

# Seasonal Variation in the Spatial Distribution of Basking Sharks (*Cetorhinus maximus*) in the Lower Bay of Fundy, Canada

Zachary A. Siders<sup>1,2\*</sup>, Andrew J. Westgate<sup>2,3</sup>, David W. Johnston<sup>4</sup>, Laurie D. Murison<sup>2</sup>, Heather N. Koopman<sup>2,3</sup>

**1** Fisheries and Aquatic Sciences, University of Florida, Gainesville, Florida, United States of America, **2** Grand Manan Whale and Seabird Research Station, Grand Manan, New Brunswick, Canada, **3** Department of Biology and Marine Biology, University of North Carolina Wilmington, Wilmington, North Carolina, United States of America, **4** Nicholas School of the Environment, Duke University, Beaufort, North Carolina, United States of America

## Abstract

The local distribution of basking sharks in the Bay of Fundy (BoF) is unknown despite frequent occurrences in the area from May to November. Defining this species' spatial habitat use is critical for accurately assessing its *Special Concern* conservation status in Atlantic Canada. We developed maximum entropy distribution models for the lower BoF and the northeast Gulf of Maine (GoM) to describe spatiotemporal variation in habitat use of basking sharks. Under the Maxent framework, we assessed model responses and distribution shifts in relation to known migratory behavior and local prey dynamics. We used 10 years (2002–2011) of basking shark surface sightings from July–October acquired during boat-based surveys in relation to chlorophyll-*a* concentration, sea surface temperature, bathymetric features, and distance to seafloor contours to assess habitat suitability. Maximum entropy estimations were selected based on AICc criterion and used to predict habitat utilizing three model-fitting routines as well as converted to binary suitable/non-suitable habitat using the maximum sensitivity and specificity threshold. All models predicted habitat better than random (AUC values >0.796). From July–September, a majority of habitat was in the BoF, in waters >100 m deep, and in the Grand Manan Basin. In October, a majority of the habitat shifted southward into the GoM and to areas >200 m deep. Model responses suggest that suitable habitat from July–October is dependent on a mix of distance to the 0, 100, 150, and 200 m contours but in some models on sea surface temperature (July) and chlorophyll-*a* (August and September). Our results reveal temporally dynamic habitat use of basking sharks within the BoF and GoM. The relative importance of predictor variables suggests that prey dynamics constrained the species distribution in the BoF. Also, suitable habitat shifted minimally from July–September providing opportunities to conserve the species during peak abundance in the region.

**Citation:** Siders ZA, Westgate AJ, Johnston DW, Murison LD, Koopman HN (2013) Seasonal Variation in the Spatial Distribution of Basking Sharks (*Cetorhinus maximus*) in the Lower Bay of Fundy, Canada. PLoS ONE 8(12): e82074. doi:10.1371/journal.pone.0082074

**Editor:** Danilo Russo, Università degli Studi di Napoli Federico II, Italy

**Received:** June 27, 2013; **Accepted:** October 29, 2013; **Published:** December 4, 2013

**Copyright:** © 2013 Siders et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The authors have no support or funding to report.

**Competing interests:** The authors have declared that no competing interests exist.

\* E-mail: zsiders@ufl.edu

## Introduction

Basking sharks (*Cetorhinus maximus*, Gunnerus 1765) are obligate filter feeders of zooplankton, primarily copepods, and are found circumglobally in temperate and tropical seas [1]. Despite being the second largest fish in the world, surprisingly little is known about their ecology [2]. Basking sharks have been subjected to high levels of anthropogenic mortality and this, combined with low reproductive rates, threatens many populations [3–5]. Basking sharks are listed globally as *Vulnerable* [6] and of *Special Concern* in Atlantic Canada [7]. Moreover, the species is listed as *Endangered* in the North Pacific and eastern North Atlantic (eNA), regions, which

historically supported large populations [8,9]. Consequently, the future recovery of basking sharks may depend on protecting remaining areas of high shark density [10,11].

Considerable seasonal changes in the regional distributions of this species [12–15] complicate conservation efforts [16]. In eNA, basking sharks exhibit spatial distribution shifts, as exemplified by the seasonal migration from the neritic zone in the summer to the offshore continental shelf in the winter [2,16]. Considerably less is known about their distribution in the western North Atlantic (wNA) where animals are rarely observed outside of inshore coastal waters in the summer [12,17–19]. Within the Gulf of Maine (GoM), Owen (1984) observed shifts in basking shark sighting records from offshore

areas during the spring to inshore areas in the summer and subsequent disappearance in the fall. In stark contrast to the eNA, Skomal et al. (2009) tracked basking sharks from Massachusetts in the summer to as far south as Brazil in the winter.

Basking sharks occur throughout the wNA, but to date their principal habitats have not been identified [19]. One suspected region is the Bay of Fundy (BoF), located between New Brunswick and Nova Scotia, Canada [19]. The BoF experiences an extreme tidal range (8–16 m) and a seasonal cyclonic gyre near its mouth [20], which promotes zooplankton retention [21] and increased densities of foraging planktivores [22–25]. For example, high densities of North Atlantic (NA) right whales (*Eubalaena glacialis*), co-occur in the BoF where the dominant zooplankton (energy-rich *Calanus finmarchicus* stage V [C5] copepodites) aggregate and comprise over 79% of the zooplankton biomass [26–28]. Similarly, basking shark sightings data positively correlate with the greatest C5 abundance (Murison, L.D., unpublished data) suggesting these megaplanktivores are also targeting this energy-dense prey source [29].

Species distribution models offer a method to predict important habitat based on sighting records and to relate environmental variables to species occurrence [30,31]. The relatively new but robust method of maximum entropy distribution modeling (Maxent) makes use of presence-only datasets [32,33], which include archival, citizen science, and unstandardized survey records. This model routine has been used recently to investigate and describe the distributions of several marine megafauna [34–40], including North Atlantic right whales [41].

To quantify the distribution of basking sharks in the lower BoF we applied Maxent analysis to surface sighting records collected between 2002 to 2011 from scientific boat surveys and commercial whale-watching vessels. We investigated how basking shark distribution changed seasonally (July–October) and responded to abiotic and biotic factors. We also present several new considerations regarding the application of this technique to marine data and discuss these over two frameworks: 1) discriminating basking shark habitat; and 2) variation due to model-fitting in large datasets.

## Materials and Methods

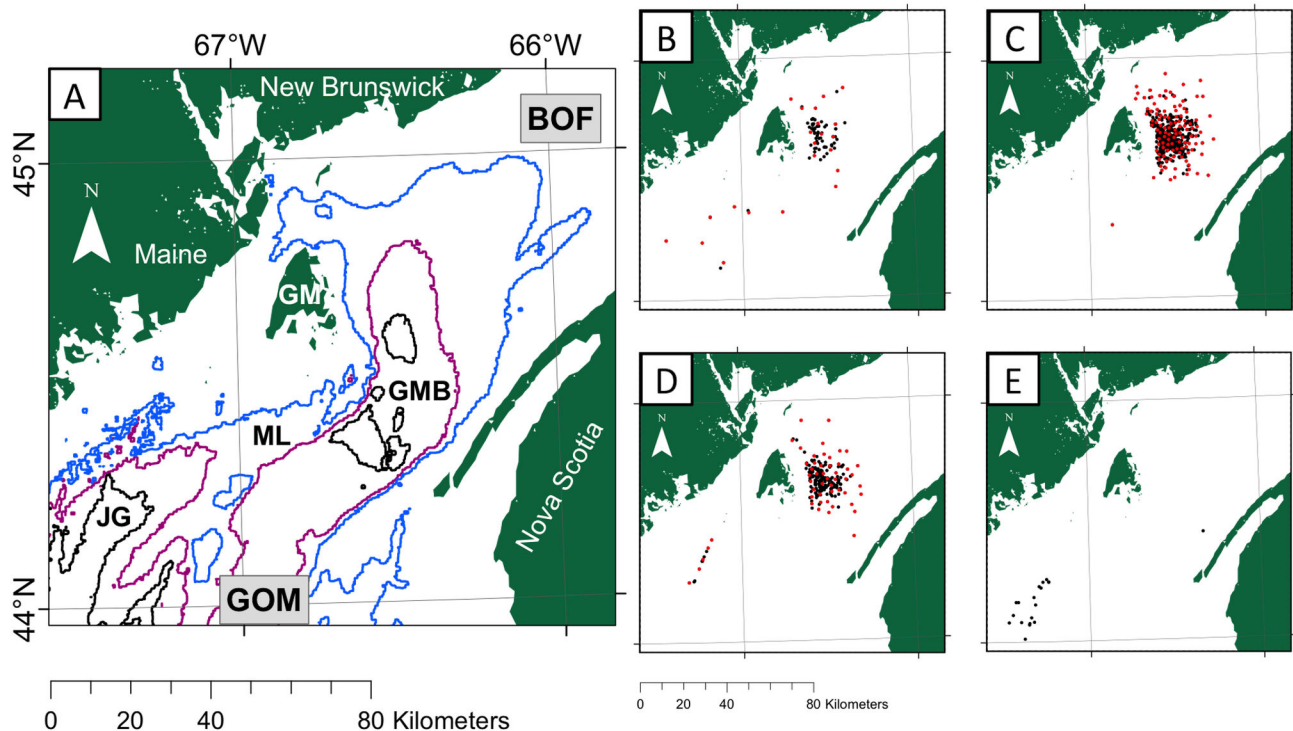
### Presence-only sighting data

Basking shark surface sightings were obtained from two long-term (2002–2011) datasets. The first was from non-standardized surveys made during commercial whale-watching trips conducted from North Head, Grand Manan (Figure 1A). The second consisted of sightings collected incidentally during surveys for NA right whales [42] conducted by the New England Aquarium (NEAq) [43]. These independent data sets encompassed 10 years and covered an area of 7539 km<sup>2</sup> in the lower BoF and northeastern GoM between 44°–45°N and 66°–67° W (Figure 1A). We collated sightings irrespective of the data source and included all animals independently to represent the number of sharks present at that location. A total of 884 sharks were sighted: 90 in July (10.18% of all sightings),

554 in August (62.67%), 221 in September (25%), and 19 in October (2.15%). Basking sharks were sighted throughout the study area in July through October but the large majority of sightings occurred off the eastern side of Grand Manan in the Grand Manan Basin (Figure 1 A, B, & D). For August, the entirety of sightings was located within the Bay of Fundy off the eastern side of Grand Manan (Figure 1 C). In October, the majority of sightings (18 out of 19) were sighted in Jones Ground in the Gulf of Maine (Figure 1E). We assumed that the locations of surface sightings mirror the locations of suitable underwater habitat. However, basking sharks move horizontally underwater, so data from basking shark diving profiles (274 dives/833 hours; Siders, unpublished data) was employed to validate this assumption. From these profiles, average horizontal distance was found by multiplying the mean vertical movement rate for descents ( $-0.22 \text{ m s}^{-1}$ ) and ascents ( $0.22 \text{ m s}^{-1}$ ) by the mean descent (189 s) and ascent time (208 s) by the tangent of 90° minus the pitch angle for descents ( $-20.54^\circ$ ) and ascents ( $+9.78^\circ$ ), resulting in an approximate horizontal distance of 110 m during a descent and 263 m in an ascent (Siders, unpublished data). These ascent and descent rates observed in the Bay of Fundy are much faster than those reported in Shepard et al. (2006) [44] for the eNA ( $0.11 - 0.37 \text{ m min}^{-1}$ ). Therefore, given the cell size of our model (2.5 x 2.5 km) and short horizontal distances during ascent and descents, it is likely that many surface-swimming sharks were closely tied to suitable habitat at depth. We acknowledge that prolonged surface swimming would confound this argument since an animal could have been at the surface swimming for an indeterminate period before the sighting took place. The dive records revealed that on average sharks were between 0–2 m of the surface  $15 \pm 10\%$  of the time. Although this may introduce some bias between sharks at the surface and at depth, we assert that the surface sightings used in our models are largely representative of suitable habitat at depth.

### Environmental variables

Environmental variables representing the bathymetry, distance to bathymetric contours, chlorophyll-*a* concentration (Chl-*a*), and sea surface temperature (SST) were chosen for our Maxent analysis. We derived six variables at a 0.05 km<sup>2</sup> resolution – depth, aspect (slope direction), slope angle (in degrees), and distance to the 0 m (shore), 50 m, 100 m, 150 m, and 200 m contours – from a digital bathymetry grid developed by the United States Geological Survey (USGS- <http://pubs.usgs.gov/of/1998/of98-801/bathy/data.htm#ArclInfo>). Two additional variables- Chl-*a* (in mg m<sup>-3</sup>) and SST (in °C) – were collated at a 16 km<sup>2</sup> resolution using the Marine Geospatial Ecology Toolbox [45] from the data products of the Moderate Resolution Imaging Spectroradiometer (MODIS-Aqua- <http://oceancolor.gsfc.nasa.gov>). The remote sensing variables are limited approximations of phytoplankton abundance (Chl-*a*), the base of the phytoplankton-copepod-basking shark food chain, and relative ocean temperature. We calculated the mean of monthly daytime measures available from July to October during 2002–2011 for MODIS-Aqua and pooled all years by month to produce individual environmental variable sets for each time period in R (2013, R Core Development Team). We



**Figure 1. Study area and sightings data of basking sharks in the Bay of Fundy.** (A) Map of the study area in the lower Bay of Fundy (BOF) and northeastern Gulf of Maine (GOM) with 100 m (blue), 150 m (purple), and 200 m (black) contour lines depicted. GM refers to Grand Manan Island, JG to Jones Ground, ML to Murr Ledges, GMB to the Grand Manan Basin. Maps depicting the sightings data (black) and the sightings left after the average nearest neighbor decimation (red) for July (B), August (C), September (D), and October (E). Sightings were collected by a commercial whale-watch organization and New England Aquarium [NEAq] from 2002 to 2011, and were used to generate *Cetorhinus maximus* distribution models.

doi: 10.1371/journal.pone.0082074.g001

recognize that year-to-year variation occurred in environmental variables, which could affect shark distributions on small scales; however we intended to generate the first general model of habitat for BoF basking sharks and thus have pooled all years by month. To ensure that significant portions of our sightings data were not excluded when eliminating data-less cells around landmasses, all variables were resampled to a 6.25 km<sup>2</sup> resolution using *Nearest-neighbor* in ArcGIS 9.3.1 (ESRI Inc., 2009; Redlands, California, USA). All environmental variables were projected on the North American Datum 1983 UTM 19N grid and clipped according to MODIS-Aqua data acquisition from 2002-2011 [imposing a near-shore exclusion zone, ranging from 4.5 km to 9 km].

### Spatial autocorrelation

We recognize that our two presence datasets were sampled under different efforts, both spatially and temporally. The commercial whale-watching operation effort was highly clustered, extending from the eastern shore of Grand Manan (Supplemental Material- Figure S1) and was primarily exerted in August and September. The NEAq effort was assumed to be uniform throughout the study area, based on the group's survey design, and in comparison to the commercial whale-

watch effort. We attempted to correct for spatial autocorrelation as a result of these sampling biases but ultimately relied on a subsampling based on the expected average nearest neighbor distance for each month in our dataset. This was determined in ArcGIS via the *Average Nearest Neighbor* routine. Our attempts at correcting for autocorrelation by decimating presence locations by the autocorrelation as detected in a semivariogram and a Moran's I-distance plot are summarized in the Material. Neither correction was applicable due to the autocorrelation structure detected in the semivariogram (Figure S2) and the decimation routine based on Moran's I-distance plot (Figure S3) that reduced our sightings by approximately 97%. The sightings for October were not subsampled due to low sample sizes ( $n \leq 5$ ) under all routines. However, the NEAq solely collected the sightings in this month and, thus, we assume they were collected in a relatively uniform manner.

### Species distribution model

**Maximum entropy modeling.** Maxent (version 3.3.3k – see <http://www.cs.princeton.edu/~schapire/maxent/>) is a machine learning species distribution modeling technique that seeks the probability distribution that maximizes entropy (i.e. closest to uniform) subject to constraints [32]; in our case to estimate

basking shark habitat in the BoF in July, August, September, and October [32,33]. Each Maxent estimation was informed using the presence-only sighting data and environmental variables. Environmental covariates are transformed into a set of “features” [32] and we utilized the feature suite available (linear, quadratic, product, threshold, and hinge) [46]. To estimate habitat, Maxent assumes a theoretical uniform distribution then iteratively adjusts the weight of constraints (derived from the empirical mean of environmental variables over the set of species occurrence locations) so that the average probability of sample points is maximized, expressed as training gain. The probability distribution of a species is estimated over the habitat area so each cell holds a probability of presence or relative suitability for the species [47]. We chose to use the cumulative output scaled from 0 to 100, which reports percent relative suitability of a grid cell by calculating the probability of that cell and all cells with equal or lower probability [32]. Maxent prevents overfitting by utilizing a regularization technique to smooth models [46]. We chose unique regularization values ( $\beta$ ) for each month following the AICc procedure outlined in Warren & Seifert (2011) [48], which we will summarize here. For each month, models were executed as described above but using  $\beta = 1, 3, 5, 7, 9, 11, 13, 15, 17$ . The suitability scores from these estimations were standardized in order to sum to 1 across the geographic space. The likelihood of a single model's data was obtained by taking the product of the standardized scores of cells containing presence points. The number of parameters for each estimation was obtained from the lambda file produced during the Maxent routine. AICc was then calculated for each month using sample sizes for each time period and for the suite of  $\beta$  values within each month. The regularization value of the model with the lowest AICc score was used in all subsequent Maxent estimations. All models were performed using a maximum of 100 replicates and a maximum of 10000 background points, which allows Maxent to estimate variance.

**Threshold suitability indices.** Following previous maximum entropy distribution models [49], we chose the maximum training sensitivity and specificity threshold values. For each time period, unique maximum training sensitivity and specificity thresholds were determined based on Receiving Operator Characteristic (ROC) curves (see Model performance evaluation).

### Model performance evaluation

**Metrics of model performance.** The most often reported measure of Maxent outputs is the threshold-independent assessment using the Area Under the Curve (AUC) metric of the ROC curve [50]. The ROC curve plots sensitivity values (true positives) against 1 minus specificity (false positive) values [32]. AUC values designate the probability that positive and negative instances are correctly classified. We evaluated AUC values following Phillips et al. (2006): <0.5 (worse than random performance), 0.5-0.7 (no discrimination), >0.7 (better than random model performance), and 1 (perfect discrimination).

We assessed variable importance using the percent contribution metric [32], which describes the relative

contribution to each model. We also assessed permutation importance or the importance of the original background values to the predicted distribution [32]. A high permutation importance indicates that changes to the original data contained within the environmental variable have a strong negative affect on the probability distribution. Finally, we used a jackknife analysis on training data to determine which variables contain the most useful information about the predicted distribution individually, as well as those that [46].

**Model-fitting comparison.** Maxent provides three methods to assess model fit: cross-validation, bootstrapping, and subsampling. The number of sightings available in our database in July, August, and September allowed us to perform empirical comparisons of these model-fitting techniques. We performed model runs for those time periods where these techniques were estimated to be more appropriate (July, August, and September) as well as in the October time period to compare a small dataset using all available variables. The random test percentage, used in subsampling, was set to 25% test and 75% training [51].

## Results

### Regularization selection

The AICc model selection procedure selected models with different regularization values for each month; July:  $\beta = 5$ ; August:  $\beta = 7$ ; September:  $\beta = 3$ ; October:  $\beta = 7$ . AICc did not follow any discernible pattern for any of the monthly Maxent models as regularization values were increased from 1 - 17.

### Model performance

**Model-fitting comparison.** We evaluated the three model-fitting approaches and found that bootstrapping had the lowest false-positive rate as assessed by AUC values, followed by cross-validation, and then by subsampling (Table 1). However, bootstrapping and subsampling produced standard deviation of AUC values 75% lower than cross-validation. The Maxent cumulative-based habitat suitability maps were similar for all models, but the relative suitability standard deviation was lowest in the cross-validation (average across grid cells and models, 0.57), followed by subsampling (1.82), and then bootstrapping (2.89). Given this, we used the cross-validation model-fitting technique. Although we incurred more variation in AUC values using cross-validation (while mean AUC values remain high), we reasoned that lower variance in habitat suitability maps is more importance for estimating areas of critical habitat with more confidence.

### Distribution shifts

All Maxent estimations had mean AUC values greater than 0.796, with the August model having the highest, followed by October, September, and then July (Table 2). The variable with the highest percent contribution was split between models: distance to shore in July, chlorophyll-a in August and September, and distance to the 200 m contour in October (Table 2). The highest permutation importance was sea surface temperature in July, distance to shore for August and

**Table 1.** Summary of the Maximum Entropy models comparing model-fitting routines: cross-validation, bootstrapping, and subsampling.

Model-fitting	Model metrics	July	August	September	October
Cross-validation	AUC ± SD	0.796 ± 0.233	0.926 ± 0.078	0.869 ± 0.171	0.892 ± 0.092
Subsampling	AUC ± SD	0.744 ± 0.089	0.925 ± 0.014	0.865 ± 0.044	0.888 ± 0.047
Bootstrapping	AUC ± SD	0.846 ± 0.071	0.938 ± 0.007	0.919 ± 0.014	0.989 ± 0.033
All routines	Top-ranked PC	Shore	Chl-a	Chl-a	200 m
	Top-ranked PI	SST*	Chl-a	Shore	200 m
	Isolation/Omitted	Shore	Chl-a	Chl-a/ Shore	200 m

All models were performed using depth, aspect, slope, distance to shore, distance to 50 m contour, distance to 100 m contour, distance to 150 m contour, distance to 200 m contour, mean chlorophyll a (Chl-a), and mean sea surface temperature. The top-ranked percent contribution (PC), permutation importance (PI), jackknife tests for variable importance of variables in isolation and variable importance when variables are omitted (Isolation/Omitted) are shown.

\*. Top-ranked PI was 150 m for the bootstrapping routine in the July model

doi: 10.1371/journal.pone.0082074.t001

**Table 2.** Summary of Maxent model for basking shark distributions in the Bay of Fundy results using crossvalidation.

Model metrics	July	August	September	October
N	22	110	51	19
ANN	4560 m	1844 m	2920 m	--
Top-ranked PC	Shore (62.2%)	Chl-a (58%)	Chl-a (50.8%)	200 m (77.2%)
2 <sup>nd</sup> ranked PC	150 m (17.0%)	150 m (12.6%)	Shore (18.6%)	100 m (22.5%)
3 <sup>rd</sup> ranked PC	SST (15.3%)	Shore (9.2%)	200 m (13.2%)	Slope (0.2%)
Top-ranked PI	SST (38.9%)	Shore (39.1%)	Shore (58%)	200m (79.9%)
1 <sup>st</sup> Isolation/Omitted	Shore	Chl-a	Chl-a/Shore	200 m/ 200 m
AUC ± SD	0.796 ± 0.233	0.926 ± 0.078	0.869 ± 0.171	0.892 ± 0.092

The numbers of sightings (N) used in the modeling routine as well as the expected average nearest neighbor distance (ANN) used to select sighting locations. The top-, 2<sup>nd</sup>-, and 3<sup>rd</sup>-ranked percent contributions (PC), top-ranked permutation importance (PI), top-ranked jackknife tests for variable importance of variables in isolation and variable importance when variables are omitted (Isolation/Omitted) are shown.

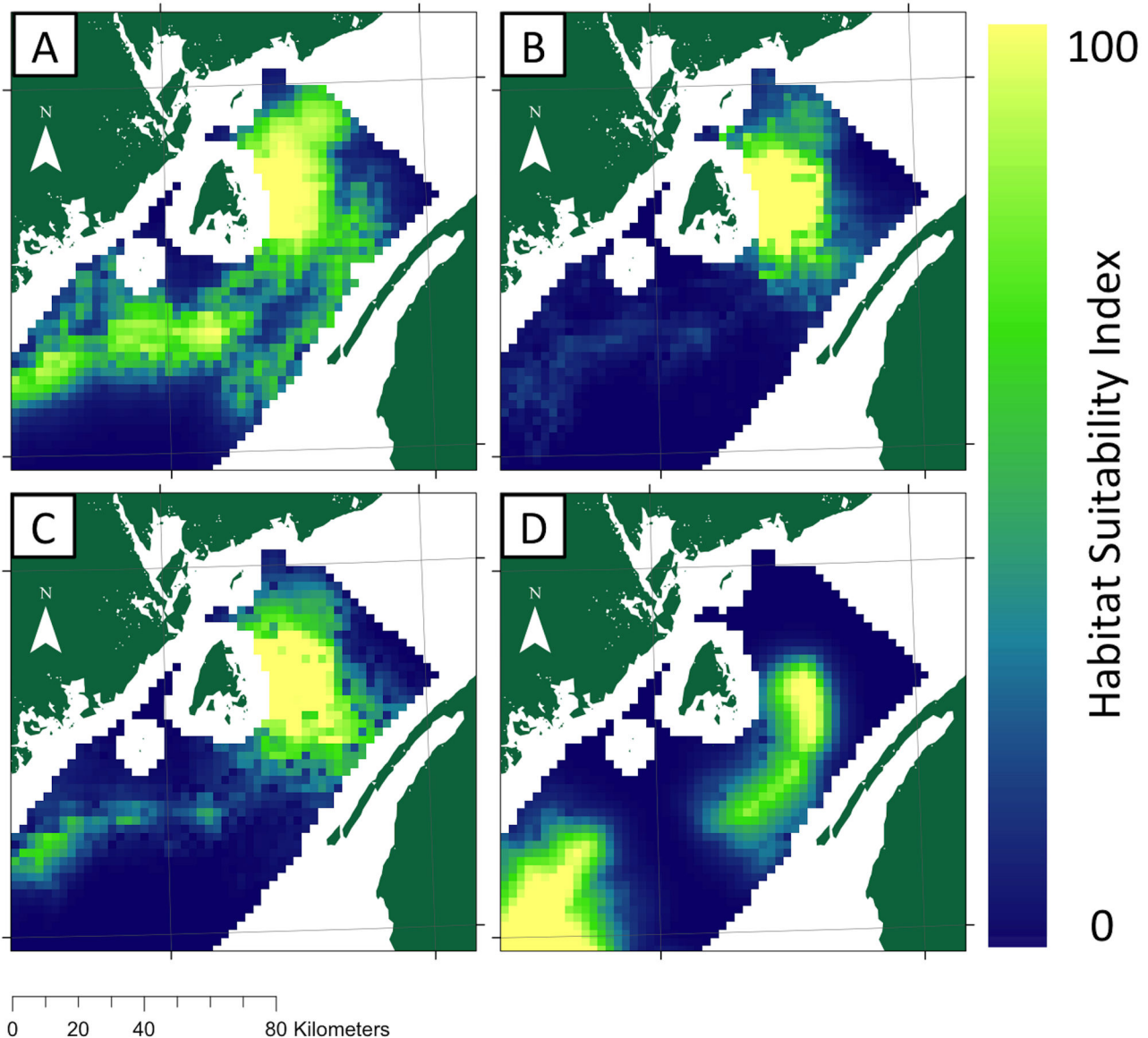
doi: 10.1371/journal.pone.0082074.t002

September, but distance to the 200 m contour in October (Table 2).

Overall, habitat of the highest suitability shifted minimally between August and September (>0.84 between each models, Brownian distance correlation coefficient (BDCC), [52]) (Figure 2). July was the next most similar distribution (>0.44 compared to August and September, BDCC) but predicted a larger extent of suitable habitat (Table 2). All three models predicted the most suitable habitat in the northeast Grand Manan Basin, straddling the 150 m contour and extending southward to the 200 m contour. The main difference between these distributions was the extent of suitable habitat predicted around Murr Ledges (see Figure 1a for location); models for July (Figure 2a) predicted this as highly suitable habitat while models for August (Figure 2b) and September (Figure 2c) did not. Models for July, August, and September all predicted suitable habitat in Jones Ground (see Figure 1a for location) in the northeastern margin of the GoM (Figure 2a-c), though considerably less was predicted in August. Habitat suitability differed in October (<0.036 compared to all other time periods, BDCC) revealing the most suitable habitat near Jones Ground as well as the southern portion of the Grand Manan Basin near the 200 m contour (Figure 2d).

Variable response curves driving the shifts between time periods varied considerably but several trends emerged (Table 2). Throughout July to September, the percent contribution from distance to the shore was in the top three but the model responses varied from month to month (1<sup>st</sup> ranked - July, < 15 km; 3<sup>rd</sup> ranked - August, < 23 km; 2<sup>nd</sup> ranked - September, < 24 km) (Figure S4). In July and August, distance to 150 m contributed the second most to these models (selecting for distances < 3 km and < 7 km from the contour, respectively) (Figure S5). Chlorophyll-a was the highest contributor to the August and September models and responded to concentrations 2.25 - 4 mg m<sup>-3</sup> and > 2 mg m<sup>-3</sup> (Figure S6). In September and October, distance to the 200 m contour was the third and first ranked contributor (selecting for distances < 20 km and < 7 km, respectively) (Figure S7). In July, sea-surface temperature was the third ranked contributor to the model and the variable response was to temperatures 13.2 - 13.8°C (Figure S8). Distance to the 100 m contour was the second ranked contributor to the October model (selecting for distances > 7 km from the contour) (Figure S9).





**Figure 2. Maps of relative habitat suitability of basking sharks in the lower Bay of Fundy.** Maxent distribution models used environmental predictors were chlorophyll-a, sea surface temperature, depth, aspect, slope, and distance to the 0, 50, 100, 150, 200 m contours for (A) July, (B) August, (C) September, and (D) October. Warmer colors indicate higher suitability.

doi: 10.1371/journal.pone.0082074.g002

### Threshold habitat values

Based on the ROC curve for each model, the maximum sensitivity and specificity cumulative suitability score threshold varied (July, 55.63; August, 14.22; September, 30.20; October, 50.33) (Table 3). These thresholds resulted in different amounts of the habitat that were above the threshold and thus likely to be highly suitable basking shark habitat (July, 25.64%; August, 16.89%; September, 14.96%; October, 15.75%) (Table 3). The locations of this habitat occurred throughout the study area in July, almost entirely in the Bay of Fundy in August and

September, then shifted in October so that a majority of habitat was in the Gulf of Maine (Figure 3).

### Discussion

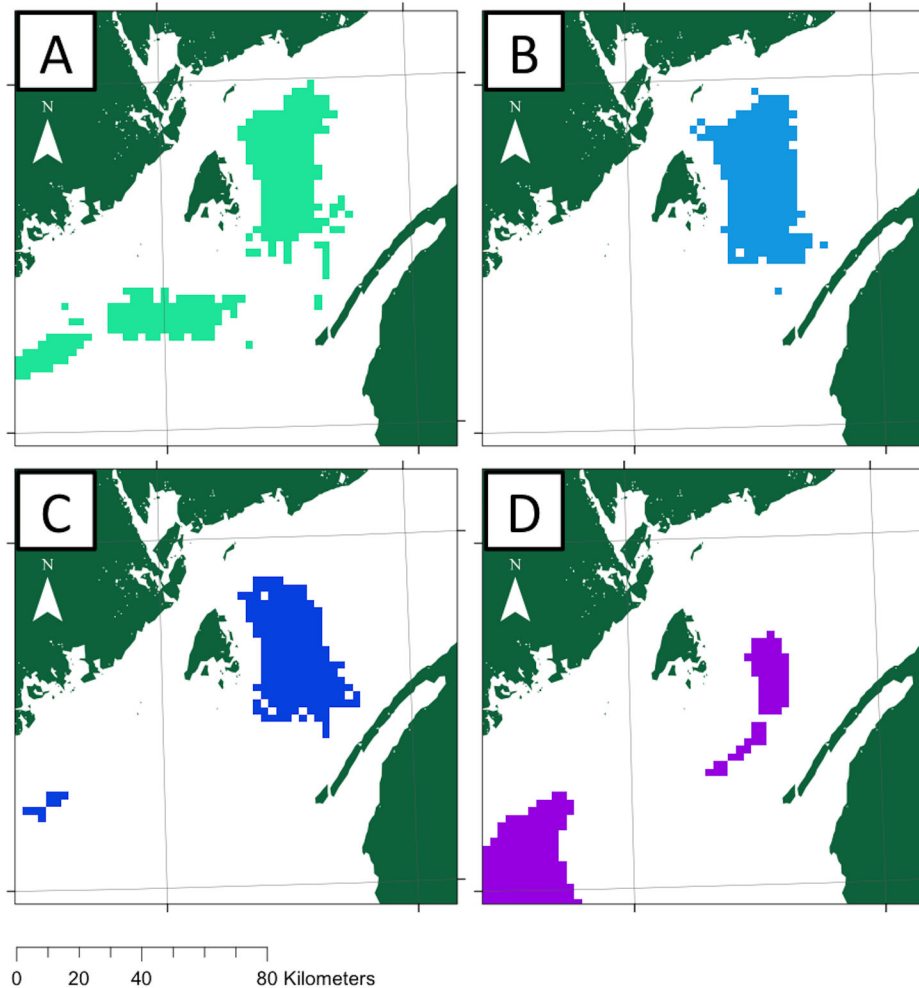
Basking sharks in the western North Atlantic inhabit coastal waters [19] but this is the first study to model their spatial distribution. All of our Maxent estimations performed well, as assessed by AUC values, despite different model-fitting techniques, different combinations of variables, and variables of different resolutions. We conclude from these results that the

**Table 3.** Summary of threshold habitat suitability evaluations.

Approach	July	August	September	October
Threshold value	55.63	14.22	30.20	50.33
Suitable area (km <sup>2</sup> )	1933	1273	1128	1187
Percent suitable habitat	25.64%	16.89%	14.96%	15.75%

The threshold value, suitable area in square kilometers, and percent suitable habitat of total habitat are given.

doi: 10.1371/journal.pone.0082074.t003



**Figure 3. Polygons of threshold defined suitable basking shark habitat.** The maximum sensitivity and specificity threshold was used to distinguish suitable (colored) and unsuitable habitat (white) from Maxent distribution model predictions for July (A), August (B), September (C), October (D).

doi: 10.1371/journal.pone.0082074.g003

maximum entropy approach is appropriate for modeling habitat suitability of basking sharks. Despite variation in the predicted habitat location, a majority of habitat occurs within the BoF from July-September. The localization of this habitat presents an opportunity to protect this species of *Special Concern* in this area [7].

**Model performance**

The results of our Maxent models reveal the consequences of using different model-fitting techniques of bootstrapping, subsampling, and cross-validation. We chose to use cross-validation because our primary focus was assessing spatial differences in basking shark distributions, not overall model performance. We show that when using other model-fitting

techniques within Maxent with large presence-only data sets, a trade-off between model performance and variation in habitat suitability predictions may occur. The trade-off is relatively simple when models perform well. However, as model performance deteriorates, sacrifices in habitat suitability confidence may be necessary to maintain adequate model function. As we demonstrated, using the bootstrapping model-fitting technique improved model performance both in mean and standard deviation of AUC values but at the sacrifice of increasing uncertainty in habitat suitability maps. Importantly, the change from cross-validation to either subsampling or bootstrapping did not change variable importance.

### Implications of spatial heterogeneity and temporal shifts

Models for July–October denoted variation in the distribution of basking sharks (Figure 2 & 3) and these distributional changes are likely driven by changes in the distribution and relative concentration of their main prey. Basking sharks in the BoF would encounter the densest swarms of C5 copepods in the Grand Manan Basin during August [26,28,53]. The rotational transit time of the seasonal BoF Gyre increases the basal residence time of zooplankton within the BoF [54] and increases the concentration of C5 copepods. If, like NA right whales, basking sharks feed on C5s only during summer, then it would be adaptive for them to be tied closely to their prey especially during times when it is more concentrated (Figure 2b).

Owen (1984) also observed evidence for habitat shifts by documenting basking shark distributions in the GoM changing from shallow water to deep water in September. Our model predicted an earlier shift to deep-water habitats in the BoF (August) that may reflect differences in the prey field between these habitats. Owen (1984) documented copepod swarms in shallow (<80 m) waters of the GoM throughout the summer and observed surface feeding sharks primarily along the western edge of Jordan Basin from June to August. These patches were likely *C. finmarchicus* C4 and C5, presumably feeding on phytoplankton prior to the C5 diapause [55,56]. Habitat shifts have also been documented in the eNA, where basking sharks moved from the neritic zone, consuming primarily shallow-water *Calanus helgolandicus*, to the continental edge where they consumed primarily deep-water *C. finmarchicus* [13].

Interestingly, the narrow bands of suitable habitat along the Murr Ledges in the July suitability map overlap with the outflow of the cyclonic Bay of Fundy Gyre [20] (Figure 2a and see Figure S10). This suggests basking sharks may follow the advection of prey via these fast currents when high concentrations of prey are not as available as in August and September. Similarly in the eNA, basking sharks have exhibited area-restrictive foraging by closely following thermal and tidal fronts as well as prey density differences [57,58].

The appearance of suitable habitat further south in September (Figure 2d) likely represents the onset of southerly movement of basking sharks out of the BoF. Limited observations of basking sharks moving from the western margins of the GoM to the northwestern margin of Georges Bank [south of the GoM] (Owen, 1984) tentatively support this.

Pop-up satellite archival tag deployments in the BoF (Westgate A.J., unpublished data) have shown that basking sharks typically leave BoF/GoM waters between mid-October and mid-November, quickly moving into much deeper water off the shelf break. Three hypotheses have been suggested for a southward distributional shift: 1) basking sharks are following the movements (both vertically and horizontally) of the most energetically favorable prey [12,59–61] (Parker & Boeseman, 1954; Owen, 1984; Sims, 1999; Sims et al., 2005); 2) basking sharks are leaving prey patches in the BoF as abundances declines [28]; or 3) ectothermic basking sharks must maintain body temperatures above some threshold [12,15] and thus move to where ocean temperatures are suitable [57].

Southerly movement of the basking shark distribution, as predicted in our October model (Figure 2d), does not exclusively support any of these hypotheses. As the environmental variables used in this study, Chl-*a* and SST, are limited to describing surface changes and cannot assess C5 *C. finmarchicus* dynamics or thermal stratification, we can only speculate on the mechanism behind distribution shifts. Owen (1984) proposed that basking sharks in the wNA may follow warm surface water as it undergoes vertical mixing because animals are thermally constrained, allowing October (or later) access to deeper-water habitats, thus eliminating basking sharks from surface sighting records. The shift to deep-water habitats in the October distribution in this study likely reflects such a movement as the model shifted from 180 m in depth in July–September to > 200 m in October–December. Additionally, response to the distance to the 200 m contour decreased from 20 km from the contour in September to less than 7 km in October. However, SST did not strongly influence the October distribution which implies that temperature is less important than prey requirements or that SST does not capture the thermal environment basking sharks are responding to [13,58,61]. Additionally, the breakdown of thermal stratification weakens the cyclonic Bay of Fundy Gyre [20], and may cause C5 ejection into the GoM, which basking sharks may follow. Since we do not understand the complexities of phytoplankton, copepod and basking shark trophodynamics in this region these findings are difficult to evaluate with any certainty.

### Threshold habitat values

The model predictions have shown that sharks regularly use select areas within the BoF. These select areas may offer opportunities for management strategies, such as shifting traffic patterns and creating protection zones for basking sharks in the northern Grand Manan Basin [19,62]. Habitat suitability as defined by the maximum sensitivity and specificity threshold significantly reduced predicted habitat area, ranging from 14.86 to 25.64% of the total habitat. This has broad implications for conservation measures, as protection of a fraction of the total habitat would translate to protection for a disproportionately greater number of sharks. It is important to note, however, that the threshold index is derived from the modeling process and associated ecological inferences, and should be validated through an independent long-term dataset, such as a decade-long aerial survey series, or via satellite telemetry tagging on a significant number of sharks.



## Conclusion

We have developed the first maximum entropy model for basking sharks and made predictions of their distribution in the BoF. Suitable habitat was shown to vary between July–October over 2002–2011, indicating heterogeneous habitat utilization. We hypothesized these distributional shifts are related to prey dynamics but further understanding of the ecological drivers is necessary to elucidate the specific mechanisms of basking shark movements in the western North Atlantic. Also, we have compared model-fitting routines and reasoned that their choice must be weighed against study objectives. Defining the spatial distribution of basking sharks, as we have done, represents a critical step in future efforts towards conserving and understanding this large but understudied species, especially in data deficient areas such as Atlantic Canada.

## Supporting Information

**Figure S1.** Map for July–October of the distribution of sampling effort throughout the study area.

(DOCX)

**Figure S2.** Semi-variograms generated from sighting locations and environmental variables showing the spatial autocorrelation in the dataset used for Maxent estimations.

(DOCX)

**Figure S3.** Moran's *I* of sightings locations and environmental variables as a function of distance used in an attempt to decimate sampling locations in order to correct for spatial autocorrelation.

(DOCX)

**Figure S4.** Model responses to distance to shore for July–September against a histogram of distance to shore in the study area.

(DOCX)

**Figure S5.** Model responses to distance to the 150 m contour for July and August against a histogram of distance to the 150 m contour in the study area.

## References

- Compagno LJV (1984) FAO species catalogue. Vol. 4. Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Part 1. Hexanchiformes to Lamniformes. FAO Fisheries Synopsis (125) Vol. 4, Pt. 1. Rome, FAO. 249 pp
- Sims DW (2008) Sieving a living: A review of the biology, ecology, and conservation status of the plankton-feeding basking shark *Cetorhinus maximus*. *Advances in Marine Biology*. Elsevier Academic Press Inc., San Diego. pp. 171–220.
- Compagno LJV (2002) Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Volume 2. Bullhead, mackerel, and carpet sharks (Heterodontiformes, Lamniformes, and Orectolobiformes). FAO Species Catalogue for Fishery Purposes, No. 1, Vol. 2. Rome, FAO. 269pp
- Fowler SL (2005) *Cetorhinus maximus*. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2. IUCN 2011.
- Witt MJ, Hardy T, Johnson L, McClellan CM, Pikesley SK et al. (2012) Basking sharks in the northeast Atlantic: spatio-temporal trends from sightings in UK waters. *Marine Ecology Progress Series*, 459: 121–134. doi:10.3354/meps09737.
- IUCN (2012). The IUCN Red List of Threatened Species. Version 2012.2.
- COSEWIC (2009) COSEWIC Assessment and Status Report on the Basking Shark *Cetorhinus maximus*, Atlantic population, in Canada.
- Sims DW, Reid PC (2002) Congruent trends in long-term zooplankton decline in the north-east Atlantic and basking shark (*Cetorhinus maximus*) fishery catches off west Ireland. *Fisheries Oceanography*, 11: 59–63. doi:10.1046/j.1365-2419.2002.00189.x.
- COSEWIC (2007). COSEWIC Assessment and Status Report on the Basking Shark *Cetorhinus maximus*, Pacific population, in Canada.
- Baum JK, Myers RA, Kehler DG, Worm B, Harley SJ et al. (2003) Collapse and conservation of shark populations in the Northwest Atlantic. *Science*, 299: 389–392. doi:10.1126/science.1079777. PubMed: 12532016.

(DOCX)

**Figure S6.** Model responses to chlorophyll-*a* for August and September against a histogram of chlorophyll-*a* in August and September in the study area.

(DOCX)

**Figure S7.** Model responses to distance to 200 m contour for September and October against a histogram of distance to 200 m contour in the study area.

(DOCX)

**Figure S8.** Model responses to sea surface temperature for July against a histogram of sea surface temperature in July in the study area.

(DOCX)

**Figure S9.** Model responses to distance to the 100 m contour for October against a histogram of distance to the 100 m contour in the study area.

(DOCX)

**Figure S10.** Map of major currents affecting the Bay of Fundy Gyre circulation.

(DOCX)

## Acknowledgements

This project would not have been possible without the support of the University of North Carolina Wilmington, the Grand Manan Whale and Seabird Research Station, the North Atlantic Right Whale Consortium, Whales n' Sails, Sea Land Adventures (Schooner D'Sonoquo) and Deane Griffin. We would like to thank J. Wilson White for helpful comments and Gina Lonati for editing assistance.

## Author Contributions

Conceived and designed the experiments: ZS DJ. Performed the experiments: ZS. Analyzed the data: ZS AW HK. Contributed reagents/materials/analysis tools: LM. Wrote the manuscript: ZS AW HK.

11. Dulvy NK, Baum JK, Clark S, Compagno LJV, Cortes E et al. (2008) You can swim but you can't hide: the global status and conservation of oceanic pelagic sharks and rays. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 18: 459-482. doi:10.1002/aqc.975.
12. Owen RE (1984) Distribution and ecology of the basking shark *Cetorhinus maximus* (Gunnerus 1765). University of Rhode Island. (thesis)
13. Sims DW, Southall EJ, Richardson AJ, Reid PC, Metcalfe JD (2003) Seasonal movements and behavior of basking sharks from archival tagging: no evidence of winter hibernation. *Marine Ecology Progress Series*, 248: 187-196. doi:10.3354/meps248187.
14. Southall EJ, Sims DW, Witt MJ, Metcalfe JD (2006) Seasonal space-use estimates of basking sharks in relation to protection and political-economic zones in the NE Atlantic Biological Conservation. 132: 33-39. doi:10.1016/j.biocon.2006.03.011.
15. Skomal GB, Zeeman SI, Chisholm JH, Summers EJ, Walsh HJ et al. (2009) Transequatorial migrations by basking sharks in the western Atlantic. *Oceans - Current Biology*, 19: 1019-1022. doi:10.1016/j.cub.2009.04.019.
16. Southall EJ, Sims DW, Metcalfe JD, Doyle JI, Fanshawe S et al. (2005) Spatial distribution patterns of basking sharks on the European shelf: preliminary comparison of satellite-tag geolocation, survey, and public sightings data. *Journal of the Marine Biological Association of the United Kingdom*, 85: 1083-1088. doi:10.1017/S0025315405012129.
17. Kenney RD, Owen RE, Winn HE (1985) Shark distributions off the Northeast United States from marine mammal surveys. *Copeia*, 1: 220-223.
18. Lien J, Fawcett L (1986) Distribution of basking sharks, *Cetorhinus maximus*, incidentally caught in inshore fishing gear in Newfoundland. *Canadian Field-Naturalist*, 100: 246-252.
19. DFO (2008) Status of Basking Sharks in Atlantic Canada. Science Advisory Report, 2008/036.
20. Aretxabaleta AL, McGillicuddy DJ, Smith KW, Lynch DR (2008) Model simulations of the Bay of Fundy Gyre: 1. Climatological results. *Journal of Geophysical Research-Oceans*, 113: C10027. doi: 10.1029/2007JC004480.
21. Fish CJ, Johnson MW (1937) The biology of the zooplankton population in the Bay of Fundy and Gulf of Maine with special reference to production and distribution. *Journal Biological Board of Canada*, 3: 189-322. doi:10.1139/f37-015.
22. Brown RGB, Gaskin DE (1988) The pelagic ecology of the Grey and Red-necked phalaropes *Phalaropus fulicarius* and *P. lobatus* in the Bay of Fundy, eastern Canada. *Ibis*, 130: 234-250.
23. Baumgartner MF, Cole TVN, Clapham PJ, Mate BR (2003) North Atlantic right whale habitat in the lower Bay of Fundy and on the SW Scotian Shelf during 1999-2001. *Marine Ecology Progress Series*, 264: 137-154. doi:10.3354/meps264137.
24. Johnston DW, Thorne LH, Read AJ (2005a) Fin whales *Balaenoptera physalus* and minke whales *Balaenoptera acutorostrata* exploit a tidally driven island wake ecosystem in the Bay of Fundy. *Marine Ecology Progress Series*, 305: 287-295. doi:10.3354/meps305287.
25. Johnston DW, Westgate AJ, Read AJ (2005b) Effects of fine-scale oceanographic features on the distribution and movements of harbour porpoise *Phocoena phocoena* in the Bay of Fundy. *Marine Ecology Progress Series*, 295: 279-295. doi:10.3354/meps295279.
26. Murison LD, Gaskin DE (1989) The distribution of right whales and zooplankton in the Bay of Fundy, Canada. *Canadian Journal of Zoology*, 67: 1411-1420. doi:10.1139/z89-200.
27. Hirche HJ (1996) Diapause in the marine copepod, *Calanus finmarchicus* -A review. *Ophelia*, 44: 129-143. doi: 10.1080/00785326.1995.10429843.
28. Michaud J, Taggart CT (2007) Lipid and gross energy content of North Atlantic right whale food, *Calanus finmarchicus*, in the Bay of Fundy. *Endangered Species Research*, 3: 77-94. doi:10.3354/esr003077.
29. McKinstry CAE, Westgate AJ, Koopman HN (2013) Annual variation in the nutritional value of stage V *Calanus finmarchicus*: implications for right whales and other copepod predators. *Endangered Species Research*. 20: 195-204. doi:10.3354/esr00497.
30. Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, 157: 101-118. doi:10.1016/S0304-3800(02)00205-3.
31. Elith J, Leathwick JR (2009). Annual Review of Ecology Evolution, and Systematics - Annual Review of Ecology, Evolution, and Systematics: 677-697.
32. Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190: 231-259. doi:10.1016/j.ecolmodel.2005.03.026.
33. Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE et al. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17: 43-57. doi:10.1111/j.1472-4642.2010.00725.x.
34. Edren SMC, Wisz MS, Teilmann J, Deitz R, Soderkvist J (2010) Modeling spatial patterns in harbour porpoise satellite telemetry data using maximum entropy. *Ecography*, 33: 698-708. doi:10.1111/j.1600-0587.2009.05901.x.
35. Friedlaender AS, Johnston DW, Fraser WR, Burns J, Halpin PN et al. (2011) Ecological niche modeling of sympatric krill predators around Marguerite Bay, Western Antarctic Peninsula. *Deep-Sea Research Part II*, 58: 1729-1740. doi:10.1016/j.dsr2.2010.11.018.
36. Sequeira A, Mellin C, Rowat D, Meekan MG, Bradshaw CJA (2011) Ocean-scale prediction of whale shark distribution. *Diversity and Distributions*, 18: 504-518.
37. Ballard G, Jongsomjit D, Veloz SD, Ainley DG (2012) Coexistence of mesopredators in an intact polar ocean ecosystem: The basis for defining a Ross Sea marine protected area. *Biological Conservation*, 156: 72-82. doi:10.1016/j.biocon.2011.11.017.
38. Thorne LH, Johnston DW, Urban DL, Tyne J, Bejder L et al. (2012) Predictive Modeling of Spinner Dolphin (*Stenella longirostris*) Resting Habitat in the Main Hawaiian Islands. *PLOS ONE*, 7: e43167. doi: 10.1371/journal.pone.0043167. PubMed: 22937022.
39. Ainley DG, Jongsomjit D, Ballard G, Thiele D, Fraser WR et al. (2012) Modeling the relationship of Antarctic minke whales to major ocean boundaries. *Polar Biology*, 35: 281-290. doi:10.1007/s00300-011-1075-1.
40. McKinney JA, Hoffmayer ER, Wu W, Fulford R, Hendon J (2012) Feeding habitat of the whale shark *Rhincodon typus* in the northern Gulf of Mexico determined using species distribution modelling. *Marine Ecology Progress Series*, 458: 199-211. doi:10.3354/meps09777.
41. Pendleton DE, Sullivan PJ, Brown MW, Cole TVN, Good CP et al. (2012) Weekly predictions of North Atlantic right whale *Eubalaena glacialis* habitat reveal influence of prey abundance and seasonality of habitat preferences. *Endangered Species Research*, 18: 147-161. doi: 10.3354/esr00433.
42. Right Whale Consortium (2011) North Atlantic Right Whale Consortium Sightings Database. 04/05/2012 (New England Aquarium, Boston, MA, U.S.A.).
43. Brown MW, Kraus SD, Slay CK, Garrison LP (2007) Surveying for Discovery, Science, and Management. The Urban Whale: North Atlantic Right Whales at the Crossroads. SD KrausRM Rolland. Harvard University Press, Cambridge, MA. pp. 105-137.
44. Shephard ELC, Ahmed MZ, Southall EJ, Witt MJ, Metcalfe JD et al. (2006) Diel and tidal rhythms in diving behaviour of pelagic sharks identified by signal processing of archival tagging data. *Marine Ecology Progress Series*. 328: 205-213. doi:10.3354/meps328205.
45. Roberts JJ, Best BD, Dunn DC, Trembl EA, Halpin PN (2010) Marine Geospatial Ecology Tools: An integrated framework for ecological geospatial processing with ArcGIS, Python, R, MATLAB, and C. *Environmental Model and Software*, 25: 1197 -1207. doi:10.1016/j.envsoft.2010.03.029.
46. Phillips SJ, Dudik M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31: 161-175. doi:10.1111/j.0906-7590.2008.5203.x.
47. Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29: 129-151. doi:10.1111/j.2006.0906-7590.04596.x.
48. Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol Appl*, 21: 335-342. doi:10.1890/10-1171.1. PubMed: 21563566.
49. Liu C, White M, Newell G (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. 40: 778-789.
50. Lobo JM, Jimenez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17: 145-151. doi:10.1111/j.1466-8238.2007.00358.x.
51. Stachura-Skierczynska K, Tumiel T, Skierczynski M (2009) Habitat prediction model for three-toed woodpecker and its implications for the conservation of biologically valuable forests. *Forest Ecology and Management*, 258: 697-703. doi:10.1016/j.foreco.2009.05.007.
52. Szekely GJ, Rizzo MJ, Bakirov NK (2007) Measuring and testing dependence by correlation of distances. *Annals of Statistics*, 35: 2769-2794. doi:10.1214/009053607000000505.
53. Baumgartner MF, Mate BR (2003) Summertime foraging ecology of North Atlantic right whales. *Marine Ecology Progress Series*, 264: 123-135. doi:10.3354/meps264123.

54. Aretxabaleta AL, McGillicuddy DJ, Smith KW, Manning JP, Lynch DR (2009) Model simulations of the Bay of Fundy Gyre: 2. Hindcasts for 2005-2007 reveal interannual variability in retentiveness. *Journal of Geophysical Research-Oceans*, 114: C09005.
55. Wishner KF, Schoenherr JR, Beardsley R, Chen CS (1995) Abundance, distribution, and population-structure of the copepod *Calanus finmarchicus* in a springtime right whale feeding area in the southwestern Gulf of Maine. *Continental Shelf Research*, 15: 475-507. doi:10.1016/0278-4343(94)00057-T.
56. Beardsley RC, Epstein AW, Chen CS, Wishner KF et al. (1996) Spatial variability in zooplankton abundance near feeding right whales in the Great South Channel. *Deep-Sea Research II*, 43: 1601-1625. doi:10.1016/S0967-0645(96)00050-1.
57. Sims DW, Quayle VA (1998) Selective foraging behavior of basking sharks on zooplankton in a small-scale front. *Nature*, 393: 460-464. doi:10.1038/30959.
58. Sims DW, Southall EJ, Merrett DA, Sanders J (2003) Effects of zooplankton density and diel period on surface-swimming duration of basking sharks. *Journal of the Marine Biological Association of the United Kingdom*, 83: 643-646. doi:10.1017/S0025315403007598h.
59. Parker HW, Boeseman M (1954) The basking shark, *Cetorhinus maximus*, in winter. *Proceedings of the Zoological Society of London*, 124: 185-194.
60. Sims DW (1999) Threshold foraging behavior of basking sharks on zooplankton: life of an energetic knife-edge? *Proceedings of the Royal Society of London-B. Biological Sciences*, 266: 1437-1443. doi:10.1098/rspb.1999.0798.
61. Sims DW, Southall EJ, Tarling GA, Metcalfe JD (2005) Habitat-specific normal and reverse diel vertical migration in the plankton-feeding basking shark. *Journal of Animal Ecology*, 74: 755-761. doi:10.1111/j.1365-2656.2005.00971.x.
62. Cotton PA, Sims DW, Fanshawe S, Chadwick M (2005) The effects of climate variability on zooplankton and basking shark (*Cetorhinus maximus*) relative abundance off southwest Britain. *Fisheries Oceanography*, 14: 151-155. doi:10.1111/j.1365-2419.2005.00331.x.