



# Environmental drivers of large-scale movements of baleen whales in the mid-North Atlantic Ocean

Sergi Pérez-Jorge<sup>1,2</sup> | Marta Tobeña<sup>1,2</sup> | Rui Prieto<sup>1,2</sup> | Frederic Vandepierre<sup>1,2</sup> | Beatriz Calmettes<sup>3</sup> | Patrick Lehodey<sup>3</sup> | Mónica A. Silva<sup>1,2,4</sup>

<sup>1</sup>Okeanos R&D Centre, Institute of Marine Research, University of the Azores, Horta, Portugal

<sup>2</sup>Marine and Environmental Sciences Centre, Faculty of Sciences of the University of Lisbon, Lisbon, Portugal

<sup>3</sup>Marine Ecosystems Modeling and Monitoring by Satellites, CLS, Ramonville, France

<sup>4</sup>Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

## Correspondence

Sergi Pérez-Jorge, Okeanos R&D Centre, Institute of Marine Research, University of the Azores, 9901-862 Horta, Portugal.

Email: sergiperezjorge@gmail.com

## Funding information

Proconvergencia Açores/EU Program; Competitiveness Factors Operational (COMPETE); Azores 2020 Operational Programme; MARE, Grant/Award Number: UID/MAR/04292/2013; QREN European Social Fund; Fundo Regional da Ciência e Tecnologia (FRCT); Fundação para a Ciência e Tecnologia (FCT), Grant/Award Number: PTDC/MAR/74071/2006 and M2.1.2/F/012/2011

Editor: George Stevens

## Abstract

**Aim:** Understanding the environmental drivers of movement and habitat use of highly migratory marine species is crucial to implement appropriate management and conservation measures. However, this requires quantitative information on their spatial and temporal presence, which is limited in the high seas. Here, we aimed to gain insights of the essential habitats of three baleen whale species around the mid-North Atlantic (NA) region, linking their large-scale movements with information on oceanographic and biological processes.

**Location:** Mid-NA Ocean.

**Methods:** We present the first study combining data from 31 satellite tracks of baleen whales (15, 10 and 6 from fin, blue and sei whales, respectively) from March to July (2008–2016) with data on remotely sensed oceanography and mid- and lower trophic level biomass derived from the spatial ecosystem and population dynamics model (SEAPODYM). A Bayesian switching state-space model was applied to obtain regular tracks and correct for location errors, and pseudo-absences were created through simulated positions using a correlated random walk model. Based on the tracks and pseudo-absences, we applied generalized additive mixed models (GAMMs) to determine the probability of occurrence and predict monthly distributions.

**Results:** This study provides the most detailed research on the spatio-temporal distribution of baleen whales in the mid-NA, showing how dynamic biophysical processes determine their habitat preference. Movement patterns were mainly influenced by the interaction of temperature and the lower trophic level biomass; however, this relationship differed substantially among species. Best-fit models suggest that movements of whales migrating towards more productive areas in northern latitudes were constrained by depth and eddy kinetic energy.

**Main conclusions:** These novel insights highlight the importance of integrating telemetry data with spatially explicit prey models to understand which factors shape the movement patterns of highly migratory species across large geographical scales. In addition, our outcomes could contribute to inform management of anthropogenic threats to baleen whales in sparsely surveyed region.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

## KEYWORDS

baleen whales, movements, North Atlantic, Prey, SEAPODYM, species distribution models, tracking

## 1 | INTRODUCTION

Migratory marine animals are increasingly vulnerable to extinction and population declines due to the cumulative anthropogenic impacts on their environment. Understanding the distribution patterns of these wide-ranging taxa is critical for effective management (Halpern et al., 2015), but often challenging, due to the difficulty of studying their movements over vast ocean regions. Advances in satellite telemetry tags have improved the knowledge on the distribution and movements of highly mobile marine species and stimulated developments in species distribution modelling (SDM) to quantify and predict animal space use across time (Block et al., 2011). However, making population-level inferences from telemetry data poses many challenges due to the presence-only observations, such as the lack of absence data and serially correlated positions. One solution is to incorporate pseudo-absences to the data (i.e. representatively selected locations where the animals could have been sampled but were not) and to use an error structure (i.e. correlation terms, random effects) that more accurately reflects the variability within and between different levels of the natural hierarchy of the sampling units (Aarts, MacKenzie, McConnell, Fedak, & Matthiopoulos, 2008; Vandeperre, Aires-da-Silva, Lennert-Cody, Serrão Santos, & Afonso, 2016).

Developing SDMs that can effectively forecast animals' distribution depend on the availability of ecologically meaningful environmental data at the appropriate spatial and temporal resolutions. Marine top predators are supported by the productivity of primary and secondary consumers, and their distribution and movements are tied closely to those of their prey (Benoit-Bird et al., 2013; Boyd et al., 2015). Yet, due to the scarcity of prey data at scales relevant for marine top predators, most SDMs have relied on proxies for prey distribution such as physiographic variables and a handful of remotely sensed or in situ oceanographic measurements (Scales et al., 2017; Tobeña, Prieto, Machete, & Silva, 2016). However, the main issue with these proxies is the potential temporal and spatial lags that may occur between physical and oceanographic processes and biological responses (i.e. prey biomass), highlighting the importance of having a good understanding of the marine food web to predict the distribution of predators (Grémillet et al., 2008; Redfern et al., 2006). Recent advances in ecosystem models provide real-time and hindcast simulations of low and mid-trophic level biomass and production over large spatio-temporal scales (Lehodey et al., 2015; Lehodey, Murtugudde, & Senina, 2010), which allow the incorporation of predicted prey distributions into SDMs and a better understanding of the ecosystem components (Lambert, Mannocci, Lehodey, & Ridoux, 2014; Mannocci, Roberts, Miller, & Halpin, 2016).

Large baleen whales are believed to make seasonal migrations from high-latitude productive waters used in summer to feed, and

low-latitude oligotrophic waters used in winter for breeding and calving (Stern, 2009). During feeding periods, these whales concentrate mainly in areas where their preferred prey (euphausiids, small epipelagic and mesopelagic schooling fish) occurs in predictable aggregations and at high densities (Croll et al., 2005; Laidre et al., 2010). In the NA Ocean, studies examining the drivers of baleen whale distribution are concentrated along the US and European coasts (Kaschner, Quick, Jewell, Williams, & Harris, 2012). In the north-west Atlantic, the spatio-temporal distribution of baleen whales was strongly related to static variables and several formulations of biological productivity, such as mesoscale fronts, and eddies, which can aggregate micronekton and zooplankton biomass (Roberts et al., 2016). In the European Atlantic shelf waters, oceanographic and physiographic features, such as sea surface temperature, depth and distance to the 2,000 m contour, were important predictors of the summer distribution of baleen whales (Macleod et al., 2009).

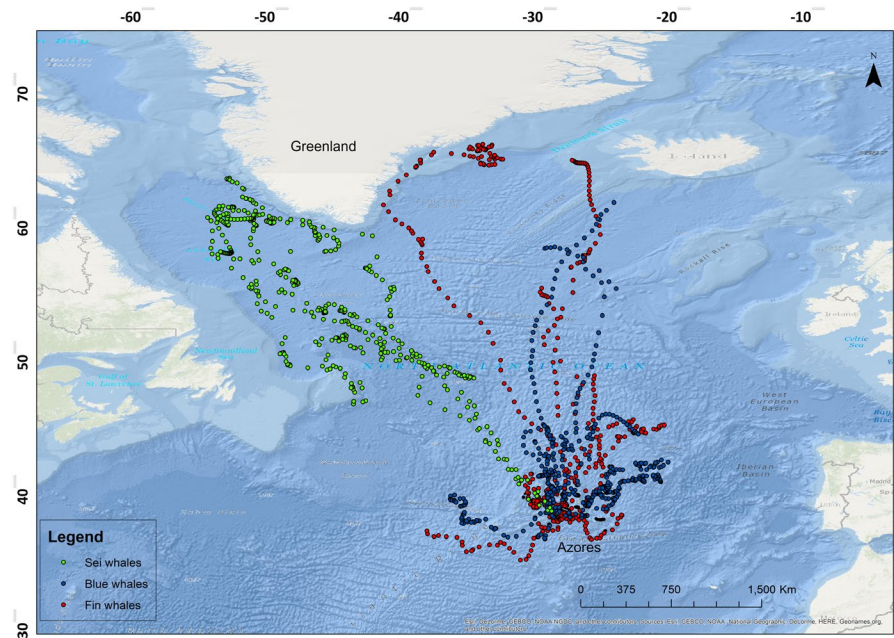
The drivers shaping baleen whale migration in the mid-NA are poorly understood due to the difficulty of observing their movements during migratory periods. Identifying the drivers of baleen whales' distribution is necessary to understand the dynamics of their populations and to forecast how these whales might respond to ecosystem variability associated with climatic changes (Hazen et al., 2013). Studies characterizing the migration patterns of these whales remain largely descriptive (Prieto, Silva, Waring, & Gonçalves, 2014; Silva, Prieto, Jonsen, Baumgartner, & Santos, 2013). Therefore, our aim was to develop habitat models based on telemetry data to investigate the distribution of three baleen whale species around the mid-NA: the blue whale (*Balaenoptera musculus*), fin whale (*B. physalus*) and sei whale (*B. borealis*). Specifically, we implemented a comprehensive ecological modelling approach to determine the effect of oceanographic conditions and prey availability as drivers of their northward migration. Based on these models, we also predicted the spatio-temporal distribution along their northward migration in order to identify important habitats for each species in the mid-NA.

## 2 | METHODS

### 2.1 | Tracking data

We deployed 31 location-only satellite tags (SPOT5 Wildlife Computers, USA) on fin, blue and sei whales (15, 10 and 6, respectively) off the Azores, between June 2008 and June 2016 (Figure 1; additional details in Table S1). Tagging procedures were conducted following the guidelines of the American Society of Mammalogists (Sikes, Gannon, & The Animal Care and Use Committee of the American Society of Mammalogists, 2011) and under research permits approved by the relevant authorities (Regional Directorate

**FIGURE 1** Map showing the original positions for blue, fin and sei whales



for Sea Affairs, Autonomous Region of the Azores). The transmitters (model SPOT5-implantable, Wildlife Computers, USA) were deployed from a 6-m rigid-hulled inflatable or a 12-m fibreglass boat using a compressed air gun (ARTS/RN, Restech, Norway) set at 8–10 bar pressure (Prieto et al., 2014; Silva, Prieto, Jonsen, et al., 2013). All tags were programmed to transmit on a daily basis, providing hourly data up to a maximum of 500 daily messages. The raw and unfiltered locations provided by the tags were fitted in a Bayesian switching state-space model (SSSM) to obtain positions at regular time steps, and to improve position estimates by incorporating the measurement errors provided by Argos (Jonsen et al., 2013). The SSSM was built in a hierarchical framework to estimate parameters jointly across multiple individual tracks. We tested multiple time steps (2, 5, 12 and 24 h) to find the most appropriate parameters to reduce the sample autocorrelation (see modelling approach workflow in Figure S1). For each model, we ran two Markov chain Monte Carlo (MCMC) chains for 50,000 iterations, dropping the first 45,000 samples as a burn-in and retaining every 5th sample from the remaining 5,000 assumed postconvergence samples. Model convergence and sample autocorrelation were visually checked in the diagnostic plots, and a time step of 12 h was selected as the most suitable due to its lower sample autocorrelation. The analysis was performed using the R software (version 2.15.3; R Development Core Team, 2015) and the package “bsam” (Jonsen, Flemming, & Myers, 2005).

## 2.2 | Simulated correlated random walks

As the track locations obtained through the SSSM only provide information on the presence of whales, we created simulated tracks (i.e. “pseudo-absences”) for each original track (case points) to characterize environmental conditions potentially available to

individuals (Aarts, Fieberg, & Matthiopoulos, 2012). The simulated tracks (control points) were built using a correlated random walk (CRW) model (Kareiva & Shigesada, 1983), randomly taking the first or last transmitted position of each satellite track as the starting point to avoid bias towards areas more intensively sampled (Phillips et al., 2009; see Figure S2). For each real track, we simulated 200 CRW tracks with the same duration and dates of the original tracks, and with turning angles and step lengths randomly sampled from the distribution of the original tracks (Willis-Norton et al., 2015). The CRW tracks from each species were simulated within a different restricted area, based on the combination of all original tracks and increased by a buffer zone accounting for the error of the satellite tags (100 km from the closest original track; Žydelis et al., 2011). Various studies highlighted the importance of choosing pseudo-absences on the same geographical extent as the tracking data (Barbet-Massin, Jiguet, H el ene Albert, & Thuiller, 2012). Moreover, selection of pseudo-absences can influence variable selection, model fit and predictions of SDMs (Wisz & Guisan, 2009). To ensure selection of an appropriate pseudo-absence dataset, we assessed the similarity between the original and the simulated tracks based on the overall direction and distance, assigning a flag value between 0 and 4 (0 being the most similar tracks and 4 the most dissimilar). This flag value was estimated as the normalized difference between the original and the simulated track length distance,  $d$ , summed with the normalized difference in net angular displacement,  $\theta$ , of the original and simulated track (Willis-Norton et al., 2015):

$$\text{Flag value} = 2 \times (\text{Distance}_{\text{original}} - \text{Distance}_{\text{sim}}) / \text{Distance}_{\text{original}} + (\text{Angle}_{\text{original}} - \text{Angle}_{\text{sim}}) / 90.$$

Simulated tracks within the upper quartile of dissimilar flag values and tracks crossing land were excluded from the analysis.

### 2.3 | Environmental and prey-related variables

Candidate variables were selected based on species' ecological preferences and regional oceanographic and physiographic characteristics (Redfern et al., 2006). Potential prey biomass distributions were obtained from the mid-trophic level spatial ecosystem and population dynamics model (SEAPODYM-MTL, Lehodey et al., 2015; Lehodey et al., 2010; Table 1). The SEAPODYM-MTL model simulates spatial and temporal dynamics of production and biomass of micronekton between the surface and 1,000 m depth. The model is driven by physical and biological variables, using a system of advection–diffusion–reaction equations (Lehodey et al., 2010). Physical variables, current and temperature (T) data (Table 1), were obtained from the eddy resolving global ocean reanalysis GLORYS-2v1 provided by Mercator Ocean (<http://www.mercator-ocean.fr/eng>), based on the ocean general circulation model configuration ORCA025 NEMO, at a spatial resolution of 0.25° for the 1992–2009 period (Bernard et al., 2006) and assimilated satellite and in situ data. From 2010, GLORYS.2v1 was updated with the equivalent Mercator Ocean operational real-time system (PSY3). Biological variables, net primary production (NPP) and euphotic depth (ZEU) data, were derived from ocean colour satellite data (<http://www.science.oregonstate.edu/ocean>), using the vertically generalized production model (VGPM) of Behrenfeld and Falkowski (1997). Both GLORYS-2v1 outputs and primary production data were interpolated onto a regular 0.25° × 0.25° grid with a weekly time step to be used as forcing for the SEAPODYM. In the model, three vertical layers are defined in relation to the euphotic depth, as micronekton diel vertical migration (DVM) is mainly induced by daylight variations, validated with biomass estimates from micronekton sampling cruises and acoustic backscatter data (Lehodey et al., 2015, 2010). The SEAPODYM-MTL model simulates six functional groups of micronekton according to the dynamics of their vertical distribution: epipelagic, migrant upper mesopelagic, upper mesopelagic, highly migrant lower mesopelagic, migrant lower mesopelagic and lower mesopelagic (Table 1). As baleen whales mainly forage on prey available within the epipelagic layer (Goldbogen et al., 2017), functional groups remaining during day and night within the lower mesopelagic layers (>400 m deep) were not included in the analysis.

We also incorporated the outputs of the lower trophic level SEAPODYM (SEAPODYM-LTL), which follow the same modelling framework of the mid-trophic level, and considered a single functional group including all mesozooplankton organisms (i.e. both holo- and mero-zooplankton). This model also had two state variables, production (LTL-prod) and biomass (LTL-bio), integrated over the epipelagic layer defined by the euphotic depth (Conchon, 2016) (Table 1). In addition to these prey-related variables, we included the surface chlorophyll-a concentration (Chl-a) used to compute the primary production and eddy kinetic energy (EKE) calculated as  $EKE = \frac{1}{2} \times (u^2 + v^2)$ , with *u* and *v*, respectively, the meridian and zonal geostrophic current components that are used as forcing variables for the epipelagic layer of SEAPODYM.

Regarding static environmental variables, depth was obtained from ETOPO1 Global Relief Model at a spatial resolution of 1/60°, provided by NOAA's National Geophysical Data Center (NGDC, <https://www.ngdc.noaa.gov/mgg/global/>). Derived static variables were slope, calculated using the elevation values of the four immediately adjacent depths (Ritter, 1987), and distance to coast (See ST1). Following the identification of time-lagged Chl-a concentration as an important predictor of the spatio-temporal distribution of baleen whales around the study area (Prieto, Tobeña, & Silva, 2016), we tested a time-lag of 1 month prior to the point location for the following variables (indicated by variable name\_1m): LTL-bio, LTL-prod, Chl-a and NPP. All environmental and prey-related variables were extracted for each original and simulated track point using the R software and the packages "raster" (Hijmans & van Ettern, 2012), "ncdf4" (Pierce, 2019) and "rgdal" (Bivand et al., 2019). Data points with incomplete information were removed from the final dataset, corresponding to 0.02%, 0.06% and 1.60% of the total dataset for fin, blue and sei whale, respectively. Prior to running the models, we examined collinearity between pairs of covariates (Zuur, Ieno, & Elphick, 2010) using the Pearson's correlation coefficient and selected those variables less correlated (<0.7) and ecologically more relevant (Dormann et al., 2013).

### 2.4 | Habitat modelling

We applied generalized additive mixed models (GAMMs) to the original and simulated track points to predict the habitat preference of each whale species. GAMMs were fit with a binomial distribution and a logit link function, using the restricted maximum likelihood (REML, mgcv 1.8.7; Wood, 2006) as an optimization method and individual IDs as a random effect. To model our data as a binary response variable, original points were assigned a value of 1 and simulated points a value of 0 (Aarts et al., 2012). In addition to single covariates, we defined two biologically interpretable interactions between covariates using two-dimensional tensor product smooths: T × LTL-bio and EKE × LTL-bio (Wood, Scheipl, & Faraway, 2013). These interactions were chosen based on current knowledge of the oceanographic conditions, as well as the relationship between these variables to understand the habitat preference of marine top predators within our study area (Vandepierre et al., 2016).

The GAMMs were built using a backward selection of relevant individual smooth terms and tensor product smooths. Nested models were compared using chi-square tests. To select the tensor product smooths, the original structure was tested to the equivalent additive structure while controlling for the degrees of freedom using the *k*-parameter (i.e. the basis dimension) (Vandepierre et al., 2016; Wood, 2006). Once the best model was chosen based on the lowest corrected AIC, the smoothing splines were limited to a maximum of 5 degrees of freedom to avoid overfitting and preserve the ecological interpretability of functional relationships. The effect of this limitation on the number of degrees of freedom was evaluated by comparing the final outputs and predictions from models with

**TABLE 1** Description of candidate variables for habitat modelling

Environmental variable (acronym, units)	Spatial resolution	Source	Biological/Ecological characteristics
Depth (Depth, m)	0.1°	Etopo 1. 1 Arc-Minute Global Relief Model	Coastal versus pelagic domains
Sea water potential temperature (T, °C)	0.25°	Mercator Ocean	Water mass distribution
Euphotic depth (ZEU, m)	0.25°	OSU (1998–2015)	Zone of effective photosynthesis
Chlorophyll-a (Chl-a, mg/m <sup>3</sup> )	0.25°	OSU (1998–2015)	Ocean productivity domains
Net primary production of carbon (NPP, mmolC/m <sup>2</sup> /d)	0.25°	CLS, VGPM algorithm (1998–2015)	Ocean productivity domains
Slope (Slope, degrees from the horizontal)	0.1°	Etopo 1. 1 Arc-Minute Global Relief Model	Influence of topographic features
Distant to coast (Coast, Km)	0.1°	Etopo 1. 1 Arc-Minute Global Relief Model	Onshore–offshore distribution patterns
Epipelagic micronekton biomass (epi_pb, gWW/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Biomass of micronekton remaining during day and night within the epipelagic layer
Epipelagic micronekton production (epi_pp, gWW/d/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Production of micronekton remaining during day and night within the epipelagic layer
Upper mesopelagic micronekton biomass (u-meso_pb, gWW/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Biomass of micronekton remaining during day and night within the upper mesopelagic layer
Upper mesopelagic micronekton production (u-meso_pp, gWW/d/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Production of micronekton remaining during day and night within the upper mesopelagic layer
Migrant upper mesopelagic micronekton biomass (mu-meso_pb, gWW/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Biomass of micronekton remaining during day within the upper mesopelagic layer and migrating at night to the epipelagic layer
Migrant upper mesopelagic micronekton production (mu-meso_pp, gWW/d/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Production of micronekton remaining during day within the upper mesopelagic layer and migrating at night to the epipelagic layer
Lower mesopelagic micronekton biomass (l-meso_pb, gWW/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Biomass of micronekton remaining during day and night within the lower mesopelagic layer
Lower mesopelagic micronekton production (l-meso_pp, gWW/d/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Production of micronekton remaining during day and night within the lower mesopelagic layer
Migrant lower mesopelagic micronekton biomass (ml-meso_pb, gWW/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Biomass of micronekton remaining during day within the lower mesopelagic layer and migrating at night to the upper mesopelagic layer
Migrant lower mesopelagic micronekton production (ml-meso_pp, gWW/d/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Production of micronekton remaining during day within the lower mesopelagic layer and migrating at night to the upper mesopelagic layer
Highly migrant lower mesopelagic micronekton biomass (hml-meso_pb, gWW/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Biomass of micronekton remaining during day within the lower mesopelagic layer and migrating at night to the epipelagic layer
Highly migrant lower mesopelagic micronekton production (hml-meso_pp, gWW/d/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Production of micronekton remaining during day within the lower mesopelagic layer and migrating at night to the epipelagic layer
Lower trophic level plankton biomass (LTL-bio, gWW/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Biomass of mesozooplankton within the epipelagic layer
Lower trophic level plankton production (LTL-prod, gWW/d/m <sup>2</sup> )	0.25°	SEAPODYM ecosystem model*	Production of mesozooplankton within the epipelagic layer

Note: The SEAPODYM defined three vertical layers in relation to the euphotic depth: Epipelagic layer as <1.5 of the euphotic depth (~150 m), upper mesopelagic layer as >1.5 and <4.5 of the euphotic depth (between 150 and 400 m) and lower mesopelagic layer as >4.5 of the euphotic depth (>400 m). Examples of species composition of functional groups are provided in Lehodey et al. (2010).

\*is to describe the seapodym ecosystem model

constrained and unconstrained smooth terms. No differences were detected between constrained and unconstrained models on variable selection and predictive performance (results not presented).

Following this habitat modelling approach, we evaluated the performance of models built with both environmental and prey-related variables, with that of models including only environmental

variables, in order to assess the contribution of modelled mid- and lower trophic level biomass and production.

A sensitivity analysis was also carried out to determine whether the simulated CRW tracks produce outputs that were robust to parameter uncertainty. We randomly selected two of the 200 CRWs tracks per whale and assessed whether the predictors from the best-fitting model were also significant for the different simulated points, repeating this procedure 40 times (Hazen et al., 2016). For each of the 40 runs, we calculated the area under the receiver operating characteristic (AUC) curve and the variance explained ( $R^2$ ). The AUC index ranges from 0 to 1, with values below 0.6 indicating a discrimination ability no better than random, values between 0.7 and 0.9 considered as reasonable and values >0.9 as excellent (DeLong, DeLong, & Clarke-Pearson, 1988).

Successive observations along the track of the same individual can be serially correlated with respect to associated habitat characteristics, which may underestimate parameter uncertainty and result in selection of inappropriate predictor variables during model selection (Fieberg, Matthiopoulos, Hebblewhite, Boyce, & Frair, 2010). To check whether the serial correlation was an issue in our data, we subsampled the original dataset by selecting every fifth data point, representing a time interval of 60 hr, which was considered sufficient for the individuals to move anywhere with overall distinct environmental conditions (Žydelis et al., 2011). GAMMs fit to the subsampled dataset were compared with the standard models through variable importance and AUCs.

## 2.5 | Model predictions

Monthly predictions of habitat preferences (scale from 0 to 1) were produced for each species based on the best GAMMs, at the spatial resolution of the fitted models (0.1°). Environmental and prey-related variables were compiled at a monthly resolution during the tracking period to obtain mean predictions for the study area and calculate the standard deviation (SD), in order to measure the stability of the predicted distribution, with stable and unstable habitat represented by low and high SD.

## 3 | RESULTS

### 3.1 | Tracking data

Tracking data were collected for the three target species between March and July, from 2008 to 2016. Two fin whale tracks shorter than 3 days were excluded from the analysis (final dataset of  $n = 29$ ; Table 2). Tracking duration and distance travelled were shortest for fin whales and longest for sei whales. Blue and fin whales generally moved north or north-east; one blue whale and one fin whale headed westwards and crossed the Mid-Atlantic Ridge. Most individuals of both species showed limited longitudinal displacement remaining within 1,000 km of the Azores. All tracked sei whales

**TABLE 2** Summary of tracking data used to fit habitat models for each whale species

Data collection	Species		
	Fin whale	Blue whale	Sei whale
Total number of tracks	13	10	6
Year interval	2009–2015	2009–2016	2008–2009
Month interval	March–July	April–July	May–July
Mean duration in days (SD)	20 (12)	27 (15)	39 (22)
Mean locations received (SD)	334 (508)	223 (162)	354 (384)
Mean distance travelled in km (SD)	2,167 (1862)	2,389 (1,366)	4,041 (2,890)

travelled north-west towards the Labrador Sea, which defined a clear migratory corridor. Once they reached these waters, individual whales exhibited different space use patterns, often associated with the shelf break off Canada and Greenland.

## 3.2 | Habitat preferences and predicted distributions

### 3.2.1 | Fin whale

The best-fitting GAMM included two additive covariates (Depth and  $\log_{10}$  EKE) and one tensor product smooth (LTL-bio  $\times$  T) (Table 3; Table S2). Fin whales preferred water depths < 2,500 m with high EKE (> 0.005  $\text{cm}^2/\text{s}^2$ ) (Figure 2). The surface of the tensor product smooth of T and LTL-bio was plotted on the scale of the response. The tensor product smooth indicated an association with cold (<10°C) productive habitats (LTL-bio > 100 gWW). Habitat preferences decreased sharply at T values between 10 and 15°C, but increased again for warmer waters with low LTL-bio. The environmental and prey-related variables of best model were significant in the large majority of the 40 models ran with different simulated points, showing the robustness of these variables to control point selection (Table 1). These models explained on average 31% of the total deviance ( $SD = 0.09$ ) and had an average AUC of 0.83 ( $SD = 0.04$ ), indicating good predictive performance. The best model using only environmental variables included five covariates (NPP\_1m, temperature, EKE, Depth and Slope) and showed a lower variance explained and predictive performance (Table S3).

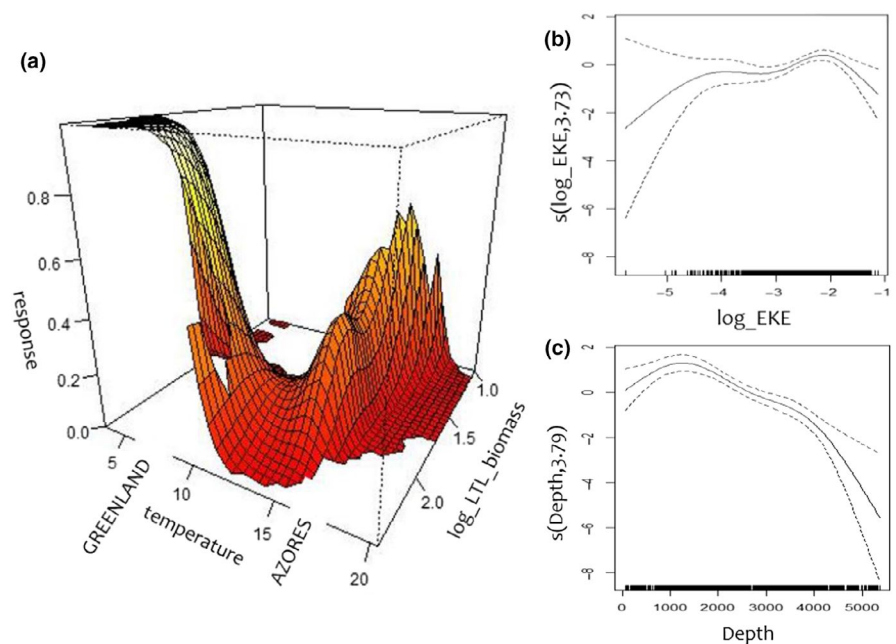
The monthly predictions captured the seasonal movement of fin whales along the mid-NA (Figure 3). In March, the highest predicted probability of presence occurred along the east coast of Greenland and south-west of the Irminger Sea, just over 50°N of latitude. This preferred habitat expanded towards the north-east over the following months, covering the whole Irminger Sea and reaching the west coast of Iceland, by June and July. The mid-Atlantic Ridge around the Azores was also a relevant area for the species, mainly between March and May. Model predictions showed a fair amount

**TABLE 3** Summary of the habitat modelling outputs for the three species: fin, blue and sei whale

Variable	edf	Ref.df	Chi-sq	n-significant (n/40)
<b>Fin whale</b>				
te(LTL-bio, T)	13.160	14.913	159.65	40/40
s(log_EKE)	3.735	3.957	14.74	30/40
s(Depth)	3.785	3.970	64.95	40/40
	$R^2$ : 0.31 (0.09)	AIC: 1261.66 (121.54)	AUC: 0.83 (0.04)	
<b>Blue whale</b>				
te(LTL-bio, T)	8.127	8.663	32.510	40/40
s(LTL-prod_1m)	2.858	3.617	9.521	34/40
s(Depth)	8.136	8.799	23.124	37/40
	$R^2$ : 0.25 (0.07)	AIC: 1672.55 (116.29)	AUC: 0.81 (0.03)	
<b>Sei whale</b>				
te(LTL-bio, T)	20.854	21.757	227.37	40/40
s(LTL-prod_1m)	3.418	3.821	16.47	8/40
s(Depth)	3.873	3.990	48.72	33/40
	$R^2$ : 0.31 (0.09)	AIC: 1353.53 (149.92)	AUC: 0.83 (0.04)	

Note: Models were run 40 times, and the significance of their variables was evaluated with p values of less than 0.01. Variances explained ( $R^2$ ), Akaike information criterion (AIC) and area under the curve (AUC) were obtained for each run, calculating a mean and SD. Variables included in the best models below: lower trophic level biomass (LTL-bio), lower trophic level production on the previous month (LTL-prod\_1m), eddy kinetic energy (log\_EKE) and depth (Depth).

**FIGURE 2** Generalized additive mixed model (GAMM) smoothers for fin whales. (a) Surfaces of the tensor product smoothers (temperature, T, and lower trophic level biomass, LTL\_bio) plotted at the scale of the response. (b) and (c) GAMMs smoother for log\_EKE (eddy kinetic energy) and Depth, respectively. Dashed lines represent the 95% confidence intervals for the fitted relationships

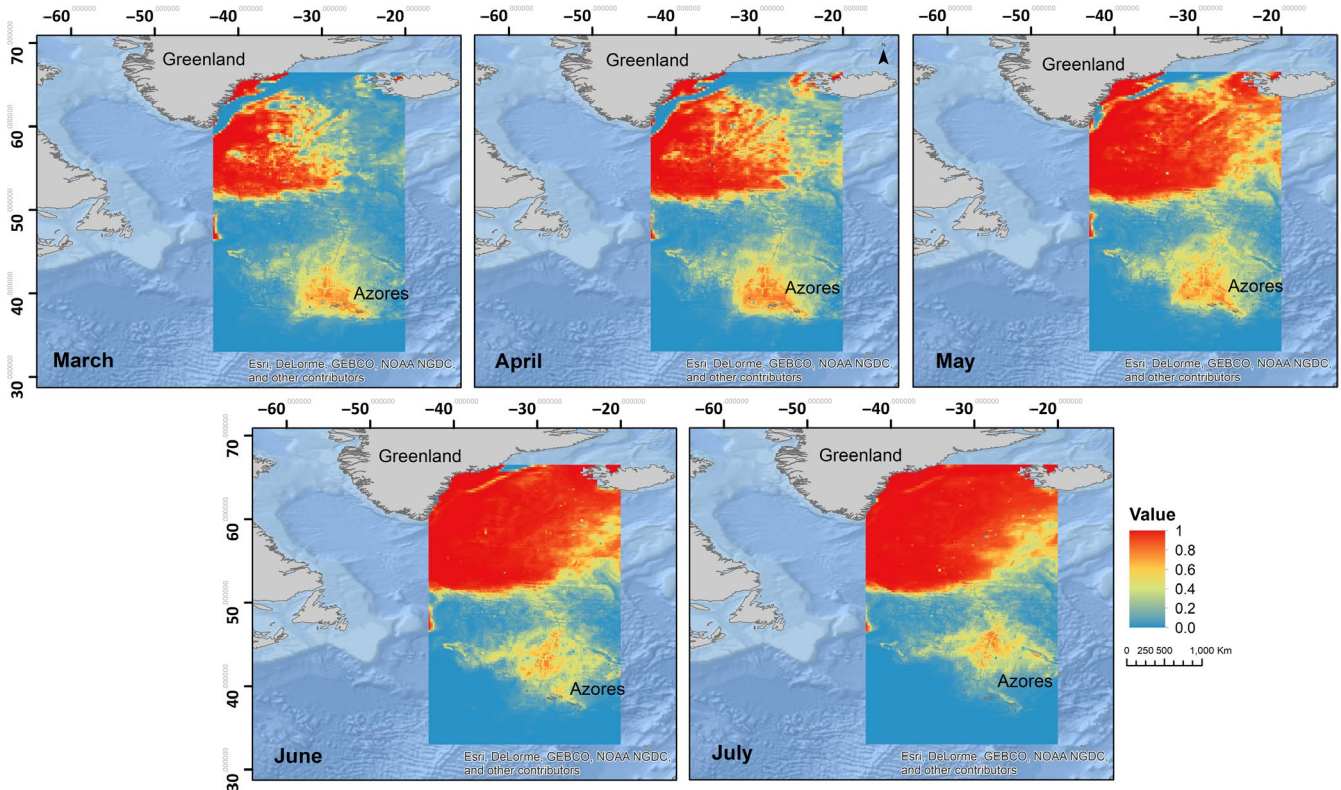


of uncertainty across the study region from March to May, with this uncertainty remaining in the southern region of the study area in June and July (Figure S3).

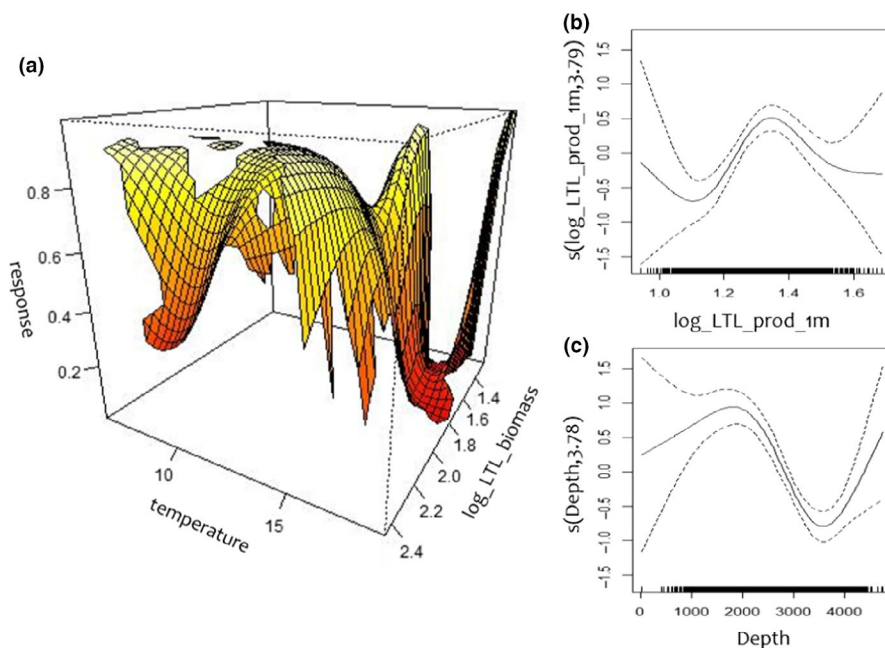
### 3.2.2 | Blue whale

The best model predicting blue whales' habitats included one tensor smooth product (LTL-bio  $\times$  T) and two additive covariates (Depth and

LTL-prod\_1m) (Table 3; Table S2). The tensor smooth product of T and LTL-bio defined a complex interaction surface (Figure 4). The species showed a bimodal response for T values between 15 and 18°C, with either high or low LTL-bio, and for T ranging from 8 to 14°C with high LTL-bio (>65 gWW). Blue whales had a preference for areas with intermediate values of LTL-prod in the previous month and water depths < 2,500 m. All variables included in the best model were significant in at least 85% of the 40 models ran with different control points, indicating that selection of pseudo-absences did not have a



**FIGURE 3** Predicted distribution of fin whales from March to July. Habitat preferences were scaled from 0 (low preference) to 1 (high preference). Spatial resolution was set to  $0.1^\circ$



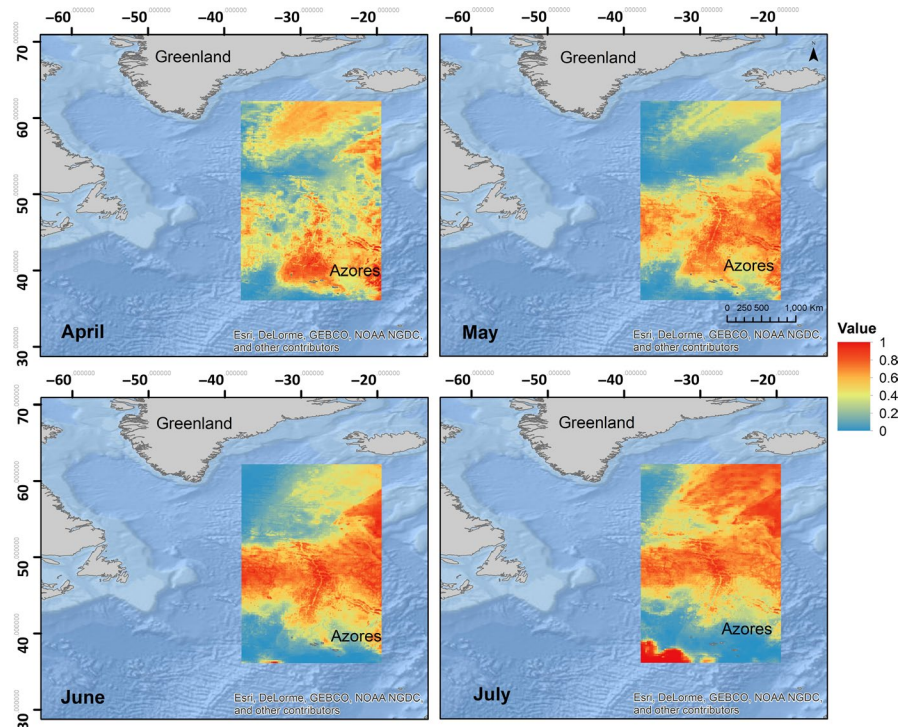
**FIGURE 4** Generalized additive mixed model (GAMM) smoothers for blue whales. (a) Surfaces of the tensor product smooths (temperature, T, and lower trophic level biomass, LTL\_bio) plotted at the scale of the response. (b) and (c) GAMMs smoother for  $\log_{LTL\_prod\_1m}$  and Depth, respectively. Dashed lines represent the 95% confidence intervals for the fitted relationships

strong influence on model results (Table 1). The models had a good predictive ability to discriminate between areas where blue whales were present versus those areas where they were absent, with an average AUC of 0.81 ( $SD = 0.03$ ), and a total deviance explained of 25% ( $SD = 0.07$ ). The best model including only environmental variables (NPP\_0m, NPP\_1m, temperature, EKE and Depth) had a moderate predictive performance, lower than the models with prey-related

variables (Table S3). Probability of blue whale presence was highest between 35 and 50°N in April and May, at north-eastern latitudes in June, and in the south-west coast of Iceland by July (Figure 5). The probability of occurrence of blue whales in the Azores decreased from April to July, reaching almost zero in July. Predictions showed a fair amount of uncertainty in the northern part of the study area in April, but decreased drastically in the following months (Figure S4).



**FIGURE 5** Predicted distribution of blue whales from April to July. Habitat preferences were scaled from 0 (low preference) to 1 (high preference). Spatial resolution was set to  $0.1^\circ$



### 3.2.3 | Sei whale

The final model for sei whales contained one tensor smooth product (LTL-bio  $\times$  T) and two additive covariates (Depth and LTL-prod\_1m) (Table 3; Table S2). Sei whales were associated with deep waters (>1,500 m) and intermediate values of LTL-prod from the previous month (25–65 gWW/day) (Figure 6). The tensor smooth product of T and LTL-bio indicates that sei whales used relatively cold waters ( $\sim 10^\circ\text{C}$ ) with average productivity values (10–100 gWW), as well as colder areas ( $\sim 5^\circ\text{C}$ ) with high productivity (LTL-bio > 100 gWW). Preference decreased sharply for warmer areas ( $>15^\circ\text{C}$ ), even when LTL-bio was high. The tensor smooth product was significant in 100% of the 40 models ran with different control points, in 80% of the models with the covariate Depth and in only 20% of models with the LTL production from the previous month. The average AUC was 0.83 ( $SD = 0.04$ ), and the average deviance explained was 31% ( $SD = 0.09$ ). The best model with environmental variables included four covariates (temperature, EKE, Depth and Slope) and had a lower predictive performance and variance explained than models with prey-related variables (Table S3).

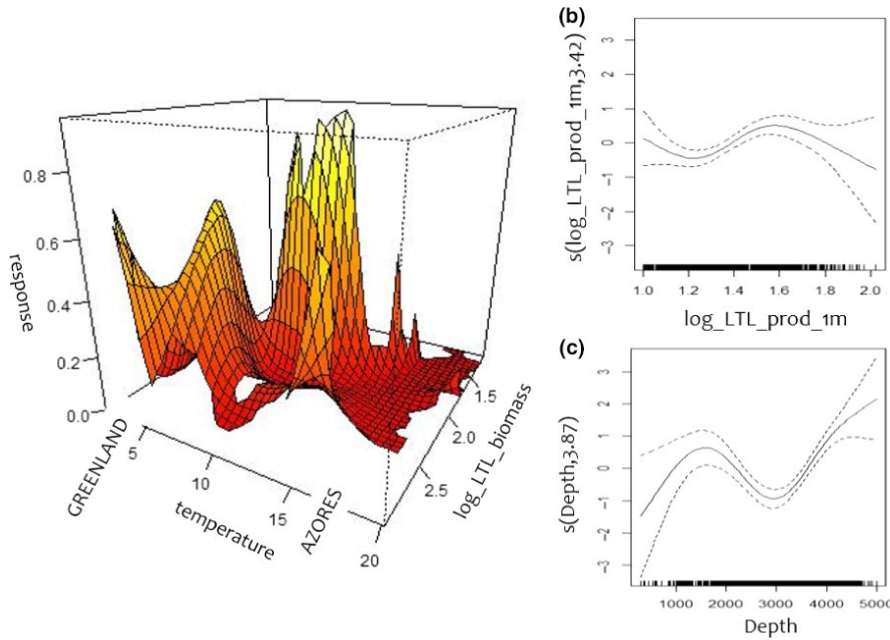
Latitudes above  $45^\circ\text{N}$ , mainly around the south-west of the Irminger Sea and in the Labrador Sea (except the shallow waters along the Greenland coast), were identified as important habitats for sei whales (Figure 7). In May, the main hotspot was near the Charlie-Gibbs Fracture Zone (CGFZ) in the mid-NA. In June and August, the predicted distribution shifted northwards and was concentrated in the entire Labrador Sea and in a few areas along the Greenland coast. Low occurrences were predicted in the area around Azores. Predictions indicated a low uncertainty across the whole study region in all sampled months (Figure S5).

### 3.3 | Serial correlation evaluation

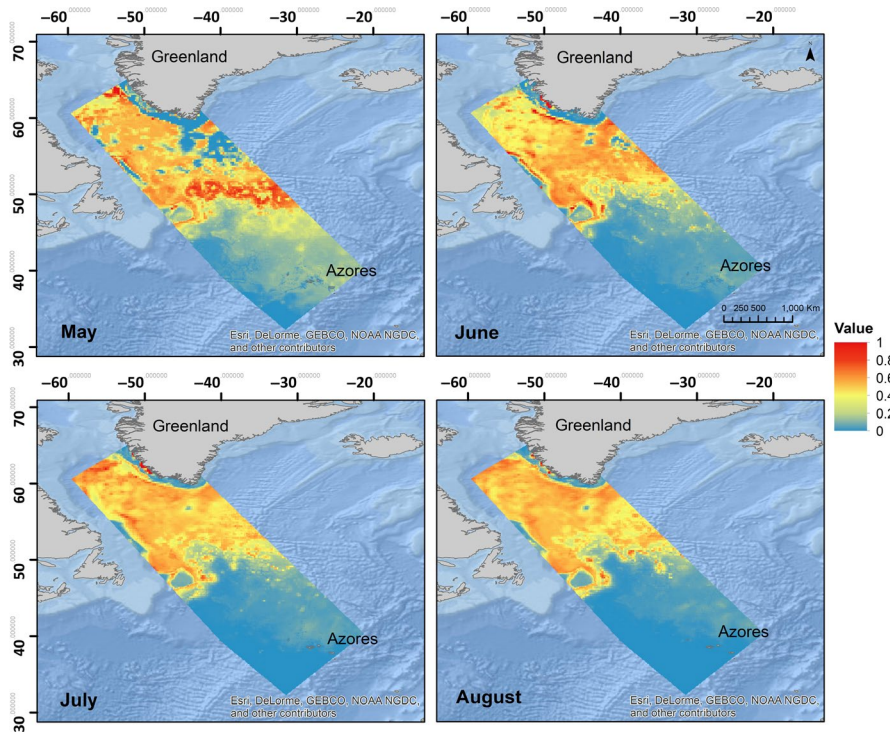
Models based on subsampled and full datasets yielded similar results for all three species in terms of importance of predictor variables and robustness to control point selection (Table S4). In the case of fin and blue whales, subsample models had slightly lower AUC values than full dataset models but AUC values were similar for sei whale models. Nonetheless, habitat preference predictions from these models were comparable to those from full datasets for all the species (Figures S6–S8). Taken together, these results indicate that our analyses and predictions were robust to serial correlation.

## 4 | DISCUSSION

This study offers novel insights into the distribution of highly migratory species across large geographical scales, and fills in a gap in knowledge of baleen whale distribution in the mid-NA. We provide spatially and temporally explicit information on the distribution of the three species of baleen whales through the combination of satellite tracking data, remotely sensed oceanography and ecosystem models, which could contribute to implement appropriate conservation and management measures. However, it should be noted that our analysis is based on a relatively small sample size, especially for sei whales, and that some of the tracks were short and did not cover the full extent of movements to the whales' presumed northern feeding grounds. Thus, our habitat preferences may not adequately reflect the relationship of whales with conditions at their feeding grounds. Moreover, inter-monthly variations may not be fully representative



**FIGURE 6** Generalized additive mixed model (GAMM) smoothers for sei whales. (a) Surfaces of the tensor product smooths (temperature, T, and lower trophic level biomass, LTL\_bio) plotted at the scale of the response. (b) and (c) GAMMs smoother for log\_LTL\_prod\_1m and Depth, respectively. Dashed lines represent the 95% confidence intervals for the fitted relationships



**FIGURE 7** Predicted distribution of sei whales from May to August. Habitat preferences were scaled from 0 (low preference) to 1 (high preference). Spatial resolution was set to 0.1°

of the seasonal movements of the species due to the limited number of tracks per month within each sampling year.

Regarding prey-related variables, SEAPODYM data also had some issues related to data estimation and validation. First, various sources of uncertainty are associated with the observations used to validate the model, due to the different techniques employed to collect zooplankton biomass data and its spatio-temporal distribution (Conchon, 2016). Second, the LTL data are based on mesozooplankton species, which constitutes only a portion of the diet of our targeted baleen whale species. While further

developments are required to optimize final output parameters, this study shows promising results on the application of modelled zooplankton biomass to study marine top predators at large scale.

#### 4.1 | Habitat preferences

Our models show that the distribution of blue, fin and sei whales is strongly associated with prey availability and with static and

dynamic environmental variables closely linked to productivity. However, habitat preferences varied among species and across the distribution range of each species. Fin whales exhibited a bimodal response to environmental variables along the northward migration route, with a marked preference for cold waters associated with high LTL-bio levels and for warmer waters, and avoidance of intermediate temperatures. Such bimodal response reflects differential exploitation of distinct water masses and productivity regimes at the ends of their studied range, one characteristic of high productivity in cold waters and the other of low productivity in temperate waters. Sei whales also showed a clear preference for productive subpolar waters in favour of temperate ones, displaying a two-phase distribution pattern during the study period, consistent with limited available observations (Sigurjónsson, Gunnlaugsson, & Payne, 1989; Waring, Nottestad, Olsen, Skov, & Víkingsson, 2008). In comparison, blue whales had a more homogenous response across the study area, targeting areas with intermediate LTL-bio values that were productive on previous months. The retention of T and LTL-bio, and the different response obtained by the three species suggest that interactions between environmental and prey-related covariates may help to predict species distribution. Past studies showed the utility of proxies such as SST and productivity to predict the occurrence, abundance and species richness of marine top predators. Although these predictors might indicate important foraging areas for these species, they were probably decoupled in time and space, and from prey availability at higher trophic levels (Grémillet et al., 2008). Incorporating the novel LTL data allowed us to shorten the temporal lags between the predictors and the predators, helping to reduce the possible spatial mismatch between predictor and response variables and representing a more direct biological relationship.

The habitat preference for the three species was significantly constrained by depth, which has been shown to be a good predictor of baleen whale distribution in multiple regions (Scales et al., 2017), including in the NA (Víkingsson et al., 2015). Blue and sei whales preferred oceanic areas and avoided coastal regions, whereas fin whales used both coastal and oceanic areas. In addition, fin whale distribution was strongly influenced by EKE, a measure of mesoscale activity. Upwelling/downwelling processes are common in areas with intense mesoscale activity, such as eddies and fronts, enhancing marine productivity and leading to the formation of predictable prey patches (Worm, Sandow, Oschlies, Lotze, & Myers, 2005). Numerous studies have linked these features to the presence of marine top predators, including fin whales (Reisinger et al., 2018; Scales et al., 2017). While fronts and eddies are very dynamic in the mid-NA and can have significant interannual variations, their persistence around the south-west Irminger Sea and CGFZ has been well documented (Volkov & Oceanic, 2005). Due to persistent thermal fronts, these latest regions are well known for being extensively exploited by a wide range of marine top predators (Skov et al., 2008).

Differences in habitat preferences among whale species should partly reflect differences in the distribution of their main prey. In the NA, sei whales primarily feed on calanoid species, especially on *Calanus finmarchicus*, and to a lesser extent on euphausiids (Sigurjónsson & Víkingsson, 1997). Blue whales feed almost exclusively on krill (e.g. *Meganyctiphanes norvegica* and *Thysanoessa spp*) (Gavrilchuk et al., 2014; Víkingsson, 1990), and fin whales feed on zooplankton and a variety of small schooling fish (Sigurjónsson & Víkingsson, 1997). Therefore, the wider dietary breadth of fin whales likely contributes to higher plasticity in foraging behaviour and enables this species to exploit a broader set of environmental conditions across their migration route.

## 4.2 | Predicting spatio-temporal distribution across the NA

Unsurprisingly, our models showed that high latitudes were the preferred habitats for all three baleen whale species during summer months, in accordance with known seasonality and abundance patterns. However, our results also highlight that their predicted summer habitats are not totally overlapping due to distinct habitat preferences. The prediction maps for fin whales displayed the highest presence probabilities above 50°N, expanding from the south-west of the Irminger Sea to Iceland from March to July. This agrees with reports of a recent range expansion of fin whales in this region, from over the continental shelf to the deep waters of the Irminger Sea, possibly linked to the warming of the area (Víkingsson et al., 2015). During this period, fin whale abundance in the area between Iceland and eastern Greenland increased, making this the densest fin whale summer habitat in the NA (NAMMCO, 2017). These findings match our fin whale predictions, showing the importance of the Irminger Sea during the month of June and July. Lack of survey data in the region during spring months prevents us from validating our predictions for the period March–April. Our prediction also identified the mid-Atlantic Ridge as an important region for the species, with a higher probability of occurrence around 40°N between March and May and around 45°N in June and July. This agrees with Visser et al. (2011), who found an increase in fin whale sightings during spring months.

The predicted preferential habitat for blue whales consists of a narrow latitudinal band at mid-latitudes (36° to 50°N) in April that expands towards north-eastern latitudes during subsequent months. A systematic survey covering the mid-Atlantic Ridge, from Reykjanes Ridge to the north of Azores, in June 2004, only encountered blue whales south of the Faraday Fracture Zone (below 45°N) (MARECO survey, Waring et al., 2008). These findings agree with our predictions for the species during this month. In July, the preferred habitat for blue whales expanded north, showing a higher preference for the eastern NA. Passive acoustic monitoring off north-western Britain and Ireland (55°N 20°W) showed few blue whale detections in April and June, with

a gradual increase in mid-July until reaching a peak in November–December (Charif & Clark, 2009). In recent years, visual surveys revealed the presence of large baleen whales during these months in the area, including blue whales, mainly around the highly productive waters of the Porcupine Seabight and Rockall Trough (Baines & Reichelt, 2017; Wall et al., 2013; Wall, O’Kelly, Whooley, & Tyndall, 2009). Our predictions in northern latitudes, close to west Iceland, largely coincide with the known distribution of this species in summer, where a population of around 1,000 blue whales was estimated. In the Azores, the predicted shift towards northern latitudes closely matches the seasonality in the sighting data, which shows a decrease from April to July (Prieto et al., 2016). During these months, blue and fin whales have a similar environmental niche with a high degree of spatial coincidence in this region, strongly influenced by the higher productivity in the region (Prieto et al., 2016).

The highest habitat predictions for sei whales occurred in areas above 45°N, from May to August. These findings agree with previous reports of the known occurrence of feeding sei whales in these waters (Anonymous, 2009; Mitchell, 1975; Sigurjónsson et al., 1989). In the eastern NA, the deep waters of the Irminger Sea, between east Greenland and west of Iceland, appear to be an important habitat for the species, while the Faroe–Shetland Channel and Norwegian Sea are used less intensively (Cattanach, Sigurjónsson, Buckland, & Gunnlaugsson, 1993). In the western NA, sei whales seem to concentrate in the northern banks at 66°N and off the southern tip of Greenland, as well as along the continental shelf waters of the Labrador and Newfoundland coasts. Recently, Roberts et al., (2016) identified the continental slope of the east coast of United States, between 35° and 45°N, as an important summer habitat for the species. Sei whales were present from April to November, with the highest densities predicted in July around the Gulf of Maine and Nova Scotia shelf (latitude 45°N). Based on the models from Roberts et al. (2016), Mannocci et al., (2016) extrapolated the sei whale distribution to the whole western NA, highlighting the Labrador Sea, mainly the offshore waters of the Labrador Current, and the previously identified US waters as potential high-density areas. The summer habitat preference in that study closely matched our predictions, with a hotspot around the Labrador Sea and the deep waters north of the NA current (latitude 45°N).

Our monthly predictions for sei whales displayed a northward expansion of their habitat preference from May to August. In May, the preferred habitat for the species occurred around the CGFZ. In this region, the interaction of the NA current with the mid-Atlantic Ridge generates intense persistent fronts leading to predictable foraging grounds for marine top predators (Doksæter, Olsen, Nøttestad, & Ferno, 2008). A high concentration of sei whales reported just north of the CGFZ was associated with the fine-scale processes interacting in the upper 100 m of the water column (Skov et al., 2008; Waring et al., 2008). Higher zooplankton biomass, particularly of *Calanus finmarchicus*, was found in the northern region of the CGFZ compared to the southern region, which is believed to

be related with retention/advection processes concentrating large zooplankton over steep topography (Gaard et al., 2008). From June to August, sei whales were predicted over the same habitats in the deep waters of the Irminger Sea and Labrador Sea, including a few hotspots along the west Greenland. Laidre et al., (2010) found that krill biomass, mainly *Meganyctiphanes norvegica* and *Thysanoessa sp.*, was the most important variable explaining the presence of baleen whales in west Greenland. Although the number of sei whale sightings ( $n = 17$ ) was limited, sighting aggregations matched well the areas with the highest densities of krill. In Azores, sei whales are seen every year, with most sightings from spring to early summer (Silva, Prieto, Cascão, et al., 2013). However, sei whale presence has not been linked to local productivity. This matches well with the lack of association with LTL biomass and the low preference for habitats in these latitudes. Although the species feeds occasionally around Azores, our results indicate that the archipelago is not an important feeding ground for sei whales, at least during spring and summer.

### 4.3 | Management and conservation implications

Highly migratory species are exposed to increasing impacts from anthropogenic activities, which can lead to the alteration of their habitats, disruption of natural behaviours or increased mortality (e.g. from ship strikes) (Maxwell et al., 2013). In the mid-NA, a major international shipping route located mainly between 35° and 45°N overlaps with the predicted preferred habitat of the three baleen whale species. Specifically, blue whales have the highest proportion of the predicted habitat overlapping with this main route, potentially increasing the risk of ship collisions, as well as exposing them to increased noise levels. Alongside these chronic forms of noise pollution, the NA has also experienced an intense warming over the last decades, which is projected to continue over this century (IPCC, 2007). These physical changes are anticipated to lead to a poleward shift of several species of copepods (Villarino et al., 2015). As a consequence, baleen whales might move northwards to track the movements of their prey, similar to predicted habitat shifts in the North Pacific (Hazen et al., 2013). These shifts may decrease usage of middle latitude habitats by baleen whales. Moreover, species with a specialized diet, such as the blue whale, may have less capacity to adapt to these climate-induced changes compared to generalist predators, such as fin whales, which may be able to switch to other prey species (Pauly, Trites, Capuli, & Christensen, 1998). To effectively mitigate increasing pressures on baleen whales in the mid-NA, there is a need for accurate knowledge on the drivers that influence their distribution. Thus, findings from this study are relevant for national and international conservation policies (e.g. Marine Strategy Framework Directive; European Commission, 2018), designation of transatlantic marine protected areas and management initiatives (e.g. modify vessel operations; Silber et al., 2012).

## 5 | CONCLUSIONS

Our study provides the most detailed investigation into the spatio-temporal distribution of large baleen whales in the mid-NA. Our results demonstrate that the movements of all three baleen whale species towards and at high latitudes were closely linked to prey availability, although species exhibited distinct ecological niches. These results support previous work for these species, suggesting that baleen whales adjust their distribution during their migrations to ocean resources (Hazen et al., 2016). In our study, the habitat preferences of baleen whales were strongly related to the LTL biomass and production within the epipelagic layer, consistent with the diet of the species, and highlighting the strength of combining spatially explicit and vertically resolved prey information with oceanographic and telemetry data. Previous studies suggested the importance of including LTL biomass to develop climatological habitat models in the north-west NA using visual observations (Roberts et al., 2016). Our study confirms the relevance of this explanatory variable for modelling the distribution of cetaceans at the scale of the NA. Future research should focus on overcoming some of the limitations of this study to develop cetacean distribution maps in near real-time to provide a more accurate and efficient conservation tool.

### ACKNOWLEDGEMENTS

We are very grateful to Cláudia Oliveira, Irma Cascão, Maria João Cruz, Miriam Romagosa and many volunteers, skilled skippers, crew and spotters that participated in the tagging fieldwork. This work was supported by Fundação para a Ciência e Tecnologia (FCT), Azores 2020 Operational Programme and Fundo Regional da Ciência e Tecnologia (FRCT) through research projects FCT-Exploratory project (IF/00943/2013/CP1199/CT0001), TRACE (PTDC/MAR/74071/2006) and MAPCET (M2.1.2/F/012/2011) co-funded by FEDER, COMPETE, QREN, POPH, ESF, ERDF, Portuguese Ministry for Science and Education, and Proconvergencia Açores/EU Program. We also acknowledge funds provided by FCT to MARE, through the strategic project UID/MAR/04292/2013. SPJ was supported by a postdoctoral grant (REF.GREENUP/001-2016), MT by a DRCT doctoral grant (M3.1.a/F/028/2015), MAS by an FCT-Investigator contract (IF/00943/2013), FV by an FCT Investigator contract (CEECIND/03469/2017) and RP by an FCT postdoctoral grant (SFRH/BPD/108007/2015). LMTL modelling work has been supported by the CMEMS Service Evolution GREENUP project, funded by Mercator Ocean. We are grateful to Elliott Hazen for offering guidance and advice, and to two anonymous referees whose comments greatly improved this work.

### DATA AVAILABILITY STATEMENT

Tracking data are available through Movebank (<http://whales.scienceontheweb.net/index.php/resources/portfolio/whale-tracks-on-movebank>).

### ORCID

Sergi Pérez-Jorge  <https://orcid.org/0000-0002-4843-0443>

Marta Tobeña  <https://orcid.org/0000-0001-9976-7962>

Rui Prieto  <https://orcid.org/0000-0002-0354-2572>

Frederic Vandeperre  <https://orcid.org/0000-0002-3947-6917>

Mónica A. Silva  <https://orcid.org/0000-0002-2683-309X>

### REFERENCES

- Aarts, G., Fieberg, J., & Matthiopoulos, J. (2012). Comparative interpretation of count, presence-absence and point methods for species distribution models. *Methods in Ecology and Evolution*, 3(1), 177–187. <https://doi.org/10.1111/j.2041-210X.2011.00141.x>
- Aarts, G., MacKenzie, M., McConnell, B., Fedak, M., & Matthiopoulos, J. (2008). Estimating space-use and habitat preference from wildlife telemetry data. *Ecography*, 31(1), 140–160. <https://doi.org/10.1111/j.2007.0906-7590.05236.x>
- Anonymous. (2009). *NAMMCO Annual Report 2007–2008*. Tromsø, Norway.
- Baines, M., & Reichelt, M. (2017). An autumn aggregation of fin (*Balaenoptera physalus*) and blue whales (*B. musculus*) in the Porcupine Seabight, southwest of Ireland. *Deep Sea Research Part II: Topical Studies in Oceanography*, 141, 168–177. <https://doi.org/10.1016/j.dsr2.2017.03.007>
- Barbet-Massin, M., Jiguet, F., Hélène Albert, C., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, 3, 327–338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>
- Behrenfeld, M., & Falkowski, P. (1997). Photosynthetic rates derived from satellite based chlorophyll concentration. *Limnology and Oceanography*, 42(1), 1–20. <https://doi.org/10.4319/lo.1997.42.1.0001>
- Benoit-Bird, K. J., Battaile, B. C., Heppell, S. A., Hoover, B., Irons, D., Jones, N., ... Trites, A. W. (2013). Prey patch patterns predict habitat use by top marine predators with diverse foraging strategies. *PLoS ONE*, 8(1), <https://doi.org/10.1371/journal.pone.0053348>
- Bernard, B., Madec, G., Penduff, T., Molines, J.-M., Treguier, A.-M., Le Sommer, J., ... De Cuevas, B. (2006). Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. *Ocean Dynamics*, 56(5–6), 543–567. <https://doi.org/10.1007/s10236-006-0082-1>
- Bivand, R., Keitt, T., Rowlingson, B., Pebesma, E., Summer, M., Hijmans, R. J., ... Rundel, C. (2019). *rgdal: Bindings for the geospatial data abstraction library*. R Package Version 1.4-7. Retrieved from <http://CRAN.R-project.org/package=rgdal>
- Block, B. A., Jonsen, I. D., Jorgensen, S. J., Winship, A. J., Shaffer, S. A., Bograd, S. J., ... Costa, D. P. (2011). Tracking apex marine predator movements in a dynamic ocean. *Nature*, 475(7354), 86–90. <https://doi.org/10.1038/nature10082>
- Boyd, C., Castillo, R., Hunt, G. L., Punt, A. E., Vanblaricom, G. R., Weimerskirch, H., & Bertrand, S. (2015). Predictive modelling of habitat selection by marine predators with respect to the abundance and depth distribution of pelagic prey. *Journal of Animal Ecology*, 84(6), 1575–1588. <https://doi.org/10.1111/1365-2656.12409>
- Cattanach, K., Sigurjónsson, J., Buckland, S. T., & Gunnlaugsson, T. (1993). Sei whale abundance in the North Atlantic, estimated from NASS-87 and NASS-89 data. *Report of the International Whaling Commission*, 43, 315–321.
- Charif, R. A., & Clark, C. W. (2009). *Acoustic monitoring of large whales in deep waters north and west of the British Isles: 1996–2005* (Preliminary Report). Ithaca, NY: Cornell University.
- Conchon, A. (2016). *Modélisation du zooplancton et du micronecton marins*. La Rochelle, France: Sciences Agricoles, Université de La Rochelle.
- Croll, D., Marinovic, B., Benson, S., Chavez, F., Black, N., Ternullo, R., & Tershy, B. (2005). From wind to whales: Trophic links in a coastal

- upwelling system. *Marine Ecology Progress Series*, 289, 117–130. <https://doi.org/10.3354/meps289117>
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44, 73–96. <https://doi.org/10.2307/2531595>
- Doksæter, L., Olsen, E., Nøttestad, L., & Ferno, A. (2008). Distribution and feeding ecology of dolphins along the Mid-Atlantic Ridge between Iceland and the Azores. *Deep Sea Research Part II: Topical Studies in Oceanography*, 55, 243–253. <https://doi.org/10.1016/j.dsr2.2007.09.009>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carr, G., ... Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(2013), 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- European Commission. (2018). *Reporting on the 2018 update of articles 8, 9 & 10 for the Marine Strategy Framework Directive* (MSFD Guidance Document 14). Brussels, Belgium: DG Environment.
- Fieberg, J., Matthiopoulos, J., Hebblewhite, M., Boyce, M. S., & Frair, J. L. (2010). Correlation and studies of habitat selection: Problem, red herring or opportunity? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550), 2233–2244. <https://doi.org/10.1098/rstb.2010.0079>
- Gaard, E., Gislason, A., Falkenhaus, T., Søiland, H., Musaeva, E., Vereshchaka, A., & Vinogradov, G. (2008). Horizontal and vertical copepod distribution and abundance on the Mid-Atlantic Ridge in June 2004. *Deep Sea Research Part II: Topical Studies in Oceanography*, 55(1-2), 59–71. <https://doi.org/10.1016/j.dsr2.2007.09.012>
- Gavrillchuk, K., Lesage, V., Ramp, C., Sears, R., Bérubé, M., Bearhop, S., & Beuplet, G. (2014). Trophic niche partitioning among sympatric baleen whale species following the collapse of groundfish stocks in the Northwest Atlantic. *Marine Ecology Progress Series*, 497, 285–301. <https://doi.org/10.3354/meps10578>
- Goldbogen, J. A., Friedlaender, A. S., Calambokidis, J., Kenna, M. F. M., Simon, M., & Douglas, P. (2017). Integrative approaches to the study of baleen whale diving behavior, feeding performance, and foraging ecology. *BioScience*, 63(2), 90–100. <https://doi.org/10.1525/bio.2013.63.2.5>
- Grémillet, D., Lewis, S., Drapeau, L., Van Der Lingen, C. D., Huggett, J. A., Coetzee, J. C., ... Ryan, P. G. (2008). Spatial match – mismatch in the Benguela upwelling zone: Should we expect chlorophyll and sea-surface temperature to predict marine predator distributions? *Journal of Applied Ecology*, 45, 610–621. <https://doi.org/10.1111/j.1365-2664.2007.01447.x>
- Halpern, B. S., Frazier, M., Potapenko, J., Casey, K. S., Koenig, K., Longo, C., ... Walbridge, S. (2015). Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nature Communications*, 6(May), 1–7. <https://doi.org/10.1038/ncomms8615>
- Hazen, E. L., Jorgensen, S., Rykaczewski, R. R., Bograd, S. J., Foley, D. G., Jonsen, I. D., ... Block, B. A. (2013). Predicted habitat shifts of Pacific top predators in a changing climate. *Nature Climate Change*, 3(3), 234–238. <https://doi.org/10.1038/nclimate1686>
- Hazen, E. L., Palacios, D. M., Forney, K. A., Howell, E. A., Becker, E., Hoover, A. L., ... Bailey, H. (2016). 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>
- Hijmans, R. J., & van Ettern, J. (2012). *raster: Geographic analysis and modeling with raster data*. R Package Version 2.0-12. Retrieved from <http://CRAN.R-project.org/package=raster>
- IPCC. (2007). *Intergovernmental panel on climate change. Fourth assessment report. Climate change 2007: The physical science basis*. Geneva, Switzerland: Author.
- Jonsen, I. D., Basson, M., Bestley, S., Bravington, M. V., Patterson, T. A., Pedersen, M. W., ... Wotherspoon, S. J. (2013). State-space models for bio-loggers: A methodological road map. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 88–89, 34–46. <https://doi.org/10.1016/j.dsr2.2012.07.008>
- Jonsen, I. D., Flemming, J. M., & Myers, R. A. (2005). Robust state-space modeling of animal movement data. *Ecology*, 86(11), 2874–2880. <https://doi.org/10.1890/04-1852>
- Kareiva, P. M., & Shigesada, N. (1983). Analyzing insect movement as a correlated random walk. *Oecologia*, 56(2–3), 234–238. <https://doi.org/10.1007/BF00379695>
- Kaschner, K., Quick, N. J., Jewell, R., Williams, R., & Harris, C. M. (2012). Global coverage of cetacean line-transect surveys: Status quo, data gaps and future. *Challenges*, 7(9), e44075. <https://doi.org/10.1371/journal.pone.0044075>
- Laidre, K. L., Heide-Jørgensen, M. P., Heagerty, P., Cossio, A., Bergström, B., & Simon, M. (2010). Spatial associations between large baleen whales and their prey in West Greenland. *Marine Ecology Progress Series*, 402, 269–284. <https://doi.org/10.3354/meps08423>
- 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(8), e105958. <https://doi.org/10.1371/journal.pone.0105958>
- Lehodey, P., Conchon, A., Senina, I., Domokos, R., Calmettes, B., Jouanno, J., ... Kloser, R. (2015). Optimization of a micronekton model with acoustic data. *ICES Journal of Marine Science*, <https://doi.org/10.1093/icesjms/fsu233>
- Lehodey, P., Murtugudde, R., & Senina, I. (2010). Bridging the gap from ocean models to population dynamics of large marine predators: A model of mid-trophic functional groups. *Progress in Oceanography*, 84(1-2), 69–84. <https://doi.org/10.1016/j.pocean.2009.09.008>
- Macleod, K., Louise, B., Cañadas, A., Lens, S., Rogan, E., Santos, B., ... Rogan, E. (2009). Distribution and Abundance of Fin whales and other baleen whales in the Distribution and Abundance of Fin whales and other baleen whales in the European. *Atlantic*, IWC, SC/61/RMP10.
- Mannocci, L., Roberts, J. J., Miller, D. L., & Halpin, P. N. (2016). Extrapolating cetacean densities to quantitatively assess human impacts on populations in the high seas. *Conservation Biology*, 31(3), 601–614. <https://doi.org/10.1111/cobi.12856>
- Maxwell, S. M., Hazen, E. L., Bograd, S. J., Halpern, B. S., Breed, G. A., Nickel, B., ... Costa, D. P. (2013). Cumulative human impacts on marine predators. *Nature Communications*, 4(1). <https://doi.org/10.1038/ncomms3688>
- Mitchell, E. D. (1975). *Preliminary report on Nova Scotia fishery for sei whales (Balaenoptera borealis)*. Reports of the International Whaling Commission.
- NAMMCO. (2017). *Report of the Scientific Committee*. Tromsø, Norway.
- Pauly, D., Trites, A. W., Capuli, E., & Christensen, V. (1998). Diet composition and trophic levels of marine mammals. *ICES Journal of Marine Science*, 55, 467–481. <https://doi.org/10.1006/jmsc.1997.0280>
- Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications*, 19(1), 181–197. <https://doi.org/10.1890/07-2153.1>
- Pierce, D. (2019). *ncdf4: Interface to unidata netCDF (version 4 or earlier) format data files*. R Package Version 1.17. Retrieved from <http://CRAN.R-project.org/package=ncdf4>
- Prieto, R., Silva, M. A., Waring, G. T., & Gonçalves, J. M. A. (2014). Sei whale movements and behaviour in the North Atlantic inferred from satellite telemetry. *Endangered Species Research*, 26(2), 103–113. <https://doi.org/10.3354/esr00630>
- Prieto, R., Tobeña, M., & Silva, M. A. (2016). Habitat preferences of baleen whales in a mid-latitude habitat. *Deep Sea Research Part II: Topical*

- Studies in Oceanography*, 141, 155–167. <https://doi.org/10.1016/j.dsr2.2016.07.015>
- R Development Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- 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>
- Reisinger, R. R., Raymond, B., Hindell, M. A., Bester, M. N., Crawford, R. J. M., Davies, D., ... Pistorius, P. A. (2018). Habitat modelling of tracking data from multiple marine predators identifies important areas in the Southern Indian Ocean. *Diversity and Distributions*, 24, 535–550. <https://doi.org/10.1111/ddi.12702>
- Ritter, P. (1987). A vector-based slope and aspect generation algorithm. *Photogrammetric Engineering and Remote Sensing*, 53(8), 1109–1111.
- 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, 1–12. <https://doi.org/10.1038/srep22615>
- Scales, K. L., Schorr, G. S., Hazen, E. L., Bograd, S. J., Miller, P. I., Andrews, R. D., ... Falcone, E. A. (2017). Should I stay or should I go? Modelling year-round habitat suitability and drivers of residency for fin whales in the California Current. *Diversity and Distributions*, 23(10), 1204–1215. <https://doi.org/10.1111/ddi.12611>
- Sigurjónsson, J., Gunnlaugsson, T., & Payne, P. M. (1989). *NASS-87: Shipboard sighting surveys in Icelandic and adjacent waters Jun-July 1987*. Report of the International Whaling Commission.
- Sigurjónsson, J., & Víkingsson, G. (1997). Seasonal abundance of and estimated food consumption by cetaceans in Icelandic and adjacent waters. *Journal of Northwest Atlantic Fishery Science*, 22, 271–287. <https://doi.org/10.2960/J.v22.a20>
- Sikes, R. S., Gannon, W. L. & The Animal Care and Use Committee of the American Society of Mammalogists. (2011). Guidelines of the American Society of Mammalogists for the use of wild mammals in research. *Journal of Mammalogy*, 92(August), 235–253. <https://doi.org/10.1644/10-MAMM-F-355.1>
- Silber, G. K., Vanderlaan, A. S. M., Tejedor Arceredillo, A., Johnson, L., Taggart, C. T., Brown, M. W., ... Sagarminaga, R. (2012). The role of the International Maritime Organization in reducing vessel threat to whales: Process, options, action and effectiveness. *Marine Policy*, 36(6), 1221–1233. <https://doi.org/10.1016/j.marpol.2012.03.008>
- Silva, M. A., Prieto, R., Cascão, I., Seabra, M. I., Machete, M., Baumgartner, M. F., & Santos, R. S. (2013). Spatial and temporal distribution of cetaceans in the mid-Atlantic waters around the Azores. *Marine Biology Research*, 10(2), 123–137. <https://doi.org/10.1080/17451000.2013.793814>
- Silva, M. A., Prieto, R., Jonsen, I., Baumgartner, M. F., & Santos, R. S. (2013). North atlantic blue and fin whales suspend their spring migration to forage in middle latitudes: Building up energy reserves for the journey? *PLoS ONE*, 8(10), <https://doi.org/10.1371/journal.pone.0076507>
- Skov, H., Gunnlaugsson, T., Budgell, W. P., Horne, J., Nøttestad, L., Olsen, E., ... Waring, G. (2008). Small-scale spatial variability of sperm and sei whales in relation to oceanographic and topographic features along the Mid-Atlantic Ridge. *Deep Sea Research Part II: Topical Studies in Oceanography*, 55, 254–268. <https://doi.org/10.1016/j.dsr2.2007.09.020>
- Stern, S. (2009). Migration and movement patterns. In: W. F. Perrin, B. Würsig, & J. G. M. Thewissen (Eds.), *Encyclopedia of marine mammals* (2nd ed., pp. 726–730). London, UK: Academic Press.
- Tobeña, M., Prieto, R., Machete, M., & Silva, M. A. (2016). Modeling the potential distribution and richness of cetaceans in the azores from fisheries observer program data. *Frontiers in Marine Science*, 3(October). <https://doi.org/10.3389/fmars.2016.00202>.
- Vandeperre, F., Aires-da-Silva, A., Lennert-Cody, C., Serrão Santos, R., & Afonso, P. (2016). Essential pelagic habitat of juvenile blue shark (*Prionace glauca*) inferred from telemetry data. *Limnology and Oceanography*, 61(5), 1605–1625. <https://doi.org/10.1002/lno.10321>
- Víkingsson, G. A. (1990). Feeding of fin whales (*Balaenoptera physalus*) off Iceland – Diurnal and seasonal variation and possible rates. *Journal of Northwest Atlantic Fishery Science*, 22, 77–89.
- Víkingsson, G. A., Pike, D. G., Valdimarsson, H., Schleimer, A., Gunnlaugsson, T., Silva, T., ... Øien, N. (2015). Distribution, abundance, and feeding ecology of baleen whales in Icelandic waters: Have recent environmental changes had an effect? *Frontiers in Ecology and Evolution*, 3(February), 1–18. <https://doi.org/10.3389/fevo.2015.00006>
- Villarino, E., Chust, G., Licandro, P., Butenschön, M., Ibaibarriaga, L., Larrañaga, A., & Irigoien, X. (2015). Modelling the future biogeography of North Atlantic zooplankton communities in response to climate change. *Marine Ecology Progress Series*, 531, 121–142. <https://doi.org/10.3354/meps11299>.
- Visser, F., Hartman, K. L., Visser, F., Hartman, K. L., Pierce, G. J., & Valavanis, V. D. (2011). Timing of migratory baleen whales at the Azores in relation to the North Atlantic spring bloom. *Marine Ecology Progress Series*, 440, 267–279. <https://doi.org/10.3354/meps09349>
- Volkov, D. L., & Oceanic, N. (2005). Interannual variability of the altimetry-derived eddy field and surface circulation in the extratropical North Atlantic Ocean in 1993–2001. *Journal of Physical Oceanography*, 35(4), 405–426. <https://doi.org/10.1175/JPO2683.1>
- Wall, D., Murray, C., Brien, J. O., Kavanagh, L., Ryan, C., Glanville, B., ... Enlander, I. (2013). *Atlas of the distribution and relative abundance of marine mammals in Irish offshore waters: 2005–2011*. Kilrush, Ireland: Irish Whale and Dolphin Group.
- Wall, D., O'Kelly, I., Whooley, P., & Tyndall, P. (2009). New records of blue whales (*Balaenoptera musculus*) with evidence of possible feeding behaviour from the continental shelf slopes to the west of Ireland. *Marine Biodiversity Records*, 2(September 2008), 15–18. <https://doi.org/10.1017/S1755267209990443>
- Waring, G. T., Nøttestad, L., Olsen, E., Skov, H., & Víkingsson, G. (2008). Distribution and density estimates of cetaceans along the mid-Atlantic Ridge during summer 2004. *Journal of Cetacean Research and Management*, 10(2), 137–146.
- Willis-Norton, E., Hazen, E. L., Fossette, S., Shillinger, G., Rykaczewski, R. R., Foley, D. G., ... Bograd, S. J. (2015). Climate change impacts on leatherback turtle pelagic habitat in the southeast pacific. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 113, 260–267. <https://doi.org/10.1016/j.dsr2.2013.12.019>
- Wisz, M. S., & Guisan, A. (2009). Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, 9(1), 8. <https://doi.org/10.1186/1472-6785-9-8>
- Wood, S. (2006). *Generalized additive models: An introduction with R*. Boca Raton, FL: CRC Press.
- Wood, S. N., Scheipl, F., & Faraway, J. J. (2013). Straightforward intermediate rank tensor product smoothing in mixed models. *Statistics and Computing*, 23(3), 341–360. <https://doi.org/10.1007/s11222-012-9314-z>
- Worm, B., Sandow, M., Oschlies, A., Lotze, H., & Myers, R. (2005). Global patterns of predator diversity in the open oceans. *Science*, 309(5739), 1365–1369. <https://doi.org/10.1126/science.1113399>
- Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1), 3–14. <https://doi.org/10.1111/j.2041-210X.2009.00001.x>
- Żydelis, R., Lewison, R. L., Shaffer, S. A., Moore, J. E., Boustany, A. M., Roberts, J. J., ... Crowder, L. B. (2011). Dynamic habitat models: Using

telemetry data to project fisheries bycatch. *Proceedings of the Royal Society B: Biological Sciences*, 278(1722), 3191–3200. <https://doi.org/10.1098/rspb.2011.0330>

#### BIOSKETCH

**Sergi Pérez-Jorge** is a postdoctoral research fellow focused on understanding how marine predators are affected by environmental and prey-related variables at multiple spatial and temporal scales, as well as assessing their population dynamics. He is part of the Azores Whale Lab (<http://whales.scienceontheweb.net>), which has carried out cetacean research off the most remote archipelago in the North Atlantic for the last 20 years and applies that knowledge to aid in their conservation.

Author contributions: S.P.J., P.L. and M.A.S. conceived the ideas; R.P., M.T. and M.A.S collected the data; S.P.J., M.T., R.P., F.V., B.C., P.L. and M.A.S. analysed the data; and S.P.J. led the writing with contributions from all the authors.

#### SUPPORTING INFORMATION

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

**How to cite this article:** Pérez-Jorge S, Tobeña M, Prieto R, et al. Environmental drivers of large-scale movements of baleen whales in the mid-North Atlantic Ocean. *Divers Distrib.* 2020;00:1–16. <https://doi.org/10.1111/ddi.13038>