Deep-Sea Research I 91 (2014) 101–114

Contents lists available at ScienceDirect

Deep-Sea Research I

journal homepage: www.elsevier.com/locate/dsri

Predictive habitat modelling of humpback (*Megaptera novaeangliae*) and Antarctic minke (*Balaenoptera bonaerensis*) whales in the Southern Ocean as a planning tool for seismic surveys

Annette Bombosch[a,*], Daniel P. Zitterbart[a,b], Ilse Van Opzeeland[a], Stephan Frickenhaus[a], Elke Burkhardt[a], Mary S. Wisz[c,d], Olaf Boebel[a]

[a] *Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Am Alten Hafen 26, 27568 Bremerhaven, Germany*
[b] *Biophysics Group, Friedrich-Alexander University of Erlangen-Nuremberg, Henkestrasse 91, Erlangen, Germany*
[c] *Arctic Research Centre, Department of Bioscience, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark*
[d] *Department of Ecology and Environment, DHI, Agern Allé 5, 2970 Hørsholm, Denmark*

ARTICLE INFO

Article history:
Received 30 November 2013
Received in revised form
12 May 2014
Accepted 23 May 2014
Available online 12 June 2014

keywords:
Seismic survey
Ocean noise
Humpback whale (*Megaptera novaeangliae*)
Antarctic minke whale (*Balaenoptera bonaerensis*)
Species distribution
Habitat preference
Antarctic
Maxent

ABSTRACT

Seismic surveys are frequently a matter of concern regarding their potentially negative impacts on marine mammals. In the Southern Ocean, which provides a critical habitat for several endangered cetacean species, seismic research activities are undertaken at a circumpolar scale. In order to minimize impacts of these surveys, pre-cruise planning requires detailed, spatio-temporally resolved knowledge on the likelihood of encountering these species in the survey area. In this publication we present predictive habitat modelling as a potential tool to support decisions for survey planning. We associated opportunistic sightings (2005–2011) of humpback (*Megaptera novaeangliae*, $N=93$) and Antarctic minke whales (*Balaenoptera bonaerensis*, $N=139$) with a range of static and dynamic environmental variables. A maximum entropy algorithm (Maxent) was used to develop habitat models and to calculate daily basinwide/circumpolar prediction maps to evaluate how species-specific habitat conditions evolved throughout the spring and summer months. For both species, prediction maps revealed considerable changes in habitat suitability throughout the season. Suitable humpback whale habitat occurred predominantly in ice-free areas, expanding southwards with the retreating sea ice edge, whereas suitable Antarctic minke whale habitat was consistently predicted within sea ice covered areas. Daily, large-scale prediction maps provide a valuable tool to design layout and timing of seismic surveys as they allow the identification and consideration of potential spatio-temporal hotspots to minimize potential impacts of seismic surveys on Antarctic cetacean species.

1. Introduction

Growing concerns regarding contingently negative effects of anthropogenic marine noise emitted by seismic surveys provide the impetus for the implementation of mitigation measures aimed at curbing such impacts (Compton et al., 2008; Nowacek et al., 2013). While the risk of direct physical harm (e.g., barotrauma) can be mitigated through shut-down or ramp-up procedures when marine mammals are visually, acoustically (e.g., JNCC, 2004), or thermographically (Zitterbart et al., 2013) detected in a seismic vessel's environs, impacts mediated through adverse behavioural responses of marine mammals defy operational surveillance as they might occur at greater distances. Rather, surveys need to be designed in space and time as to minimize the chance of such impacts on sensible species, i.e. to minimize what is called e.g., ‘Level B harassment’ in the U.S. Marine Mammal Protection Act.

Such pre-cruise planning requires detailed, spatio-temporally resolved knowledge on the likelihood of encountering the species of concern in the study area. While such knowledge exists to some degree, particularly for highly frequented and hence well studied areas, such as the Gulf of Mexico (e.g., Baumgartner et al., 2001; Best et al., 2012) or the North-East Atlantic (e.g., Booth et al., 2013; MacLeod et al., 2007), it becomes increasingly sparse towards the poles. In the Antarctic, data collection is hampered by its limited (seasonal) accessibility and logistic constraints. So far, the most ambitious dedicated cetacean sighting programme, providing large-scale information about cetacean distributions and habitat

---

* Corresponding author.
E-mail addresses: annette.bombosch@gmail.com (A. Bombosch), Olaf.Boebel@awi.de (O. Boebel).

http://dx.doi.org/10.1016/j.dsr.2014.05.017

preferences in the Southern Ocean, was conducted by the International Whaling Commission (IWC). It is known as the International Decade of Cetacean Research (IDCR) and its follow-up programme the Southern Ocean Whale and Ecosystem Research (SOWER) (1978/79–2008/09), circumnavigating the Southern Ocean three times. Additionally, more local cetacean surveys are conducted by various nations as part of multidisciplinary research cruises (e.g., Friedlaender et al., 2006; Širović and Hildebrand, 2011; Thiele et al., 2000). However, despite these efforts, due to the Southern Ocean's vastness and coverage by sea ice, many areas have received only little survey effort to date.

Seismic surveys in the Antarctic occur solely in the context of academic research, investigating topics related to e.g., plate tectonics or the palaeoceanographic and climate history of the Southern Ocean (Breitzke, 2014). As geophysical research occurs in the Antarctic at a circumpolar scale and throughout the austral summer, knowledge (and tools) are required to predict the likelihood of encounters with marine mammals at similar scales. To be able to do justice to both, species protection and scientific needs, spatio-temporally resolved, reliable and up-to-date information of the likely distribution of marine mammals in the Southern Ocean is needed.

Such information can be obtained from habitat models, which investigate cetacean-habitat preferences by quantifying relationships between a species and its environment. Their strength is to allow predictions of habitat suitability for environmentally sampled, yet geographically unsampled regions (Guisan and Zimmermann, 2000; Redfern et al., 2006). In order to obtain meaningful models and predictions, a major prerequisite of the input data is that they have been acquired across the Southern Ocean's different environmental strata. With IWC efforts having been largely restricted to open water areas and current surveys being conducted on small- to meso-scales, these datasets fall short of meeting this requirement. Instead, we utilized the widely distributed tracks of the German Antarctic research icebreaker RV *Polarstern*, which regularly operates in the Southern Ocean and which logistic and research duties result in a quite extensive coverage of the Southern Ocean's different environmental strata.

In this study we developed habitat models and calculated predictions maps as a tool to support risk assessment and aid decisions of seismic survey planning to minimize the impact of anthropogenic noise on marine mammals. We correlated opportunistic sightings of humpback (*Megaptera novaeangliae*) and Antarctic minke whales (*Balaenoptera bonaerensis*) obtained between 2005 and 2011 with a comprehensive set of remotely sensed environmental variables on a circumpolar scale at daily resolution. We predicted daily habitat suitability from November to April to investigate how habitat suitability for both species evolved throughout the season. Model robustness and reliability were assessed by comparing model results to the current knowledge of humpback and Antarctic minke whale ecology and distribution.

2. Materials and methods

2.1. Cetacean sightings data

Opportunistic cetacean sightings were collected during 14 multidisciplinary research expeditions of RV *Polarstern* to the Southern Ocean (defined here as waters south of the Polar Front) from January 2005 to January 2011. The nautical officers systematically logged cetacean encounters using a customized logging software called 'Walog' (Burkhardt, 2009a). The data comprise time and location (approximated by the ship's geographical position) of cetacean sightings (i.e. presence-only data), identified to the lowest taxonomic level possible (ranging from 'unidentified whale' to species level), the associated certainty level in the species identification (i.e. 'possible', 'probable', 'definite'), as well as the number of animals sighted. The spatial separation required for animals to be considered to belong to different groups was left to the discretion of the observer. All sightings are stored in the publicly accessible database PANGAEA (http://www.pangaea.de/, e.g., Burkhardt, 2009b). For modelling, we used humpback and Antarctic minke whale sightings with a certainty level 'definite' or 'probable' to minimize false positives due to potential species misidentification. Furthermore, as required by Maxent, each cetacean sighting was treated as a single presence record, independent of the number of animals sighted.

2.2. Environmental data

To study circumpolar/basinwide habitat preferences of humpback and Antarctic minke whales we relied on remote sensing datasets, which were selected based on their potential ecological relevance in influencing each species' distribution (Table 1).

Depth was obtained from the Earth Topography Digital Dataset 1 Arc-Minute Global Relief Model (ETOPO1) as 'grid-centred Bedrock' version (Amante and Eatkins, 2009). Daily **Sea Surface Temperature** (SST), measured by the Advanced Very-High Resolution Radiometer (AVHRR) instrument on board the three NOAA TIROS-N series of polar-orbiting satellites (NOAA-11 to NOAA-18) were obtained from the Modular Ocean Data Assimilation System (MODAS) at a resolution of $1/8^\circ \times 1/8^\circ$ (Barron and Kara, 2006). Daily, absolute **Sea Surface Height** (SSH) data, based on altimetry instruments on board of up to four satellites (Jason-1, Topex/Poseidon, Envisat and GFO) were downloaded as 'updated delayed time products of absolute dynamic topography' on a $1/3^\circ \times 1/3^\circ$ Mercator grid from Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO). Daily **sea ice concentration** values were derived from the Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E) instrument on board the Aqua satellite with a spatial resolution of 6.25 km × 6.25 km (Kaleschke et al., 2001; Spreen et al., 2008). Daily **chlorophyll-a** (chl-a) concentration values (Level 3 Standard Mapped Image products) with a spatial resolution of $0.83^\circ \times 0.83^\circ$ were derived from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS). However, due to large data gaps chl-a was excluded from the list of candidate environmental variables (see Section 2.3).

Based on these environmental variables several additional parameters were derived. Bathymetric **slope** as well as **spatial gradients** for SST and SSH were derived from depth, SST and SSH, respectively, using the *gradientm* function in Matlab®. **Sea ice variance** was obtained from sea ice concentration by calculating the variance of sea ice concentration over the 14 days prior to the cetacean sighting. **Time-lagged sea ice concentrations** were derived by using the sea ice concentration at the 7th and 14th day previous to the cetacean sighting. The **distance to the sea ice edge**, defined at 15% sea ice concentration (Tynan and Thiele, 2003) was calculated as an 'effective' distance d_{eff} to the sea ice edge, which is based on the weighted average of distances to a subset of suitably selected points from the sea ice edge. This returns an ecologically relevant average distance to the surrounding, possibly fractured, sea ice edge (see Supplementary Appendix A for calculation of d_{eff}). Our approach differs from the frequently calculated distance between the cetacean sighting and the nearest point of the sea ice edge (e.g., Beekmans et al., 2010; Friedlaender et al., 2011), which might result in erratic measures when the cetacean sighting occurs in the vicinity of small ice floes (with ice concentration values $<15\%$) while the overall sea ice edge is located at a larger distance. **Length of day** was calculated as the amount of minutes that the angle of the sun was $\geq -6^\circ$ (according to the definition of civil twilight).

2.3. Habitat modelling

Maxent, a maximum entropy algorithm, which was specifically developed for the use with presence-only data (Phillips et al., 2006, 2004), was used to develop independent habitat models for humpback and Antarctic minke whales. Maxent is widely applied in the habitat modelling community and has seen over 1000 applications so far (Merow et al., 2013). It offers the unique advantage of combining presence-only data with a powerful model algorithm (Elith et al., 2006; Wisz et al., 2008) and shows excellent predictive ability even when sample size is small (Hernandez et al., 2006; Wisz et al., 2008). Maxent furthermore allows accounting for potential sampling biases inherent in the sightings data by selecting background data with the same bias as prevalent in the presence data (Phillips et al., 2009).

Maxent estimates, based on the location of cetacean sightings (presence-only data) and a set of associated environmental data (background data), the probability distribution of species occurrence by finding the probability distribution of maximum entropy, i.e. the distribution that is closest to uniform across the study area. This probability distribution is subject to a set of constraints, which represent the available yet incomplete information about the target distribution (Phillips et al., 2006). These constraints are governed by the distribution of the environmental variable values at the sighting locations and are expressed in terms of simple functions known as features, which can be based on continuous and categorical environmental data as well as on interactions between different environmental variables (Phillips et al., 2006; Phillips and Dudík, 2008). We used Maxent version 3.3.3k for modelling, which offers 6 feature classes: linear, quadratic, product, hinge, threshold and category indicator (http://www.cs.princeton.edu/~schapire/maxent/).

To compensate for potential regional and strata biases in the sightings data, background data were selected randomly from all cruise tracks covered by RV *Polarstern* from January 2005 to January 2011, i.e. the period during which cetacean sightings were collected. Cruise tracks, along which cetacean sightings were recorded, were determined by multidisciplinary research programmes, rather than by survey designs optimized for cetacean surveys. The coverage of the study area was therefore spatially biased with preferences for open water as well as a frequently repeated transect along the Greenwich meridian and across the Weddell Sea (Fig. 1). By selecting background data based on the cruise tracks, the background data were subject to the same spatial bias as the sightings data, a method that has been shown to considerably improve model performance (Phillips et al., 2009).

Each cetacean sighting and background data point was subsequently merged with the corresponding set of candidate environmental variables (Table 1) by selecting the geographically closest environmental variable value for the same day as the cetacean sighting or background data point. This resulted in two ‘samples with data’ (SWD) datasets, which were subsequently used as input files for Maxent. If a daily file was missing, the respective environmental variable values of the temporally closest available previous and subsequent days (with a maximum of six previous/subsequent days) were linearly interpolated and used instead. Despite interpolation, chl-a values could only be assigned to a small number of cetacean sightings, which proved insufficient to develop habitat models. Therefore chl-a was excluded as a candidate environmental variable. SSH values were unavailable for sea ice covered areas, resulting in SSH values missing for several grid points in the southern part of the study area. These points were assigned an average SSH value, based on all available SSH values (1992–2011). This appeared a sensible approximation, as only small seasonal SSH variability is to be expected in the far South (except for the region of the coastal current).

For each SWD dataset, an exploratory data analysis was conducted prior to modelling to obtain a first overview of the

Table 1
Candidate environmental variables.

Environmental variable	Data source	Resolution spatial/temporal	Unit	% missing/interpolated values
Environmental variables (data source=data provider)				
Depth	National Geophysical Data Center (NGDC), National Oceanic and Atmospheric Administration (NOAA) http://www.ngdc.noaa.gov/mgg/global/global.html	1/60° Static	m	0
Sea surface temperature (SST)	Modular Ocean Data Assimilation System (MODAS), United States Naval Research Laboratory (NRL) http://www7320.nrlssc.navy.mil/modas	1/8° Daily	°C	0
Sea surface height (SSH)	Archiving, Validation & Interpretation of Satellite Oceanographic data (AVISO) http://www.aviso.oceanobs.com/en	1/3° Mercator grid daily	cm	16
Sea ice concentration	Integrated Climate Data Center (ICDC), University of Hamburg http://icdc.zmaw.de/cryosphere.html?&L=1	6.25 km polar stereographic grid daily	%	0
Chlorophyll-a concentration[a]	National Aeronautics and Space Administration (NASA) Goddard Space Flight Center's Ocean Data Processing System (ODPS) http://oceancolor.gsfc.nasa.gov	0.083° Daily	mg/m^3	92.67
Length of day	Matlab sun_position function www.mathworks.com/matlabcentral/fileexchange	Daily	minutes	0
Derived environmental variables (data source=respective environmental variable above)				
Slope	Depth	1/60° Static	deg	0
SST gradient	SST	1/8° Daily	mK/m	0
SSH gradient	SSH	1/3° Mercator grid daily	mm/m	18.53
Sea ice variance	Sea ice concentration	6.25 km polar stereographic grid daily	$\%^2$	9.05
Time-lagged sea ice concentration	Sea ice concentration	6.25 km polar stereographic grid daily	%	1.29
Effective distance to the sea ice edge	Sea ice concentration	6.25 km polar stereographic grid daily	km	0

[a] Due to large data gaps chl-a was excluded from the final list of candidate environmental variables.

distribution of the sightings data across the range of each of the candidate environmental variables and to test for potential outliers (Supplementary Appendix B, Table B.1). Only one Antarctic minke whale sighting deviated substantially in 5 of the associated environmental variable values (> 3-fold the SD) compared to the other sightings for this species and was therefore excluded from further analyses. After data processing the final cetacean dataset used for modelling comprised 93 humpback and 139 Antarctic minke whale sightings (Fig. 1). Additionally, we calculated correlation coefficients for the candidate environmental variables for each SWD dataset. As correlation coefficients did not exceed a value of 0.6, all variables were included in the Maxent model runs.

2.3.1. *Model runs and settings*

Maxent runs were based on hinge features only, which have been shown to provide more succinct approximations of the species' true distribution probability, exhibit high predictive ability, while not significantly increasing model complexity (Phillips and Dudík, 2008).

Response curves for both humpback and Antarctic minke whale models were built using default settings for the regularization parameters and were visually inspected for signs of overfitting, which would express itself in complex or ecologically unrealistic structures of the response curves. Overfitted models, while predicting training data well, generally fail in predicting independent data and are hence less transferable to unsampled areas of the study site. The response curves of neither model showed signs of overfitting, suggesting that the default settings might be applied expediently.

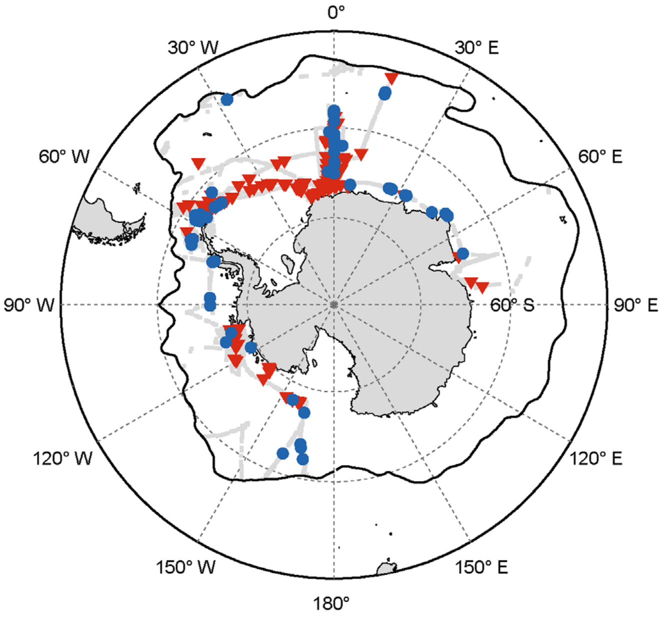

Fig. 1. Humpback (blue circle) and Antarctic minke (red triangle) whale sightings from 2005–2011 as used for Maxent modelling. The respective background dataset used for the models is depicted in light grey. The black line marks the climatological mean position of the Polar Front (Harris and Orsi, 2001, updated 2008).

To assess the relative importance of the candidate environmental variables within each Maxent model, jackknife tests of environmental variable importance were conducted. Results were assessed on the basis of AUC values (area under receiver operating characteristic (ROC) curve) (see Fielding and Bell, 1997) calculated for test data. This test provides an indication of how well the models perform when an environmental variables is omitted and additionally how each variable contributes to the model individually. Based on this test we selected the set of 'most relevant environmental variables' for each model. All environmental variables that, when omitted, resulted in an increased AUC value were excluded from subsequent model runs. Starting with the variable which, when omitted, yielded the largest increase in the AUC value, this procedure was repeated until the AUC value could not be increased any further.

2.3.2. *Spatial prediction maps*

We calculated spatial prediction maps of habitat suitability based on Maxent's logistic output, which depicts habitat suitability across the study area with values ranging from 0 to 1, whereby values are scaled such that a value of 0.5 corresponds to sites exhibiting typical conditions for the species (see Elith et al., 2011; Phillips and Dudík, 2008 for further details). We resampled all environmental variables to a consistent spatial resolution of 0.25° for the prediction maps. For humpback whales, circumpolar prediction maps were calculated from November to April on a daily basis for the years 2005 through 2011 (only until January in 2011 due to availability limitations of SSH values), i.e. the same time frame as the sightings data. As patterns of habitat suitability changed steadily throughout the season, the 1st and 15th of each month were selected to provide a representative overview of seasonal change in habitat suitability. For Antarctic minke whales the high computational load caused by the environmental variable d_{eff} (chosen only by the Antarctic minke whale model, see Table 2) restricted the calculation of the prediction maps to 60° W–60° E and the 1st and 15th of each month.

2.3.3. *Intrinsic model evaluation*

To assess model performance, sightings data were split into training and test datasets based on n -fold cross-validation ($n=9$ for humpback whales and $n=13$ for Antarctic minke whales). As the aim of this study was to create large-scale prediction maps, we projected habitat suitability also to unsampled sites of the study area. When transferring model results to novel geographic areas, model predictions can become uncertain when

Table 2
Environmental variables selected by Maxent models for humpback and Antarctic minke whales. Permutation importance of the single environmental variables is given, with variables contributing most to each model highlighted in bold.

Species	Number of sightings	Mean AUC ± SD	Number of cross-validation folds	Environmental variables (permutation importance)								
				Length of day	Depth	Slope	Ice conc[a]	Ice ret14[b]	d_{eff}[c]	SSH	SST	SST gradient
Humpback whale	93	0.844 ± 0.062	9	**38.7%**	8.5%	6.9%	2.3%	**10.9%**	–	**19.6%**	–	**13%**
Antarctic minke whale	139	0.715 ± 0.072	13	2.2%	**18.4%**	6.7%	–	–	**39.6%**	9.1%	**24%**	–

[a] Ice conc = sea ice concentration.
[b] Ice ret14 = 14-day time-lagged sea ice concentration.
[c] d_{eff} = effective distance to sea ice edge.

environmental conditions at these sites are outside the model's training range. We used Maxent's MESS (multivariate similarity surfaces) (Elith et al., 2010) and MoD (most dissimilar variable) map functions to test how similar the environmental conditions of the prediction area are to the those encountered during model training (MESS) and which environmental variable deviates most at any given point of the prediction map (MoD).

2.3.4. Model evaluation using independent data

Models were further assessed using independent data of humpback and Antarctic minke whale distributions taken from catch records (Allison, 2013; Zemsky et al., 1995) and the IWC's IDCR/SOWER sighting surveys. Both datasets are available at a circumpolar and long-term scale (catch records: humpback whales (1929–1967, November–April), Antarctic minke whales (1955–2012, November–April); IWC sightings (1978–2009, January–February)).

3. Results

3.1. Maxent models and spatial prediction maps

Spatial prediction maps calculated for humpback and Antarctic minke whales from 2005 through 2011 show that spatio-temporal distribution patterns for both species are largely consistent between the different years (Supplementary Appendix C, Figs. C.1 and C.2). Some inter-annual variability in the distribution of favourable habitats is observed, e.g., for humpback whales near Maud Rise (0–30° E) – a well-known area of early polynia formation – between 15 December 2007 and 2010 (Fig. 2). Inter-annual variability in the prediction maps is driven by variations in the spatio-temporal interaction of the static and dynamic environmental variables selected by each model between the different years. For humpback whales, the maximum spatial extent of habitat suitability occurs in January, for Antarctic minke whales in November. On an intra-annual scale, however, – from the spring to the late summer months – the distribution of suitable habitat conditions changes considerably for both cetacean species. Here we use the IWC's division of the Southern Ocean into 6 management areas to facilitate the discussion of spatial patterns of predicted habitat suitability: Area I: 120° W–60° W, Area II: 60° W–0°, Area III: 0°–70° E, Area IV: 70° E–130° E, Area V: 130° E–170° W, Area VI: 170° W–120° W (Donovan, 1991).

3.1.1. Humpback whales

Fig. 3 (left) depicts circumpolar spatial prediction maps for humpback whale habitat suitability for the 1st and 15th of each month exemplarily from November 2006 to April 2007. Beginning in November and persisting throughout much of the season, the model indicates 5 patches of higher habitat suitability around 30° E, 90° E, 155° E, 135° E, 60° W. Early in the season, these patches are spatially restricted to a relatively narrow latitudinal band. As the season progresses, areas with favourable conditions form a circumpolar band expanding southward towards the Antarctic continent with an interruption in the Amundsen and Bellingshausen Seas (IWC Area I) for much of the season. This local ‘interruption’ is a result of relatively unfavourable habitat conditions in the western part of IWC Area I (120–~75° W). Here habitat conditions for humpback whales become favourable only for a short period from mid-January to mid-February, covering a narrow latitudinal band.

Habitat suitability reaches its maximal spatial extent around January 15. Throughout February, most favourable conditions contract into a fairly narrow band surrounding the Antarctic continent. An exception is IWC Area VI where habitat suitability spreads to its maximal spatial extent covering areas from the Polar Front to the Antarctic coast. By mid-February, overall habitat suitability starts to decrease rapidly throughout the study area. From the beginning of March, habitat suitability is low throughout the Southern Ocean and decreases further until the end of the season.

Areas remaining relatively unsuitable throughout the season occur close to the coast in IWC Areas I and VI, as well as in the inner Weddell and Ross Seas between 150° E and 170° E. Areas that consistently indicate very high habitat suitability from November to January/February, occur in IWC Area II around the South Sandwich Islands, near 30° E and 130° W.

3.1.2. Antarctic minke whales

Fig. 3 (right) depicts spatial prediction maps for Antarctic minke whale habitat suitability from 60° W–60°E exemplarily for the 1st and 15th of November 2006 to April 2007. Throughout November and early December, favourable conditions for Antarctic minke whales extend over a wide latitudinal range. Patches of higher suitability occur in IWC Area II westwards of the South Sandwich Trench and in the eastern part of IWC Area III. The maximum spatial extent occurs at the beginning of November. Highly favourable conditions are also predicted for a region around 70° S (next to an area of missing values indicated in grey) and for highly localized, scattered single grid cells that persist throughout the season, which appear to be model artefacts.

By mid-December, habitat suitability starts to decrease rapidly in latitudinal extent and concentrates towards coastal areas for the remaining months of the season. Despite the decrease in latitudinal extent, for all months favourable habitat conditions exist for a broader latitudinal range for IWC Area II than for IWC Area III. By mid-March, habitat suitability reaches is spatial minimum and starts extending in latitudinal range further north until mid-April.

An area of distinctly low suitability throughout the season occurs around the South Sandwich Islands.

4. Discussion

4.1. Reliability of model predictions

The results of this study provide large-scale predicted patterns of habitat suitability for humpback and Antarctic minke whales, by ‘extrapolating’ knowledge of sightings from sampled to unsampled areas. Through habitat modelling, this apparent extrapolation is the result of an interpolation in environmental space followed by a projection from environmental to geographic space. Hence, while the Maxent habitat models make predictions for areas outside the spatially sampled area, these predictions are educated by the knowledge of environmental conditions at the sighting locations, at least as long as they are within the environmental space sampled (see discussion on the use of MESS and MoD plots below). By operating in environmental space, the habitat models make use of the underlying assumption that a species' distribution likely is the result of the species' specific response to its environment (Barry and Elith, 2006, Elith and Leathwick, 2009). Therefore, the most reliable model predictions are to be expected for those regions that are (a) within the range of the overall known distribution of the respective species, (b) associated with previously sampled environmental spaces, and (c) that represent areas with the same ecological function. Our spatial projections would for example fail should we use our humpback whale model (which is trained on data from high latitude feeding areas) to project habitat suitability in the lower latitude breeding areas, as these are not comprised in the environmental space sampled (e.g., much higher sea surface temperatures, lower depth values) and are also likely to differ in terms of species-specific responses to the environment.

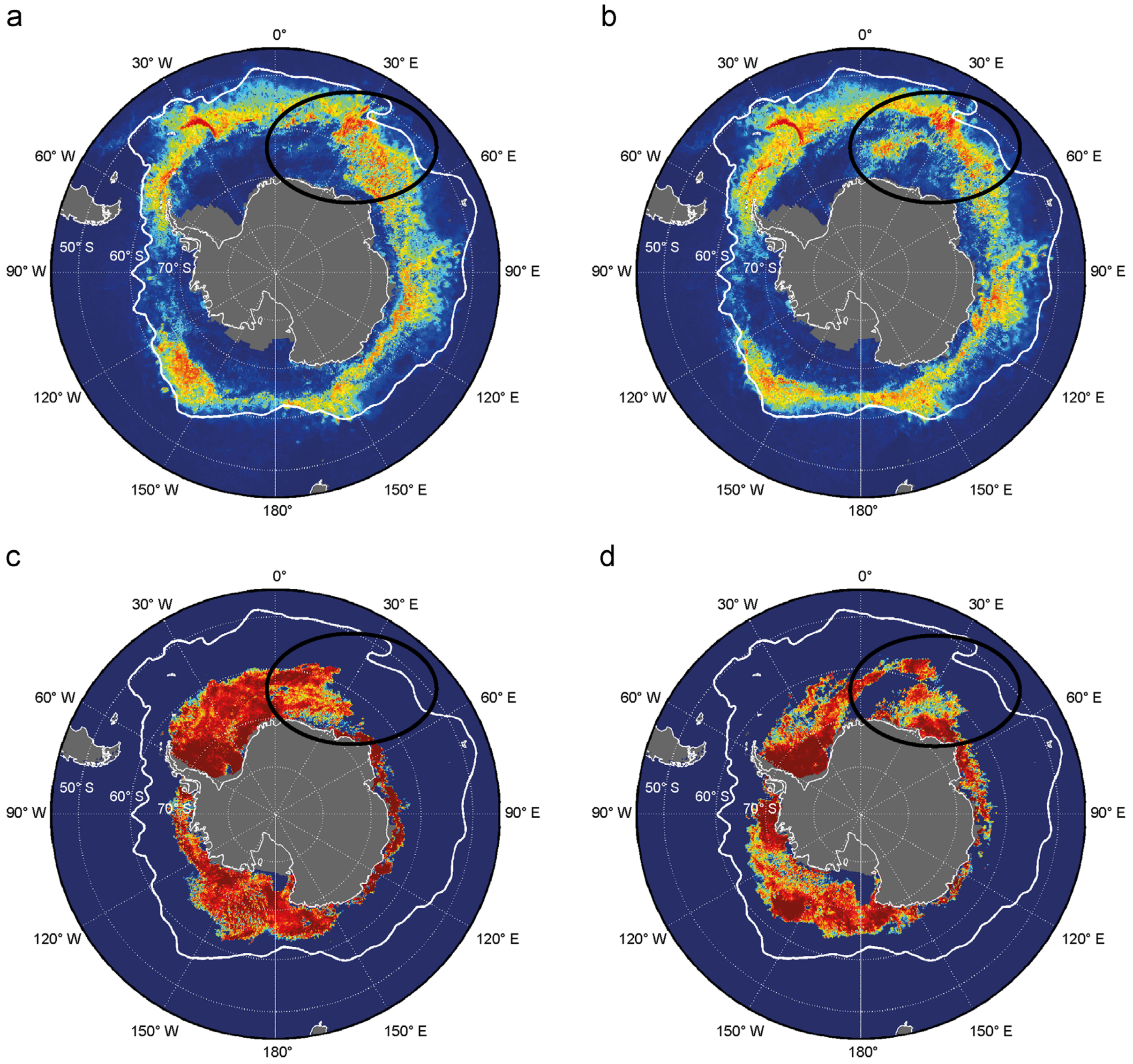

Fig. 2. Example of inter-annual variability in spatial predictions for humpback whales (a and b) and environmental variables (here sea ice concentration, c and d) for 15 December 2007 (a and c) and 2010 (b and d). The white line indicates the Polar Front (Harris and Orsi, 2001, updated 2008).

Humpback whales occur throughout the Southern Ocean and based on what is known on their behaviour and habitat usage, there is no indication that differences exist in feeding ecology between stocks or regions of the Southern Ocean. Our humpback habitat suitability maps therefore describe how habitats (as defined through Maxent via the environmental parameters), where humpback whales have been observed, are distributed throughout the entire Southern Ocean. The Antarctic minke whale model is per-se constrained in geographic space to more or less the Weddell Gyre (where also the majority of sightings occurred) and consequently also in the ecology it represents.

Nevertheless, in order to evaluate the models usefulness as a pre-cruise planning tool for seismic surveys, the reliability of model projections to unsampled parts of the study area needs to be assessed. To this end, internal consistency is analysed and comparisons of habitat suitability maps with independent catch records and IWC sightings are conducted below.

4.1.1. *Internal model consistency checks*

A first check considers the standard deviation of the average AUC values calculated for the *n* replicate cross-validation runs for each species' habitat model. In this study, the standard deviation is small: 0.844 ± 0.062 for the humpback whale model, 0.715 ± 0.072 for the Antarctic minke whale model (Table 2). This indicates consistency between folds and hence reliability of the models as such.

Additionally, we derived the relative error calculated as the ratio of the standard deviation of the Maxent prediction maps to the mean habitat suitability value (based on the *n*-fold cross-validation). For both species the relative error is small (22.4% for humpback and 15.3% for Antarctic minke whales). During the peak season (Nov to early March) when most cetacean sightings were collected, the relative error is smaller than 25% for the majority of the study area ($>70\%$ for humpback and for Antarctic minke whales, see Supplementary Appendix D, Fig. D.1), whereas the relative error increases towards the end of the season. It is interesting to note, that the largest relative error generally occurs in areas showing habitat suitability values lower than average conditions (see Fig. 3 and Supplementary Appendix D, Figs. D.2 and D.3), indicating that predictions in suitable areas are of greater robustness.

As model predictions to unsampled geographic areas may become unreliable when the environmental conditions at these sites are outside the model's training range (referred to as novel

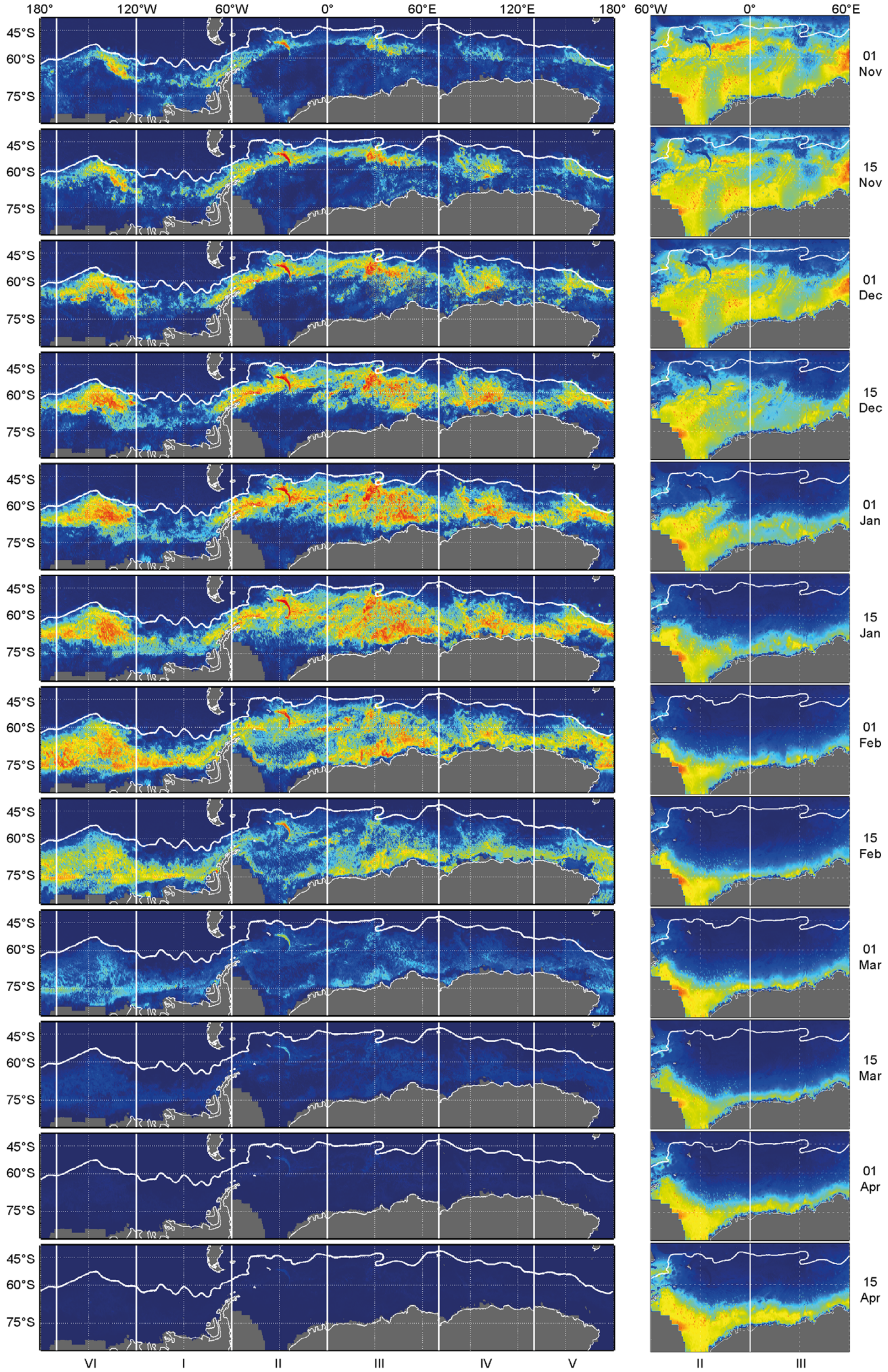

Fig. 3. Maxent spatial prediction maps for humpback (left) and Antarctic minke whales (right) for the 1st and 15th of each month from November 2006 to April 2007. Habitat suitability is colour-coded with blue colours indicating less suitable to unsuitable habitat, greenish colours depicting 'typical' conditions and red colours indicating more suitable to highly suitable habitat conditions. Typical habitat conditions are model specific, i.e. for each species Maxent will only show relative suitability in relation to the species' typical conditions. Relative suitability values are therefore not directly comparable between species. The white line represents the climatological mean position of the Polar Front (Harris and Orsi, 2001, updated 2008). Grey areas indicate land areas or regions for which values for one of the environmental variables are missing. The 6 IWC management areas are indicated by the solid white lines (Donovan, 1991).

environmental conditions), we assessed each model's reliability based on Maxent's MESS and MoD map functions (Fig. 4). For both humpback and Antarctic minke whales the MESS maps reveal that the environmental variables' ranges encountered throughout the unsampled parts of the study area by and large overlap with the training range. Exceptions are the area of the South Sandwich Trench, where depth values exceed the sampled depth range and in the southern part of the study area around mid-April when length of day values are smaller compared to those encountered during model training.

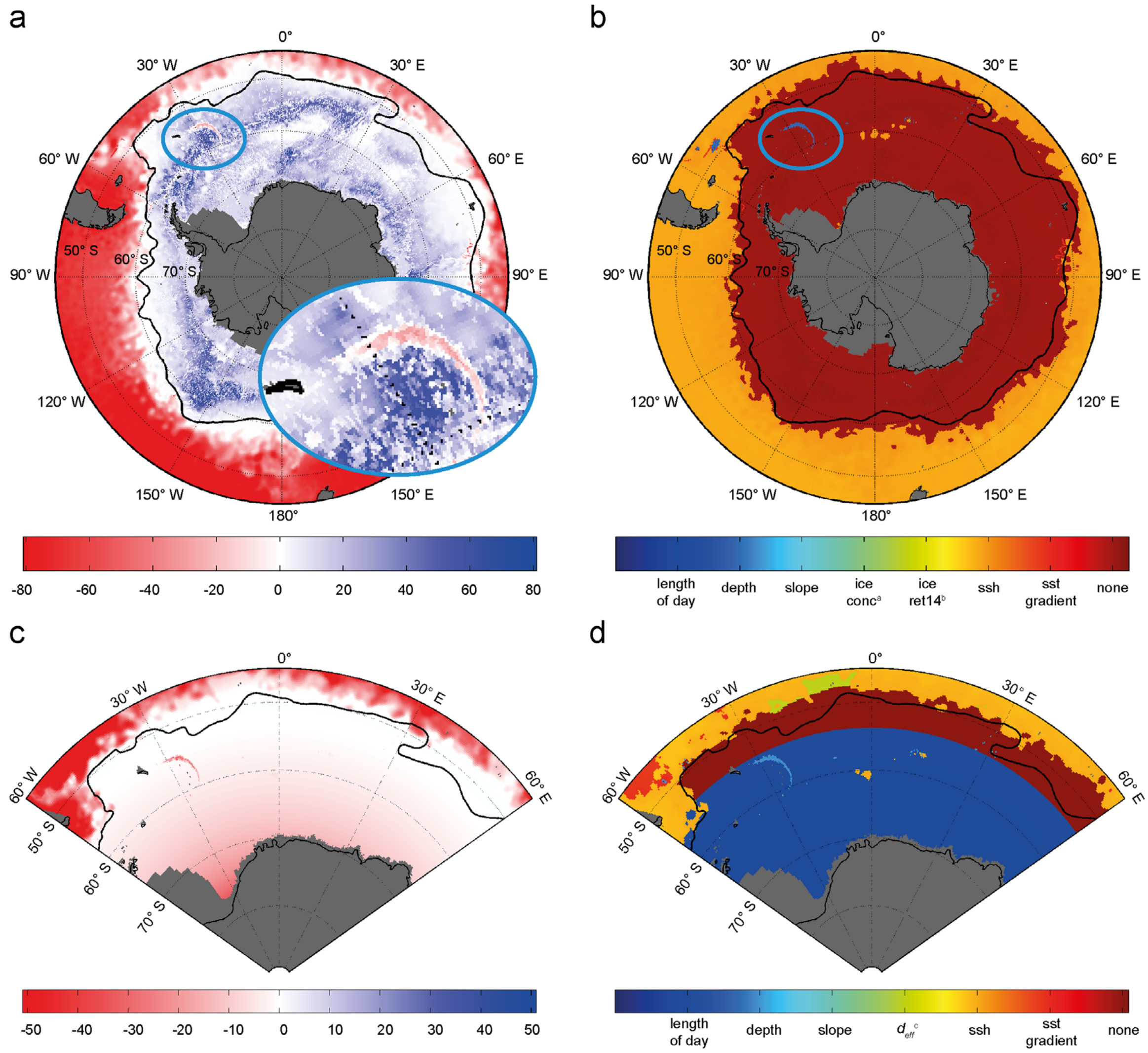

Fig. 4. MESS (multivariate similarity surface, (a and c)) and MoD (most dissimilar variable, (b and d)) maps for humpback whales for 15 November 2006 (a and b) and for Antarctic minke whales for 15 April 2007 (c and d). Areas within the study area for which novel environmental conditions are encountered appear in red in the MESS map. The respective 'out-of-range' environmental variable is indicated in the MoD map on the right. Top left inset (a): magnification of the South Sandwich Island region. The black line indicates the Polar Front (Harris and Orsi, 2001, updated 2008). ^{a}ice conc = sea ice concentration; ^{b}ice ret14 = 14-day time-lagged sea ice concentration; $^{c}d_{eff}$ = effective distance to sea ice edge.

For Antarctic minke whales, novel environmental conditions occur also along a latitudinal band centred near 60° S throughout November, indicating that values for d_{eff} are larger here compared to the sampled study area. Given that the reliability of model predictions in areas of novel environmental conditions is questionable, predictions in these areas need to be considered with extreme caution.

The high overlap of environmental conditions in the prediction area with those encountered during model training likely results from the adequate coverage of the study area's environmental strata. Areas or times with novel environmental conditions should form the basis for future surveys to increase the models' predictive reliability.

4.1.2. *Model assessment based on external data*

Additionally, we compared prediction maps with known distribution patterns derived from historic catch and IWC sightings data. When comparing model predictions with these independent data, several environmental and temporal biases have to be considered. First, pelagic whaling did not primarily target humpback and Antarctic minke whales, but rather blue and fin whales, which likely biased catch effort towards regions with high densities of the latter two species. Second, catch effort was spatially not equally distributed. Relatively little catch effort was, e.g., directed towards IWC Area I and the eastern part of IWC Area VI (70–160°W) as this area was designated as a cetacean sanctuary from 1937 to 1955 (Tønnessen and Johnsen, 1982). The area around South Georgia had already been heavily exploited during the land-based whaling period from 1905 to the late 1920s, with no positional information on catches available. This likely explains the relative absence of catches recorded after 1929 in this region (Tønnessen and Johnsen, 1982) (see Fig. 5). Third, both the catch records and the IWC sightings are limited to open water areas due to the incapability of the whaling and sighting vessels to operate in sea ice covered areas, restricting model comparisons to areas of open water and the marginal sea ice zone. Finally, the catch records and the environmental variables used in our study differ by several decades. Environmental changes, such as a warming of the Southern Ocean (Gille, 2002) and regional changes in sea ice extent (Stammerjohn et al., 2008) have been observed over the last decades and might have affected cetacean distribution patterns. However, these datasets form the only long-term and circumpolar sources on humpback and Antarctic minke whale

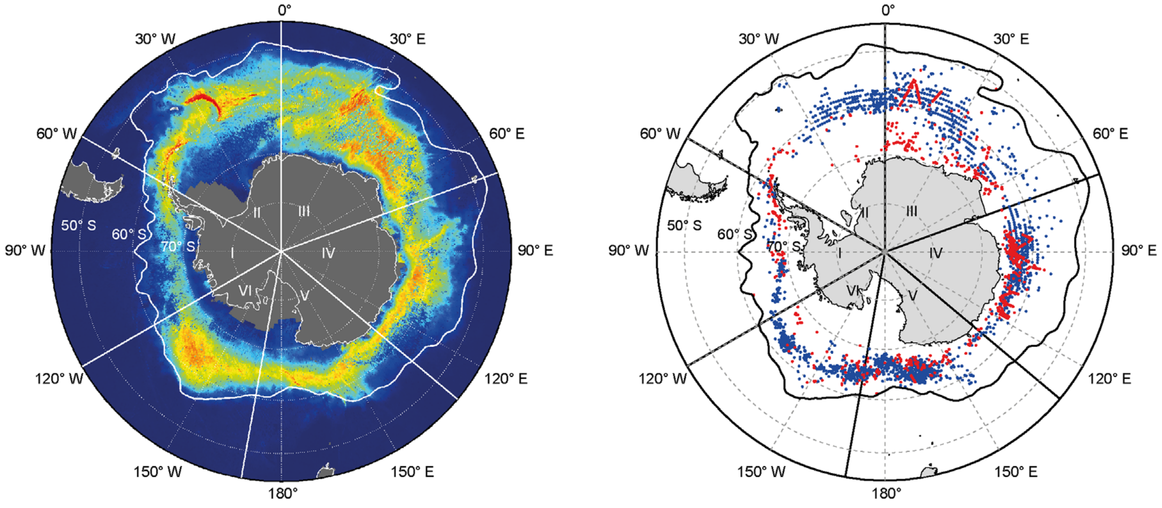

Fig. 5. Predicted habitat suitability for humpback whales averaged for January 2005–2011 (left) and corresponding monthly catch records (1929–1967, blue) and IWC sightings (1978–2009, red) (right). Grey areas indicate land areas or regions for which values for one of the environmental variables are missing. The white/black line indicates the climatological mean position of the Polar Front (Harris and Orsi, 2001, updated 2008). The white/black lines extending from the South Pole indicate the 6 IWC management areas (Donovan, 1991). Catch records from (Allison, 2013; Zemsky et al., 1995). IWC sightings data acquired from the office of the IWC (http://iwc.int/).

distributions and were used (with appropriate caveats and assumptions) to compare large-scale distribution patterns with habitat suitabilities predicted by the models.

4.1.2.1. Humpback whales. The spatio-temporal change in predicted humpback whale habitat suitability is also evident in the catch records and IWC sightings data (Fig. 5). Catches in November were relatively sparse but match to a large extent the distinct areas predicted by the model for this month. From December to January, catches occurred over wider latitudinal and longitudinal ranges. At the end of February, when the model predicts most favourable conditions closer to the Antarctic continent, whalers caught humpback whales also in the vicinity to the coast. Nevertheless, catches remained spread out between 0° and 45° E and also extended in latitudinal range in IWC Area VI. These are both regions where favourable habitat conditions are predicted over wider latitudinal areas. During March, the overall habitat suitability as predicted by the model is low and the number of whales caught also decreased considerably. Throughout April, the model predicts highly unsuitable habitat conditions for the full extent of the study area in accordance with low catch records for this month.

The IWC data also indicate that humpback whales were frequently sighted during January and February (Fig. 5), whereby sightings occurred closer towards the Antarctic coast in February compared to the wider latitudinal range (especially from 0° to 45° E and in IWC Area VI) in January. The same pattern is evident in Kasamatsu et al.'s (1996) analysis of IWC sightings and additional data from Japanese sighting surveys, showing an increase in encounter rates from early November to late December. High humpback whale densities were encountered over a wide latitudinal range throughout January, which steadily decreased throughout February.

Throughout the season, predicted habitat conditions remain highly favourable within distinct longitudinal sectors (e.g., 30° E, 90° E, 130° W). This pattern matches well with reports on longitudinally heterogeneous humpback whale distribution patterns based on historic catch and more recent sightings data. Omura (1973) showed that most humpback whales were caught between 20° E and 30° E, 80° E and 120° E and 140° E and 160° E. High occurrences of humpback whales were also found by Kasamatsu et al. (1996) between 0° and 60° E, 80° E and 120° E, 150°E and 160° E, 120° W and 180° and between 40° W and 80° W and between 70° E and 130° E by Matsuoka et al. (2011) and Thiele et al. (2000).

We additionally used the IWC sightings data (described in Section 2.3.4), yet confined to those collected during the same period as this model's opportunistic humpback whale sightings data, to conduct a quantitative test of model performance by using the IWC data as independent test data for the Maxent model. This resulted in an AUC value of 0.877, which matches well with the AUC values obtained from the 9-fold cross-validation (0.844 ± 0.062, see Table 2). Recognizing that the AUC value for this entirely independent dataset lies within the range of AUCs obtained by the cross-validation suggests, that the model avoids overfitting the training data (in this case, the IWC data-based AUC value should have dropped significantly).

A second quantitative test of model performance was conducted by comparing the distribution of Maxent suitability values for the IWC's humpback whale sightings (Fig. 6, red, $N=1136$) with the distribution of Maxent suitability values at arbitrarily chosen positions (blue). The latter were extracted for the very same days as those of the IWC sightings (i.e. $N=1136$ as well), hence factoring in the temporal evolution of habitat suitability within the Southern Ocean. This process was repeated 1000-fold and the results averaged, providing an average distribution of Maxent suitability values at random locations (called ‘random distribution’ hereinafter) as shown in Fig. 6.

The distribution of Maxent values for the IWC humpback whale sightings is significantly different from the random distribution. In particular, the IWC sightings show a clear preference for Maxent scores between 0.65 and 0.75, which is not reproduced by the random distribution's scores (the difference between the IWC's value and the random value is 36-times the standard deviation of the random distribution). Contrastingly, the random distribution peaks at lowest scores (≤ 0.15) for which only few IWC sightings occurred. These differences indicate that the model aptly identified areas with preferential sighting probability for humpback whales during independent IWC surveys.

4.1.2.2. Antarctic minke whales. For the area between 60° W and 60° E, only few Antarctic minke whale catch data are available, especially for IWC Area II, reflecting the low catch priority of Antarctic minke whales during the commercial whaling period, which started after the depletion of the larger baleen whale stocks around the 1970s (Tønnessen and Johnsen, 1982). The spatio-temporal pattern in catch locations (including catch data from IWC Areas I, IV–VI), is nevertheless fairly consistent with model predictions; Antarctic minke whales were caught over a wider latitudinal range during November and December and in regions close to the Antarctic coast during the remaining season (Fig. 7).

In contrast to the catch records, the IWC sightings database for Antarctic minke whales is fairly extensive and shows Antarctic minke whale sightings to occur predominantly close to the Antarctic coast during January and February, consistent with the model predictions (Fig. 7). However, IWC Antarctic minke whale sightings occurred over a broader latitudinal range when compared to model predictions, possibly as a result of the across-year variability in sea ice extent between survey and model years, affecting the area that could effectively be covered during the IWC surveys. Kasamatsu et al. (1996) noted that encounter rates for Antarctic minke whales were highest south of around 65° S in January and February, which is consistent with the suitable habitat predicted by the model mainly in the southern extent of the study area.

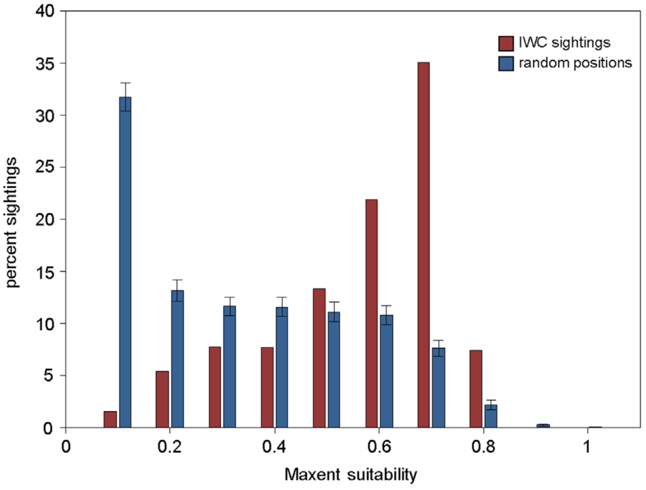

Fig. 6. Distribution of Maxent suitability values (binned) for the IWC's humpback whale sightings (selected for the same time period as the opportunistic sightings used for modelling, red, $N=1136$), as well as the average distribution of randomly positioned Maxent suitability values for days of IWC sightings (blue, $N=1136$ as well). Mean and standard deviation were obtained from 1000 independent realizations.

Given that the IWC's Antarctic minke whale sightings were collected in open water areas, we could not perform a similar quantitative assessment of model performance as we did for the humpback whale model. This is the aim of further research, especially in the context of the current debates regarding the importance of sea ice areas as suitable Antarctic minke whale habitat and related efforts in estimating Antarctic minke whale abundances (Branch, 2006a,2006b).

4.1.3. *Ecological agreement*

The spatio-temporal patterns in habitat suitability predicted by each model agree well with the current knowledge on the spatial ecology of humpback and Antarctic minke whales.

The humpback whale model predicts suitable habitat conditions consistently north of the sea ice edge, which is in accordance with humpback whales being described as an open water species (Ainley et al., 2003). On the circumpolar scale, Tynan (1998) showed that humpback whale distribution during December and January coincides with the climatological mean position of the southern boundary of the ACC. Humpback whales are assumed to follow the southern retreat of the sea ice edge, thereby taking advantage of highly productive areas associated with the continued development of the marginal sea ice zone throughout the season and move closer to the Antarctic continent by February (Tynan, 1998). The same pattern is evident in our model predictions. Predicted suitability is especially high in areas where the front is deflected (around 30° E, 90° E, 130° W and the western part of IWC Area II). These regions have also been identified as humpback whale 'hotspots' in previous studies (Matsuoka et al., 2011; Thiele et al., 2000; Tynan, 1997, 1998).

Antarctic minke whales are a pagophilic (i.e. ice-loving) species, having been observed within heavily sea ice covered areas (Ainley et al., 2007; Scheidat et al., 2011). Antarctic minke whales have further been associated with low SST temperatures (Kasamatsu et al., 1998; Ribic et al., 1991) and continental shelf areas (e.g., Ballard et al., 2012; Beekmans et al., 2010; Kasamatsu et al., 2000), which occur throughout the southernmost parts of the study area. These observations match with our model predictions, which

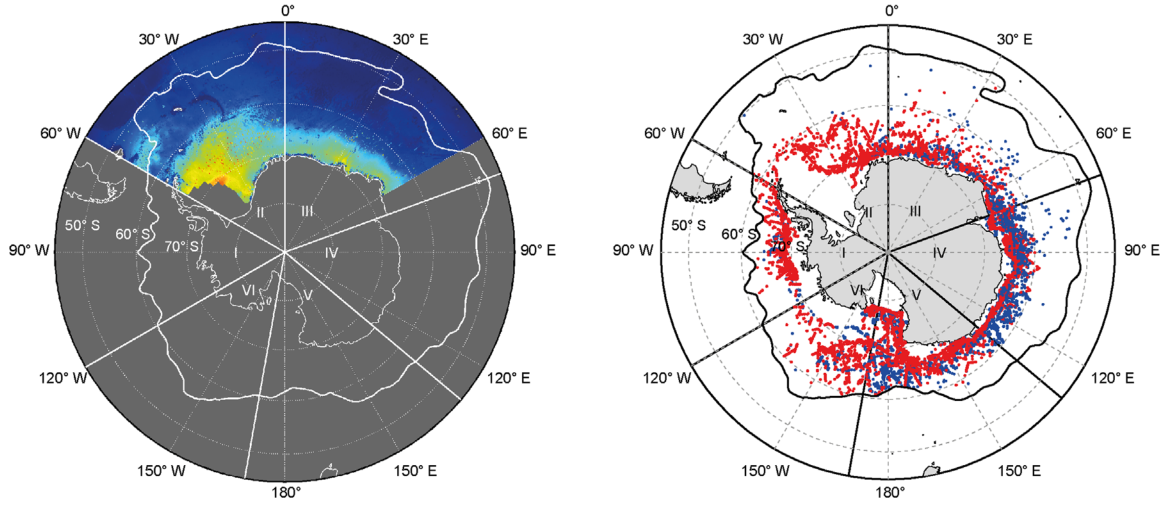

Fig. 7. Predicted habitat suitability for Antarctic minke whales averaged for January 2005–2011 (left) and corresponding monthly catch records (1955–2012, blue) and IWC sightings (1978–2009, red) (right). Grey areas indicate land areas or regions for which values for one of the environmental variables are missing and for which predictions were not calculated. The white/black line indicates the climatological mean position of the Polar Front (Harris and Orsi, 2001, updated 2008). The white/black lines extending from the South Pole indicate the 6 IWC management areas (Donovan, 1991). Catch records from (Allison, 2013; Zemsky et al., 1995). IWC sightings data acquired from the office of the IWC (http://iwc.int/).

show that high Antarctic minke whale habitat suitability is closely associated with sea ice covered areas.

Comparing model predictions between humpback and Antarctic minke whales for IWC Areas II and III indicates only little spatio-temporal overlap of suitable habitats for much of the season (Fig. 8). The sea ice edge thereby seems to act as a major ‘separator’ between each species' main habitat. Spatial habitat separations between humpback and minke whales have also been described by Kasamatsu et al. (1996). Scheidat et al. (2011) further noted a latitudinal separation of humpback and Antarctic minke whales during summer in the Weddell Sea, with Antarctic minke whales sighted further to the South and close to or within the sea ice.

Given that the majority of baleen whale species feed almost exclusively on Antarctic krill (*Euphausia superba*) (Kawamura, 1994), spatially different habitats may have evolved as a result to reduce interspecific competition for food. Murase et al. (2002) suggested that humpback and Antarctic minke whales could reduce competition for food by utilizing different areas. Due to their comparatively small size, Antarctic minke whales seem ideally adapted to exploit food resources within the often narrow confines of the sea ice environment (Thiele et al., 2000, Ainley et al., 2003). Additionally, Antarctic minke whales are known to occasionally consume other species: in the southern part of the Ross Sea, Antarctic minke whales also feed on ice krill (*Euphausia crystallorophias*) and Antarctic silverfish (*Pleuragramma antarcticum*) (Ichii et al., 1998). The preference of Antarctic minke whales for sea ice covered regions has also been related to predation avoidance, as these areas likely offer refuges from sea ice avoiding Type A killer whales, which are known to prey on Antarctic minke whales (Pitman and Ensor, 2003).

Some spatial overlap in humpback and Antarctic minke whale habitat, however, occurs in the vicinity of the sea ice edge. Potential inter-specific competition for krill in this area might be resolved by resource partitioning as observed on the western Antarctic Peninsula, where humpback and Antarctic minke whales target different krill length-maturity classes (Santora et al., 2010) and also at different depths (Friedlaender et al., 2009). Unravelling the mechanisms that led to the formation of each species' ecological niche, however, requires small-scale habitat studies, investigating both krill distribution and associated cetacean foraging behaviour as well as predator avoidance strategies, which are beyond the scope of this study.

The species-specific habitat preferences of humpback and Antarctic minke whales are also reflected in the set of environmental variables selected by each species' final model, which differ both in the composition and number of environmental variables, as well as the permutation importance of the respective variables (Table 2). For humpback whales, the most important environmental variables are length of day contributing most to the explanatory power of the model (38.7%), followed by SSH (19.6%), SST gradient (13%) and 14-day time-lagged sea ice concentration (10.9%). For Antarctic minke whales, the most important variables for the model are effective distance to the sea ice edge (39.6%), SST (24%) and depth (18.4%). The relevance of this selection of variables is further supported by the AUC jackknife tests.

4.2. *Model predictions as a planning tool*

The habitat models presented in this study are the first attempt of a (partly) circumpolar representation of each species'habitat during the Antarctic spring and summer months. Habitat models can only estimate that part of a species' niche that was actually captured by the sightings data and their applicability to unsampled areas is therefore dependent on the representativeness of environmental conditions inherent in the input data. Overall, both models reproduce well-known areas of higher humpback and Antarctic minke whale occurrences, also in unsampled areas. This likely results from the wide range of environmental strata across which the sightings data were sampled. As a next step the prediction maps should be correlated with local, spatio-temporally resolved density estimates, ideally sampled across a range of different suitability values (high/low). This correlation would allow calculating estimates of likely ship-whale encounter rates, a critical parameter in assessing the number of animals and potentially the percentage of the population which might become affected during a planned seismic survey.

Given the good performance of the models, it is instructive to explore the potential of the prediction maps at hand as a planning tool for seismic surveys. Particularly the fine temporal resolution of the maps allows identifying potential hotspots throughout the season, which could be avoided by knowledgeable pre-cruise planning. Avoidance of important habitats does not only minimize potential impacts on cetaceans, but also benefits the seismic campaign as shut-downs are less likely when surveys are conducted in areas or at times of low habitat suitability in the survey region, reducing overall anthropogenic acoustic exposure of the area. Fig. 9 depicts the layout of a hypothetical seismic survey, comprising a North–South trackline along the Greenwich meridian from 50° S to the Antarctic coast (around 70° S) and a second South–North trackline for the same latitudes at 1° E. The question at hand is to select the most suitable time and starting point for this cruise while aiming to minimize the likelihood of encountering whales by avoiding ship presence in areas that have high habitat suitability.

We tested two different temporal scenarios for the year 2007, the first starting on January 15 finishing mid-February and the

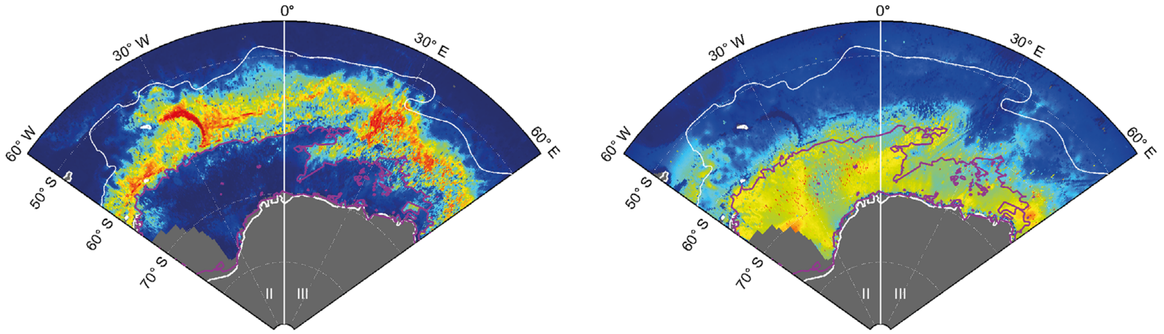

Fig. 8. Predicted habitat suitability for humpback (left) and Antarctic minke whales (right) for 15 December 2009. The purple line indicates the sea ice edge. Grey areas indicate land areas or regions for which values for one of the environmental variables are missing. The white line depicts the climatological mean position of the Polar Front (Harris and Orsi, 2001, updated 2008). The IWC management areas are indicated by the white line extending from the South Pole (Donovan, 1991).

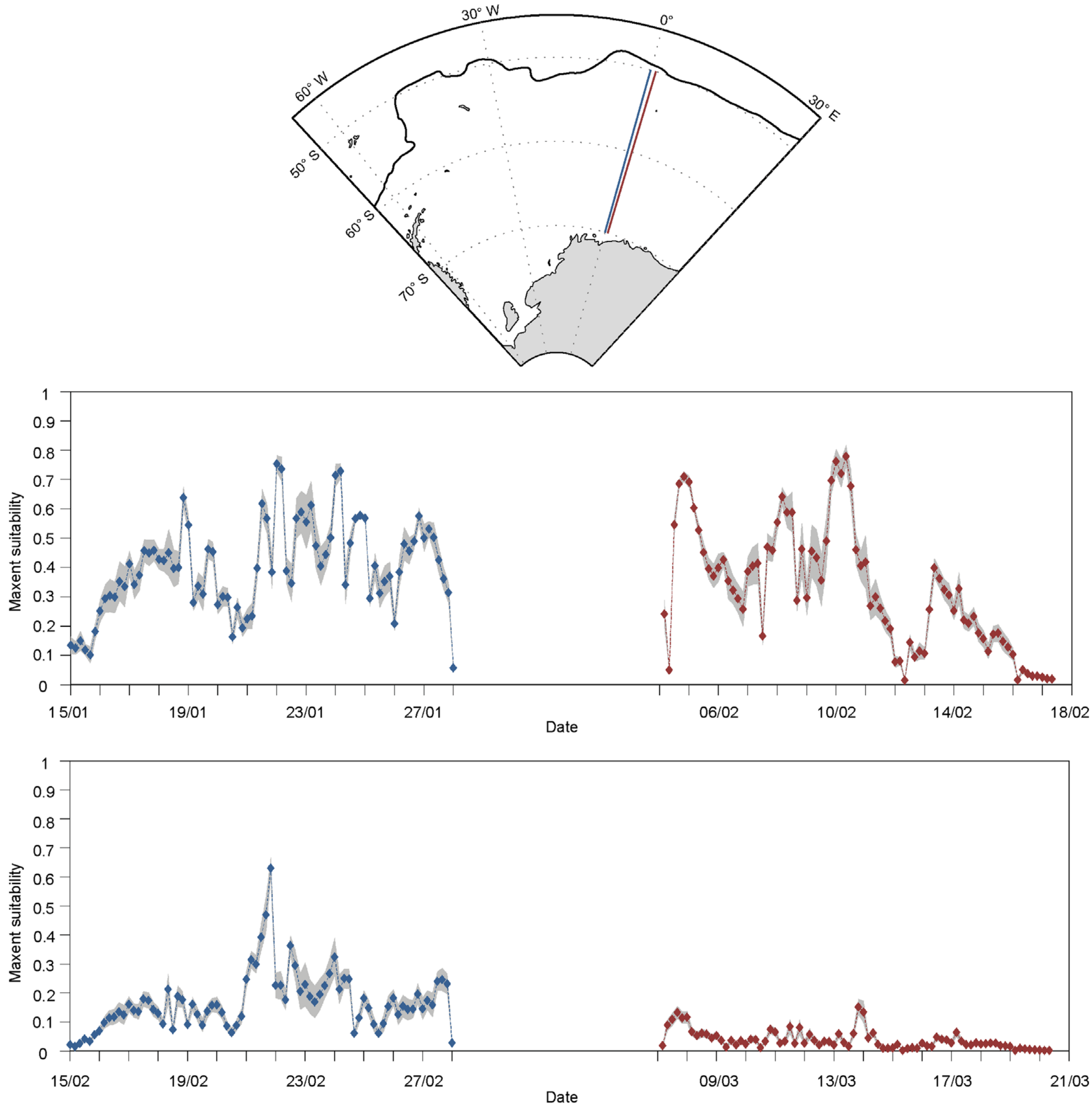

Fig. 9. Top: hypothetical layout of a seismic survey conducted along two transect lines from North to South (blue) and South to North (red). Middle: habitat suitability values for humpback whales at the time and location of the hypothetical ship for a survey from January 15 to February 17, 2007. Bottom: habitat suitability values for humpback whales at the time and location of the hypothetical ship for a survey from February 15 to March 20, 2007. The standard deviation is depicted by the grey shading.

second starting on February 15 and returning mid-March, respectively (assumed survey speed ca. 4 kn), with a hypothetical operational gap of 1 week between the North-South/South-North transects. For each scenario we derived the habitat suitability values at the time and location of the hypothetical ship based on our prediction maps for humpback whales. This exercise requires daily prediction maps and could therefore only be executed for the humpback whales at this time. As Fig. 9 indicates, overall habitat suitability values encountered during a survey planned according to the second scenario are considerably lower compared to the first scenario. This is also supported when calculating the mean habitat suitability for each survey along with its standard error according to

$$SE = \frac{\sqrt{\Sigma_{i=1}^{N} std_i^2}}{\sqrt{N}} \tag{1}$$

where by std_i is the respective standard deviation of the i-th Maxent suitability value as based on the n -fold cross-validation and N the number of grid cells virtually 'transected' during the planned survey. The mean habitat suitability for the second survey is significantly lower (0.10 ± 0.03) compared to the first survey (0.36 ± 0.05). In addition, higher quality habitats (≥ 0.65) are encountered less frequently (0 vs. 12) during the second survey.

Hence, planning our hypothetical survey towards the end of the summer months offers a possibility to avoid ship presence in areas of higher mean habitat suitability as well as in localized areas of higher habitat suitability (potential hotspots) to minimize potentially negative effects of seismic surveys on humpback whales in this part of the Southern Ocean.

When varying the prospective survey design, the lower average habitat suitability value identifies the approach that is likely to minimize encounter rates for the respective species. Making such assessments in the context of an EIA (environmental impact assessment) would represent 'best possible practise' as long as other research does not to provide contradicting information, which of course requires appropriate discussion in any contingent EIA.

This relatively simple example illustrates how seismic surveys in offshore areas could be designed as to minimize their contingent ecological impact. The general trend in our prediction maps suggests that offshore surveys should be planned towards the end of the season, when habitat suitability for both species is highest in coastal areas and hence potential impacts on humpback and Antarctic minke whales would be minimized. When North–South transects are conducted, transect lines could e.g., be planned such that they follow the southward retreat of humpback whales within a certain time delay. In coastal areas, seismic surveys are temporally restricted to ice-free periods and overlap with the occurrence of suitable humpback and Antarctic minke whale habitats. Nevertheless, the prediction maps indicate areas with lesser habitat suitability, which – depending on the targeted geographic research area – might allow focusing surveys on areas with a lesser likelihood to encounter humpback and Antarctic minke whales.

Of course, adequate risk assessment of seismic surveys needs to consider the full range of cetacean and seal species occurring in the Southern Ocean, at least to the extent feasible, and not only the humpback and Antarctic minke whales studied herein. Based on the results from this study, the use of opportunistic sightings in conjunction with suitable habitat models appears a promising approach in order to accelerate our knowledge on Antarctic marine mammals. The data logging software developed for this project lends itself to an easy transfer to other research or tourist vessels navigating the Southern Ocean, allowing the generation of multi-species data across large temporal and spatial scales and hence various environmental strata. This way, a more comprehensive assessment of the risk posed to a wider range of Antarctic marine mammal species by seismic research activities will become possible.

5. Conclusion

This study demonstrates the potential of habitat modelling to derive large-scale information on the habitat suitability patterns of humpback and Antarctic minke whales in the Southern Ocean. For both cetacean species, prediction maps identified rather opposed distributions of suitable habitat conditions, with suitable humpback whale habitats occurring north of those for Antarctic minke whales. The fine temporal resolution used for the habitat models and prediction maps revealed considerable changes in habitat suitabilities throughout the spring and summer months for both cetacean species. This highlights the importance to account for dynamic processes that structure the Southern Ocean ecosystem throughout the seasons, which in turn allows a more detailed risk assessment and planning of seismic research activities.

The spatial prediction maps as presented here provide not only a valuable planning tool in the context of seismic activities, but could also be used to direct more detailed cetacean-habitat studies. Given that visual data are biased towards the summer months, the maps may be used to identify suitable positions for passive acoustic monitoring moorings to obtain year-round information on cetacean presence. In some cases, specific call types have been associated with behaviours, which would allow associating spatial habitat information with specific habitat usage.

Pre-cruise planning (based on prediction maps) should of course not be considered as an alternative to operational mitigation measures, such as shut-downs, when whales enter the environs of an acoustic source. These approaches are primarily aimed at minimizing the risk of injuring individuals that might get too close to the acoustic source. Rather, prediction map based pre-cruise planning allows minimizing the likelihood of ship-whale encounters based on avoiding activity in high suitability areas and is hence of relevance at the population instead of the individual level. Further extended applications of the maps as a pre-cruise planning tool may allow the inclusion of an ‘impact radius’ when calculating encounter likelihoods, by considering information on distances at which species-specific behavioural responses are likely to occur.

Acknowledgements

We thank the nautical officers on board RV *Polarstern* for their enthusiastic collection of sightings data. We furthermore like to acknowledge the International Whaling Commission (www.iwc.int) for allowing use of the data from its series of International Decade of Cetacean Research (IDCR) and Southern Ocean Whale and Ecosystem Research (SOWER) programmes, and for providing the historical catch data. We would also like to thank two anonymous reviewers and the associate editor for their insightful comments, which greatly helped improving the manuscript.

This project was funded by the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (Grant no. BMU 370891101-01). The responsibility for the content of this publication rests solely with the authors.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dsr.2014.05.017.

References

Ainley, D., Tynan, C., Stirling, I., 2003. Sea ice: a critical habitat for polar marine mammals and birds. In: Thomas, D.N., Dieckmann, G.S. (Eds.), Sea ice: an introduction to its physics, chemistry, biology and geology. Blackwell Publishing, Oxford, pp. 240–266.

Ainley, D.G., Dugger, K.M., Toniolo, V., Gaffney, I., 2007. Cetacean occurrence patterns in the Amundsen and southern Bellingshausen Sea sector, Southern Ocean. Mar. Mamm. Sci. 23, 287–305.

Allison, C., 12 February 2013. IWC individual catch database Version 5.5.

Amante, C., Eatkins, B.W., 2009. ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24. March, 19 pp.

Aviso. The altimeter products were produced by Ssalto/*Duacs* and distributed by *Aviso*, with support from *Cnes* 〈http:/www.aviso.oceanobs.com/duacs/〉.

Ballard, G., Jongsomjit, D., Veloz, S.D., Ainley, D.G., 2012. Coexistence of mesopredators in an intact polar ocean ecosystem: The basis for defining a Ross Sea marine protected area. Biol. Conserv. 156, 72–82.

Barron, C.N., Kara, A.B., 2006. Satellite-based daily SSTs over the global ocean. Geophys. Res. Lett. 33, L15603.

Barry, S., Elith, J., 2006. Error and uncertainty in habitat models. J. Appl. Ecol. 43, 413–423.

Baumgartner, M.F., Mullin, K.D., May, L.N., Leming, T.D., 2001. Cetacean habitats in the northern Gulf of Mexico. Fish B – NOAA 99, 219–239.

Beekmans, B.W.P.M., Forcada, J., Murphy, E.J., De Baar, H.J.W., Bathmann, U.V., Fleming, A.H., 2010. Generalised additive models to investigate environmental drivers of Antarctic minke whale (*Balaenoptera bonaerensis*) spatial density in austral summer. J. Cetacean Res. Manag. 11, 115–129.

Best, B.D., Halpin, P.N., Read, A.J., Fujioka, E., Good, C.P., LaBrecque, E.A., Schick, R.S., Roberts, J.J., Hazen, L.J., Qian, S.S., Palka, D.L., Garrison, L.P., McLellan, W.A., 2012. Online cetacean habitat modeling system for the US east coast and Gulf of Mexico. Endanger. Species Res. 18, 1–15.

Booth, C.G., Embling, C., Gordon, J., Calderan, S.V., Hammond, P.S., 2013. Habitat preferences and distribution of the harbour porpoise Phocoena phocoena west of Scotland. Mar. Ecol. Prog. Ser. 478, 273–285.

Branch, T.A., 2006a. Abundance estimates for Antarctic minke whales from three completed circumpolar sets of surveys, 1978/79 to 2003/04. Paper: SC/58/IA18, Submitted to the Scientific Committee of the International Whaling Commission, Paper Available from the Office of the IWC, 28 pp.

Branch, T.A., 2006b. Possible reasons for the appreciable decrease in abundance estimates for Antarctic minke whales from the IDCR/SOWER surveys between the second and third circumpolar sets of cruises. Paper SC/58/IA4, Submitted to the Scientific Committee of the International Whaling Commission, Paper Available from the Office of the IWC, 8 pp.

Breitzke, M., 2014. Overview of seismic research activities in the Southern Ocean – quantifying the environmental impact. Antartic Sci. 26, 80–92.

Burkhardt, E., 2009a. 〈http:/epic.awi.de/29131/1/Bur2009a.pdf〉.

Burkhardt, E., 2009b. Whale sightings during POLARSTERN cruise ANT-XXII/3. Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven, http://dx.doi.org/10.1594/PANGAEA.729027.

Compton, R., Goodwin, L., Handy, R., Abbott, V., 2008. A critical examination of worldwide guidelines for minimising the disturbance to marine mammals during seismic surveys. Mar. Policy 32, 255–262.

Donovan, G.P., 1991. A review of IWC stock boundaries. Rep. Int. Whal. Comm. (Spec. Issue) 13, 39–68.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151.

Elith, J., Kearney, M., Phillips, S., 2010. The art of modelling range-shifting species. Methods Ecol. Evol. 1, 330–342.

Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Divers. Distrib. 17, 43–57.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24, 38–49.

Friedlaender, A.S., Halpin, P.N., Qian, S.S., Lawson, G.L., Wiebe, P.H., Thiele, D., Read, A.J., 2006. Whale distribution in relation to prey abundance and oceanographic processes in shelf waters of the Western Antarctic Peninsula. Mar. Ecol. Prog. Ser. 317, 297–310.

Friedlaender, A.S., Hazen, E.L., Nowacek, D.P., Halpin, P.N., Ware, C.R., Weinrich, M.T., Hurst, T., Wiley, D., 2009. Diel changes in humpback whale Megaptera novaeangliae feeding behavior in response to sand lance Ammodytes spp. behavior and distribution. Mar. Ecol. Prog. Ser. 395, 91–100.

Friedlaender, A.S., Johnston, D.W., Fraser, W.R., Burns, J., Halpin, P.N., Costa, D.P., 2011. Ecological niche modeling of sympatric krill predators around Marguerite Bay, Western Antarctic Peninsula. Deep Sea Res. Part II 58, 1729–1740.

Gille, S.T., 2002. Warming of the Southern Ocean since the 1950s. Science 295, 1275–1277.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135, 147–186.

Harris, U., Orsi, A.H., 2001. Locations of the various fronts in the Southern Ocean Australian Antarctic Data Centre – CAASM Metadata, updated 2008. 〈https:/data.aad.gov.au/aadc/metadata/metadata_redirect.cfm?md=/AMD/AU/southern_ocean_fronts〉.

Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L., 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29, 773–785.

Ichii, T., Shinohara, N., Fujise, Y., Nishiwaki, S., Matsuoka, K., 1998. Interannual changes in body fat condition index of minke whales in the Antarctic. Mar. Ecol. Prog. Ser. 175, 1–12.

JNCC, 2004. Guidelines for Minimising Acoustic Disturbance to Marine Mammals from Seismic Surveys. Joint Nature Conservation Committee, Aberdeen.

Kawamura, A., 1994. A review of baleen whale feeding in the Southern Ocean. Rep. Int. Whal. Comm. 44, 261–271.

Kaleschke, L., Lupkes, C., Vihma, T., Haarpaintner, J., Bochert, A., Hartmann, J., Heygster, G., 2001. SSM/I sea ice remote sensing for mesoscale ocean-atmosphere interaction analysis. Can. J. Remote Sens. 27, 526–537.

Kasamatsu, F., Ensor, P., Joyce, G.G., Kimura, N., 2000. Distribution of minke whales in the Bellingshausen and Amundsen Seas (60° W-120° W), with special reference to environmental/physiographic variables. Fish Oceanogr. 9, 214–223.

Kasamatsu, F., Joyce, G.G., Ensor, P., Kimura, N., 1998. Distribution of minke whales in the Weddell Sea in relation to the sea-ice and sea surface temperature. Bull. Jpn. Soc. Sci. Fish. 62, 334–342.

Kasamatsu, F., Joyce, G.G., Ensor, P., Memorez, J., 1996. Current occurence of baleen whales in Antarctic waters. Rep. Int. Whal. Comm. 46, 293–303.

MacLeod, C.D., Weir, C.R., Pierpoint, C., Harland, E.J., 2007. The habitat preferences of marine mammals west of Scotland (UK). J. Mar. Biol. Assoc. UK 87, 157–164.

Matsuoka, K., Hakamada, T., Kiwada, H., Murase, H., Nishiwaki, S., 2011. Abundance estimates and trends for humpback whales (*Megaptera novaeangliae*) in Antarctic areas IV and V based on JARPA sightings data. J. Cetacean Res. Manag. (Spec. Issue) 3, 74–94.

Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36, 1058–1069.

Murase, H., Matsuoka, K., Ichii, T., Nishiwaki, S., 2002. Relationship between the distribution of euphausiids and baleen whales in the Antarctic (35° E-145° W). Polar Biol. 25, 135–145.

Nowacek, D.P., Bröker, K., Donovan, G., Gailey, G., Racca, R., Reeves, R.R., Vedenev, A.I., Weller, D.W., Southall, B.L., 2013. Responsible practices for minimizing and monitoring environmental impacts of marine seismic surveys with an emphasis on marine mammals. Aquat. Mamm. 39, 356–377.

Omura, H., 1973. A review of pelagic whaling operations in the Antarctic based on the effort and catch data in 10° squares of latitude and longitude. Sci. Rep. Whales Res. Inst. 25, 105–121.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231–259.

Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175.

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. Ecol. Appl. 19, 181–197.

Phillips, S.J., Dudík, M., Schapire, R.E., 2004. A maximum entropy approach to species distribution modeling. In: Proceedings of the Twenty-First International Conference on Machine Learning,ACM, Banff, Alberta, Canada. pp. 655–662.

Pitman, R.L., Ensor, P., 2003. Three forms of killer whales in Antarctic waters. J. Cetacean Res. Manag. 5, 1–9.

Redfern, J.V., Ferguson, M.C., Becker, E.A., Hyrenbach, K.D., Good, C., Barlow, J., Kaschner, K., Baumgartner, M.F., Forney, K.A., Ballance, L.T., Fauchald, P., Halpin, P., Hamazaki, T., Pershing, A.J., Qian, S.S., Read, A., Reilly, S.B., Torres, L., Werner, F., 2006. Techniques for cetacean-habitat modeling. Mar. Ecol. Prog. Ser. 310, 271–295.

Ribic, C.A., Ainley, D.G., Fraser, W.R., 1991. Habitat selection by marine mammals in the marginal ice-zone. Antartic Sci. 3, 181–186.

Santora, J.A., Reiss, C.S., Loeb, V.J., Veit, R.R., 2010. Spatial association between hotspots of baleen whales and demographic patterns of Antarctic krill Euphausia superba suggests size-dependent predation. Mar. Ecol. Prog. Ser. 405, 255–269.

Scheidat, M., Friedlaender, A., Kock, K.H., Lehnert, L., Boebel, O., Roberts, J., Williams, R., 2011. Cetacean surveys in the Southern Ocean using icebreaker-supported helicopters. Polar Biol. 34, 1513–1522.

Spreen, G., Kaleschke, L., Heygster, G., 2008. Sea ice remote sensing using AMSR-E 89-GHz channels. J. Geophys. Res. – Oceans 113, C02S03, http://dx.doi.org/10.1029/2005JC003384.

Stammerjohn, S.E., Martinson, D.G., Smith, R.C., Yuan, X., Rind, D., 2008. Trends in Antarctic annual sea ice retreat and advance and their relation to El Niño–Southern Oscillation and Southern Annular Mode variability. J. Geophys. Res. 113, C03S90.

Širović, A., Hildebrand, J.A., 2011. Using passive acoustics to model blue whale habitat off the Western Antarctic Peninsula. Deep Sea Res. Part II 58, 1719–1728.

Thiele, D., Chester, E.T., Gill, P.C., 2000. Cetacean distribution off Eastern Antarctica (80–150° E) during the Austral summer of 1995/1996. Deep Sea Res. Part II 47, 2543–2572.

Tønnessen, J.N., Johnsen, A.O., 1982. The history of modern whaling. C. Hurst & Co. (Publishers) Ltd., London.

Tynan, C., Thiele, D., 2003. Report on Antarctic ice edge definition by the ad hoc working group on ice data collection in the Antarctic. Paper: SC/55/19, Submitted to the Scientific Committee of the International Whaling Commission, Paper Available from the Office of the IWC, 1 p.

Tynan, C.T., 1997. Cetacean distributions and oceanographic features near the Kerguelen Plateau. Geophys. Res. Lett. 24, 2793–2796.

Tynan, C.T., 1998. Ecological importance of the southern boundary of the Antarctic circumpolar current. Nature 392, 708–710.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., Distribut, N.P.S., 2008. Effects of sample size on the performance of species distribution models. Divers. Distrib. 14, 763–773.

Zemsky, V.A., Berzin, A.A., Mikhaliev, Y.A., Tormosov, D.D., 1995. Soviet Antarctic Whaling Data (1947–1972), 1st ed. Center for Russian Environment Policy, Moscow p. 321.

Zitterbart, D.P., Kindermann, L., Burkhardt, E., Boebel, O., 2013. Automatic round-the-clock detection of whales for mitigation from underwater noise impacts. PLoS ONE 8, e71217.