HOME IS WHERE THE HABITAT IS: MODELING SHORTFIN MAKO HABITAT SUITABILITY VIA MACHINE LEARNING METHODS

BY

JULIAN GARRISON

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

BIOLOGICAL AND ENVIRONMENTAL SCIENCES

UNIVERSITY OF RHODE ISLAND

2023

MASTER OF SCIENCE THESIS

OF

JULIAN GARRISON

APPROVED:

Thesis Committee:

Major Professor Bradley Wetherbee

Jeremy Collie

Marco Alvarez

Madison Jones

Brenton DeBoef
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2023

ABSTRACT

Given the mounting threats of species overexploitation, climate change, and other anthropogenic stressors to global biodiversity, there is a growing need for conservation and management efforts informed by the life history and ecology of target species. Apex marine predators such as the shortfin mako shark (*Isurus oxyrinchus*) are especially vulnerable owing to their life history traits, but accurately mapping habitat preferences remains challenging. Using a novel framework that combines multiple analytical techniques, I report on nearly a decade of habitat preferences of 106 shortfin makos in the Gulf of Mexico (GoM) and western North Atlantic Ocean (NAO). I leverage the predictive power of machine learning (ML) to generate region-specific habitat suitability models based on satellite telemetry and remote sensed environmental data. Ensemble-based models performed best in predicting shortfin mako habitat suitability, and variables indicating coastal proximity were consistently the most important for model predictions at broad scales. In the GoM, sharks concentrated their residency behaviors around the Yucatán Peninsula during the late winter and early spring but expanded home ranges to include much of the GoM during the summer. In contrast, NAO sharks concentrated their residency behaviors off the northeastern U.S. coast during the summer, whereas winter habitats were more diffuse and located further south along the U.S. East Coast and in the open western NAO. Predicted habitat suitability from ML models aligned well with these observed contrasting patterns in seasonal shortfin mako movements, while also demonstrating considerable interannual variability.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Dr. Bradley Wetherbee, for the opportunity to conduct this research and for all the support, guidance, and encouragement throughout my time at the University of Rhode Island. Similarly, I would like to acknowledge the members of my committee, Dr. Marco Alvarez and Dr. Jeremy Collie, for their insights and feedback, especially as I became more familiar with machine learning. I am grateful to the Guy Harvey Foundation and Guy Harvey Research Institute for providing funding. This project could not have been completed without the collective tagging efforts of numerous individuals from the University of Rhode Island, Nova Southeastern University, and Captain Anthony Mendillo of Keen M Blue Water Encounters, as well as Captain Mark Sampson of Fish Finder Adventures. I would also like to acknowledge members of the Wetherbee lab at URI and the Shivji lab at NSU for their support and technical assistance. I am particularly grateful to Daniel Daye (URI), whose parallel research on whale sharks afforded numerous opportunities for mutual consultation and collaboration as we both learned new skills in shark tagging, R, and machine learning. Finally, I could not have completed this research without the considerable support of family and friends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE PAGE

CHAPTER 1

INTRODUCTION

Given the mounting threats of species overexploitation, climate change, and other anthropogenic stressors to global biodiversity (Johnson et al. 2017), there is a growing need for conservation and management efforts informed by the life history and ecology of target species. In particular, understanding species' movement patterns and habitat preferences will be critical in identifying and protecting key habitat areas, as well as predicting climate change impacts (Hays et al. 2016). This knowledge is especially important for highly mobile species and apex predators for a number of reasons. For example, apex predators directly promote biodiversity through resource facilitation and trophic cascades (Sergio et al. 2008; Ritchie and Johnson 2009). Apex predators can also serve as biodiversity indicators owing to their dependence on ecosystem productivity, sensitivity to disturbance, habitat selection, and links to multiple trophic levels (Sergio et al. 2008; Ritchie and Johnson 2009). Moreover, their charismatic nature makes apex predators popular targets for conservation and ecotourism (Ordiz et al. 2013; Macdonald et al. 2017). However, many of the traits that characterize apex predators also render them highly vulnerable to anthropogenic stressors (Ordiz et al. 2013). Highly mobile marine species are additionally challenging to manage, as their subsurface movements are not readily visible, their large ranges increase their exposure to a variety of threats, and these ranges often span multiple management jurisdictions (Shiffman and Hammerschlag 2016; Dulvy et al. 2017; Manz 2021).

The shortfin mako shark (*Isurus oxyrinchus*; hereafter, “shortfin mako”) is a large, pelagic species that exemplifies many of the conservation needs and challenges described above. Shortfin makos are widely distributed throughout the world’s temperate and tropical oceans (Stevens 2008; Lohe et al. 2022) and are known to travel long distances. As apex predators, shortfin makos influence ecosystem structure and function directly through predation and indirectly via behavioral effects on prey species, which include teleosts, cephalopods, other smaller sharks, cetaceans, and crustaceans (Ferretti et al. 2010; Dulvy et al. 2014; Lohe et al. 2022). Through their far-ranging movements, shortfin makos can also facilitate horizontal nutrient and energy transfer between ecosystems (Shiffman and Hammerschlag 2016).

However, many of their life history traits also leave shortfin makos particularly susceptible to fishing pressure. Due to their high mobility and broad distribution, shortfin mako ranges frequently overlap with longline fisheries (Oliver et al. 2015; Queiroz et al. 2016; Vaudo et al. 2017; Lohe et al. 2022), where they are caught as incidental bycatch. Shortfin makos are also popular in recreational fisheries, owing to their size and power. Unfortunately, due to their late age-at-maturity and low fecundity, shortfin mako populations recover slowly and are especially vulnerable to fishing pressures (Shiffman and Hammerschlag 2016; Lohe et al. 2022).

Consequently, shortfin makos have experienced substantial population declines in the Atlantic Ocean (Dulvy et al. 2014), and there is growing recognition of the need for improved conservation and management measures. For example, in 2018 the International Union for Conservation of Nature (IUCN) listed shortfin makos as globally “Endangered” (Rigby et al. 2019), while in 2019 shortfin makos were added to the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) Appendix II (Sellheim 2020). In 2021, the International Commission for the Conservation of Atlantic Tunas (ICCAT) implemented a two-year retention ban on shortfin makos, based on assessments in 2017 and 2019 that suggested a high likelihood of the North Atlantic shortfin mako stock being overfished (ICCAT 2021). However, shortfin mako ranges overlap with multiple regional and international jurisdictions (Vaudo et al. 2017; Gibson et al. 2021; Manz 2021), which complicates conservation and management if sufficient protection is not enacted across the entire range. In addition, climate change will likely shift, and in some cases decrease, suitable habitat areas for shortfin makos (Robinson et al. 2015; Birkmanis, Freer, et al. 2020). Therefore, understanding shortfin mako movement patterns and habitat preferences will be critical for informing future conservation efforts and promoting recovery of their stocks (Shiffman and Hammerschlag 2016; Dulvy et al. 2017; Birkmanis, Partridge, et al. 2020).

Shortfin makos are known to inhabit a broad range of habitats, and their distributions have been extensively tracked worldwide through a combination of satellite telemetry, tag-and-recapture, and fisheries observer efforts (Lohe et al. 2022). Prior studies have found sea surface temperature (Vaudo et al. 2017), depth (Byrne et al. 2019; Francis et al. 2019; Gibson et al. 2021), and primary productivity (Byrne et al. 2019; Nasby-Lucas et al. 2019) to be important indicators of shortfin mako habitat preference, though seasonality, regional differences, and size segregation in behaviors contribute additional variability (Byrne et al. 2019; Nasby-Lucas et al. 2019; Gibson et al. 2021). However, knowledge gaps remain surrounding the relative influences of environmental conditions in driving shortfin mako habitat preferences, and most studies have examined only a handful of environmental variables at a time. Given that marine environments are extremely complex, dynamic, and expansive, shortfin makos likely make habitat decisions based on a multitude of sensory information. Thus, accurately predicting shortfin mako habitat preferences requires extensive data not only on the sharks' movements, but also the underlying environmental conditions associated with those habitats (Beyan and Browman 2020).

A large, ongoing tagging effort has allowed us to gather relatively high-resolution and long-term satellite telemetry data from 106 sharks within the Atlantic shortfin mako population. Coupled with the increasing availability of high resolution remote-sensing information, these data afforded me the opportunity to explore relationships between observed shortfin mako movements and their potential environmental drivers. Here, I report on nearly a decade of movement patterns and habitat preferences of 106 shortfin makos tagged in the Gulf of Mexico (GoM) and the western North Atlantic Ocean (NAO). Notably, I leverage the predictive power of machine learning (ML) and demonstrate its application in movement ecology through a novel framework that evaluates a suite of ML algorithms in predicting habitat suitability from remotely sensed environmental variables. The aims of this study were three-fold: (1) Evaluate the individual and ensemble performance of 11 ML algorithms to assess their suitability to shark telemetry data; (2) Identify which of 17 environmental variables contribute most to model predictions; and (3) Compare shortfin mako habitat suitability predictions between two ecoregions (NAO and GoM).

CHAPTER 2

METHODS

To understand the movement patterns and habitat preferences of shortfin mako sharks in the Gulf of Mexico (GoM) and open western North Atlantic Ocean (NAO), I analyzed a long-term satellite telemetry dataset using a variety of techniques that combined animal-borne telemetry, remote sensed environmental data, spatial and movement modeling techniques, and machine learning. Broadly, these methods can be divided into four sections: (1) Movement Patterns, which includes satellite telemetry, kernel utilization distributions, and state-space and movement-persistence modeling; (2) Environmental Variables, which includes acquisition, transformation, and imputation of remote sensing environmental data; (3) Machine Learning, which includes model training, evaluation, and selection as well as ranking of environmental variables; and (4) Habitat Suitability Predictions, where final models were used to predict on gridded ocean data. These are described in further detail below. All data processing, analysis, and figure creation were performed using R (R Core Team 2021), unless specified otherwise.

Movement Patterns

Satellite telemetry—Shortfin mako sharks have been extensively tracked with satellite telemetry in the GoM and NAO (e.g., Vaudo et al. 2017; Byrne et al. 2019). Tagging efforts typically commenced in the late spring (April) through the summer season (mid-October), depending on the tagging locations, which included coastal waters off Rhode Island, USA, Maryland, USA, and Cancun, Mexico. As of this study, 106 sharks have been tagged since 2012, comprising 58 males and 48 females (Table S1, S2; Fig. S1).

I followed similar tagging procedures as described in previous studies (Vaudo et al. 2017; Francis et al. 2019; Logan et al. 2020). Shortfin mako sharks were caught by chumming the waters from boats and landing them with baited circle hooks. Once hooked, sharks were restrained alongside the vessel or brought aboard, in which case a saltwater hose placed in the mouth supplied flowing seawater over the gills. Each shark was sexed and had its precaudal, fork, and total lengths (PCL, FL, and TL) measured; hereafter, size refers to FL in accordance with prior size-at-maturity analyses (Natanson et al. 2006; Natanson et al. 2020). A Smart Position and Temperature tag (SPOT tag; Wildlife Computers, Redmond, WA, USA) was then attached to the shark's dorsal fin. These transmitters communicated directly with the ARGOS satellite tracking system (argos-system.org) each time the dorsal fin broke the sea surface long enough to trigger a wet/dry sensor on the tag, yielding location data that ARGOS classified based on its accuracy (i.e., location class; Vaudo et al. 2017; Francis et al. 2019). Location classes ranged from LC 3 (best quality; error < 250m) to LC B (error unknown; CLS 2016). Hereafter, "location" will refer to a latitude-longitude pair unless stated otherwise. All landing, measurement, and tagging were performed in such a way to minimize unnecessary stress and injury to the animals, as per the Institutional Animal Care and Use Committee and the Institutional Review Board protocols (IACUC protocol AN1617-020).

Prior to analyses, I pre-processed the shortfin mako satellite telemetry data, as recommend by Vaudo et al. (2017) and Logan et al. (2020). To reduce the impact of post-release stress on shortfin mako movements, I removed the first 10 days of location data from each shark. I also removed very short tracks (< 20 transmissions or < 20 days) and anomalous transmissions (e.g., on-land). As per Vaudo et al. (2017), I applied a speed-distance-angle filter to remove improbable ARGOS locations requiring turn angles ≥ 165° and 155° for distances > 5 and 8 km, respectively (see Freitas et al. 2008 for further details). Pre-processing removed 5,652 locations (8.28% of all locations) and omitted tracks from 11 sharks (10.38% of all sharks), reducing the final sample size to 95 sharks. Table 1 summarizes the final, pre-processed track data for each region.

Kernel utilization distributions—To visualize population-level shortfin mako home ranges, I constructed Kernel Utilization Distributions (KUDs) from the pre-processed satellite telemetry track data using functions from the *adehabitatHR* package in R (Calenge 2006). Following Vaudo et al. (2017) and Manz (2021), I used bivariate normal kernel density estimation with the reference bandwidth as the smoothing parameter (Calenge 2015). Monthly KUDs were generated using the satellite locations of all sharks for each tagging region (GoM and NAO) and depicted as percent density isopleths (25%, 50%, 75%, 95%), each containing the given percentage of satellite locations for that month and region (Calenge 2015). Typically, the 50% isopleth is considered an animal's core area within the 95% isopleth home range (Vaudo et al. 2017; Manz 2021). These KUDs also allowed me to visually ground truth the predictions from my ML habitat suitability models; the KUDs were generated from actual shark locations, whereas the habitat suitability predictions were generated based on gridded environmental and temporal data.

State-space & move persistence modeling—In addition to KUDs, I fit hierarchical state-space models (hSSMs) and move persistence models (MPMs) to the shortfin mako telemetry data using functions from the R package *foieGras* (Jonsen et al. 2019; Jonsen et al. 2020; Jonsen and Patterson 2020). The hSSMs interpolate smoothed, continuous tracks by estimating shark locations (± 95% confidence intervals [CIs]) at regular intervals from the intermittent ARGOS telemetry data (hereafter, "location estimates" or "interpolated locations"). Hierarchical models can also reduce uncertainty and improve location interpolation because they estimate one set of movement parameters for all sharks simultaneously, rather than individually (Jonsen 2016). Based on previously documented regional differences in shortfin mako behaviors (Byrne et al. 2019) and examination of the present study's transmission data (Fig. S2; Table S1), I fit separate hSSMs to the GoM and NAO sharks. I generated hSSM location estimates at 12-hour and 8-hour intervals, respectively, which matches the average reporting interval for each ecoregion (Table S1; Vaudo et al. 2017). Additionally, because large gaps in ARGOS transmissions can increase uncertainty in hSSM location estimates, I