and chl *a* gradient (chl_grad; computed as the difference in chl *a* concentration in the same pixel over a 7 d period).

The 2 subsurface variables included temperature at 100 m depth (SST₁₀₀) and mixed layer depth (MLD). SST₁₀₀ and MLD have proven to be helpful to improve SDMs for large pelagic species (Brodie et al. 2018) and help describe the 2-dimensional (i.e. vertical and horizontal space) structure of the water column properties.

The 2 static variables included bathymetry (Depth) and the distance

to land (LandDistance). These variables were extracted from the Global Marine Environmental Data-sets (Basher et al. preprint, doi:10.5194/essd-2018-64) and MARSPEC Ocean Climate Database (Sbrocco & Barber 2013), respectively, and have been shown to be important in defining leatherback turtle habitat (e.g. Hazen et al. 2018, Robinson et al. 2016, Willis-Norton et al. 2015).

2.3. Model development

In the interest of robustness and to inform comparisons, we took a hierarchical multi-model approach, building 6 presence–absence (catch vs. zero catch per set, binary response) model configurations with each set of variables, from the simplest to the most complex models. The following model configurations were established with all presence and absence data:

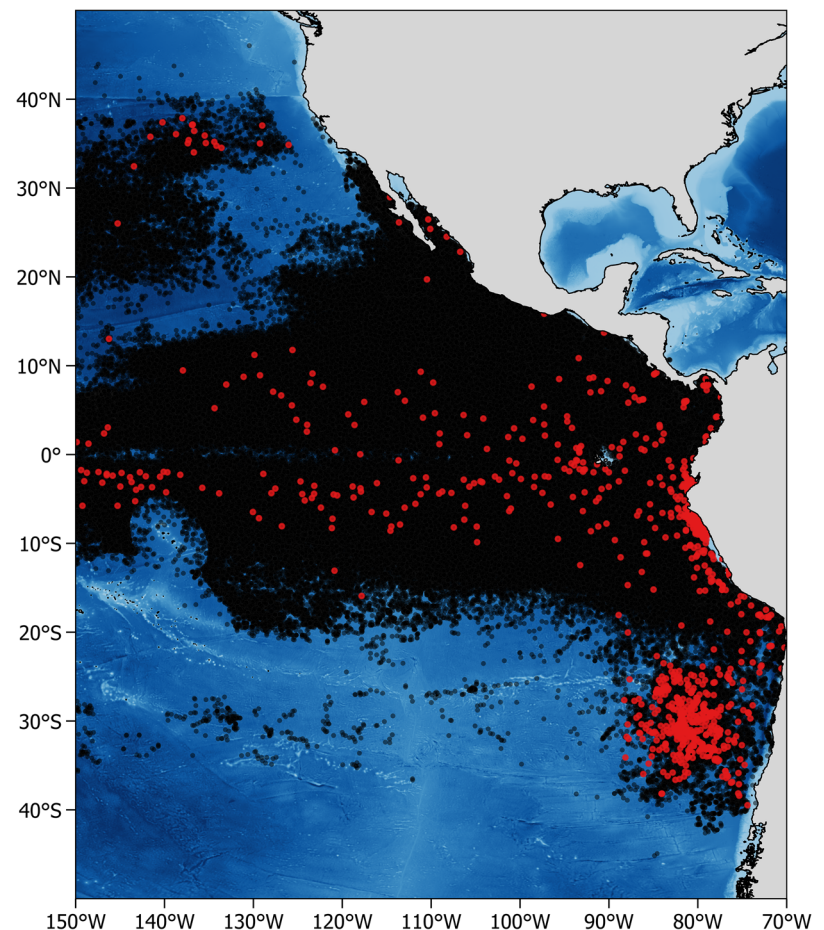


Fig. 1. Distribution of leatherback sea turtle presences (red points) and absences (black points) collected by on-board observers and logbooks for industrial longline and purse-seine fisheries and small-scale longline and surface gillnet fisheries in the Eastern Pacific Ocean between 1995 and 2020

Table 2. Comparing data sources and spatio-temporal resolutions for variables used in species distribution models. D: daily; W: weekly

Variable	Unit	Spatial resolution	Temporal resolution	Source
Spatio-temporal				
Latitude	Decimal degree	GPS	—	—
Longitude	Decimal degree	GPS	—	—
Day of the year	Calendar day	—	D	Processed
Environmental surface				
Sea surface temperature	°C	1/4°	D	CMEMS
Sea surface temperature gradient	°C	1/4°	D	Processed
Salinity	g kg ⁻¹	1/4°	D	CMEMS
Sea surface height	m	1/4°	D	CMEMS
Current speed	m s ⁻¹	1/4°	D	Processed
Current direction	Degrees	1/4°	D	Processed
Eddy kinetic energy	cm ² s ⁻²	1/4°	D	Processed
Finite size Lyapunov exponents	d ⁻¹	1/25°	D	AVISO
Front index	Count of front pixels per cell for the 7 d window	1/4°	W	Processed
Chlorophyll	mg m ⁻³	1/4°	W	CMEMS
Chlorophyll gradient	mg m ⁻³	1/4°	W	Processed
Environmental subsurface				
Temperature at 100 m	°C	1/4°	D	CMEMS
Mixed layer depth		1/4°	D	CMEMS
Static				
Bathymetry		1/12°	—	GMED
Distance to coast		1/120°	—	MARSPEC
				(Sbrocco & Barber 2013)

(1) spatio-temporal, (2) surface, (3) subsurface, (4) environmental (surface + subsurface), (5) static, and (6) full (environmental and static) (Fig. 2). Because the full model had the best performance metrics (see Table 3), subsequent models using different proportions of presence to absence data were only established with this set of variables (full; environmental and static variables) (see Section 2.3.3 for details).

2.3.1. Model building

BRTs are a flexible classification and regression algorithm based on machine-learning principles (Elith et al. 2006, De'ath 2007). Consequently, some of the caveats of more commonly used techniques, such as generalized linear models (GLMMs) or generalized additive mixed models (GAMM), are not applicable. BRTs have the particular advantage of being tolerant of missing values, outliers, correlation, collinearity, non-independence, and allowing for the inclusion of irrelevant predictors (Leathwick et al. 2006). BRTs are also designed to accommodate non-linear relationships, large high-dimensional data sets, imbalanced classes, and limited species occurrences (Elith et al.

2008, Mi et al. 2017). While GLMMs and GAMMs seek to fit the most parsimonious model to a data set, BRTs combine stochastic predictions of many simple models (i.e. many shallow classification trees) to maximize robustness and predictive performance to reduce associated error (Scales et al. 2017b). Accordingly, we fitted BRTs with all available sets of covariates. In the past, authors also fitted GAMM and random forest (RF) models to presence–absence data for other species to compare and better understand consistency and interpretation between algorithms (e.g. Lopez et al. 2019). In these cases, BRTs performed better than GAMMs and had very similar performance to the RFs. As such, we decided to use BRTs to build all the models in this study, which were implemented using the R package 'dismo' (Hijmans et al. 2017).

In fitting BRTs, we adapted the protocols outlined by Elith et al. (2008), Scales et al. (2017b), Brodie et al. (2018), and Hazen et al. (2018). Presence–absence models were built with a binomial (Bernoulli) distribution. We used a tree complexity of 3, a bag fraction of 0.7, and conducted sensitivity analyses on learning rate ('shrinkage') for each model set, aiming for at least 1000 trees in final model configurations. The sensitivity runs determined 0.01 as the learning rate to be used in all the models, except for the model with

the same number of presences and absences (i.e. 50:50 presence–absence ratio; see Section 2.3.3), where a value of 0.005 was used. Tree complexity refers to the number of nodes in a tree, which constrains the maximum size of each of the regression trees that together make up a boosted regression tree model. By controlling the number of nodes or branches, tree complexity also sets the maximum number of interactions between predictor variables that are possible (i.e. 3 in this case, as 2- and perhaps 3-way interactions among variables may be important, but higher orders are unnecessary in fisheries contexts; Soykan et al. 2014). Bag fraction refers to the proportion of the data that is randomly used for model building at each step, which usually ranges between 0.6 and 0.75 (Elith et al. 2008). The stochasticity that this step provides to the model-building process improves model performance (Soykan et al. 2014).

The potential for model simplification was evaluated with the function 'gbm.simplify'. Simplified models were fitted by re-running models without those variables that gave no evidence of improving predictive performance. Deviance explained, variable importance, and interactions between variables were also estimated for all models using the function 'gbm.interactions'. Each of these configuration settings and the performance procedures are described in detail by Elith & Leathwick (2017), Elith et al. (2008), Hazen et al. (2018), Scales et al. (2017b), and Soykan et al. (2014).

2.3.2. Model validation

Patterns derived from SDMs—particularly those produced by machine-learning techniques—are susceptible to bias if the original data is biased (see Lee-Yaw et al. 2022), which may be the case with our highly imbalanced data set from different sources (note that the leatherback turtle is not a target species but a bycatch species for all fisheries included in this study; hence, the data is expected to be less influenced by fishing strategies or other variables that can affect data collection and bias). In an attempt to address this issue, we conducted extensive model validation exercises. Independent data sets such as leatherback turtle space use derived from satellite tracking data (e.g. Shillinger et al. 2008, Bailey et al. 2012) were not made available for our analysis; such tests of how well SDM predictions agree with independent estimates of distributions are rare (Lee-Yaw et al. 2022). However, we employed multiple validation approaches as well as an examination of uncertainty or confidence around predictions by using subsets of our data set. We acknowledge that such internal validation approaches do not explicitly address biased predictions resulting from biased underlying data sets. Two cross-validation methods were used to evaluate the reliability and the predictive performance of the final models: *k*-fold cross-validation (main method) and a hold-out cross-validation (additional complementary method). These methods consist of using independent data sets for model

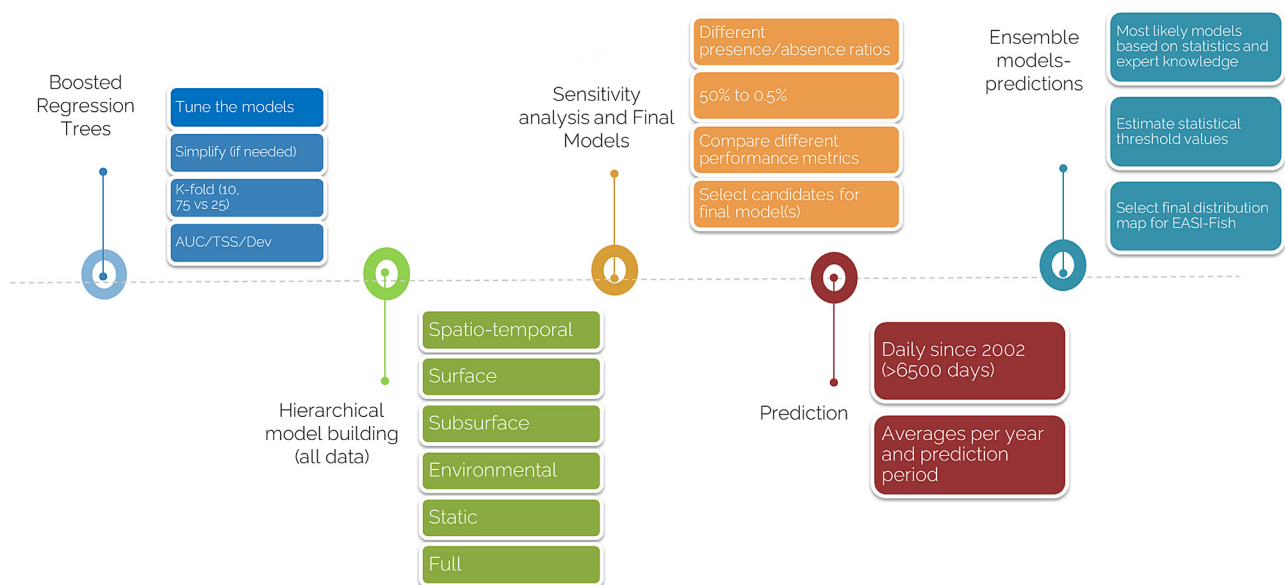


Fig. 2. Infographic description of the methods (sequential from left to right) used to develop the species distribution models. See Section 2.3 for details

building (i.e. the training data) and model validation (i.e. the test data), where data are partitioned into k equally sized segments or folds through random resampling (k -fold cross-validation) or are intentionally partitioned into different segments based on spatial-temporal properties (hold-out cross-validation). Model performance is assessed by successively removing each subset, rebuilding the model on the retained data, and predicting on the omitted data (Elith & Leathwick 2009). In this study, a $k = 4$ partitioning method was used, meaning that 75% of the observations were used for model building, and the other 25% were used for model cross-validation. In the hold-out cross-validation, there is no random shuffling of observations. Instead, data is partitioned into blocks of interest for model training, and testing is done with the remaining contrasting block of data (e.g. Becker et al. 2020). As such, data from neutral years was used to train the model and test on El Niño–La Niña data. Similarly, data was also partitioned into inshore and offshore blocks, where inshore data was used to train the model and testing offshore, and vice versa (2 different limits were explored to separate inshore and offshore data; 90° W and 100° W). Both k -fold and hold-out cross-validations avoid the overlap between training data and test data, yielding a more accurate estimate of the generalization performance of the algorithm (Villarino et al. 2015).

The predictive power and the stability (i.e. uncertainty or confidence) of the model were assessed by computing a set of diagnostic metrics. The mean area under the receiver-operating curve (AUC) (Hanley & McNeil 1982) and the mean true skill statistic (TSS) (Allouche et al. 2006) were calculated for each iteration from each confusion matrix to evaluate the predictive performance of the models, and the coefficient of variation (CV) of the predictions was used to evaluate their stability (e.g. Wang et al. 2019, Montoya-Jiménez et al. 2022, Borokini et al. 2023, Roberts et al. 2023). The AUC provides a single measure of overall model accuracy that is threshold-independent, with an AUC value of 0.5 indicating that the prediction is as good as random, whereas $AUC = 1$ indicates perfect prediction (Fielding & Bell 1997). AUC has been extensively used in SDMs and measures the ability of the model to correctly predict where a species is present or absent (Elith et al. 2006). An AUC value of >0.75 is considered to have good predictive power and is acceptable for conservation planning (Pearce & Ferrier 2000). TSS is an alternative measure of model accuracy that is threshold-dependent and not affected by the size of the validation set,

and it is an appropriate evaluative tool in cases where model predictions are formulated as presence–absence maps (Allouche et al. 2006). TSS is on a scale from -1 to $+1$, with 0 representing no predictive skill, and is calculated from the confusion matrix outputs as sensitivity plus specificity minus 1 (i.e. $TSS = sensitivity + specificity - 1$). Threshold independent and dependent statistics, such as AUC and TSS, respectively, should be used in combination when evaluating the predictive power of an SDM (Pearson et al. 2006). The model uncertainty or confidence in the resulting probability estimates was quantified using the CV, which measures the percentage of variation around the arithmetic mean of a series (i.e. the ratio of the standard error to the predicted value per observation) and was mapped to describe uncertainty alongside the predictions.

2.3.3. Sensitivity analyses

The model utilizing all data (hereafter called the 'original model') contained 573 883 observations (1088 presences; 0.19%). To determine the effect of using different proportions of presence on model performance, the relationship between the response variable and the covariates (i.e. environmental and static), as well as the generated predictions, a multi-model approach was conducted using 10 data sets, each having a different presence–absence ratio. The presence–absence ratio in the data used to build the final models was incrementally decreased from 50 to 0.5%. In each case, all presence observations were included with a variable number of randomly selected absences. For example, the 50:50 model included 1088 presences and 1088 absences while the 0.5:99.5 model included 1088 presences and 216 512 absences (see Table 4). All models (hereafter called 'final models') were run using all environmental and static variables, as per the full model, and followed the same model building and validation procedures mentioned above.

2.4. Predictions

2.4.1. Daily predictions

Daily predictions of the probability of occurrence of leatherback turtles across the IATTC convention area were conducted for 2002–2020 (i.e. 6935 daily predictions). A series of time-matched environmental data fields (both surface and subsurface as well as static

variables) were used to generate daily predictions based on the 11 final models and their best number of trees using function 'predict' in the R package 'raster' (Hijmans & van Etten 2015). Therefore, 11 different predictions were computed for each day of the time series to inform consistency and interpretation, and to visualize the effect of accounting for different proportions of presences on the predictions. The spatial resolution of the predictive surface was set to the lowest common resolution of environmental data fields (0.25°).

2.4.2. Prediction averaging and ensemble

Daily predictions ($n = 6935$) were averaged for 2002–2020 for each of the 11 final models, at the full regional scale averaged across the study period (Figs. S1 & S2 in the Supplement at www.int-res.com/articles/suppl/n053p271_supp.pdf), quarterly (Fig. S3 in the Supplement), and by ENSO phase (i.e. neutral, La Niña, El Niño; Figs. S4 & S5 in the Supplement). Different ENSO regimes were defined following NOAA's Earth System Research Laboratories Oceanic Niño Index, where index values of +0.5 or higher indicate El Niño and values of -0.5 or lower indicate La Niña phases (<https://www.ncei.noaa.gov/access/monitoring/enso/sst#oni>). Visual inspection of predictions and exploration of performance metrics and the relationships between the response variable and covariates suggested 2 groups of similar models: (1) 6 models with a presence–absence ratio ranging from 50 to 10% and (2) 5 models with a presence–absence ratio ranging from 5 to 0.19%. We concluded that the models of the first group best reflected the full suite of existing leatherback turtle distribution data (e.g. Shillinger et al. 2008, Quiñones et al. 2021) and SDMs (Hoover et al. 2019, Degenford et al. 2021, Liang et al. 2023) in the region. Therefore, an ensemble model was created using the average predictions from all models having a presence–absence ratio of 50 to 10%. Similarly, the model with a 25% presence–absence ratio was also identified by both expert opinion and performance metrics as a plausible model (hereafter called the 'reference model'). Therefore, an ensemble model was developed using the average predictions from the upper (33:66), lower (20:80), and intermediate (i.e. reference) (25:75) models. Generating the 2 ensemble predictions allowed comparisons between candidates and the exploration of the potential effects of model selection on the final prediction (Fig. S6 in the Supplement).

2.4.3. Probability-of-occupancy threshold

The first stage of EASI-Fish in estimating the vulnerability of a species to fisheries impacts is to determine the number of grid cells in which the species is considered to be present and which also contain fishing effort. Because SDMs estimate the probability of a species to occur in each grid cell, a probability-of-occupancy threshold value (ψ) must be used to define whether the species is present or absent in each cell. However, the spatial extent of the species' distribution increases and decreases with lower and higher ψ values, respectively, and thus influences the proportion of the species' population that is exposed to fishing. To account for this uncertainty, 3 ψ values (0.1, 0.2, and 0.3) were applied to each 0.25° cell, based on statistically determined thresholds and expert evaluation of prediction maps. This range was statistically determined by overlaying the distribution of predicted probability of presence with that of predicted probability of absence. The ψ value where these 2 distributions intersected was selected to define the most probable species distribution, and upper and lower bounds were selected by visual inspection of the 2 distributions on either side of the intersection point (Lopez et al. 2020).

Given the Critically Endangered status of EP leatherbacks, we selected relatively low ψ values to conservatively include areas where experts considered leatherbacks likely to occur, even if in relatively low numbers and for limited periods of time, based on documented patterns of spatio-temporal habitat use (Shillinger et al. 2008, 2011, Donoso & Dutton 2010, Bailey et al. 2012, Quiñones et al. 2021) and previously published SDM maps (Hoover et al. 2019, Degenford et al. 2021). This was important to ensure that EASI-Fish would be precautionary in its calculations of fishery impacts on leatherbacks throughout their distribution and across industrial and small-scale fisheries known to interact with the species (see Griffiths & Wallace et al. 2024).

3. RESULTS

3.1. Model performance

Models that included all data and environmental and static variables (e.g. Model 6—full) performed better according to the diagnostic measures we used (deviance explained, AUC, and TSS) (Table 3). In general, complex models (e.g. Models 4 and 6) had better performance than simpler models including

sets of variables individually (e.g. Models 1, 2, 3, and 5). These models explained between 19.71 and 40.57% of the deviance in the data, had AUC values between 0.79 and 0.94, and had TSS values that ranged between 0.51 and 0.76 (Table 3). The number of trees created by these models ranged between 3600 and 7000. The comparison in model performance led us to recommend the use of the full model configuration (Model 6—full, environmental and static variables) for further exploration on the sensitivity analysis.

Models including different proportions of presence ratios with environmental and static variables (i.e. Model 6—full) showed similar but also different performance under the diagnostic measures that we used (Tables 4 & S1 in the Supplement). In general, models with balanced or slightly imbalanced data sets (i.e. models 50:50, 33:66, 25:75, 20:80, 15:85, and 10:90) had better performance than models that used highly imbalanced data sets (i.e. models 5:95, 2.5:97.5, 1:99, 0.5:99.5, and original—0.19%). These models explained between 40.57 and 61.54% of the deviance in the data, had AUC values between 0.92 and 0.96, and had TSS values that ranged between 0.76 and 0.81 (Table 4). Models built for the hold-out additional complementary cross-validation also performed well (Table S1). In particular, models trained without data pertaining to specific periods and then tested against them (e.g. El Niño, La Niña) showed high AUC, TSS, and deviance explained values (AUC ~0.9–0.85, TSS ~0.8–0.60, 50–60% deviance explained) (Table S1). Similarly, models trained with inshore or offshore data (i.e. limits at 90W and 100W explored) and tested against the remaining data showed reasonable performance metrics (AUC ~0.7–0.6, TSS ~0.3–0.2, ~30–60% deviance explained) (Table S1). Based on these diagnostic measures, the model using a 25% presence–absence ratio was identified as the reference model.

Table 3. Hierarchical model performance metrics with all data (n = 573889). Fit: fitted model; Simp: simplified model; ST: spatio-temporal variables; SUR: environmental surface variables; SUB: environmental subsurface variables; ENV: environmental surface and subsurface variables; STA: static variables; FULL: all environmental variables and static variables. lr: learning rate; n.trees: number of trees generated by model; Dev: deviance; AUC: area under the curve; TSS: true skill statistics; Drop: Dropped variables after model simplification; EkE: eddy kinetic energy

		lr	n.trees	Dev	AUC	TSS	Drop
1	Fit	0.01	4500	38.64	0.89	0.66	
ST	Simp	—	—	—	—	—	No drop
2	Fit	0.01	5150	32.28	—	—	
SUR	Simp	—	5150	32.40	0.92	0.71	EkE
3	Fit	0.01	3600	23.62	0.88	0.65	
SUB	Simp	—	—	—	—	—	No drop
4	Fit	0.01	5600	38.14	0.94	0.74	
ENV	Simp	—	—	—	—	—	No drop
5	Fit	0.01	4400	19.71	0.79	0.51	
STA	Simp	—	—	—	—	—	No drop
6	Fit	0.01	7000	40.57	0.94	0.76	
FULL	Simp	—	—	—	—	—	No drop

Table 4. Performance metrics of the models conducted during the sensitivity analyses for full models (all environmental and static variables) with different presence–absence ratios (e.g. 50:50 presence to absence ratio = 50% presence data and 50% absence data in the model). Fit: fitted model; Simp: simplified model; Drop: Dropped variables after model simplification

		lr	n.trees	Dev	AUC	TSS	Drop
50:50	Fit	0.005	2900	58.10	—	—	
n = 2176	Simp	0.005	3650	59.37	0.92	0.71	7 variables
33:66	Fit	0.01	2050	59.16	—	—	
n = 3264	Simp	0.01	2650	61.54	0.94	0.76	5 variables
25:75	Fit	0.01	2700	59.70	—	—	
n = 4352	Simp	0.01	3250	60.42	0.96	0.81	7 variables
20:80	Fit	0.01	3200	60.29	—	—	
n = 5440	Simp	0.01	3350	59.14	0.94	0.77	6 variables
15:85	Fit	0.01	3950	60.21	—	—	
n = 7250	Simp	0.01	3650	58.96	0.94	0.76	3 variables
10:90	Fit	0.01	5050	60.93	—	—	
n = 10880	Simp	0.01	4800	60.16	0.95	0.78	2 variables
5:95	Fit	0.01	5300	56.36	0.95	0.78	
n = 21760	Simp	—	—	—	—	—	No drop
2.5:97.5	Fit	0.01	6400	54.67	—	—	
n = 43520	Simp	0.01	5650	52.08	0.95	0.77	7 variables
1:99	Fit	0.01	6650	49.24	0.94	0.78	
n = 108800	Simp	—	—	—	—	—	No drop
0.5:99.5	Fit	0.01	6100	45.23	0.94	0.76	
n = 217600	Simp	—	—	—	—	—	No drop
0.19:99.81	Fit	0.01	7000	40.57	0.94	0.76	
n = 573889	Simp	—	—	—	—	—	No drop

3.2. Drivers of leatherback turtle presence

An examination of the relationships between species and the environmental and static variables showed a range of interesting patterns for each of the

11 final models, based on variable importance analysis (Fig. 3) and partial dependence plots (Fig. S7). These indicators suggested 2 groups of similar models: (1) models with a presence–absence ratio ranging from 50 to 10% and (2) models with a presence–absence ratio ranging from 5 to 0.19%. The first group showed higher variable importance with SST, SST₁₀₀, chl, MLD, FrontIndex, chl_grad and Dir, whereas the second group showed higher values for Depth, Sal, LandDistance, FSLE and SSH.

After simplification of the reference model, 8 variables were included in the final model, for which relative variable importance was 3.3–49.9%. EkE, FSLE, LandDistance, SSH, Vel, SST_grad, and FrontIndex were dropped from the final model as they did not improve predictive performance. With the exception of chl_grad (3.3%), all variables contributed more than 5%: SST (49.9%), Depth (12.5%), MLD (9%), SST₁₀₀ (7.5%), Dir (6.7%), chl (6.2%), and Sal (5%) (Fig. S7).

The model identified higher probabilities of leatherback turtle presence around SST values of 16–20°C and in waters shallower than 1000 m. MLD values of < 100 m were associated with higher probabilities of leatherback turtle presence, whereas SST₁₀₀ showed a positive relationship. Currents with southwest direction showed an evident negative relationship with leatherback turtle presence. A positive relationship was also observed between the leatherback turtle presence probability and CHL, while the opposite relationship was observed for Sal (i.e. salinities higher than 30 PSU). Similarly, the model showed higher probabilities of leatherback turtle presence at positive chl_grad values.

3.3. Predictions

Final models were used to predict species habitat suitability in the convention area for 2002–2020



Fig. 3. Relative variable importance in the series of species distribution models generated using different proportions of presence versus absence data; i.e. the ratio of presences to absences used to build the final model (e.g. 50 refers to a 50:50 presence to absence ratio; 33 refers to a 33:66 presence to absence ratio, etc.). Variables included (see Section 2.3 for more details): sea surface temperature (SST) and its gradient within the same pixel over a 7 d period (SST_grad), salinity (Sal), sea surface height (SSH), current speed (Vel), current direction (Dir), eddy kinetic energy (EkE), finite size Lyapunov exponents (FSLE), front index (FrontIndex), chlorophyll *a* (chl), chlorophyll *a* gradient (chl_grad), temperature at 100 m depth (SST₁₀₀), mixed layer depth (MLD), bathymetry (Depth), and the distance to land (LandDistance)

(Figs. S1 & S2), as well as by quarter (Fig. S3), and ENSO phase (Figs. S4 & S5). Predictions revealed spatial differences among models, with, in general, higher probabilities and lower CV of leatherback turtles predicted by models with presence–absence ratios ranging from 50 to 10%. The predictions of this group of models highlighted several areas expected to support higher residence or occurrence by leatherback turtles (Fig. 4). These include coastal areas near nesting beaches in Mexico and Central America, nearshore foraging areas in southern latitudes (Quiñones et al. 2021), and some high-seas areas through which leatherbacks transit and in which they are presumed to forage (e.g. Bailey et al. 2012). These areas also showed relatively low CV values (Figs. 5 & S2). In particular, the continental shelf and adjacent high-seas areas within South American EEZs (specifically from southern Ecuador to northern Chile) as well as the higher latitude subtropical convergence zone extending from south-central Chile clearly and consistently supported higher probabilities of leatherback turtle presence, with relatively low associated CV values (Figs. 5 & S2). These patterns were also reflected in maps of the ensemble predictions and after applying the probability of occupancy thresholds mentioned above (Fig. 4).

While these spatial patterns were generally consistent by quarter (Figs. 6 & S3) and ENSO phase (Figs. 7 & S4), the relative area and distribution of predicted habitat varied at these scales. Predicted leatherback habitat in the Northern Hemisphere, particularly off Mexico and Central America, was most available during Q1 and reduced in other quarters, roughly opposite the trends for predicted habitat off South America (Fig. 6). In addition, predicted leatherback habitat generally increased during the La Niña phase relative to neutral and El Niño phases, particularly along the equator and in the Humboldt Current region (Figs. 7, 8, S4 & S5). However, there were some areas where predicted habitat increased during the El Niño phase, such as the high-latitude region between 25° and 40° S (Figs. 7, 8, S4 & S5).

4. DISCUSSION

Understanding the spatial distributions of marine species is becoming increasingly important as international and regional oceans and fisheries management instruments evolve further toward ecosystem-based approaches to conservation and management of marine resources (Aburto et al. 2012, Kirkfeldt 2019). In a fisheries context, having a reliable prediction of a spe-

cies' distribution that can be compared with that of fishing effort can allow managers to determine the extent of overlap and implement conservation and management measures (CMMs), such as spatial closures, that can reduce fishery interactions and fishing mortality and thus ensure long-term population sustainability.

4.1. Model performance and relevant features

SDMs have traditionally required a large number of species presence locations at a fine resolution to be useful for fisheries management in jurisdictions that typically span spatial scales of hundreds or thousands of kilometers. Furthermore, because the largest quantity of data is usually collected for species of high commercial importance (i.e. target species), the application of SDMs has primarily been constrained to these species (Melo-Merino et al. 2020). Unfortunately, reliable catch or even simple occurrence data is often lacking for bycatch species for a variety of reasons, such as being of lower economic importance, the absence of observer programs and policies mandating fishers to record all species interactions in catch logbooks, policies requiring rapid release and no-retention of species, the typically low frequency of interactions of these non-target species, and issues pertaining to species identification. As the present study showed, the incidence of observations of the Critically Endangered EP leatherback turtle population in EPO fisheries was extremely low—1088 occurrences in over half a million observations. Regardless of whether the low frequency of interactions is due to a naturally low density of leatherback turtles in the EPO, rarity of the species following its documented population decline (Laúd OPO Network 2020), or low gear selectivity by EPO fisheries, the data available are insufficient to undertake conventional stock assessments. Therefore, simpler ecological risk assessment approaches that are designed to assess the vulnerability of data-limited species (e.g. EASI-Fish) are highly reliant on estimating the 3-dimensional overlap between the species and fisheries, and thus reliable SDMs.

Although BRTs have been applied to several marine species, they are typically used with larger numbers of occurrences than we had available for this study. Furthermore, validation of our model predictions using a truly 'independent' data set (e.g. satellite telemetry data) was not possible because such data were not made available for our project. This required a deep exploration of specific model diagnostics and sensitivity analyses to determine the utility of BRTs in such data-limited settings. The present study showed the im-

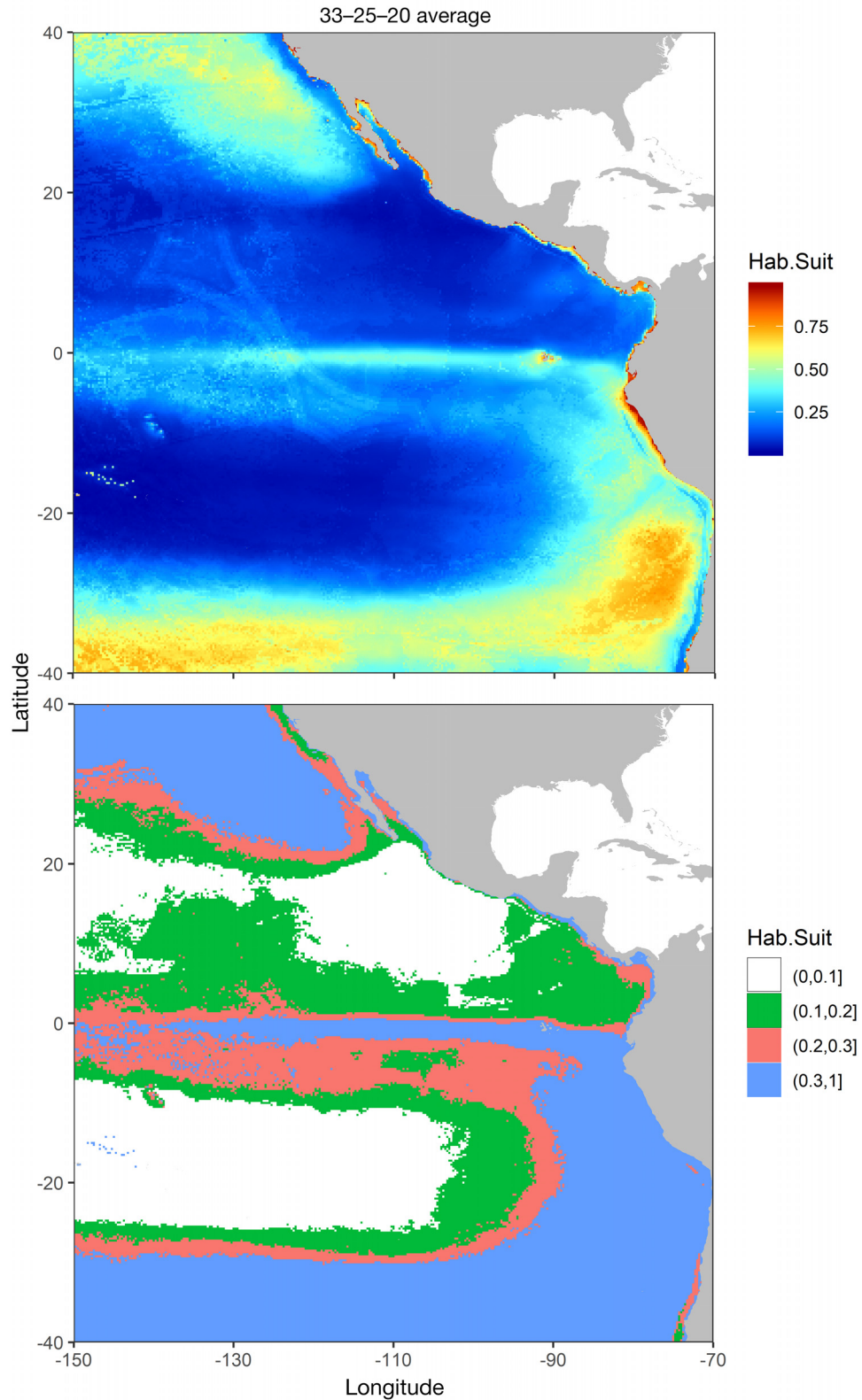


Fig. 4. (a) Average predictions from an ensemble of species distribution models (SDMs) where the ratio of presence to absences were 33–25–20 for 2002–2020; warmer colors: higher probabilities. (b) The final appearance of the SDM predictions using 3 threshold values (minimum = 0.3, blue; most probable = 0.2, red; maximum = 0.1, green) upon which the predicted probability of presence is used to create binary values of species presence. For example, at a threshold of 0.2, predicted probabilities of presence above and below 0.2 are predicted to be absence and presence records, respectively

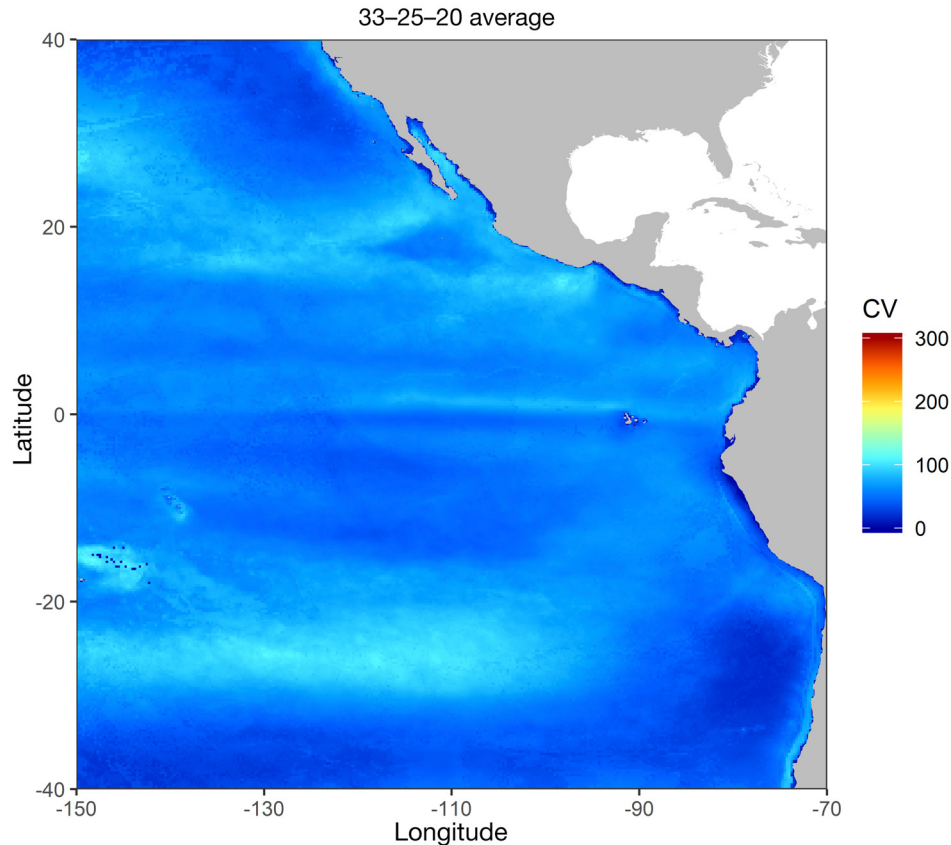


Fig. 5. Reference model (i.e. 33–25–20 ensemble) coefficient of variation (CV) in predictions (averages for daily predictions for the period 2002–2020). See the CV of the rest of the model-specific predictions in Fig. S3

proved capability of BRTs to characterize the habitat preferences of a critically endangered species for which very few occurrence records existed in industrial purse-seine and longline fisheries in the EPO (Hall & Roman 2013, Griffiths & Duffy 2017, Lezama-Ochoa et al. 2017). Our SDM is spatially comprehensive with respect to the full distribution of the EP leatherback population and to the distributions of various fisheries that are known to interact with the species in the EPO. Performance metrics for the BRTs (i.e. AUC, TSS, deviance explained) as well as the uncertainty estimates via CV showed that the models were somewhat robust to the extreme imbalance between leatherback presences (~1000) and absences (>570 000). The performance metrics of our BRTs are in line with, or even surpass, the values shown in other BRT models developed for several large pelagic species, including sharks, tuna, swordfish, and leatherback turtles, among others (Scales et al. 2017b, Brodie et al. 2018, Hazen et al. 2018). Thus, we suggest that our modeling approach has the potential to be applied to other such data-limited species and populations. However, we recognize that for many species, researchers might not have access to absence data when generating SDMs (Liang et al. 2023) and will have to rely on presence-only approaches (e.g. MaxEnt).

In the present case, although our model relied on fisheries-dependent observation data to produce the SDM, it used presences and absences from several different fisheries (industrial and small-scale), each of which use different gear types (e.g. longlines, purse seines, and nets) in different regions (i.e. from Mexico to Chile, and international waters to 150° W), different depths (i.e. surface to ~300 m), and in different habitat types (i.e. temperate to tropical, coastal to high seas). Therefore, our EPO leatherback distribution data set was derived from a diverse and extensive suite of sampling platforms. In particular, the fact that we were able to include presences as well as absences from the same sampling platforms (i.e. fisheries) improved the strength of our inferences about the magnitude and direction of environmental covariate influence on predicted leatherback occurrence.

4.2. Predicted patterns of leatherback occurrence in the EPO

Our model successfully highlighted several areas known or expected to support high residence or occurrence by leatherback turtles (Figs. 4–8), many

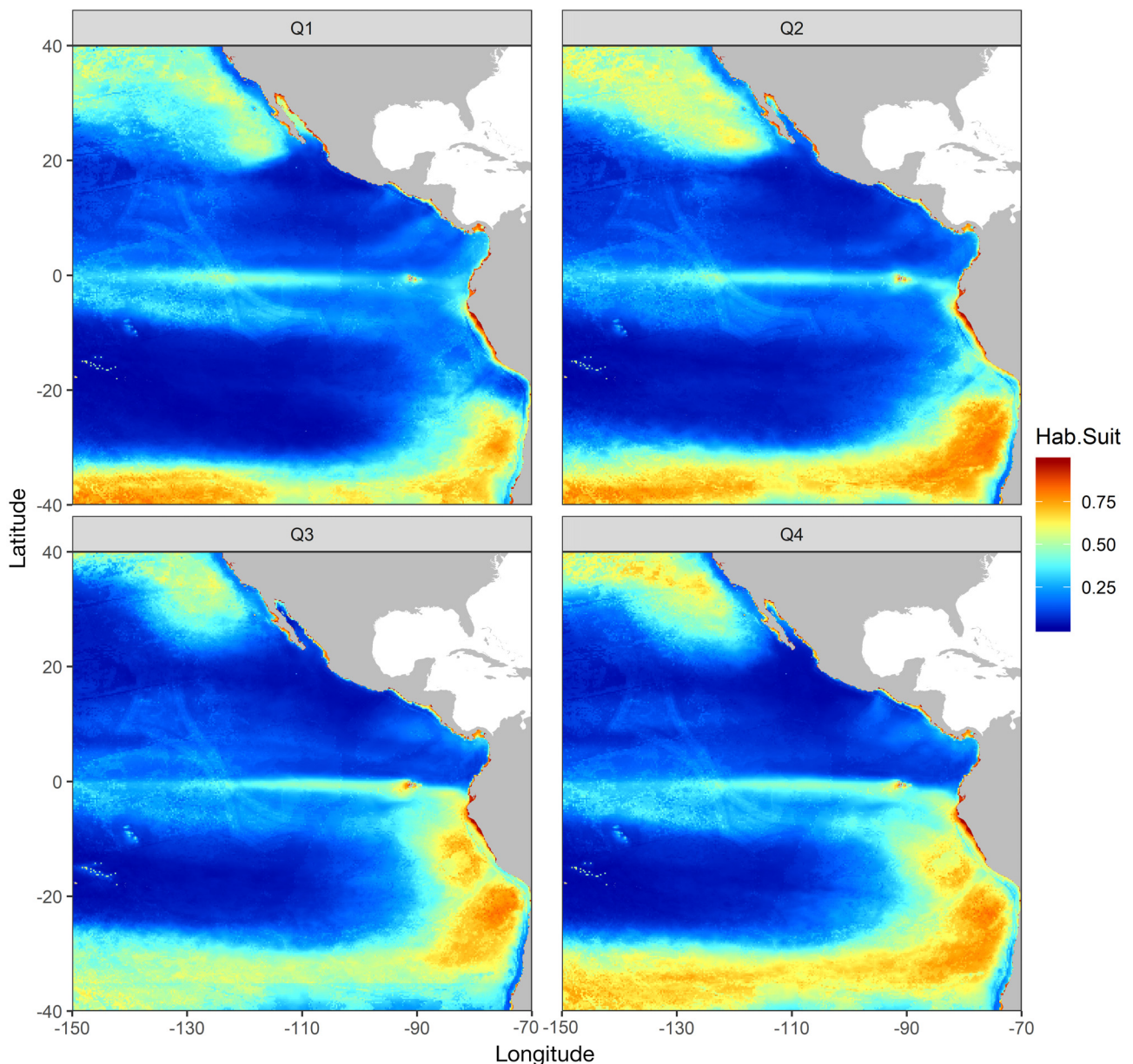


Fig. 6. Average mean quarterly predictions (Q1–Q4) from an ensemble of species distribution models where the ratio of presence to absences were 33–25–20 for 2002–2020

of which are relatively close to shore (Quiñones et al. 2021, Liang et al. 2023) (Figs. 3, S1 & S6). These include coastal areas near nesting beaches in Mexico and Central America (e.g. Shillinger et al. 2011, Laúd OPO Network 2020), high-seas areas through which leatherbacks transit and in which they are presumed to forage (Shillinger et al. 2008, Donoso & Dutton 2010, Bailey et al. 2012, Hoover et al. 2019), and near-shore foraging areas in southern latitudes (Alfaro-Shigueto et al. 2007, Degenford et al. 2021, Quiñones et al. 2021).

Patterns of leatherback occurrence also appeared to be driven by SST and bathymetry (Figs. 3 & S7), particularly when models included near parity between numbers of presences and absences (Fig. S1). The influence of SST on leatherback occurrence has been described by telemetry (e.g. Jonsen et al. 2007, Sherrill-Mix et al. 2008, Shillinger et al. 2008, 2011) and modeling studies (e.g. Hoover et al. 2019), which suggest seasonal movements that track preferred SST ranges, particularly in foraging areas. In the EPO, for example, leatherbacks are thought to exploit

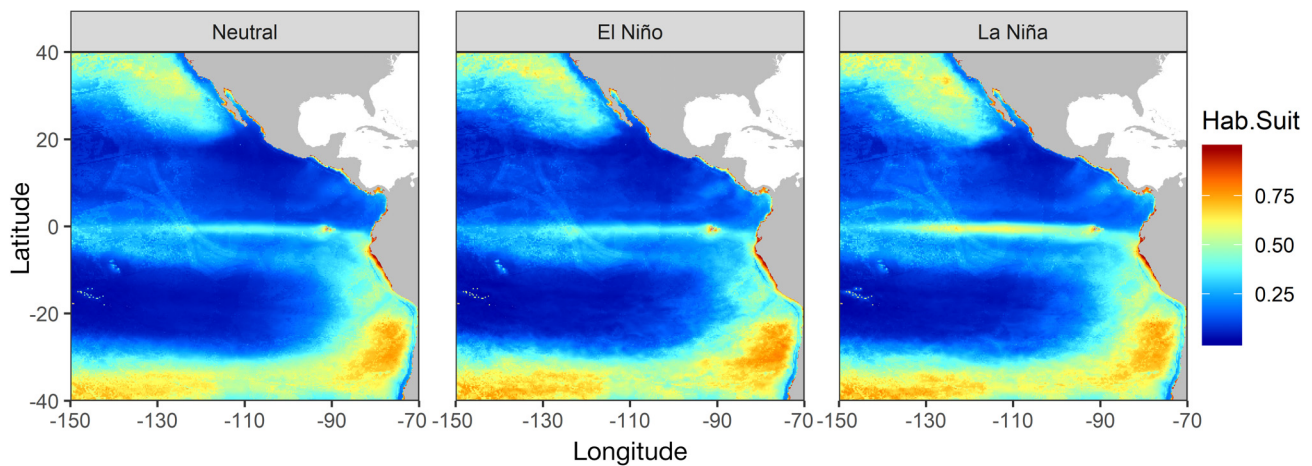


Fig. 7. Predicted leatherback habitat for neutral, La Niña, and El Niño conditions using the ensemble of species distribution models where the ratio of presence to absences were 33–25–20 during 2002–2020

high-latitude foraging areas in the South Pacific Subtropical Convergence during the austral summer (December–April) when SST is higher in this region, later moving northward toward tropical and subtropical latitudes in austral winter (May–November) (Saba et al. 2008, Hoover et al. 2019). Leatherbacks in the Northwest Atlantic Ocean follow a similar general pattern of high-latitude occurrence in summer followed by movement to lower latitudes in winter, tracking SSTs above approximately 18–20°C (Sherrill-Mix et al. 2008). In apparent support for this pattern, our results appear to show higher probabilities of leatherback occurrence in the South Pacific Subtropical Convergence zone during Q1 and Q2 (January–June) than in Q3 and Q4 (July–December) (Figs. 6 & S3).

The seasonal patterns displayed in our model results support inferences from satellite telemetry data and other published studies about how leatherbacks may purposefully depart from certain areas to arrive in other areas when conditions are becoming more favorable for resource availability. For example, leatherbacks might depart high-latitude areas when prey abundance declines (perhaps related to temperature changes), rather than specifically because of decreased water temperatures, considering their capacity for effective thermoregulation (Sherrill-Mix et al. 2008, Wallace & Jones 2008). Similarly, seasonal variations in predicted habitat might imply differential foraging habitat use by leatherbacks; for instance, from oceanic, high-latitude areas in Q1 to more coastal, lower-latitude areas in other times of year (Figs. 6

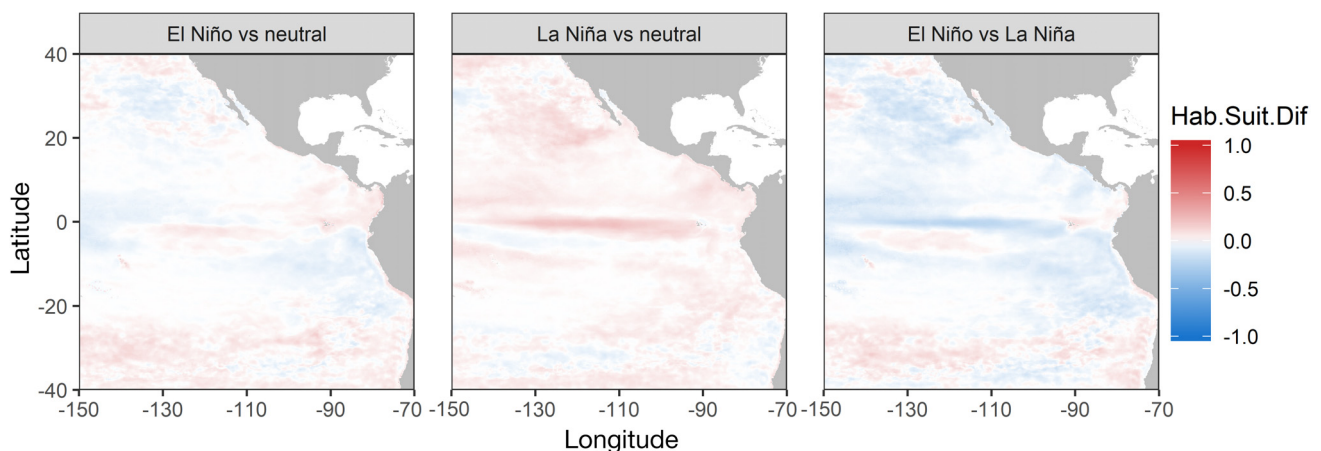


Fig. 8. Relative differences in predicted leatherback habitat between El Niño and La Niña conditions compared to neutral years (left and central panels, respectively), and El Niño compared to La Niña using the ensemble of species distribution models where the ratio of presence to absences were 33–25–20 for 2002–2020. Red pixels represent net gain of predicted habitat while blue pixels represent net loss of predicted habitat of one phase relative to the other

& S3). Furthermore, predicted leatherback habitat offshore from Mexico and Central America is greatest in Q1, when predicted habitat off Peru and northern Chile is lowest (Figs. 6 & S3), which coincides with the southward migration of leatherbacks from reproduction areas in the Northern Hemisphere toward foraging areas south of the equator (Shillinger et al. 2008). This suggests that adult leatherbacks depart breeding areas in Q1 to arrive in feeding areas off South America later in the year when conditions are more suitable for the species (Figs. 6 & S3), a notion supported by observations of correlated increases in leatherback abundance and the abundance of their preferred prey (i.e. gelatinous zooplankton) during Q2 (Quiñones et al. 2021). However, sub-adult or large immature leatherbacks are observed during Q1 as well as Q2 in nearshore Peruvian waters, demonstrating potentially important ontogenetic differences in spatio-temporal habitat use that are generalized in our model predictions (Quiñones et al. 2021). Overall, our seasonal predictions of leatherback occurrence—interpreted together with results of telemetry studies—provide useful heuristics for understanding the timing and patterns of leatherback migrations throughout the EPO.

ENSO strongly influences resource availability and thus energy budgets and reproduction in leatherbacks (Saba et al. 2007, 2008, Wallace & Saba 2009) and other organisms in the EPO (Suryan et al. 2009). Our model results showed variations in predicted leatherback habitat based on ENSO phases (Figs. 7 & S4), with generally more habitat available during La Niña phases (Figs. 8 & S5). In particular, predicted leatherback habitat increased in the northern Humboldt Current region off Peru—a well-documented foraging area for EP leatherbacks—as well as along the Equatorial Counter-Current during La Niña phases, in comparison to neutral and El Niño phases. Although predicted habitat in these areas was generally reduced during El Niño, some areas showed increases in predicted leatherback habitat. For example, predicted leatherback habitat increased during El Niño phases near the South Equatorial Current between 100° and 140° W and between 25° and 35° S across the southeastern Pacific. These patterns illustrate the phenomenon described in previous studies in which stochastic oceanographic conditions related to ENSO drive variations in EP leatherback habitat availability and quality, with consequences for leatherback life history (Saba et al. 2007, 2008, Wallace & Saba 2009). Similar to the seasonal patterns described above (Figs. 6 & S3), our results provide hypothetical habitat use patterns by leatherbacks during different ENSO phases that, if further investigated and ground-truthed,

could provide guidance to managers about relative risk of overlaps between fishing operations and leatherbacks based on long-term oceanographic conditions. Examples of similar applications include the ENSO-related closure of Pacific Loggerhead Conservation Area (US National Marine Fisheries Service, Federal Register 72 FR 31756) and the seasonal closure of the Pacific Leatherback Conservation Area (NMFS 2001; Federal Register 66 FR 44549) of the California (USA) driftnet fisheries to avoid loggerhead turtle *Caretta caretta* and leatherback bycatch, respectively. However, it is worth noting that our model can also predict at daily scale; short-term predictions are the basis for dynamic ocean management (Hazen et al. 2018, Abrahms et al. 2019), an adaptive and flexible management mechanism to reduce incidental catch, which has proven to be useful for other vulnerable species and fisheries (see Section 4.4).

Our SDMs predicted similar areas of relatively high probability of occurrence to those highlighted in previous modeling efforts using satellite telemetry data (Hoover et al. 2019), presence-only observation data in national fisheries, mainly from Peru and Chile (Degenford et al. 2021), and both telemetry and fisheries observer data (Liang et al. 2023). For example, the continental shelf and adjacent high-seas areas within South American EEZs, as well as the higher-latitude subtropical convergence zone extending from south-central Chile clearly and consistently supported high probabilities of occurrence in our SDM. This result is generally supported by long-term monitoring observations (Donoso & Dutton 2010, Quiñones et al. 2021), telemetry data (Shillinger et al. 2008), and presence-only distribution models (Hoover et al. 2019, Degenford et al. 2021, Liang et al. 2023). These comparable patterns are encouraging, particularly when considering the distinctly different analytical methods, data, and spatial and temporal scales used in each study, which were each constrained by the amount and types of information available.

4.3. Enhanced data collection and reporting

The scarcity of leatherback observation data, and their Critically Endangered status, motivated our conservative selection of relatively low bounds as threshold values for defining 'high' probabilities of occurrence to parameterize the EASI-Fish model to assess vulnerability (Griffiths & Wallace et al. 2024). While noting other potential factors, low observer coverage in IATTC tuna longline fisheries and small-scale fisheries throughout the region is an

obvious reason for the lack of available data. Thus, occurrence probabilities in areas with low observer coverage (e.g. $>20^{\circ}$ S and $>90^{\circ}$ W) should be interpreted with caution. Because sea turtle interactions are not generally required to be recorded in catch logbooks in any IATTC tuna fishery or fishery within national jurisdictions of EPO coastal states, their occurrence is generally only recorded by onboard observers when they interact with the gear or are sighted in the vicinity of the vessel. Although the AIDCP/IATTC observer program covers 100% of sets of all large (Class 6; >363 mt) purse-seine vessels in the EPO, very few sea turtles—especially leatherbacks—are taken by this fleet (Hall & Roman 2013, Lezama-Ochoa et al. 2017). By contrast, observer coverage is very low for most other fleets, especially longline, where sea turtles are frequently hooked or entangled (Swimmer et al. 2017). For example, IATTC Resolution C-19-08 requires only 5% of the fishing effort of vessels greater than 20 m length overall to be observed, but even this low level of coverage is frequently not met by some nations, or the extent of coverage is unknown because countries do not provide the information in annual reports to the IATTC (e.g. IATTC 2021a). Although the quality of species-specific reporting by the industrial fleet has improved in recent years with IATTC Resolution C-19-08 mandating reporting of operational-level data, the reported information is not considered representative of the fleets in space and time (Griffiths et al. 2021), and therefore there may be some important habitats for leatherback turtles where data is not currently collected. Given that IATTC scientists have proposed to the Commission Members an increase in observer coverage to at least 20% for several years (e.g. IATTC 2021b), and the recently approved IAC Resolutions CIT-COP10-2022-R7 and CIT-COP10-2002-R6 that urge countries to strengthen existing fisheries monitoring programs to ensure systematic collection of statistically reliable bycatch data, it is hoped that longline observer data will further improve in the near future, which would be particularly important for data-limited species.

Improving data collection programs, both fisheries-dependent and fisheries-independent, is one of the most important considerations for improving SDMs, not only to better characterize a species' distribution with an increased number of occurrence points but also to disentangle some of the aforementioned issues pertaining to interpretation of increasing or decreasing incidences of the capture of vulnerable species that can guide the subsequent development of appropriate management action. To optimize the utility of

data collection programs—for SDM development in particular—they need to be carefully designed to cover the widest possible spatial extent of the species of interest that cover the full gradient of environmental variables in which the species is naturally exposed (Araújo & Guisan 2006, Grenouillet et al. 2011). This maximizes the performance of SDMs since strong and contrasted species–environmental relationships allow the model to better learn from these relationships and interpolate into unsampled areas or environmental conditions (Miller 2010). However, data collection programs can be a costly and logistically difficult proposition for highly migratory wide-ranging pelagic species, such as the leatherback turtle, and for fisheries that cover thousands of square kilometers of ocean. Consequently, fishery-dependent data is often the primary source of information from which species occurrence locations can be derived. Although with cooperation and proper coordination with fishers and their national management organizations, these programs can be run cost-effectively, such data may not cover the breadth of habitats occupied by the species of interest.

4.4. Future directions for SDMs

Though we are confident that our EPO-wide SDM provides robust results to evaluate CMM efficacy within the EASI-Fish framework as well as other conservation applications (see below), we recognize additional potential directions for this SDM work. In particular, considering that the development of EP leatherback SDMs has received notable attention to date (Hoover et al. 2019, Degenford et al. 2021, this study), integration of fisheries-independent data with fisheries-dependent data to validate observation-based models and generate a single SDM for the species would be beneficial (Liang et al. 2023). Our model expanded on the fishery-based observation data used by Degenford et al. (2021) to also include the high-seas and higher latitudes, among others, but EP leatherback satellite telemetry data (Shillinger et al. 2008, Hoover et al. 2019) were not made available for this study. Given the similarities in estimated residence times and occurrence probabilities among the existing studies, it is unclear whether one SDM that included all available fishery-dependent and fishery-independent data would produce significantly different, region-wide results than those presented here. Nonetheless, a single, integrated SDM would avoid confusion among users about which SDM to use in their own work, and thus ensure consistency in research and conservation applications.

On that note, we encourage applications of the EP leatherback SDM presented here beyond the current specific application to the EASI-Fish vulnerability assessment and evaluation of CMM efficacy. Improved SDMs can inform dynamic ocean management approaches to identify and manage spatio-temporal overlaps between high-use areas and potential threats, including fisheries activities. For example, Turtle-Watch identifies potential turtle 'hotspots' based on environmental characteristics, such as favorable SSTs and the presence of current fronts (Howell et al. 2008, 2015). However, this model relies heavily upon data collected by observers who have observed 100% of shallow sets in the Hawaiian longline fishery since 2004 (Sippel et al. 2015), providing turtle catch counts as well as information on turtle absences. A conceptually comparable tool is being developed for EP leatherback turtles, but it relies heavily on habitat use data from > 15 yr ago that was limited to post-nesting females from a single nesting population (Hoover et al. 2019, Liang et al. 2023). Furthermore, moving from SDMs for individual species to integrated, multi-species predictions could inform the development of ocean planning conservation tools to highlight areas in which effective management of fisheries and other activities (e.g. ship traffic) would be most beneficial (Hazen et al. 2018, Abrahms et al. 2019). In addition, simply making SDM outputs publicly available to researchers and conservation practitioners could unlock several potential and yet unforeseen applications of these important resources to various issues at different scales. To this end, we have made the products of this model (e.g. maps, feature classes) publicly available (<https://seamap.env.duke.edu>) to facilitate other conservation and management applications. In this vein, we see value in implementing participatory-inclusive approaches with stakeholders, and fishers in particular, to discuss results and bycatch mitigation ideas, define or propose practical management plans, as well as reinforce capacity development with regards to sensitive species, and the leatherback turtle in particular.

5. CONCLUSIONS

This study highlighted several ways in which a flexible machine-learning SDM framework may be used to improve our understanding of the spatio-temporal distributional dynamics of marine species and how this knowledge may translate into improved assessment and management of vulnerable and data-limited species at multiple spatial and temporal scales.

Furthermore, this effort was the product of an international collaboration under an MoU between 2 inter-governmental conventions—one established to protect and conserve marine turtles (IAC) and the other to sustainably manage commercially valuable tuna and tuna-like stocks (IATTC). Because this collaborative effort brought together several individual representatives of member countries as well as experts in turtle biology and fisheries operating at multiple scales, the resulting product was able to integrate proprietary data sets whose combined value far exceeds their individual value to understanding comprehensive patterns of leatherback distribution throughout the EPO region. Our results are encouraging, as they suggest that ample suitable habitat exists in the EPO for leatherback turtles to occupy, which theoretically could support improved population growth if conservation efforts can successfully reduce fisheries bycatch—the primary remaining source of mortality affecting this population (Laúd OPO Network 2020).

Acknowledgements. The authors thank La Red para la Conservación de la Tortuga Laúd del Océano Pacífico Oriental (Laúd OPO Network) and the IAC members for sharing data and expert knowledge. We also thank Ann Marie Lauritsen, Yonat Swimmer, Barbara Schroeder, and Brian Stacy from the US National Oceanic and Atmospheric Administration for their input during this project. The authors also thank Stephanie Brodie, Jason Hartog, and Michael Jacox for assistance and advice with data preparation and the statistical methodology applied in this study. This research was presented to the 11th IATTC Bycatch Working Group meeting as a technical report (https://iattc.org/GetAttachment/ec8a1a22-3e24-4d0f-bf12-69634843b314/BYC-11-01_Species-distribution-modelling-leatherback-turtle.pdf).

LITERATURE CITED

- ✦ Abrahms B, Welch H, Brodie S, Jacox MG and others (2019) Dynamic ensemble models to predict distributions and anthropogenic risk exposure for highly mobile species. *Divers Distrib* 25:1182–1193
- Aburto MO, De Los Angeles Carvajal M, Barr B, Barbier EB and others (2012) *Ecosystem-based management for the oceans*. Island Press, Washington, DC
- ✦ Alfaro-Shigueto J, Dutton PH, Van-Bressem M, Mangel J (2007) Interactions between leatherback turtles and Peruvian artisanal fisheries. *Chelonian Conserv Biol* 6: 129–134
- ✦ Alfaro-Shigueto J, Mangel JC, Bernedo F, Dutton PH, Seminoff JA, Godley BJ (2011) Small-scale fisheries of Peru: a major sink for marine turtles in the Pacific. *J Appl Ecol* 48:1432–1440
- ✦ Alfaro-Shigueto J, Mangel JC, Darquea J, Donoso M, Baquero A, Doherty PD, Godley BJ (2018) Untangling the impacts of nets in the southeastern Pacific: rapid assessment of marine turtle bycatch to set conservation priorities in small-scale fisheries. *Fish Res* 206:185–192

- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43:1223–1232
- Andraka S, Mug M, Hall M, Pons M and others (2013) Circle hooks: developing better fishing practices in the artisanal longline fisheries of the Eastern Pacific Ocean. *Biol Conserv* 160:214–224
- Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *J Biogeogr* 33:1677–1688
- Bailey H, Benson SR, Shillinger GL, Bograd SJ and others (2012) Identification of distinct movement patterns in Pacific leatherback turtle populations influenced by ocean conditions. *Ecol Appl* 22:735–747
- Becker EA, Carretta JV, Forney KA, Barlow J and others (2020) Performance evaluation of cetacean species distribution models developed using generalized additive models and boosted regression trees. *Ecol Evol* 10:5759–5784
- Borokini IT, Nussear K, Petitpierre B, Dilts TE, Weisberg PJ (2023) Iterative species distribution modeling results in the discovery of novel populations of a rare cold desert perennial. *Endang Species Res* 50:47–62
- Brodie S, Jacox MG, Bograd SJ, Welch H and others (2018) Integrating dynamic subsurface habitat metrics into species distribution models. *Front Mar Sci* 5:219
- Casale P, Tucker AD (2017) *Caretta caretta* (amended version of 2015 assessment). The IUCN Red List of Threatened Species 2017:e.T3897A119333622. <https://dx.doi.org/10.2305/IUCN.UK.2017-2.RLTS.T3897A119333622.en> (accessed 6 Feb 2023)
- Chaloupka MY, Bjorndal KA, Balazs GH, Bolten AB and others (2008) Encouraging outlook for a once severely exploited marine megaherbivore. *Glob Ecol Biogeogr* 17:297–304
- De'ath G (2007) Boosted trees for ecological modelling and prediction. *Ecology* 88:243–251
- Degenford JH, Liang D, Bailey H, Hoover AL and others (2021) Using fisheries observation data to develop a predictive species distribution model for endangered sea turtles. *Conserv Sci Pract* 3:e349
- Donoso M, Dutton PH (2010) Sea turtle bycatch in the Chilean pelagic longline fishery in the southeastern Pacific: opportunities for conservation. *Biol Conserv* 143:2672–2684
- Eckert KL, Wallace BP, Frazier JG, Eckert SA, Pritchard PCH (2012) Synopsis of the biological data on the leatherback sea turtle, *Dermochelys coriacea* (Vandelli, 1761). US Fish and Wildlife Service, Jacksonville, FL
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol Evol Syst* 40:677–697
- Elith J, Leathwick J (2017) Boosted Regression Trees for ecological modeling. https://rspatial.org/raster/sdm/9_sdm_brt.html
- Elith J, Graham CH, Anderson RP, Dudík M and others (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151
- Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. *J Anim Ecol* 77:802–813
- FAO Fisheries Department (2009) Guidelines to reduce sea turtle mortality in fishing operations. FAO, Rome
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- Fossette S, Witt MJ, Miller P, Nalovic MA and others (2014) Pan-Atlantic analysis of the overlap of a highly migratory species, the leatherback turtle, with pelagic longline fisheries. *Proc R Soc B* 281:20133065
- Frazier J, Brito JL (1990) Incidental capture of marine turtles by the swordfish fishery at San Antonio, Chile. *Mar Turtle Newsl* 49:8–13
- Grenouillet G, Buisson L, Casajus N, Lek S (2011) Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography* 34:9–17
- Griffiths SP, Duffy L (2017) A preliminary metadata analysis of large-scale tuna longline fishery data in the eastern Pacific Ocean: a precursor to Ecological Risk Assessment (8th Meeting of the Scientific Advisory Committee of the Inter-American Tropical Tuna Commission, 8–12 May 2017). Document SAC-08-07b. IATTC, La Jolla, CA
- Griffiths SP, Kesner-Reyes K, Garilao C, Duffy LM, Román MH (2019) Ecological Assessment of the Sustainable Impacts of Fisheries (EASI-Fish): a flexible vulnerability assessment approach to quantify the cumulative impacts of fishing in data-limited settings. *Mar Ecol Prog Ser* 625:89–113
- Griffiths SP, Lennert-Cody C, Wiley B, Fuller L (2021) Update on operational longline observer data required under resolution C-19-08 and a preliminary assessment of data reliability for estimating total catch for bycatch species in the eastern Pacific Ocean (10th meeting of the IATTC Working Group on Bycatch, 5 May 2021). Document BYC-10 INF-D. IATTC, La Jolla, CA
- Griffiths SP, Wallace B, Cáceres V, Rodríguez LH and others (2024) Vulnerability of the Critically Endangered leatherback turtle to fisheries bycatch in the eastern Pacific Ocean. II. Assessment of mitigation measures. *Endang Species Res* 53:295–326
- Hall M, Roman M (2013) Bycatch and non-tuna catch in the tropical tuna purse seine fisheries of the world. Fisheries and Aquaculture Technical Paper No. 568. FAO, Rome
- Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143:29–36
- Hazen EL, Scales KL, Maxwell SM, Briscoe DK and others (2018) A dynamic ocean management tool to reduce bycatch and support sustainable fisheries. *Sci Adv* 4:eaar3001
- Hijmans RJ, van Etten J (2015) raster: Geographic analysis and modeling with raster data. R package version 2.0-12. <https://CRAN.R-project.org/package=raster>
- Hijmans RJ, Phillips S, Leathwick J, Elith J (2017) dismo: Species Distribution Modeling. R package version 1.0-12. <https://CRAN.R-project.org/package=dismo>
- Hoover AL, Liang D, Alfaro-Shigueto J, Mangel JC and others (2019) Predicting residence time using a continuous-time discrete-space model of leatherback turtle satellite telemetry data. *Ecosphere* 10:e02644
- Howell EA, Kobayashi DR, Parker DM, Balazs GH, Polovina JJ (2008) TurtleWatch: a tool to aid in the bycatch reduction of loggerhead turtles *Caretta caretta* in the Hawaii-based pelagic longline fishery. *Endang Species Res* 5:267–278
- Howell EA, Hoover A, Benson SR, Bailey H, Polovina JJ, Seminoff JA, Dutton PH (2015) Enhancing the Turtle-Watch product for leatherback sea turtles, a dynamic habitat model for ecosystem-based management. *Fish Oceanogr* 24:57–69
- Inter-American Tropical Tuna Commission (IATTC) (2020) The tuna fishery in the eastern Pacific Ocean in 2019. In: 10th Meeting of the Scientific Advisory Committee of the IATTC, 11–15 May 2020. Document SAC-11-03. IATTC, La Jolla, CA

- IATTC (2021a) Synopsis of longline observer data reporting pursuant to Resolution C-19-08. In: 12th meeting of the Scientific Advisory Committee of the IATTC, 10–14 May 2020. Document SAC-12-04. IATTC, La Jolla, CA
- IATTC (2021b) Staff recommendations for management and data collection, 2021. In: 12th meeting of the Scientific Advisory Committee of the IATTC, 10–14 May 2020. Document SAC-12-04. IATTC, La Jolla, CA
- James MC, Ottensmeyer CA, Myers RA (2005) Identification of high-use habitat and threats to leatherback sea turtles in northern waters: new directions for conservation. *Ecol Lett* 8:195–201
- Jonsen ID, Myers RA, James MC (2007) Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. *Mar Ecol Prog Ser* 337:255–264
- Kirkfeldt TS (2019) An ocean of concepts: why choosing between ecosystem-based management, ecosystem-based approach and ecosystem approach makes a difference. *Mar Pol* 106:103541
- Laúd OPO Network (2020) Enhanced, coordinated conservation efforts required to avoid extinction of critically endangered eastern Pacific leatherback turtles. *Sci Rep* 10:4772
- Lee-Yaw JA, McCune JL, Pironon S, Sheth SN (2022) Species distribution models rarely predict the biology of real populations. *Ecography* 2022:e05877
- Lezama-Ochoa N, Murua H, Hall M, Román M and others (2017) Biodiversity and habitat characteristics of the bycatch assemblages in fish aggregating devices (FADs) and school sets in the eastern Pacific Ocean. *Front Mar Sci* 4:265
- Lezama-Ochoa N, Pennino MG, Hall MA, Lopez J, Murua H (2020) Using a Bayesian modelling approach (INLSPDE) to predict the occurrence of the spinytail devil ray (*Mobular* [sic] *mobular*). *Sci Rep* 10:18822
- Liang D, Bailey H, Hoover AL, Eckert S and others (2023) Integrating telemetry and point observations to inform management and conservation of migratory marine species. *Ecosphere* 14:e4375
- Lopez J, Lennert-Cody CE, Maunder MN, Xu H, Brodie S, Jacox M, Hartog J (2019) Developing alternative conservation measures for bigeye tuna in the eastern Pacific Ocean: a dynamic ocean management approach. In: 10th meeting of the IATTC Scientific Advisory Committee, 13–17 May 2019. Document SAC-10-INF-D. IATTC, San Diego, CA
- Lopez J, Alvarez-Berastegui D, Soto M, Murua H (2020) Using fisheries data to model the oceanic habitats of juvenile silky shark (*Carcharhinus falciformis*) in the tropical eastern Atlantic Ocean. *Biodivers Conserv* 29:2377–2397
- Mazaris AD, Schofield G, Gkazinou C, Almpandou V, Hays GC (2017) Global sea turtle conservation successes. *Sci Adv* 3:e1600730
- Melo-Merino SM, Reyes-Bonilla H, Lira-Noriega A (2020) Ecological niche models and species distribution models in marine environments: a literature review and spatial analysis of evidence. *Ecol Model* 415:108837
- Mi C, Huettmann F, Guo Y, Han X, Wen L (2017) Why choose random forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ* 5:e2849
- Miller J (2010) Species distribution modeling. *Geogr Compass* 4:490–509
- Montoya-Jiménez JC, Valdez-Lazalde JR, Ángeles-Perez G, De Los Santos-Posadas HM, Cruz-Cárdenas G (2022) Predictive capacity of nine algorithms and an ensemble model to determine the geographic distribution of tree species. *iForest* 15:363–371
- NMFS (National Marine Fisheries Service) (2001) Endangered and threatened wildlife; sea turtle conservation requirements; taking of threatened or endangered species incidental to commercial fishing operations. *Fed Regist* 66:44549–44552
- Northwest Atlantic Leatherback Working Group (2019) *Dermodochelys coriacea* (Northwest Atlantic Ocean subpopulation) (errata version published in 2020). The IUCN Red List of Threatened Species 2019:e.T46967827A184748440, <https://dx.doi.org/10.2305/IUCN.UK.2019-2.RLTS.T46967827A184748440.en> (accessed 6 Feb 2023)
- Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecol Model* 133:225–245
- Pearson RG, Thuiller W, Araújo MB, Martinez-Meyer E and others (2006) Model-based uncertainty in species range prediction. *J Biogeogr* 33:1704–1711
- Pons M, Watson JT, Ovando D, Andracka S and others (2022) Trade-offs between bycatch and target catches in static versus dynamic fishery closures. *Proc Nat Acad Sci USA* 119:e2114508119
- Quiñones J, Quispe S, Zeballos J (2021) Recopilación histórica de la captura de laúd (*Dermodochelys coriacea*) en Perú y su relación con la medusa Scyphozoa *Chrysaora plocamia*. *Bol Inst Mar Perú* 36:140–155
- R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.r-project.com
- Roberts KE, Stepanuk JEF, Kim H, Thorne LH, Chong-Montenegro C, Nye JA (2023) Developing a subseasonal ecological forecast to reduce fisheries bycatch in the Northeast US *Prog Oceanogr* 213:103021
- Robinson NJ, Morreale SJ, Nel R, Paladino FV (2016) Coastal leatherback turtles reveal conservation hotspot. *Sci Rep* 6:37851
- Saba VS, Santidrián PT, Reina RD, Spotila JR, Musick JA, Evans DA, Paladino FV (2007) The effect of the El Niño Southern Oscillation on the reproductive frequency of eastern Pacific leatherback turtles. *J Appl Ecol* 44:395–404
- Saba VS, Spotil JR, Chavez FP, Musick JA (2008) Bottom-up and climatic forcing on the worldwide population of leatherback turtles. *Ecology* 89:1414–1427
- Sbrocco EJ, Barber PH (2013) MARSPEC: ocean climate layers for marine spatial ecology. *Ecology* 94:979
- Scales KL, Hazen EL, Jacox MG, Edwards CA, Boustany AM, Oliver MJ, Bograd SJ (2017a) Scale of inference: on the sensitivity of habitat models for wide-ranging marine predators to the resolution of environmental data. *Ecography* 40:210–220
- Scales KL, Hazen EL, Maxwell SM, Dewar H and others (2017b) Fit to predict? Eco-informatics for predicting the catchability of a pelagic fish in near real time. *Ecol Appl* 27:2313–2329
- Sequeira AMM, Mellin C, Fordham DA, Meekan MG, Bradshaw CJA (2014) Predicting current and future global distributions of whale sharks. *Glob Change Biol* 20:778–789
- Sherrill-Mix SA, James MC, Myers RA (2008) Migration cues and timing in leatherback sea turtles. *Behav Ecol* 19:231–236

- Shillinger GL, Palacios DM, Bailey H, Bograd SJ and others (2008) Persistent leatherback turtle migrations present opportunities for conservation. *PLOS Biol* 6:e171
- Shillinger GL, Palacios DM, Bailey H, Bograd SJ and others (2010) Identification of high-use interesting habitats for eastern Pacific leatherback turtles: role of the environment and implications for conservation. *Endang Species Res* 10:215–232
- Shillinger GL, Swithenbank AM, Bailey H, Bograd SJ and others (2011) Vertical and horizontal habitat preferences of post-nesting leatherback turtles in the South Pacific Ocean. *Mar Ecol Prog Ser* 422:275–289
- Sippel T, Paige Eveson J, Galuardi B, Lam C and others (2015) Using movement data from electronic tags in fisheries stock assessment: a review of models, technology and experimental design. *Fish Res* 163:152–160
- Soykan CU, Eguchi T, Kohin S, Dewar H (2014) Prediction of fishing effort distributions using boosted regression trees. *Ecol Appl* 24:71–83
- Suryan RM, Saba VS, Wallace BP, Hatch SA, Frederiksen M, Wanless S (2009) Environmental forcing on life history strategies: multi-trophic level response at ocean basin scales. *Prog Oceanogr* 81:214–222
- Swimmer Y, Gutierrez A, Bigelow K, Barceló C and others (2017) Sea turtle bycatch mitigation in US longline fisheries. *Front Mar Sci* 4:e260
- Tiwari M, Wallace BP, Girondot M (2013) *Dermochelys coriacea* (West Pacific Ocean subpopulation). The IUCN Red List of Threatened Species 2013:e.T46967817A46967821, <https://dx.doi.org/10.2305/IUCN.UK.2013-2.RLTS.T46967817A46967821.en> (accessed 6 May 2022)
- Villarino E, Chust G, Licandro P, Butenschön M, Ibaibarriaga L, Larrañaga A, Irigoien X (2015) Modelling the future biogeography of North Atlantic zooplankton communities in response to climate change. *Mar Ecol Prog Ser* 531:121–142
- Wallace BP, Jones TT (2008) What makes marine turtles go: a review of metabolic rates and their consequences. *J Exp Mar Biol Ecol* 356:8–24
- Wallace BP, Saba VS (2009) Environmental and anthropogenic impacts on intra-specific variation in leatherback turtles: opportunities for targeted research and conservation. *Endang Species Res* 7:11–21
- Wallace BP, DiMatteo AD, Bolten AB, Chaloupka MY and others (2011) Global conservation priorities for marine turtles. *PLOS ONE* 6:e24510
- Wallace BP, Tiwari M, Girondot M (2013) *Dermochelys coriacea* (East Pacific Ocean subpopulation). The IUCN Red List of Threatened Species 2013:e.T46967807A46967809, <https://dx.doi.org/10.2305/IUCN.UK.2013-2.RLTS.T46967807A46967809.en>
- Wallace BP, Posnik ZA, Hurley BJ, DiMatteo AD and others (2023) Marine turtle regional management units 2.0: an updated framework for conservation and research of wide-ranging megafauna species. *Endang Species Res* 52:209–223
- Wang S, Zhuang Q, Yang Z, Yu N, Jin X (2019) Temporal and spatial changes of soil organic carbon stocks in the forest area of northeastern China. *Forests* 10:1023
- Wibbels T, Bevan E (2019) *Lepidochelys kempii* (errata version published in 2019). The IUCN Red List of Threatened Species 2019:e.T11533A155057916, <https://dx.doi.org/10.2305/IUCN.UK.2019-2.RLTS.T11533A155057916.en> (accessed 6 Feb 2023)
- Willis-Norton E, Hazen EL, Fossette S, Shillinger G and others (2015) Climate change impacts on leatherback turtle pelagic habitat in the Southeast Pacific. *Deep Sea Res II* 113:260–267

Appendix. Authors and affiliations

**Jon Lopez^{1,*}, Shane Griffiths¹, Bryan Wallace^{2,3,4}, Verónica Cáceres⁴,
Luz Helena Rodríguez⁴, Marino Abrego⁵, Joanna Alfaro-Shigueto^{6,7,8},
Sandra Andraka⁹, María José Brito¹⁰, Leslie Camila Bustos¹¹, Iliá Cari¹²,
José Miguel Carvajal¹³, Ljubitza Clavijo¹², Luis Cocas¹¹, Nelly de Paz¹⁴,
Marco Herrera¹⁰, Jeffrey C. Mangel^{7,8}, Miguel Pérez-Huaripata¹⁵,
Rotney Piedra¹⁶, Javier Antonio Quiñones Dávila¹⁵, Liliana Rendón⁹,
Juan M. Rguez-Baron^{17,18}, Heriberto Santana¹⁹, Jenifer Suárez²⁰,
Callie Veelenturf²¹, Rodrigo Vega¹², Patricia Zárate¹²**

¹Inter-American Tropical Tuna Commission, 8901 La Jolla Shores Drive, La Jolla, CA 92037, USA

²Ecolibrium, Inc., 5343 Aztec Drive, Boulder, CO 80303, USA

³Department of Ecology and Evolutionary Biology, University of Colorado Boulder, Boulder, CO 80310, USA

⁴Inter-American Convention for the Protection and Conservation of Sea Turtles, Falls Church, VA 22046, USA

⁵Ministerio de Ambiente, Panama City C-0843-00793, Panamá

⁶Carrera de Biología Marina, Universidad Científica del Sur, Lima 15067 Perú

⁷ProDelphinus, Jose Galvez 780E, Lima 10680 Perú

⁸School of Biosciences, University of Exeter, Cornwall Campus, Penryn, Cornwall TR10 9EZ, UK

⁹EcoPacífico+, San José 11801, Costa Rica

¹⁰Instituto Público de Investigación de Acuicultura y Pesca, Guayaquil 090314, Ecuador

¹¹Subsecretaría de Pesca y Acuicultura, Valparaíso 2340000, Chile

¹²Instituto de Fomento Pesquero, Valparaíso 2340000, Chile

¹³Instituto Nacional Costarricense de Pesca y Acuicultura, Puntarenas 60101, Costa Rica

¹⁴Áreas Costeras y Recursos Marinos, Pisco 11600, Perú

¹⁵Instituto del Mar del Perú, Callao 07021, Peru

¹⁶Sistema Nacional de Áreas de Conservación, Nicoya 50201, Costa Rica

¹⁷JUSTSEA Foundation, Bogotá 1100111, Colombia

¹⁸University of North Carolina Wilmington, Wilmington, NC 28403, USA

¹⁹Instituto Nacional de Pesca y Acuicultura, Manzanillo, Colima 28200, Mexico

²⁰Parque Nacional Galápagos, Puerto Ayora, Galápagos Islands 200101, Ecuador

²¹The Leatherback Project, Norfolk, MA 02056, USA

Editorial responsibility: Paolo Casale,

Pisa, Italy

Reviewed by: A. Dujon and 2 anonymous referees

Submitted: October 7, 2022

Accepted: October 23, 2023

Proofs received from author(s): February 16, 2024