

Using modelled prey to predict the distribution of a highly mobile marine mammal

Daniel E. Pendleton¹  | Elizabeth E. Holmes²  | Jessica Redfern¹  | Jinlun Zhang³

¹Anderson Cabot Center for Ocean Life at the New England Aquarium, Boston, MA, USA

²Northwest Fisheries Science Center, National Marine Fisheries Service, Seattle, WA, USA

³Applied Physics Laboratory, University of Washington, Seattle, WA, USA

Correspondence

Daniel E. Pendleton, Anderson Cabot Center for Ocean Life at the New England Aquarium, Boston, MA 02110, USA.
Email: dpendleton@neaq.org

Funding information

National Aeronautics and Space Administration, Grant/Award Number: NNX11AO93G and NNX15AG68G

Editor: Luigi Maiorano

Abstract

Aim: Species distribution models (SDMs) are a widely used tool to estimate and map habitat suitability for wildlife populations. Most studies that model marine mammal density or distributions use oceanographic proxies for marine mammal prey. However, proxies could be a problem for forecasting because the relationships between the proxies and prey may change in a changing climate. We examined the use of model-derived prey estimates in SDMs using an iconic species, the western Arctic bowhead whale (*Balaena mysticetus*).

Location: Western Beaufort Sea, Alaska, USA.

Methods: We used Biology Ice Ocean Modeling and Assimilation System (BIOMAS) to simulate ocean conditions important to western Arctic bowhead whales, including important prey species. Using both static and dynamic predictors, we applied Maxent and boosted regression tree (BRT) SDMs to predict bowhead whale habitat suitability on an 8-day timescale. We compared results from models that used bathymetry with those that used only BIOMAS simulated variables.

Results: The best model included bathymetry and BIOMAS variables. Inclusion of dynamic variables in SDMs produced predictions that reflected temporal dynamics evident from the survey data. Bathymetry was the most influential variable in models that included that variable. Zooplankton was the most important variable for models that did not include bathymetry. Models with bathymetry performed slightly better than models with only BIOMAS derived variables.

Main conclusions: Bathymetry and modelled zooplankton were the most important predictor variables in bowhead whale distribution models. Our predictions reflected within-year variability in bowhead whale habitat suitability. Using modelled prey availability, rather than oceanographic proxies, could be important for forecasting species distributions. Predictor variables used in our study were derived from a biophysical ocean model with demonstrated ability to project future ocean conditions. A natural next step is to use output from our biophysical ocean model to understand the effects of Arctic climate change.

KEYWORDS

Arctic, boosted regression tree, bowhead whale, forecast, hindcast, maxent, ocean model, species distribution model, zooplankton

1 | INTRODUCTION

Species distribution models (SDMs) are a widely used tool to estimate and map habitat suitability for wildlife populations. Correlative SDMs associate species presence or presence–absence data with environmental predictor variables, fit a relationship between predictor and response, and make predictions for an area of interest (Redfern et al., 2006). These types of SDMs have been used to conduct near real-time monitoring (Abrahms et al., 2019; Hazen et al., 2017) and address a number of conservation challenges, including evaluating and mitigating ship-strike risk (Redfern, Becker, & Moore, 2020; Redfern et al., 2019), and they have shown potential in forecasting species distributions (Becker et al., 2018).

Many studies that model marine animal density or distributions use oceanographic proxies for prey (e.g., Becker et al., 2016; Redfern et al., 2006; Sequeira, Mellin, Rowat, Meekan, & Bradshaw, 2012), which is expected to be the true driver of species' migratory phenology and their distribution on foraging grounds. Oceanographic proxies are used because prey data are often unavailable at the spatial and temporal extents, and resolutions, needed to develop species distribution models. This lack of availability can be attributed to the logistical difficulties and costs associated with comprehensive sampling in marine environments. Oceanographic proxies may include static geographic features (e.g. depth, slope and distance from the coast), climatologically averaged environmental data (e.g. World Ocean Atlas (Locarnini et al., 2013)) and dynamic variables (e.g. sea surface temperature, salinity and chlorophyll) from in situ sampling, remote-sensing platforms or ocean models. However, the use of oceanographic proxies could be a problem for forecasting species distributions because the relationships between the proxies and prey distribution or density may change in a changing climate (GREG, Baumgartner, Laidre, & Palacios, 2013; Silber et al., 2017). Further complications may arise from climate-induced shifts in the distribution and phenology of predators and their prey (Carroll et al., 2019). Finally, the use of proxies obtained from limited in situ and modelled oceanographic data may not represent the spatial and temporal scales of processes important to foraging, such as processes that aggregate prey into concentrations sufficient to trigger predator feeding activity (e.g. Carroll et al., 2017; Mayo & Marx, 1990).

To overcome the limitations associated with oceanographic proxies, we need to use environmental variables known to have a causal relationship with the distribution and concentration of prey species. Several studies have demonstrated that the use of prey data improved the performance of ecological models for fishes (Lezama-Ochoa, Boyra, Goñi, Arrizabalaga, & Bertrand, 2010; Schick & Lutcavage, 2009), sharks (Wirsing, Heithaus, & Dill, 2007), birds (Miller, Carlile, Scutt Phillips, McDuie, & Congdon, 2018) and cetaceans (Lambert, Mannocci, Lehoudey, & Ridoux, 2014; Pendleton

et al., 2012). Conversely, some studies (e.g. Torres, Read, & Halpin, 2008) did not find prey to be a useful predictor in comparison to proxies.

In the present study, we examined the use of model-derived prey estimates in SDMs of an iconic marine species, the western Arctic bowhead whale (*Balaena mysticetus*). Western Arctic bowhead whales engage in seasonal migrations between the Bering, Chukchi and Beaufort seas (Moore & Reeves, 1993) and, as such, are also widely known as the Bering–Chukchi–Beaufort population. They spend winter in the Bering Sea. In April–May they migrate to the Chukchi Sea. Most whales then migrate eastward to feeding grounds in the Canadian Beaufort Sea, arriving in May–June (Moore & Reeves, 1993; Quakenbush et al., 2012). From mid-July to October, whales migrate westward and return to the Bering Sea (Mate, Krutzikowsky, & Winsor, 2000; Quakenbush et al., 2012; Quakenbush, Small, & Citta, 2013). Our study focuses on this late summer/early fall period when bowhead whales are migrating westward, offshore of the north coast of Alaska, USA, where they largely feed on copepods and euphausiids (Lowry, 1993; Lowry, Sheffield, & George, 2004).

We used SDMs to model the bowhead whale distribution in this region during the late summer/early fall season from 1988–2012. Extensive aerial surveys have focused effort in this area during this season because it is an important bowhead whale migration and foraging area. We paired bowhead whale presence–absence data from aerial surveys with bathymetry data and ocean conditions derived from the Biology Ice Ocean Modeling and Assimilation System (BIOMAS) Arctic Ocean model (Zhang et al., 2010, 2015). BIOMAS is a pan-Arctic biophysical model which estimates many physical and biological variables that are important to bowhead whales, including sea temperature, sea ice thickness, primary productivity and zooplankton groups known to be consumed to bowhead whales, namely copepods and predator zooplankton (e.g. euphausiids). We trained and tested Maxent and boosted regression tree (BRT) SDMs with data from mutually exclusive training and testing years and used the models to predict bowhead whale habitat suitability for each 8-day period from late August through late October of each test year.

The objectives of our study were to: (a) Assess the relative utility of bathymetry and model-derived predictor variables, including prey availability (from BIOMAS), to drive SDMs that correctly categorize bowhead whale habitat use in space and time, and (b) Assess the value of modelled zooplankton prey versus physical variables for predicting bowhead whale habitat use. These two objectives provide the foundational research needed to successfully use SDMs to forecast bowhead whale habitat under different climate change scenarios. More generally, these objectives contribute to our ability to predict how marine mammal distributions will change in response to climate-driven changes in oceanographic conditions.

2 | METHODS

2.1 | Study area

Our study area is the western Beaufort Sea, from 140–160° W and from the Alaska coast to 72°N (Figure 1a). Our study period begins in 1988 and ends in 2012. For each year, we analysed data from 21 August–31 October. The region and dates of our study were determined by the dates and area covered by the bowhead whale survey programs used in this study.

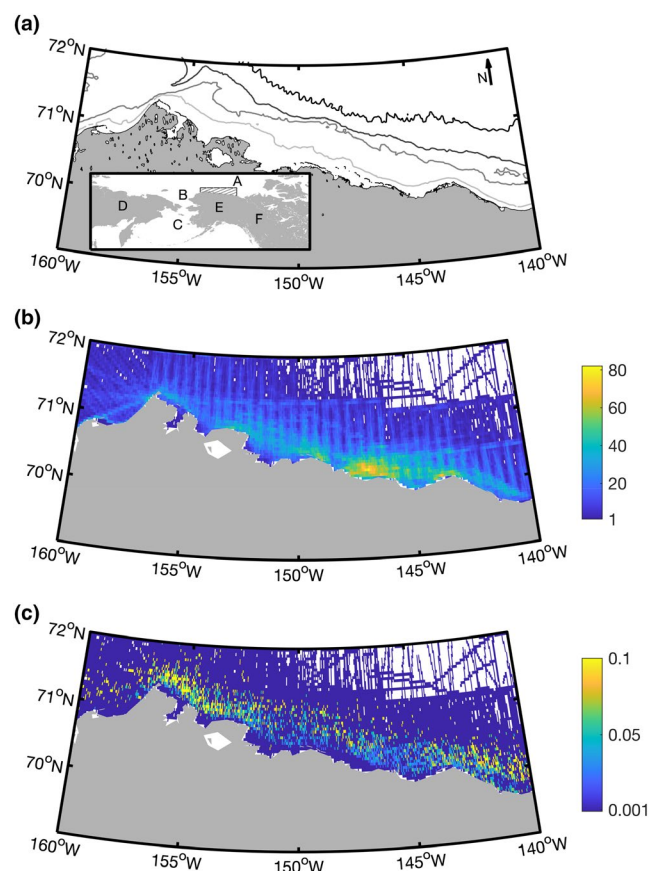


FIGURE 1 Study area with Alaska shown in grey. Inset shows the larger region including our study area (hatched), the Beaufort (A), Chukchi (B) and Bering (C) seas, and Russia (D), Alaska USA (E) and Canada (F). Bathymetric contours at -20, -50, -200 and -2,000 m are shown from light to dark. Point Barrow is the northernmost point of land. (a). Survey effort shown as the number of visits to each grid cell, with a maximum of one visit per 8-day period per year in the time frame 21 August–31 October between the years 1988–2012 (b). Bowhead whale sighting rate calculated as the number of bowhead whale presence records (grid cells with ≥ 1 bowhead whale) divided by the number visits to each grid cell after binning survey data by 8-day period within years. Colour scale compressed to (0 0.1] to highlight cells with bowhead whale sightings (c)

2.2 | Presence-absence data

Observations of bowhead whale presence and absence were collected during aerial surveys of the Alaskan Beaufort and Chukchi seas. Our study used data from both areas and data collection programs. These surveys were originally intended to provide real-time information and data synthesis products to the U.S. Department of the Interior and other agencies regarding the fall migration of bowhead whales. Several U.S. government entities and their contractors have managed the aerial surveys since inception in 1979 (Clarke, Brower, Ferguson, & Willoughby, 2017). Since 2008 the surveys have been conducted by the U.S. Department of Commerce National Marine Fisheries Service (NMFS) and co-managed by NMFS and the Bureau of Ocean Energy Management. In aggregate, data collected during the surveys described here constitute one of the longest continuously running marine mammal monitoring programs in existence.

During 1988–2012, the years used in our study, the broader survey area spanned a large region which was partitioned into survey blocks of various sizes between 140–169°W, and between the Bering Strait (65°40'N) and 73°N. Surveys in most years were conducted between early September and late October, with surveys beginning earlier and ending later in some years. Additional details on the history of these surveys and their methodologies are provided in Appendix S1 and Figures S1.1 and 1.2.

Prior to 2006, the role and position of observers, primary versus secondary and port versus starboard, was not well documented for all surveys. However, each flight carried a minimum of two observers, a data recorder and two pilots executing the survey design described in Appendix S1. Improved documentation of in-flight observation protocols began in 2006, with a primary observer stationed at port and starboard bubble windows, and a dedicated data recorder that served as a secondary observer (Clarke, Christman, Brower, & Ferguson, 2012). A few recent studies have analysed data from surveys after 2008, when survey protocols were more homogeneous (Brower, Ferguson, Schonberg, Jewett, & Clarke, 2017; Clarke, Ferguson, Willoughby, & Brower, 2018). However, because our study included survey data from 1988–2012, a period of time when data were collected with different aerial survey platforms and without comprehensive documentation of observer configurations and data collection protocols, we treated the entire dataset as a presence-absence dataset.

2.3 | Environmental predictor variables

We derived a suite of environmental predictor variables from BIOMAS (Zhang et al., 2010). Output from ocean models such as BIOMAS are contiguous in space and time and do not contain data gaps due to cloud cover, common in observations from earth-orbiting satellites. This feature of ocean models makes their output attractive for use as predictors in SDMs. BIOMAS couples three sub-models for sea ice, ocean circulation, and the pelagic ecosystem, respectively. The ecosystem model is adapted to the Arctic Ocean

based on an 11-component Nutrient-Phytoplankton-Zooplankton-Detritus (NPZD) biogeochemical model (Kishi et al., 2007). With nitrogen and silicon as model currencies (Zhang et al., 2010, table 1), this model includes two phytoplankton components (diatom and flagellate), three zooplankton components (microzooplankton, copepods and predator zooplankton), dissolved organic nitrogen, detrital particulate organic nitrogen, particulate organic silica, nitrate, ammonium and silicate. Additional details on BIOMAS are provided in the Appendix S1. Comparisons of BIOMAS with satellite-observed variables demonstrate that it provides a good representation of the interannual variation sea ice extent, primary productivity, and surface-air temperatures. Comparisons of the model with in situ data show that it captures the timing and magnitude of the spring bloom in the Chukchi and Beaufort seas.

From the suite of BIOMAS outputs, we selected variables we deemed to have greatest potential influence on bowhead whale distribution: sea ice thickness, sea temperature, diatoms, flagellates, copepods and predator zooplankton. We included sea ice thickness and sea temperature because seasonal fluctuations in temperature and ice are dominant features of the Arctic environment, and bowhead whales are an ice-adapted animal whose migrations track seasonal changes in the environment. We used sea ice and sea temperature variables from BIOMAS rather than from satellite observations because BIOMAS assimilates satellite observations of sea ice concentration and sea surface temperature, and modelled values for these parameters reflect the combined effects of satellite observations and model dynamics. We included diatoms and flagellates because they represent the phytoplankton, and chlorophyll is often used as a proxy for phytoplankton and primary productivity in SDMs. We included copepod and predator zooplankton groups because bowhead whales are known to feed on organisms in both of these general zooplankton categories within our study area during the season under study (Lowry, 1993). We included bathymetry because previous studies have identified it as an important feature of bowhead whale habitat in the Beaufort Sea (e.g. Clarke et al., 2018). We acquired bathymetry data from the International Bathymetric Chart of the Arctic Ocean version 3.0 (Jakobsson et al., 2012).

2.4 | Data selection and pre-processing

Aerial survey data and environmental predictor variables were temporally binned into 8-day periods and were spatially gridded to 4.6 km to conform to the native resolution of BIOMAS in the Beaufort and Chukchi seas (Zhang et al., 2010), NOAA PATHFINDER v5.2 (Casey, Brandon, Cornillon, & Evans, 2010) and to other datasets available at these common spatial and temporal resolutions. Depth resolved BIOMAS outputs (sea temperature, diatoms, flagellates, copepods and predator zooplankton) were summarized by calculating the mean value within the upper 100 m of the water column (summarized in Figure S1.3). A depth resolved representation of environmental variables is appropriate because bowhead whales exploit resources throughout the water column and are known to feed

on prey located at various depths (Heide-Jørgensen, Laidre, Nielsen, Hansen, & Røstad, 2013), with feeding at the ocean floor being inferred from visual observations of mud on the heads of surfacing whales (Moore, George, Sheffield, Bacon, & Ashjian, 2010).

Our study used survey data from transects, crosslegs, circling and search legs to and from primary areas that were surveyed. To produce the bowhead whale presence dataset for this study, we assigned a 1 to each date (8-day period) and location (grid cell) at which one or more bowhead whales were sighted. We compiled an observed absence dataset, representing where whales were not sighted, by assigning a 0 to every date and location surveyed, after removing bowhead whale presence records.

Using the temporally binned and spatially gridded presence-absence data, we estimated survey effort and bowhead whale sighting rate. We estimated effort by summing the number of visits to each cell by the survey aircraft. We estimated sighting rate by dividing the number of presences at each grid cell by the number of visits to that grid cell. After temporal binning, there was a maximum of one visit and one bowhead whale presence for each grid cell within each 8-day period and year. Thus, the maximum number of presences or visits (i.e. effort) attributed to each grid cell over the whole study period was 225 (nine 8-day periods per year multiplied by 25 years). This representation of survey effort and sighting rate is consistent with the data used in our species distribution models. After temporally and spatially binning survey data as described above, there was an average of 7,753 (std = 1934) grid cells surveyed each year, and an average of 135 (std = 89) grid cells at which a bowhead whale presence was recorded each year (Table S1.1).

Highly correlated variables may hinder interpretation of variable importance. To address this issue, we calculated Pearson's correlation coefficient between each pair of variables on a cell-by-cell basis. Due to the large number of grid cells in our temporal and geographic domain (>4,000,000), we used a randomly selected subset of 100,000 grid cells for our correlation analysis. This sample size provided stable correlation coefficients, as determined by performing correlations with smaller and larger numbers of grid cells. Following the guidance of Dormann et al. (2013), we removed or modified variables with a correlation coefficient of $|r| \geq .70$.

2.5 | Species distribution models

Species distribution models use species occurrence (response) and environmental data (predictors) to characterize the relationship between a species and its environment. Once derived, the fitted relationship is typically applied to environmental predictor data from a more expansive region than the species occurrence data alone, producing a continuous map of habitat suitability over some region of interest. Values in the final prediction map are usually scaled from 0 to 1 to represent habitat suitability ranging from poor to excellent, and to facilitate comparisons between prediction maps. There are many algorithms that may be used to derive the fitted relationship between species and environment (Elith et al., 2006). We used two

SDM algorithms: Maxent and boosted regression trees. We selected these two modelling algorithms because multi-model comparisons have highlighted each as having high accuracy (e.g. Elith et al., 2006). Each algorithm is commonly used in applied SDM studies, yet each operates on different principles. Both Maxent and BRT algorithms are capable of fitting complex nonlinear relationships. Maxent (Phillips, Anderson, Dudík, Schapire, & Blair, 2017; Phillips, Anderson, & Schapire, 2006) operates by selecting a distribution that maximizes entropy subject to a set of constraints drawn from environmental data across the study area or at the location of absences. See Elith et al., (2011) for a detailed yet accessible explanation of this algorithm. We used MAXENT (v 3.4.0) with default settings. Boosted regression trees combine regression trees with boosting, a technique used to combine individual tree models in an adaptive fashion (Elith, Leathwick, & Hastie, 2008). We implemented BRTs with the "gbm" package in R (v 3.5.2, R Development Core Team, 2018) with code adapted from Elith et al., (2008). In fitting BRTs, there are three important settings: learning rate, tree complexity and number of trees. To determine the optimal settings for our dataset, we ran a 10-fold cross-validation procedure and selected settings that minimized predictive deviance. Based upon this analysis we selected a learning rate of 0.01 and tree complexity of 9 (Table S1.2). The optimal number of trees determined by cross-validation was 2,200; however, we found little difference in deviance scores and habitat maps produced with 2,200 versus 1,000 trees. Therefore, to reduce computation time we used 1,000 trees, suggested as minimum by Elith et al. (2008). The bag fraction was set to 0.5. We fitted all SDMs using environmental predictor variable information from the same 8-day period and year of each presence-absence record.

2.6 | Predictive performance tests

We evaluated Maxent and BRT models built with (a) BIOMAS variables and bathymetry and (b) only BIOMAS variables (no bathymetry). For each set of variables, we withheld 1 to 24 of the 25 available years for model evaluation. Test years withheld for model evaluation were selected randomly without replacement over all cross-validation folds. For example, models that were tested with two years were trained with 23 years, and there were 12 cross-validation folds. Test years for the first cross-validation fold were selected randomly from 1988–2012. Test years for the second cross-validation fold were selected randomly from 1988–2012, excluding the two test years from the first fold. This selection procedure was repeated to produce additional cross-validation folds until there were no more sets of two unused test years. Training data for each cross-validation fold consisted of the non-intersecting set of 1988–2012 with the randomly selected test years (Table S1.3 and Appendix S1). This type of cross-validation uses the heterogeneity within the dataset as a surrogate for heterogeneity among datasets (Wenger & Olden, 2012), and it is suited to testing model performance in an unsampled time period (e.g. the past or the future) (Vaughan & Ormerod, 2005). Absence data used to train each SDM was a random selection of 10,000 grid

cells drawn from the set of grid cells surveyed during the training years at which no bowhead whale was recorded. After discarding absences that did not overlap the spatial domain of predictor variables, we had approximately 9,870 absences to use in each SDM (Table S1.3). Identical presence and absence datasets were used in corresponding Maxent and BRT SDMs.

We used each fitted SDM to make predictions for all nine 8-day periods (21 August–31 October) for each test year. We calculated one area under the receiver operator characteristic (ROC) curve (AUC) (Fielding & Bell, 1997; Hanley & McNeil, 1982) for each set of test years and model (Maxent and BRT) by concatenating predictions from all 8-day periods in the test years and evaluating them as if they were one prediction (Appendix S1). We also computed another model performance metric, the true skill statistic (TSS) (Allouche, Tsoar, & Kadmon, 2006). It is necessary to select a threshold to compute TSS, and we used the sensitivity-specificity sum maximization approach (Liu, Berry, Dawson, & Pearson, 2005) to select thresholds. We assessed the relative contribution of each variable using output from MAXENT software and the R "gbm" package.

We also fit models using an additional 15 variable combinations. These model configurations consisted of 14 additional possible combinations of BIOMAS variables and the case of using only bathymetry. We fit and evaluated these models using the procedures described above for all other models.

2.7 | Habitat maps

To synthesize and present a concise representation of predicted habitat suitability, we generated maps of mean predicted habitat suitability in our study area for each 8-day period. We created the maps using results from the set of models having 5 test years, 20 training years and 5 cross-validation folds. From these models we generated one prediction map for each of the nine 8-day periods, each of the 25 years, and both Maxent and BRT models, resulting in 50 predictions of habitat suitability for each 8-day period. To generate the mean predicted habitat suitability for each 8-day period, we calculated the mean predicted habitat suitability value for corresponding grid cells from all 50 predictions.

3 | RESULTS

3.1 | Data filtering

After subsetting the full bowhead whale dataset with our data and locations filters, there were 3,378 bowhead whale presence records and 527,021 absence records. Due to high correlation between copepod and predator zooplankton, we combined those variables by computing their average to represent bowhead whale prey (hereafter zooplankton). Diatoms were highly correlated with both copepods and zooplankton and were therefore excluded from our analysis (Table S1.4). The final BIOMAS variables used in our study were

zooplankton, sea temperature, sea ice and flagellates. These four variables BIOMAS variables were used for all experiments presented in our study, except where we fit models using the 15 combinations of BIOMAS variables and bathymetry.

3.2 | SDM performance

Using 1–17 test years, mean performance scores for the models that included bathymetry and BIOMAS variables remained above 0.8 and 0.5 for AUC and TSS, respectively (Figure 2a, Table S1.3). Overall, mean scores for the models built with only BIOMAS variables were lower than for models that included bathymetry and BIOMAS variables (Figure 2b, Table S1.3). A notable decline in AUC and TSS scores for both model configurations (with and without bathymetry) occurred when the number of test years was greater than 21.

Since model performance scores were relatively insensitive to the number of test years less than or equal to 17, we elected to fit models using 5 test years and 20 training years to compare performance of models with different variable combinations. Using 5 test years allowed for prediction onto each of the 25 test years, allowed for 5 cross-validation folds, and allowed for an 80/20 training/testing data split (Table S1.5). Comparison of performance scores from models built with various variable combinations revealed a clear pattern (Figure 3, Table S1.6). Models that included bathymetry scored

the highest (mean AUC = 0.810, mean TSS = 0.512). Models that included zooplankton but not bathymetry scored second highest (mean AUC = 0.764, mean TSS = 0.437). Models that included temperature and excluded bathymetry and zooplankton scored third highest (mean AUC = 0.714, mean TSS = 0.366). Models that included flagellates and excluded bathymetry, zooplankton and temperature scored fourth highest (mean AUC = 0.670, mean TSS = 0.264), and the sea ice only models had the lowest scores (mean AUC = 0.632, mean TSS = 0.256).

3.3 | Environmental drivers

Overall, bathymetry was the most important variable in models that used that variable, while zooplankton was the most important variable in models that did not include bathymetry (Figure 4). Variable importance measures for the Maxent SDMs that included bathymetry and BIOMAS variables suggested that bathymetry was the most important predictor variable (67.44%) followed by sea ice (19.64%), zooplankton (4.91%), flagellates (4.86%) and sea temperature (3.15%). Importance of zooplankton, flagellates and sea temperature were all within a standard deviation of one another. For BRT SDMs, bathymetry was again the most important variable (38.08%), followed by sea ice (17.22%), sea temperature (16.13%), zooplankton (15.12%) and flagellates (13.45%). With the exception of sea ice

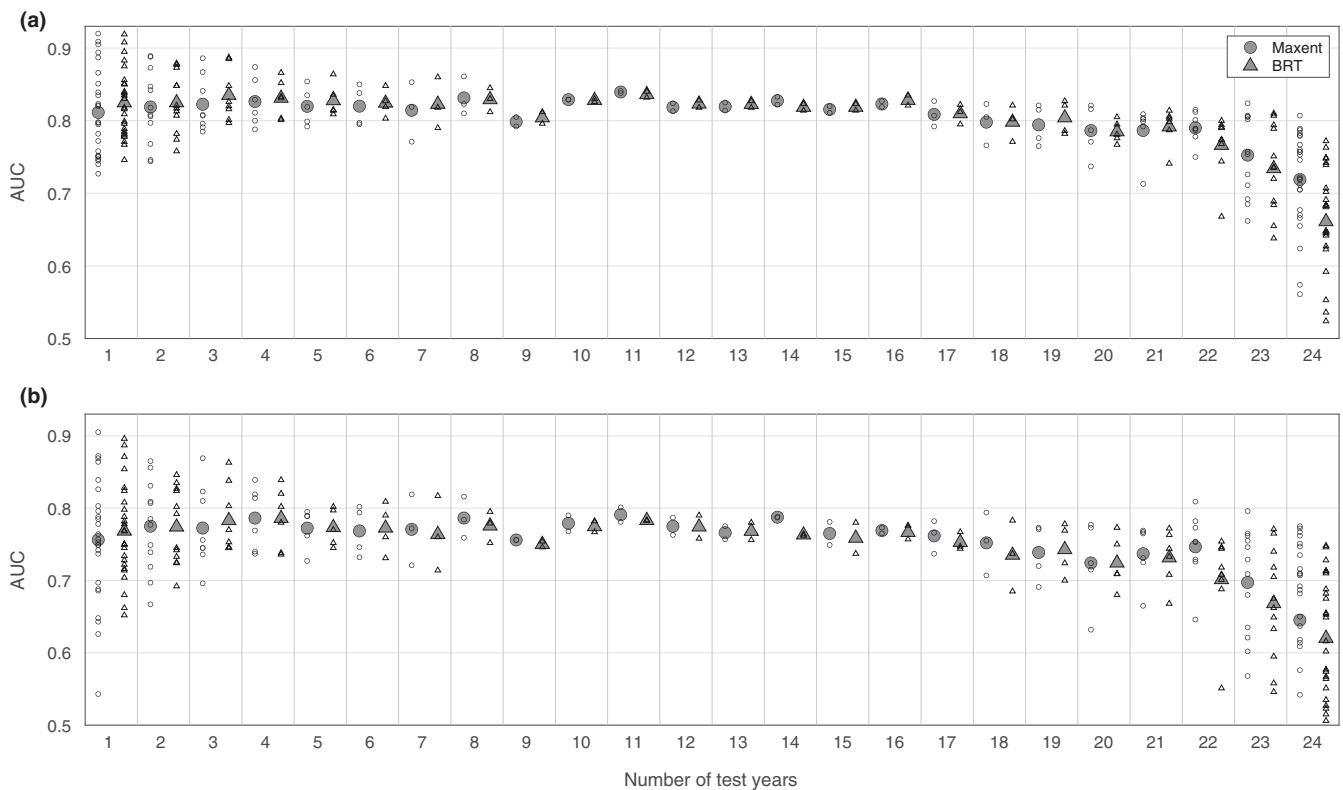


FIGURE 2 Area under the curve (AUC) scores from Maxent (circle) and BRT (triangle) models as the number of independent test years varies from 1 to 24. Mean AUC score across all cross-validation folds is shown by the large marker. AUC scores for each fold are shown above and below the mean. The test years were mutually exclusive between folds for each test year bin. Model built with bathymetry and BIOMAS variables(a). Model built with only BIOMAS variables (b)

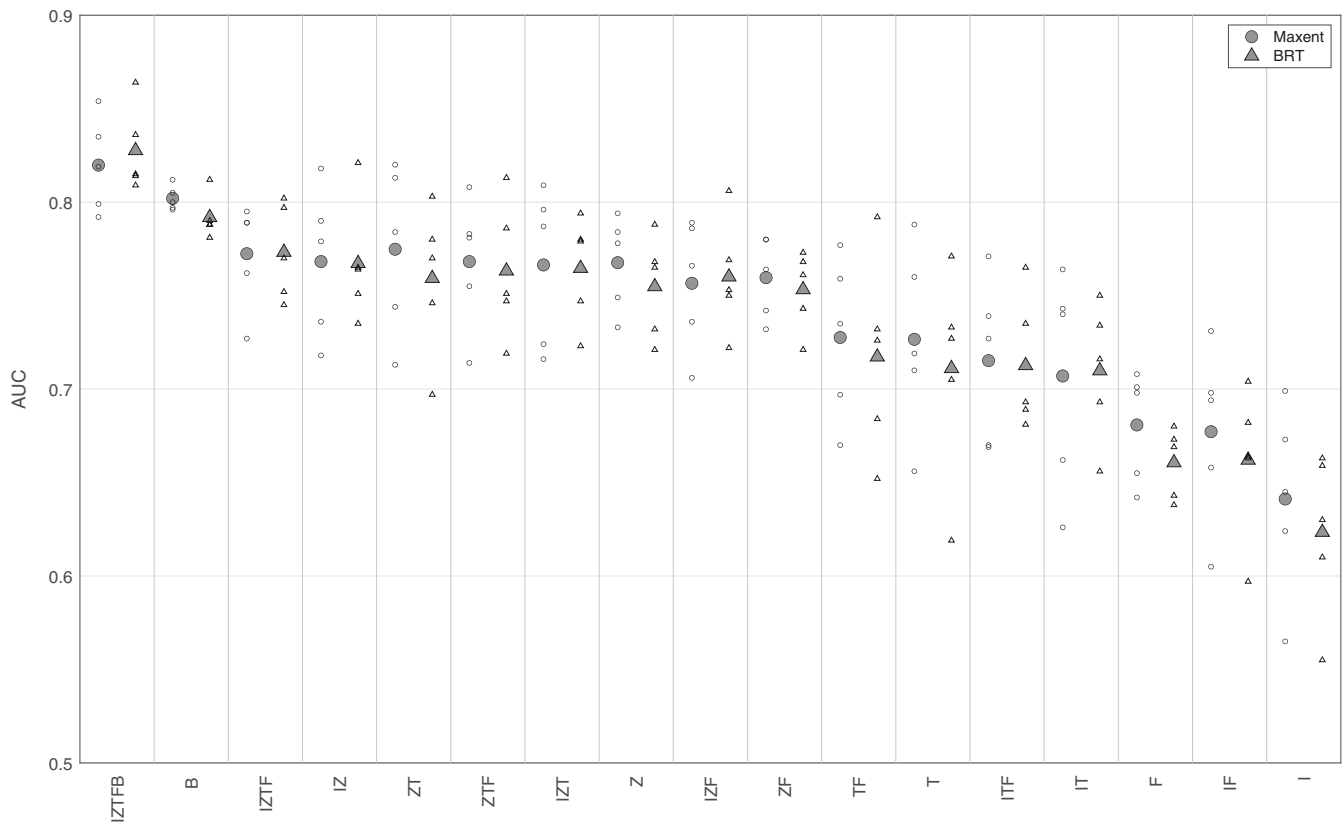


FIGURE 3 AUC scores from Maxent (circle) and BRT (triangle) models fit with combinations of variables. All models used 5 test years and 20 training years. Mean AUC score across all folds is shown by the large marker. AUC scores for each fold are shown above and below the mean. B, bathymetry; F, flagellates; I, sea ice; T, temperature; Z, zooplankton

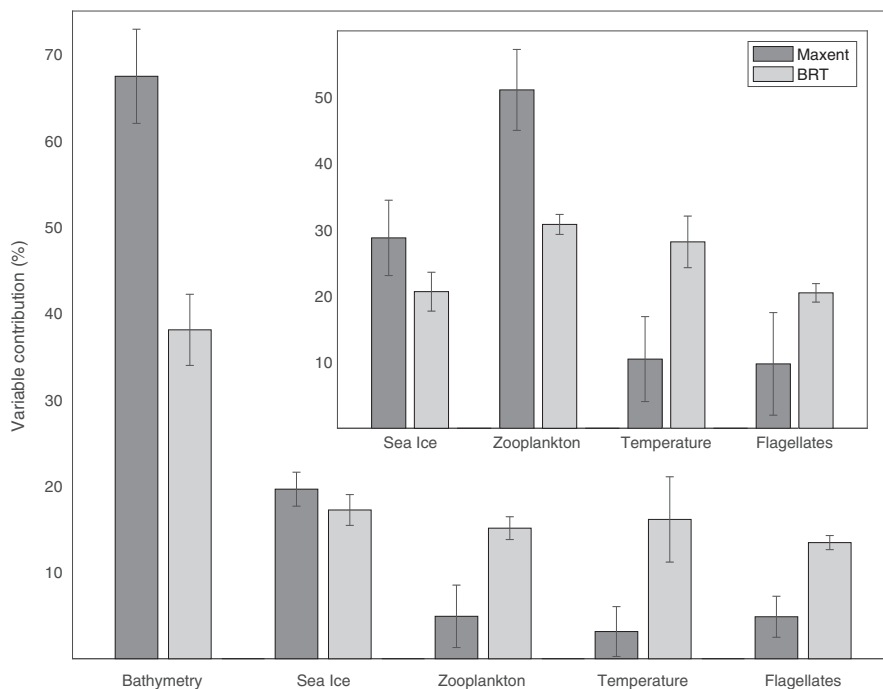


FIGURE 4 Mean variable importance from Maxent and BRT models using bathymetry and BIOMAS variables, and using only BIOMAS variables (inset). Whiskers show standard deviation. Results shown for models that used 5 test years and 20 training years

and flagellates, the importance of all BIOMAS variables was within a standard deviation of one another.

For models that included only BIOMAS variables (no bathymetry) (Figure 4 inset), zooplankton was the most important predictor for

both Maxent and BRT models. Variable importance for Maxent models was ordered as follows: zooplankton (51.10%), sea ice (28.75%), sea temperature (10.44%) and flagellates (9.72%). Sea temperature and flagellates were within a standard deviation of one another.

Variable importance for BRT models was ordered as follows: zooplankton (30.78%), sea temperature (28.14%), sea ice (20.63%) and flagellates (20.45%). Sea ice and flagellates were within a standard deviation of one another, as were zooplankton and sea temperature. Differences in the contribution of BIOMAS variables after bathymetry was removed from the models was greatest for zooplankton, whose contribution increased by 46.19% for Maxent models and 15.66% for BRT models.

Response curves for both Maxent and BRT SDMs (Figure 5) provide a heuristic measure of bowhead whale habitat preferences. For the models that included bathymetry and BIOMAS variables (Figure 5a), bowhead whale presence was associated with zooplankton most strongly between 0.5–1.0 mmol-N/m³. In contrast, for models that included only BIOMAS variables (Figure 5b), the response of zooplankton was more well defined, with the region between 0–0.5 mmol-N/m³ being associated with low predicted habitat suitability, and the region between 0.5–1.0 mmol-N/m³ associated with higher suitability. All other BIOMAS variable responses were similar to their counterparts in models with versus without bathymetry (Figure 5a vs. Figure 5b). The most highly preferred sea temperatures (averaged over 0–100 m depth) were between 4

and 12°C, with relatively high preference for temperatures in that range. There was a strong preference for low sea ice thickness. High predictions of bowhead whale habitat suitability were associated with low levels of flagellates, between 0 and 0.3 mmol-N/m³. The response to bathymetry was weak for bathymetry values less than approximately –500 m, and the response increased rapidly for bathymetry values of approximately –500 m to near 0 m.

3.4 | Habitat maps

We summarized habitat suitability predictions by averaging each 8-day prediction from Maxent and BRT models, and we used the models fit with 5 test years and 20 training years for this calculation. Results are shown for models that used BIOMAS variables and bathymetry (Figure 6a, Table S1.5) and only BIOMAS variables (Figure 6b). Predictions of mean habitat suitability for each 8-day period suggested that the most suitable areas were located over the continental shelf. The general range of highly suitable habitat concurred with the spatial pattern of sighting rates derived from the survey data (Figure 1c). Predictions for each 8-day period were

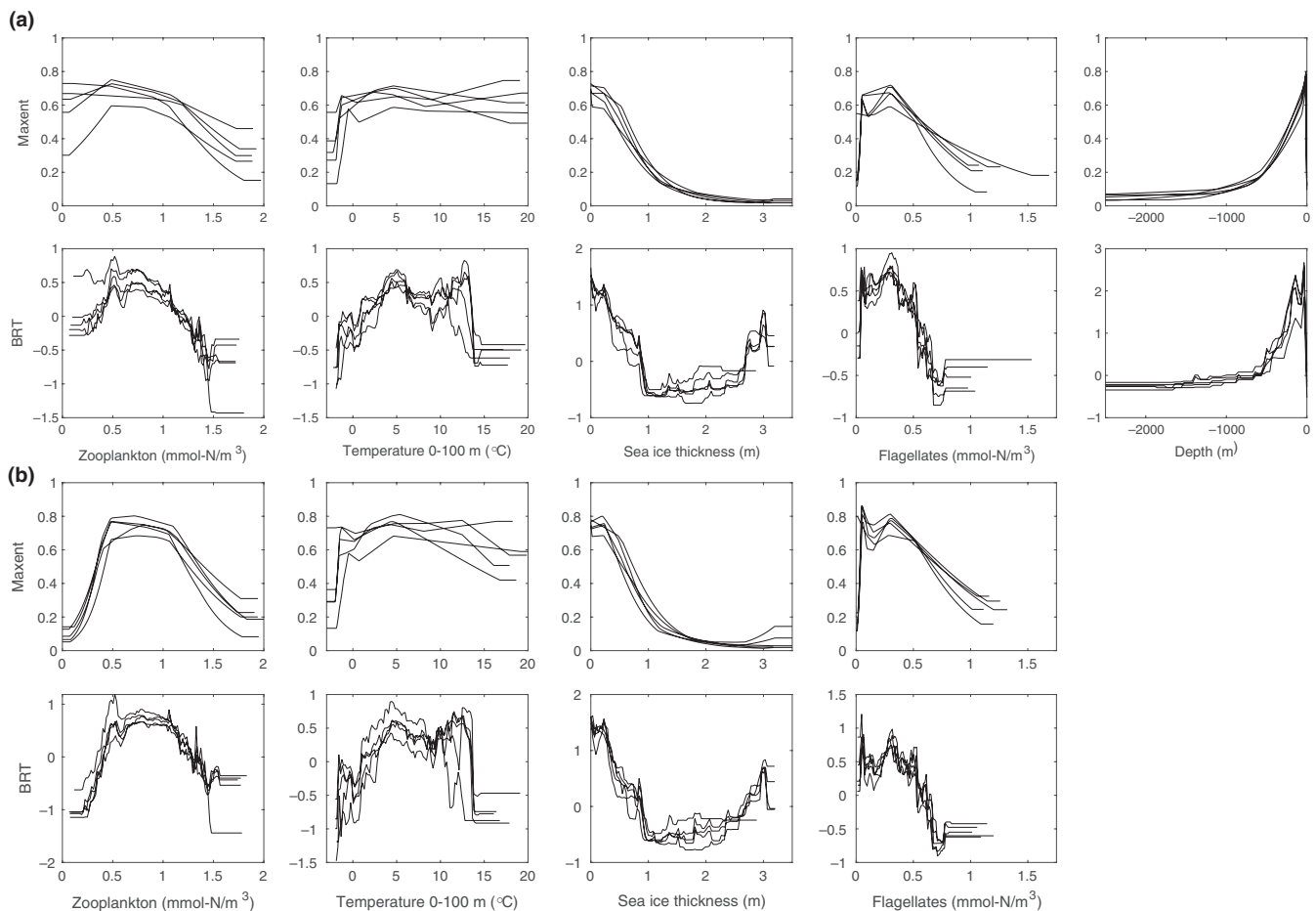


FIGURE 5 Response curves showing the marginal effect from Maxent (upper rows) and BRT (lower rows) SDMs fit with BIOMAS and bathymetry predictor variables (a) and for SDMs fit only with BIOMAS variables (b). Each panel shows one response curve for each cross-validation fold. Y-axis for Maxent SDMs shows predicted probability of presence. Y-axis for BRT SDMs shows the logit of the presence-absence response. Results shown for models that used 5 test years and 20 training years

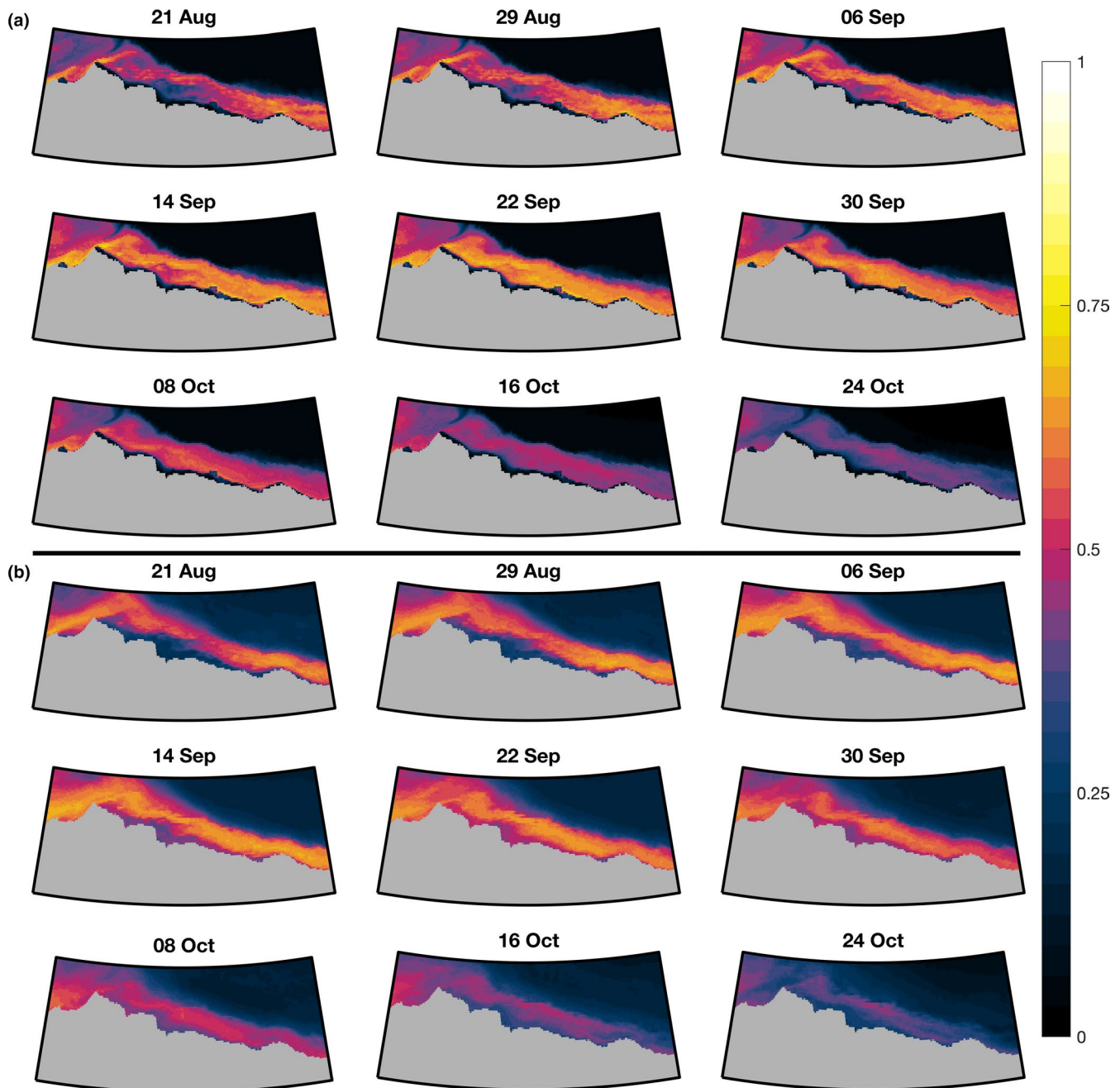


FIGURE 6 Mean of 8-day predictions of habitat suitability across years (1988–2012) from Maxent and BRT models using bathymetry and BIOMAS variables (a) and from the models using only BIOMAS variables (b). All models were trained and tested with the same data. Dates above plots are the first day of the 8-day prediction. Habitat suitability varies from low (0) to high (1) per the colour bar. Grey is the land of Alaska, USA. Boundaries for each subplot are 140–160°W and 69–72°N. Results shown for models that used 5 test years and 20 training years

similar between the models fit with BIOMAS variables and bathymetry (Figure 6a) and the models fit with only BIOMAS variables (Figure 6b). However, there are a few notable differences. Models with bathymetry predicted low habitat suitability in the narrow band of very shallow coastal water, which corresponds with lower sighting rates in this region during the surveys. Models that excluded bathymetry did not identify these nearshore waters as an area of low habitat suitability. Predicted habitat suitability from models with bathymetry resolved habitat suitability at a finer-grain, owing

to differences in bottom topography. In comparison, predicted habitat suitability from models without bathymetry showed a smoother prediction surface. Finally, predicted habitat suitability in the final three 8-day periods (8–31 October) was noticeably higher for models that included bathymetry than for models that did not include bathymetry.

Mean habitat suitability for each 8-day period (Figure 7a) predicted by the models with bathymetry and BIOMAS variables was unimodal with a maximum occurring during the week of 22–29

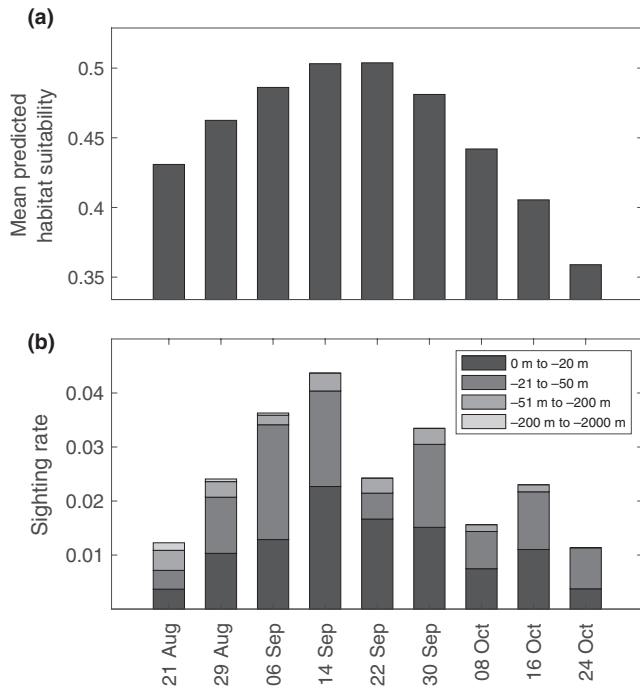


FIGURE 7 Mean 1988–2012 predicted habitat suitability between 0.10 and 1.0 from Maxent and BRT models using bathymetry and BIOMAS variables from models that used 5 test years and 20 training years. (a). Mean 1988–2012 sighting rates calculated from survey data used in SDMs (presence and absence binned by 8-day period) in depth bins from inshore to offshore (dark to light shading) (b). Dates on x-axis are the first day of the 8-day period

September, with the previous week 14–21 September being almost as high (difference of <1%). The predicted minimum habitat suitability occurred during the last 8-day period (24–31 October). Cells with habitat suitability values below 0.1 were excluded from these calculations to present a representative picture of within-year dynamics of habitat suitability relevant to bowhead whale presence.

Predictions of habitat suitability for corresponding grid cells, 8-day periods, and years, from Maxent and BRT were very similar (Figure S1.4). The overall Pearson correlation coefficient between predictions from Maxent and BRT models that included BIOMAS predictor variables and bathymetry was $r = .93$, and for models that included only BIOMAS predictor variables it was $r = .89$.

3.5 | Sighting rates

Calculation of presence–absence sighting rate from the raw survey data (Figure 7b), consistent with the presence–absence data used in SDMs, revealed waters received more survey effort than deeper waters to the northeast and a generally unimodal pattern with the maximum presence–absence sighting rate in 14–21 September, and the minimum in 24–31 October. Sighting rate for the first 8-day period was within 1% of the rate for 24–31 October.

4 | DISCUSSION

We found that including physical and biological modelled environmental data, especially modelled prey, is useful for building SDMs that accurately represent seasonal dynamics of species occurrence. Our best performing models included static and modelled dynamic environmental variables whose values were contemporaneous with species' presence–absence records. Our results suggest that static variables may carry additional weight for species whose distributions that have not changed with respect to those static features and that geographic stability in a species' distribution, as in the case of the western Arctic bowhead whale in the western Beaufort Sea, may allow for SDMs trained on a small percentage of the data to achieve a high degree of predictive performance.

4.1 | Bowhead whale distributions

Predictions of habitat suitability from the model built with BIOMAS variables and bathymetry (Figure 6a) and the model built with only BIOMAS variables (Figure 6b) were generally similar. Our predicted habitat maps show concurrence with areas known to be important for bowhead whales. For example, our predictions show the area around Point Barrow as highly suitable habitat. This area was identified as a late August–late October bowhead whale core-use area by Citta et al. (2015, figure 6a) from satellite tagged whales. Our results also show highly suitable habitat on the continental shelf that generally align in time and space with Biologically Important Areas identified by Clarke, Ferguson, Curtice, and Harrison (2015, figures 1b–d, 8.2). Importantly, predictions from our SDMs show poor habitat suitability in the northeast quadrant of our study area, much of which is known to be a spring (April–May) migratory corridor (Clarke et al., 2015, fig 8.3) and is not considered to be important during the fall migration period (late August–October) considered in this study. Due to the aerial survey objective of documenting the majority of bowhead whales during their fall migration, coastal waters received more survey effort than deeper waters to the northeast (Clarke, Christman, Brower, & Ferguson, 2013; Treacy, 2000) (Figures 1b–c; Figures S1.1). We do not know if lower levels of effort in the northeast quadrant contributed to it being predicted to have low habitat suitability.

4.2 | The importance of bathymetry in bowhead whale distribution models

Predictions from models built with BIOMAS variables and bathymetry, versus those built with only BIOMAS variables, are more realistic in the narrow band of very shallow inshore waters (Figure 6a vs. Figure 6b), where bowhead whales are not typically seen. Comparisons among variable combinations used to build models (Figure 3, Table S1.6) also emphasized the importance of the bathymetry variable. These comparisons and the variable importance plots

showed that the two best models were (1) BIOMAS variables and bathymetry, and (2) bathymetry alone. The importance of bathymetry highlights the spatial stability in bowhead whale distributions and suggests there is likely a strong relationship between bathymetry and biophysical processes that determine the distribution of bowhead whales. When bathymetry was removed from the models, the importance of zooplankton increased more than each of the other variables (+46.19% for Maxent and +15.66% for BRT SDMs). Modelled zooplankton declined seasonally by over 50% on average during our late summer/early fall within-year study period. Although bathymetry is a static variable whose correlation with zooplankton changes on a seasonal basis, it may be a reasonable proxy for prey in the time and area of our study. It is possible that the phenology of zooplankton prey has not changed substantively with respect to bowhead whale distributions in the western Beaufort Sea across the years and within the season of our study.

Bowhead whales have consistently occurred in inner shelf waters since long-term monitoring began in the 1980s. Clarke et al. (2018) compared bowhead whale sighting rates from 2009–2016 in the western Beaufort Sea with those from 1982–1991 (Moore, DeMaster, & Dayton, 2000). In both time periods, bowhead whale sighting rates during September and October were highest in the inner shelf habitat (0 to –50 m). We confirmed this result by comparing sighting rates from our presence–absence data in 1988–1995 versus 2005–2012, and we found that the highest sighting rates between 29 August and 31 October occurred over the inner shelf for both sets of years. Bowhead whales in our study area have been documented in deeper water over the outer shelf (–50 to –200 m), continental slope (–200 to –2,000 m) and basin (> –2,000 m) during July and August, largely before the within-year term of our study. It is likely that this spatial stability in bowhead whale distributions resulted in the small number of years needed to train successful models. Specifically, we examined how model performance scores, AUC and TSS, varied as the number of years used for model testing increased from 1 to 24 (Figure 2, Table S1.3). We found that models with as many as 17 test years, and only 8 training years, maintained scores in general agreement with models that were tested with 1 year and trained with 24 years (Figure 2).

4.3 | The importance of modelled prey data in bowhead whale distribution models

Bathymetry has been found to be an important variable in other marine species distribution modelling studies (Ainley et al., 2011; Burgos et al., 2020; Gomez et al., 2017; Sequeira et al., 2012). Although our model with bathymetry alone performed well, it is not dynamic and produces the same prediction for every 8-day period (Figure S1.5). Consequently, it misses important temporal variability in the distribution of bowhead whales. Specifically, sighting rates are lower at the beginning and the end (Figure 7b) of the within-year study period (21 August–31 October), likely reflecting the seasonal dynamics of bowhead whale migration through the region. The

models fit with only BIOMAS variables, and those with BIOMAS variables and bathymetry correctly predicted lower habitat suitability at the beginning and end of the study period (Figure 6). There were only small differences in the predictive capacity between the models that were trained with only bathymetry versus those trained with BIOMAS variables and bathymetry (Figure 3, Table S1.6). This was surprising because models trained with only bathymetry could not reproduce the within-season temporal dynamics seen in models that included BIOMAS predictor variables (Figures 6 and 7, S1.5). This may be an artefact of the temporal scale at which we measure AUC and TSS. Because AUC and TSS were calculated using test year(s) across 8-day periods, rather than evaluating 8-day periods across test year(s), the model with bathymetry alone was not penalized for its inability to capture the within-year changes that are seen in the sighting rates (Figure 7b). However, the within-year temporal variability in bowhead distribution shows that it is critical to include dynamic variables in the model (compare Figure 6 with Figure S1.5 and Figure 7). Scores from all models that included zooplankton were second in rank to those that included bathymetry (Figure 3, Table S1.6). The stability of scores across the combinations of variables that included zooplankton provides strong support for the positive utility of our particular zooplankton product (a combination of copepods and predator zooplankton) for predicting the distribution of zooplanktivores in the western Beaufort Sea.

The Canadian Beaufort Sea, east of our study area, and the Point Barrow, Alaska area at the western end of our study area are important feeding areas. The coastal shelf waters between these habitats are often thought of as a migratory corridor, yet feeding is known to take place in these waters (Clarke et al., 2015, 2018; Lowry et al., 2004). Western Arctic bowhead whales feed on zooplankton, primarily euphausiids and copepods, and to a lesser extent amphipods (Lowry, 1993; Lowry et al., 2004; Moore et al., 2010). Energetic calculations (Walkusz et al., 2012), data from time–depth recorders (Laidre, Heide-Jørgensen, & Nielsen, 2007), and opportunistic observations (Moore et al., 2000) suggest that bowhead whales need to locate and feed on dense patches of prey in order to meet their caloric demands. We did not analyse behavioural data that could have identified feeding activity for individual bowhead whale presence records. Therefore, we are making the implicit assumption that feeding takes place and is important in the area and time of year of our study. This assumption has support, as feeding is known to take place in our study area (Clarke et al., 2015; Lowry et al., 2004). The predictive strength of the modelled zooplankton variable relative to other BIOMAS variables supports the findings of Pendleton et al., (2009) and Pershing et al., (2009). They showed that regional-scale mean copepod concentrations, in situ and modelled, were good predictors of the relative abundance of North Atlantic right whales (*Eubalaena glacialis*), a species with similar feeding ecology to bowhead whales.

The comparatively weak performance of models that did not use bathymetry or zooplankton, and used temperature, flagellates and/or sea ice was not surprising due to their limited direct impact on the ecology of bowhead whales. Specifically, the comparatively low

variable importance for flagellates, representing the phytoplankton in our study, makes sense because bowhead whales do not target phytoplankton. Additionally, flagellate concentrations decline when zooplankton concentrations increase, as phytoplankton are consumed by zooplankton. This effect can be seen in the complementary response curves of flagellates versus zooplankton (Figure 5). Our results suggested moderate importance of the sea ice variable, and the response curve suggests that bowhead whales are more likely to be found where sea ice thickness is low. It is important to note that BIOMAS models the thickness of ice, rather than the ice extent which is more commonly reported by remote observing platforms, such as satellites.

4.4 | Future research and conclusions

It is important to acknowledge that the generalized migration of western Arctic bowhead whales includes the Bering, Beaufort and Chukchi seas (Moore & Reeves, 1993). We intentionally limited our analysis to the areas in the western Beaufort Sea because it had the longest and most consistent data collection (Clarke et al., 2013). As such, our results may not apply in other portions of the western Arctic bowhead whales' range. For example, the bathymetry variable may be less meaningful in habitats where differences in bottom topography are less pronounced, such as the Chukchi Sea. Likewise, modelled zooplankton may have limited utility in areas where feeding is less important. Current and future research on model transferability (Torres et al., 2015; Yates et al., 2018) and extrapolation (Mannocci, Roberts, Miller, & Halpin, 2017) will be critical to understanding how confidently SDMs can be used to make future predictions. Further study is required to understand how the models and variables used in this study would perform in other bowhead whale habitats and at other times of year.

The research presented here demonstrates the applicability of presence-absence datasets (e.g. observer bias not estimable, multiple or under-documented survey protocols) to generate predictions of habitat suitability using modelled environmental conditions. Due to the high cost of implementing advanced survey protocols, there is a growing need to understand the nuances of Maxent, BRT, and other presence-absence modelling algorithms when used to make out-of-sample hindcasted or forecasted predictions. Many monitoring programs and opportunistic platforms collect data in a fashion that is not suitable for state-of-the-art density modelling (e.g. Roberts et al., 2016). Therefore, research about our ability to incorporate these data into non-extrapolative SDMs is needed to lay the groundwork for future studies on how extrapolation, transferability and forecasting can be done with confidence (Derville, Torres, Iovan, & Garrigue, 2018; Fiedler et al., 2018; Mannocci et al., 2017) using presence-absence data.

Our models show the possibility of improving SDM performance using modelled dynamic predictor variables that estimate expected prey availability. Using more direct estimates of prey availability, rather than oceanographic proxies for prey, could be important

for forecasting species distributions because the relationships between the proxies and prey may change in a changing climate (GREGG et al., 2013; Silber et al., 2017). Our results complement the understanding of bowhead whale ecology and show how presence-absence data can be used in combination with both static and spatio-temporally varying modelled environmental data to predict bowhead whale distributions.

The BIOMAS predictor variables were derived from a dynamic data assimilative three-dimensional ocean model with the demonstrated ability to project future ocean conditions (Banas et al., 2016), and a natural next step for study of this system is to use the BIOMAS variables to explore the potential effects of Arctic climate change on bowhead whale habitat use. Changes in the ocean climate are altering species distributions (Nye, Link, Hare, & Overholtz, 2009; Ramp, Delarue, Palsbøll, Sears, & Hammond, 2015) as well as human commercial activities (Young et al., 2019). Effectively monitoring migratory whale populations is challenging, particularly in regions that are undergoing environmental change, such as the western Beaufort Sea. An important focus of future research efforts would be to identify and measure emerging conflicts between Arctic wildlife and human activities (Reeves et al., 2014; Reeves, Rosa, George, Sheffield, & Moore, 2012) as the Arctic Ocean changes.

ACKNOWLEDGEMENTS

This research was funded by NASA Biodiversity and Ecological Forecasting Program grant NNX11AO93G and Cryosphere Program grant NNX15AG68G. We thank Eduardo Carrillo-Rubio, Janet Clarke and Megan Ferguson for critical reviews and thoughtful comments on this manuscript. We thank the many marine mammal observers and flight crew members, without whom this work would not have been possible. Finally, we thank the U.S. Bureau of Ocean Energy Management, Alaska OCS Region, for funding the aerial surveys and NOAA Marine Mammal Laboratory for the collection, curation and dissemination of marine mammal survey data used in this study.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/ddi.13149>.

DATA AVAILABILITY STATEMENT

Marine mammal survey data used in this study were downloaded from <https://www.afsc.noaa.gov/nmml/cetacean/bwasp/>.

ORCID

Daniel E. Pendleton  <https://orcid.org/0000-0001-9341-8841>
Elizabeth E. Holmes  <https://orcid.org/0000-0001-9128-8393>
Jessica Redfern  <https://orcid.org/0000-0003-4042-3668>

REFERENCES

Abrahms, B., Welch, H., Brodie, S., Jacox, M. G., Becker, E. A., Bograd, S. J., ... Hazen, E. L. (2019). Dynamic ensemble models to predict distributions and anthropogenic risk exposure for highly mobile species.

- Diversity and Distributions*, 25, 1182–1193. <https://doi.org/10.1111/ddi.12940>
- Ainley, D. G., Jongsomjit, D., Ballard, G., Thiele, D., Fraser, W. R., & Tynan, C. T. (2011). Modeling the relationship of Antarctic minke whales to major ocean boundaries. *Polar Biology*, 35, 281–290. <https://doi.org/10.1007/s00300-011-1075-1>
- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS): Assessing the accuracy of distribution models. *Journal of Applied Ecology*, 43(6), 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Banas, N. S., Zhang, J., Campbell, R. G., Sambrotto, R. N., Lomas, M. W., Sherr, E., ... Lessard, E. J. (2016). Spring plankton dynamics in the Eastern Bering Sea, 1971–2050: Mechanisms of interannual variability diagnosed with a numerical model: Bering Sea Spring Plankton Dynamics. *Journal of Geophysical Research: Oceans*, 121(2), 1476–1501. <https://doi.org/10.1002/2015JC011449>
- Becker, E., Forney, K., Fiedler, P., Barlow, J., Chivers, S., Edwards, C., ... Redfern, J. (2016). Moving Towards Dynamic Ocean Management: How well do modeled ocean products predict species distributions? *Remote Sensing*, 8, 149. <https://doi.org/10.3390/rs8020149>
- Becker, E. A., Forney, K. A., Redfern, J. V., Barlow, J., Jacox, M. G., Roberts, J. J., & Palacios, D. M. (2018). Predicting cetacean abundance and distribution in a changing climate. *Diversity and Distributions*, 25(4), 626–643. <https://doi.org/10.1111/ddi.12867>
- Brower, A. A., Ferguson, M. C., Schonberg, S. V., Jewett, S. C., & Clarke, J. T. (2017). Gray whale distribution relative to benthic invertebrate biomass and abundance: Northeastern Chukchi Sea 2009–2012. *Deep Sea Research Part II: Topical Studies in Oceanography*, 144, 156–174. <https://doi.org/10.1016/j.dsr2.2016.12.007>
- Burgos, J. M., Buhl-Mortensen, L., Buhl-Mortensen, P., Ólafsdóttir, S. H., Steingrund, P., Ragnarsson, S. Á., & Skagseth, Ø. (2020). Predicting the distribution of indicator taxa of vulnerable marine ecosystems in the Arctic and Sub-arctic Waters of the Nordic Seas. *Frontiers in Marine Science*, 7, 131. <https://doi.org/10.3389/fmars.2020.00131>
- Carroll, G., Cox, M., Harcourt, R., Pitcher, B. J., Slip, D., & Jonsen, I. (2017). Hierarchical influences of prey distribution on patterns of prey capture by a marine predator. *Functional Ecology*, 31, 1750–1760. <https://doi.org/10.1111/1365-2435.12873>
- Carroll, G., Holsman, K. K., Brodie, S., Thorson, J. T., Hazen, E. L., Bograd, S. J., ... Selden, R. L. (2019). A review of methods for quantifying spatial predator–prey overlap. *Global Ecology and Biogeography*, 28, 1561–1577. <https://doi.org/10.1111/geb.12984>
- Casey, K. S., Brandon, T. B., Cornillon, P., & Evans, R. (2010). The past, present, and future of the AVHRR Pathfinder SST Program. In V. Barale, J. F. R. Gower, & L. Alberotanza (Eds.), *Oceanography from space* (pp. 273–287). Dordrecht, The Netherlands: Springer.
- Citta, J. J., Quakenbush, L. T., Okkonen, S. R., Druckenmiller, M. L., Maslowski, W., Clement-Kinney, J., ... Heide-Jørgensen, M. P. (2015). Ecological characteristics of core-use areas used by Bering–Chukchi–Beaufort (BCB) bowhead whales, 2006–2012. *Progress in Oceanography*, 136, 201–222. <https://doi.org/10.1016/j.pocan.2014.08.012>
- Clarke, J. T., Brower, A. A., Ferguson, M. C., & Willoughby, A. L. (2017). *Distribution and Relative Abundance of Marine Mammals in the Eastern Chukchi and Western Beaufort Seas*, 2016. (Annual Report, OCS Study BOEM 2017–078.). Marine Mammal Laboratory, Alaska Fisheries Science Center, NMFS, NOAA, 7600 Sand Point Way NE, F/AKC3, Seattle, WA 98115–6349.
- Clarke, J. T., Christman, C. L., Brower, A. A., & Ferguson, M. C. (2012). *Distribution and Relative Abundance of Marine Mammals in the Alaskan Chukchi and Beaufort Seas*, 2011. (Annual Report, OCS Study BOEM 2012–009.). National Marine Mammal Laboratory, Alaska Fisheries Science Center, NMFS, NOAA, 7600 Sand Point Way NE, F/AKC3, Seattle, WA 98115–6349.
- Clarke, J. T., Christman, C. L., Brower, A. A., & Ferguson, M. C. (2013). *Distribution and Relative Abundance of Marine Mammals in the Northeastern Chukchi and Western Beaufort Seas*, 2012. (Annual Report, OCS Study BOEM 2013–00117.). National Marine Mammal Laboratory, Alaska Fisheries Science Center, NMFS, NOAA, 7600 Sand Point Way NE, F/AKC3, Seattle, WA 98115–6349.
- Clarke, J. T., Ferguson, M. C., Curtice, C., & Harrison, J. (2015). 8. Biologically important areas for Cetaceans within U.S. Waters – Arctic Region. *Aquatic Mammals*, 41(1), 94–103. <https://doi.org/10.1578/AM.41.1.2015.94>
- Clarke, J. T., Ferguson, M. C., Willoughby, A. L., & Brower, A. A. (2018). Bowhead and Beluga Whale Distributions, Sighting Rates, and Habitat Associations in the Western Beaufort Sea in Summer and Fall 2009–16, with Comparison to 1982–91. *Arctic*, 71(2), 115. <https://doi.org/10.14430/arctic4713>
- Derville, S., Torres, L. G., Iovan, C., & Garrigue, C. (2018). Finding the right fit: Comparative cetacean distribution models using multiple data sources and statistical approaches. *Diversity and Distributions*, 24, 1657–1673. <https://doi.org/10.1111/ddi.12782>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., ... E. Zimmermann, N. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129–151. <https://doi.org/10.1111/j.2006.0906-7590.04596.x>
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802–813. <https://doi.org/10.1111/j.1365-2656.2008.01390.x>
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1), 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>
- Fiedler, P. C., Redfern, J. V., Forney, K. A., Palacios, D. M., Sheredy, C., Rasmussen, K., ... Ballance, L. T. (2018). Prediction of Large Whale Distributions: A comparison of presence-absence and presence-only modeling techniques. *Frontiers in Marine Science*, 5, 419. <https://doi.org/10.3389/fmars.2018.00419>
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(01), 38–49. <https://doi.org/10.1017/S0376892997000088>
- Gomez, C., Lawson, J., Kouwenberg, A. L., Moors-Murphy, H., Buren, A., Fuentes-Yaco, C., ... Wimmer, T. (2017). Predicted distribution of whales at risk: Identifying priority areas to enhance cetacean monitoring in the Northwest Atlantic Ocean. *Endangered Species Research*, 32, 437–458. <https://doi.org/10.3354/esr00823>
- Gregg, E., Baumgartner, M., Laidre, K., & Palacios, D. (2013). Marine mammal habitat models come of age: The emergence of ecological and management relevance. *Endangered Species Research*, 22(3), 205–212. <https://doi.org/10.3354/esr00476>
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29–36. <https://doi.org/10.1148/radiology.143.1.7063747>
- Hazen, E. L., Palacios, D. M., Forney, K. A., Howell, E. A., Becker, E., Hoover, A. L., ... Bailey, H. (2017). WhaleWatch: A dynamic management tool for predicting blue whale density in the California Current. *Journal of Applied Ecology*, 54(5), 1415–1428. <https://doi.org/10.1111/1365-2664.12820>
- Heide-Jørgensen, M., Laidre, K. L., Nielsen, N. H., Hansen, R. G., & Røstad, A. (2013). Winter and spring diving behavior of bowhead

- whales relative to prey. *Animal Biotelemetry*, 1(1), 15. <https://doi.org/10.1186/2050-3385-1-15>
- Jakobsson, M., Mayer, L., Coakley, B., Dowdeswell, J. A., Forbes, S., Fridman, B., ... Weatherall, P. (2012). The International Bathymetric Chart of the Arctic Ocean (IBCAO) Version 3.0: IBCAO VERSION 3.0. *Geophysical Research Letters*, 39(12), L12609–L12609. <https://doi.org/10.1029/2012GL052219>
- Kishi, M. J., Kashiwai, M., Ware, D. M., Megrey, B. A., Eslinger, D. L., Werner, F. E., ... Zvalinsky, V. I. (2007). NEMURO—a lower trophic level model for the North Pacific marine ecosystem. *Ecological Modelling*, 202(1–2), 12–25. <https://doi.org/10.1016/j.ecolmodel.2006.08.021>
- Laidre, K. L., Heide-Jørgensen, M. P., & Nielsen, T. G. (2007). Role of the bowhead whale as a predator in West Greenland. *Marine Ecology Progress Series*, 346, 285–297. <https://doi.org/10.3354/meps06995>
- Lambert, C., Mannocci, L., Lehodey, P., & Ridoux, V. (2014). Predicting Cetacean Habitats from their energetic needs and the distribution of their prey in two contrasted Tropical Regions. *PLoS One*, 9, e105958. <https://doi.org/10.1371/journal.pone.0105958>
- Lezama-Ochoa, A., Boyra, G., Goñi, N., Arrizabalaga, H., & Bertrand, A. (2010). Investigating relationships between albacore tuna (*Thunnus alalunga*) CPUE and prey distribution in the Bay of Biscay. *Progress in Oceanography*, 86, 105–114. <https://doi.org/10.1016/j.pcean.2010.04.006>
- Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3), 385–393. <https://doi.org/10.1111/j.0906-7590.2005.03957.x>
- Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., Paver, C. R., Reagan, J. R., Johnson, D. R., Hamilton, M., & Seidov, D. (2013 & World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus & A. Mishonov, NOAA atlas NESDIS; 73. <http://www.nodc.noaa.gov/OC5/indprod.html>
- Lowry, L. F. (1993). Foods and feeding ecology. In J. J. Burns, J. J. Montague, & C. J. Cowles (Eds.), *The Bowhead Whale* (pp. 201–238). Lawrence, KS: Allen Press.
- Lowry, L. F., Sheffield, G., & George, J. C. (2004). Bowhead whale feeding in the Alaskan Beaufort Sea, based on stomach contents analyses. *Journal of Cetacean Research and Management*, 6(3), 215–223.
- Mannocci, L., Roberts, J. J., Miller, D. L., & Halpin, P. N. (2017). Extrapolating cetacean densities to quantitatively assess human impacts on populations in the high seas: Cetacean densities in the high seas. *Conservation Biology*, 31(3), 601–614. <https://doi.org/10.1111/cobi.12856>
- Mate, B. R., Krutzikowsky, G. K., & Winsor, M. H. (2000). Satellite-monitored movements of radio-tagged bowhead whales in the Beaufort and Chukchi seas during the late-summer feeding season and fall migration. *Canadian Journal of Zoology*, 78(7), 1168–1181. <https://doi.org/10.1139/z00-045>
- Mayo, C. A., & Marx, M. K. (1990). Surface foraging behaviour of the North Atlantic right whale, *Eubalaena glacialis*, and associated zooplankton characteristics. *Canadian Journal of Zoology*, 68, 2214–2220.
- Miller, M., Carlile, N., Scutt Phillips, J., McDuire, F., & Congdon, B. (2018). Importance of tropical tuna for seabird foraging over a marine productivity gradient. *Marine Ecology Progress Series*, 586, 233–249. <https://doi.org/10.3354/meps12376>
- Moore, S. E., DeMaster, D. P., & Dayton, P. K. (2000). Cetacean habitat selection in the Alaskan Arctic during summer and autumn. *Arctic*, 53(4), 432–447. <https://doi.org/10.14430/arctic873>
- Moore, S. E., George, J. C., Sheffield, G., Bacon, J., & Ashjian, C. J. (2010). Bowhead whale distribution and feeding near Barrow, Alaska, in late summer 2005–06. *Arctic*, 63(2), 195–205. <https://doi.org/10.14430/arctic974>
- Moore, S. E., & Reeves, R. R. (1993). Distribution and movement. In J. J. Burns, J. J. Montague, & C. J. Cowles (Eds.), *The Bowhead Whale* (pp. 313–386). Lawrence, KS: Allen Press.
- Nye, J., Link, J., Hare, J., & Overholtz, W. (2009). Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. *Marine Ecology Progress Series*, 393, 111–129. <https://doi.org/10.3354/meps08220>
- Pendleton, D., Pershing, A., Brown, M., Mayo, C., Kenney, R., Record, N., & Cole, T. (2009). Regional-scale mean copepod concentration indicates relative abundance of North Atlantic right whales. *Marine Ecology Progress Series*, 378, 211–225. <https://doi.org/10.3354/meps07832>
- Pendleton, D. E., Sullivan, P. J., Brown, M. W., Cole, T., Good, C. P., Mayo, C. A., ... Pershing, A. J. (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(2), 147–161. <https://doi.org/10.3354/esr00433>
- Pershing, A., Record, N., Monger, B., Mayo, C., Brown, M., Cole, T., ... Woodard, L. (2009). Model-based estimates of right whale habitat use in the Gulf of Maine. *Marine Ecology Progress Series*, 378, 245–257. <https://doi.org/10.3354/meps07829>
- Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: An open-source release of Maxent. *Ecography*, 40(7), 887–893. <https://doi.org/10.1111/ecog.03049>
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3–4), 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>
- Quakenbush, L., Citta, J., George, J. C., Heide-Jørgensen, M. P., Small, R., Brower, H., ... Pokiak, J. (2012). Seasonal movements of the Bering-Chukchi-Beaufort stock of bowhead whales: 2006–2011. *Satellite Telemetry Results*. Int Whal Comm Sci Report SC/64/BRG1.
- Quakenbush, L. T., Small, R. J., & Citta, J. J. (2013). *Satellite Tracking of Bowhead Whales: Movements and analysis from 2006 to 2012*. Final Report. U.S. Dept. of the Interior, Bureau of Ocean Energy Management, Alaska Outer Continental Shelf Region, Anchorage, AK. OCS Study BOEM 2013–01110, 123.
- R Development Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Ramp, C., Delarue, J., Palsbøll, P. J., Sears, R., & Hammond, P. S. (2015). Adapting to a Warmer Ocean—Seasonal Shift of Baleen Whale movements over three decades. *PLoS One*, 10(3), e0121374. <https://doi.org/10.1371/journal.pone.0121374>
- Redfern, J. V., Becker, E. A., & Moore, T. J. (2020). Effects of variability in ship traffic and whale distributions on the risk of ships striking whales. *Frontiers in Marine Science*, 6, 793. <https://doi.org/10.3389/fmars.2019.00793>
- Redfern, J. V., Ferguson, M. C., Becker, E. A., Hyrenbach, K. D., Good, C., Barlow, J., ... Werne, F. (2006). Techniques for cetacean-habitat modeling. *Marine Ecology Progress Series*, 310, 271–295. <https://doi.org/10.3354/meps310271>
- Redfern, J. V., Moore, T. J., Becker, E. A., Calambokidis, J., Hastings, S. P., Irvine, L. M., ... Palacios, D. M. (2019). Evaluating stakeholder-derived strategies to reduce the risk of ships striking whales. *Diversity and Distributions*, 25(10), 1575–1585. <https://doi.org/10.1111/ddi.12958>
- Reeves, R. R., Ewins, P. J., Agbayani, S., Heide-Jørgensen, M. P., Kovacs, K. M., Lydersen, C., ... Blijleven, R. (2014). Distribution of endemic cetaceans in relation to hydrocarbon development and commercial shipping in a warming Arctic. *Marine Policy*, 44, 375–389. <https://doi.org/10.1016/j.marpol.2013.10.005>
- Reeves, R., Rosa, C., George, J. C., Sheffield, G., & Moore, M. (2012). Implications of Arctic industrial growth and strategies to mitigate future vessel and fishing gear impacts on bowhead whales. *Marine Policy*, 36, 454–462. <https://doi.org/10.1016/j.marpol.2011.08.005>
- Roberts, J. J., Best, B. D., Mannocci, L., Fujioka, E. I., Halpin, P. N., Palka, D. L., ... Lockhart, G. G. (2016). Habitat-based cetacean density

- models for the U.S. Atlantic and Gulf of Mexico. *Scientific Reports*, 6, 22615. <https://doi.org/10.1038/srep22615>
- Schick, R. S., & Lutcavage, M. E. (2009). Inclusion of prey data improves prediction of bluefin tuna (*Thunnus thynnus*) distribution. *Fisheries Oceanography*, 18, 77–81.
- Sequeira, A., Mellin, C., Rowat, D., Meekan, M. G., & Bradshaw, C. J. A. (2012). Ocean-scale prediction of whale shark distribution: Environmental drivers of whale shark distribution. *Diversity and Distributions*, 18, 504–518. <https://doi.org/10.1111/j.1472-4642.2011.00853.x>
- Silber, G. K., Lettrich, M. D., Thomas, P. O., Baker, J. D., Baumgartner, M., Becker, E. A., ... Waples, R. S. (2017). Projecting marine mammal distribution in a changing climate. *Frontiers in Marine Science*, 4, 413. <https://doi.org/10.3389/fmars.2017.00413>
- Torres, L. G., Read, A. J., & Halpin, P. (2008). Fine-scale habitat modeling of a top marine predator: do prey data improve predictive capacity. *Ecological Applications*, 18, 1702–1717. <https://doi.org/10.1890/07-1455.1>
- Torres, L. G., Sutton, P. J. H., Thompson, D. R., Delord, K., Weimerskirch, H., Sagar, P. M., ... Phillips, R. A. (2015). Poor transferability of species distribution models for a pelagic predator, the grey petrel, indicates contrasting habitat preferences across ocean basins. *PLoS One*, 10(3), e0120014. <https://doi.org/10.1371/journal.pone.0120014>
- Treacy, S. D. (2000). *Aerial surveys of Endangered Whales in the Beaufort Sea, Fall 1998–1999*. Final Report to the U.S. Minerals Management Service, Alaska OCS Region.
- Vaughan, I. P., & Ormerod, S. J. (2005). The continuing challenges of testing species distribution models: Testing distribution models. *Journal of Applied Ecology*, 42(4), 720–730. <https://doi.org/10.1111/j.1365-2664.2005.01052.x>
- Walkusz, W., Williams, W. J., Harwood, L. A., Moore, S. E., Stewart, B. E., & Kwasniewski, S. (2012). Composition, biomass and energetic content of biota in the vicinity of feeding bowhead whales (*Balaena mysticetus*) in the Cape Bathurst upwelling region (south eastern Beaufort Sea). *Deep-Sea Research I: Oceanographic Research Papers*, 69, 25–35. <https://doi.org/10.1016/j.dsr.2012.05.016>
- Wenger, S. J., & Olden, J. D. (2012). Assessing transferability of ecological models: An underappreciated aspect of statistical validation: Model transferability. *Methods in Ecology and Evolution*, 3(2), 260–267. <https://doi.org/10.1111/j.2041-210X.2011.00170.x>
- Wirsing, A. J., Heithaus, M. R., & Dill, L. M. (2007). Can measures of prey availability improve our ability to predict the abundance of large marine predators? *Oecologia*, 153, 563–568. <https://doi.org/10.1007/s00442-007-0769-0>
- Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., ... Sequeira, A. M. M. (2018). Outstanding challenges in the transferability of Ecological Models. *Trends in Ecology & Evolution*, 33(10), 790–802. <https://doi.org/10.1016/j.tree.2018.08.001>
- Young, T., Fuller, E. C., Provost, M. M., Coleman, K. E., St. Martin, K., McCay, B. J., & Pinsky, M. L. (2019). Adaptation strategies of coastal fishing communities as species shift poleward. *ICES Journal of Marine Science*, 76(1), 93–103. <https://doi.org/10.1093/icesjms/fsy140>
- Zhang, J. L., Ashjian, C., Campbell, R., Spitz, Y. H., Steele, M., & Hill, V. (2015). The influence of sea ice and snow cover and nutrient availability on the formation of massive under-ice phytoplankton blooms in the Chukchi Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, 118, 122–135. <https://doi.org/10.1016/j.dsr2.2015.02.008>
- Zhang, J. L., Spitz, Y. H., Steele, M., Ashjian, C., Campbell, R., Berline, L., & Matrai, P. (2010). Modeling the impact of declining sea ice on the Arctic marine planktonic ecosystem. *Journal of Geophysical Research-Oceans*, 115, C10015. <https://doi.org/10.1029/2009JC005387>

BIOSKETCH

Daniel E. Pendleton is a research scientist at the Anderson Cabot Center for Ocean Life at the New England Aquarium (USA). His research is focused on measuring and modelling the effects of climate change and human activities on marine mammals. This research team is focused on developing and using ocean models to improve our understanding of marine mammal species distributions, with a focus on executing modelling studies to address contemporary challenges in marine mammal conservation.

Author contributions: D.E.P. conceived the study design, implemented and evaluated the species distribution models, and led the writing. E.E.H. led the grant writing which funded the study and contributed to the study design, interpretation of results and writing. J.Z. conducted BIOMAS simulation and analysis. J.V.R. contributed to the interpretation of results and writing.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Pendleton DE, Holmes EE, Redfern J, Zhang J. Using modelled prey to predict the distribution of a highly mobile marine mammal. *Divers Distrib*. 2020;26:1612–1626. <https://doi.org/10.1111/ddi.13149>