SC/67A/CMP/15

Preliminary ensemble ecological niche modelling of Arabian Sea humpback whale vessel sightings and satellite telemetry data

Andrew Willson, Robert Baldwin, Tim Collins, Brendan Godley, Gianna Minton, Suaad al Harthi, Stephen Pikesley and Matt Witt.

INTERNATIONAL WHALING COMMISSION

Preliminary ensemble ecological niche modelling of Arabian Sea humpback whale vessel sightings and satellite telemetry data

A. Willson[1,2,3], R. Baldwin[3] T. Collins[4], B. Godley[2], G. Minton[5], Suaad al-Harthi[6], S. Pikesley[1], M. Witt[1].

[1] *Environment and Sustainability Institute, University of Exeter, Penryn Campus, Cornwall, UK*
[2]*Centre for Ecology and Conservation University of Exeter, Exeter, UK*
[3]*Five Oceans Environmental Services, PO Box 660, PC131, Ruwi, Sultanate of Oman*
[4]*Wildlife Conservation Society, Ocean Giants Program, 2300 Southern Blvd, Bronx, NY 10460-1099, USA*
[5]*Megaptera Marine Conservation, Netherlands.*
[6]*Environment Society of Oman, PO Box 3955, PC 112, Ruwi, Sultanate of Oman*

ABSTRACT

Ensemble ecological niche modelling (EENM) can provide insight into the relationship between marine mammals and their environment and can predict distribution beyond the range of observed locations. The technique can be used to identify sites for future field research and guide conservation and management activities. The spatial ecology of Arabian Sea humpback whales (ASHWs) has been described off the coast of Oman, although a paucity of information exists from which to describe their distribution across the rest of their potential range. Here we present an ensemble ecological niche modelling framework to predict habitat suitability of ASHWs across the north Indian Ocean. Sightings data from Oman-based small vessel surveys (2003-2014) and satellite telemetry records (2014-2016) were used along with environmental co-variate data from a season between December and May. Net primary productivity featured as the only co-variate with a strong influence on models for both datasets. Model test evaluation metrics scored >0.9, and mapped outputs of likely distribution highlighted spatial similarity across multiple models. Telemetry data predicted suitable habitat to be further offshore than the models derived from sightings data. All resulting distribution maps described areas of high suitability (index value <0.75) along the southern and central coast of Oman and of the northern Arabian Sea between the Gulf of Kutch and sub-marine canyon features off the Indus delta. There was good spatial concordance between ensemble model predictions with actual locations of Soviet catches of humpback whales in the northern Indian Ocean between 1964 and 1966. Both the telemetry and the sightings data were temporally sporadic in their coverage (across months) and biologically biased (towards males) and as such results from our preliminary efforts should be considered in light of these caveats. However, these preliminary results are valuable and indicate likely co-occurrence with high density shipping traffic routes in the region and target additional areas for focussed field surveys. Results from this study should be considered together with results of recent north Indian Ocean blue whale ENM studies to help guide future research and conservation management objectives in the region.

INTRODUCTION

Spatio-temporal resolution of the occurrence of marine mammals derived from quantifying relationships between species and the environment through habitat modelling is considered a useful approach where sparse sightings data exists to confirm their distribution (Guisan and Zimmermann, 2000 and Redfern et al. 2006). Mapped outputs from habitat modelling on potential species distributions have been widely applied to risk analysis and the evaluation of relevant co-occurrence of anthropogenic activities (Kaschner *et al.* 2006, Becker *et al.* 2012, Redfern *et al.* 2013, Hazen *et al.* 2016). Unlike other well studied populations of humpback whales, the Arabian Sea humpback whale (ASHW) received relatively little dedicated field based investigation regarding the species until the initiation of dedicated small boat surveys off Oman in 2000, (Minton *et al.* 2011). Prior to this knowledge was based on opportunistic records (Brown, 1957; Slijper *et al*. 1964; Wray and Martin, 1983) and Soviet whaling records (n=238 takes) described by Mikhalev (1997). A review of encounters described in a review paper by Reeves *el al*. (1991) supported a hypothesis of an isolated non-migratory population across the Northern Indian Ocean (NIO).

Photo-identification work conducted from small vessel surveys off Oman (2000-2003) provided mark-recapture population estimate of 82 individuals (95% CI 60-111; Minton *et al.* 2008) and supported designation of this population as 'Endangered' on the International Union for the Conservation of Nature (IUCN) Red List. Recent genetic analysis on biopsy samples obtained from the same surveys support the hypothesis of the isolated status of these whales and indicates divergence time from Southern Hemisphere populations of ~70,000 yrs BP (Pomilla, Amaral *et al*. 2014).

Distribution and seasonality of ASHW presence off the coast of Oman has also been derived from small vessel as described by Minton *et al*. (2011), with further habitat modelling work performed by Corkeron *et al.* (2011). Results of both studies highlighted the importance of two areas in Oman as humpback whale habitat including the Gulf of Masirah and the Hallaniyats Bay. These findings have further been confirmed by satellite tagging studies of ASHW conducted between 2014 and 2015 revealing tagged whales spending 35% and 22% of their time in these areas respectively (n=9).

The absence of ASHW photographic data beyond Oman was noted as a constraint of the IUCN Red List population assessment (Minton et al. 2008) and apart from data generated by Soviet whaling activities there remains a paucity of published sightings data across the wider NIO.

In recent years the ASHW population has been described as extremely vulnerable to anthropogenic threats (Baldwin *et al.* 1999, Minton *et al.* 2008; Baldwin *et al.* 2010), with evidence of fishing, commercial vessel activity and hydrocarbon exploration escalating within habitats associated with highest sighting densities (Corkeron *et al*. 2012; Willson *et al*. 2016a; Willson *et al.* 2016b). Understanding the wider distribution of ASHW in the NIO has been identified as a priority by the Arabian Sea Whale Network (Minton *et al.* 2015) and endorsed by the IWC Scientific Committee (2015). Redfern *et al.* (2017) addressed a similar problem for predicting habitat of blue whales in the NIO through the application of Eastern Pacific and California Current whale models to the northern Indian Ocean using a Generalised Additive Model framework. Single model frameworks can have associated biases that reduce comparability of results and limit predictive capacity (Scales *et al*. 2015). Ensemble ecological niche models (EENM; Araujo & New, 2007), address these issues by combining the outputs of multiple models into a single predictive surface. EENM has been used to map sea turtle distribution in relation to fisheries responsible for bycatch (Pikesley *et al.* 2013), predict where dedicated abundance surveys should be conducted for seabirds (Oppel *et al.* 2012) and conservation efforts focused for grey-headed albatrosses, *Thalassarche chrysostoma* (Scales *et al*. 2015). Where resources are limited the application of robust modelling efforts can provide guidance where future research and conservation efforts should be directed.

Here we use sightings data from small vessel surveys collected between 2003 and 2014 together with humpback whale satellite telemetry data from 2014 to 2016 within an EENM framework as a mixed modelling approach to predict the habitat utilisation of ASHWs across the NIO. The intention of this preliminary investigation is to a) inform where ASHW may be found outside of Oman to guide emerging survey efforts b) provide information to high level risk assessments (including shipping) and c) evaluate suitable methods for predictive habitat model refinement as a greater range of data sources becomes available from the region.

MATERIALS AND METHODS

Data Collection & Selection

Sightings of humpback whales used in this study were collected by small vessel surveys predominantly located in the Gulf of Masirah, Dhofar and off Muscat (Figure 1). Surveys followed methods as described by Minton *et al.* (2011). These surveys were conducted by 6.5m rigid hull inflatable boat. Haphazard line transect surveys were bound by constraints including distance from anchorage (40nm), distance offshore (30nm), and seasonality of access to study areas based on safe operational conditions at sea. To define seasonality and distribution of whales, survey effort track-lines followed irregular saw-tooth patterns along the coast and were traversed at speeds between 12 to 15 knots (Corkeron *et al*. 2011). Suspected sightings resulted in suspension of the effort to allow close approach of the vessel to confirm species identification and acquire photographic images and biopsy samples. Summer monsoon winds prohibited small vessel surveys and sea state conditions below Beaufort Force 4 for months between May and September (Willson *et al.* 2016).

Selection of the sightings dataset for the modelling from 2003 to 2014 was determined by availability of environmental covariate data. Survey effort was un-equally distributed between the three principle areas in Dhofar , Gulf of Masirah and Muscat (Appendix Table 4 and Table 5). A total of 274 days was spent surveying, representing a cumulative 2375 survey hours. The coast off Muscat was only surveyed for the period 2003-2006. Survey effort in Dhofar (884 hours) was approximately twice that of the Gulf of Masirah (428 hours).

Surveys conducted in 2014, 2015a and 2016 were designed to provide opportunity for instrumentation of whales with Wildlife Computers SPOT5 and SPLASH10 satellite tags (Redmond, WA, USA). Satellite telemetry data was sourced from the ARGOS system (CLS, 2011). The implantable tags anchored in the fascia and just forwards of the dorsal were deployed by pneumatic tag application system (a modified version of the Air Rocket Transmitter system 'ARTs', HeideJørgensen et al. 2001). Satellite tags were deployed in the Gulf of Masirah (n=3) and Hallaniyats Bay (n=6) between February 2014 and November 2015 (Figure 1), providing mean track durations of 56±21 days (mean ± SD) (Appendix Table 8).

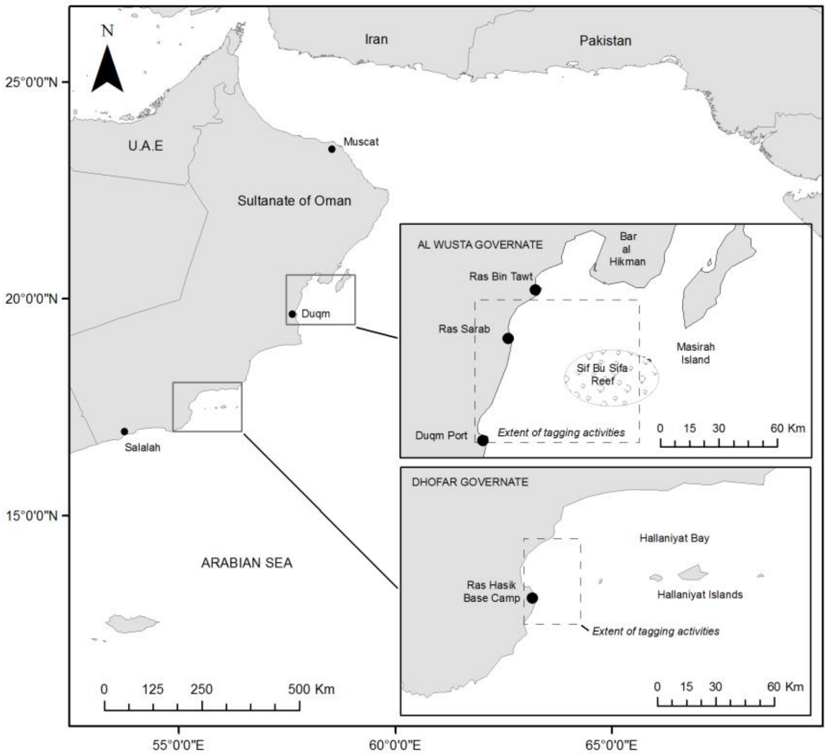

Figure 1. Locations of tagging areas in Dhofar and the Gulf of Masirah (2014 - 2015) and vessel surveys in Dhofar, the Gulf of Masirah and Muscat (2003- 2014).

Geoprocessing

Sightings data was filtered according to encounters made during dedicated surveys vessel only. For the telemetry location data we choose to use the dataset previously processed according to Willson et al. (2016) using a switching space state model (SSSM) developed by Jonson (2005) and Breed (2009). This mechanistic model was applied to address serial autocorrelation and reduce the influence of detectability (of telemetry signals) caused by behavioural shifts of animals and environmental screening of the data in different habitats as described by Aaral et al. (2008).

Selection of Environmental Co-Variate Data

The spatial extent of environmental co-variate data used for the north west Indian Ocean as 31˚N, 32˚E, 83˚W and 2˚S and was guided by Soviet catch positions occuring between 1964 and 1966 (Mikhalev, 1997) and by acoustic encounters reported by Whitehead (1985) to the north west of Sri Lanka. Historical sightings off the north Somali coast have been documented by Brown (1957) and in the northern area of the Red Sea (Notarbartolo di Sciara et al. 2014). It is not yet known if these wider ranging records of humpback whales in the NIO represent seasonal migrants, vagrants or the most southerly extent of ASHWs.

We selected physical and biological co-variate data based on the availability of datasets that were spatially and temporally concordant with sightings and telemetry data and also reflective of known local ecology. Previous studies indicated tagged whales spent 72% of their time in water less than 200m (Willson et al. 2016). Selection of depth is consistent with studies of other humpback whale populations where humpbacks have been found associated with the centre of the shelf along the 50m contour (Moore et al. 2002).

Co-variates related to prey availability were also considered for the study. Sardines and euphasiids were found in the stomachs of whales examined in the Soviet catch (n=85) (Mikhalev, 1997). Both of these prey items occupy different habitats. Sardine species found in Oman predominantly feed on phytoplankton (Randal, 1995). Piontkovski (2014) identified monthly fluctuations in sardine landings associated with remotely sensed chlorophyll-*a* concentration from MODIS-Aqua and SeaWIFS archives off the northern coast of Oman. Warm water surface extensions (streamers) of 10km width have been associated with migration of sardines towards coastal areas and concentration of dense schools at the head of these streamers (Sugimoto and Tameishi, 1992; Tameishi et al., 1994) indicating that both temperature differential and frontal activity should be considered as an important component in analysis. The distribution of euphausiids in relation to environmental co-variates is less well studied in the region although the proximity to shelf edge is considered an important variable in the concentration of krill (Harris et al. 2014).

Data relevant to breeding and rearing of young was also considered for inclusion in the list of co-variates. The sightings database documents a small number of sightings of mothers and calves (n=5), all of which were found in close proximity to the coast <2nm. Minton *et al.* (2011) reports on the occurrence of male song synchronous with the breeding season detected primarily in Dhofar (Hallaniyats Bay). The bay is characterised as an area of mixed shelf and steep sloping bathymetry. The strongest covariates used to describe humpback whale breeding and nursery habitats on the Great Barrier Reef included sea-surface temperature, water depth, slope, distance to reef and distance to coast (Smith et al. 2012).

Co-variate Data Sources and Processing

Bathymetry data were sourced through UK Hydrographic Office data portal (GEBCO) and resolved to a 9km resolution. Seabed slope was derived from bathymetry data in ArcGIS. Bathymetry data were also used to create a distance to shelf edge data layer. For this study the shelf edge was taken to be 200m depth.

Satellite co-variate data were accessed from MODIS Aqua (Moderate-Resolution Imaging Spectrometer) L3 using MATLAB (The MathWorks, Natick, MA, USA) to extract and archive monthly TIF files for the study area (NASA) Ocean Color Group (http://oceancolor.gsfc.nasa.gov). Selected environmental data included chlorophyll (Chl; mg m^{-3}), net primary productivity (NPP; mg C m^{-2} day^{-1}), and night sea surface temperature (NSST; ˚C) for the period between 2003 and 2016. Further processing of NSST monthly TIF files was performed to detect oceanic frontal activity as defined by the Cayula-Cornillion (1992) algorithm using Marine Geospatial Ecological Tools v0.8a43 (MGET; Roberts et al., 2010). To perform this NSST temperature data was initially converted to integers. We defined fronts to exist where a horizontal surface temperature differential > 0.5 °C existed between adjacent cells. This was performed for each of the monthly NSST rasters. Archived MODIS data described above was managed with R (R Development Core Team 2013) to create monthly climatologies of mean, sum, minimum and maximum values for each co-variate from the 10-year period.

Selection of Location Data

We defined seasonal extents from the observed biological characteristics of ASHWs, into two seasons based on oceanographic conditions of the north Indian Ocean and availability of sightings and telemetry data. Embryonic development of ASHWs documented in Soviet whaling records indicates the breeding season to extend from January to late May, with an eleven- month gestation period, (Mikhalev 1997). ASHW whale song, a behaviour considered to coincide with breeding cycles predominantly falls between November and May off Oman (Minton et al. 2011; Willson et al. 2016; Cerchio et al. 2016). Temporal aspects of this breeding season also align within a period of the year when oceanographic processes are influenced by the north eastly winter monsoon season between December and April (Bruce et al, 1994).

Sightings data were classified according to survey area, year, season (Appendix Table 6) and survey effort (Appendix Table 7). Sightings records were only included from dedicated vessel surveys made by visual detection (sightings classification 1-3). Third party sightings, acoustic and shore based sightings were excluded. The majority of sightings (82%) were made in the Dhofar area (n=99), with the remaining 18% falling outside of this area. Eighty five percent of sightings were also concentrated into the season between December and May (n=103) representing 164 individual whale encounters of which 26 were females, 94 males and 44 unknown (2017). Vessel surveys between 2003 and 2014 were principally conducted between the months of November and April (Minton *et al,* 2011; Willson *et al*. 2012; Willson *et al*. 2013). The limited number of sightings available between June and November (n=18) discounted the use of data from this second season (Appendix 9).

Satellite telemetry data collected between February 2014 and February 2016 were collected from eight males and one female (Appendix Table 8). The average operational period of tags was 56 days (SD=44; n=9) with a range from 1 to 164 days. Data processed with a SSSM framework using a 12-hour time step resulted in 913 locations and revealed animals spending an average of 88% of time in ‘Area Restricted Search’ behaviour mode (SD=12%) and 2% in transiting mode (SD=3%). A total of 382 telemetry days of data was derived from tags between December and May. Climatologies were only produced for these months given the paucity of sightings and telemetry data for this period.

All bathymetry and satellite data were geo-spatially aligned to a 9km grid. Values of depth (Figure 2a), slope (Figure 2b), sea surface temperature and net primary productivity were extracted at each SSSM location. Rasters indicating the locations of fronts were eliminated from modelling after poor performance in early EENM analysis.

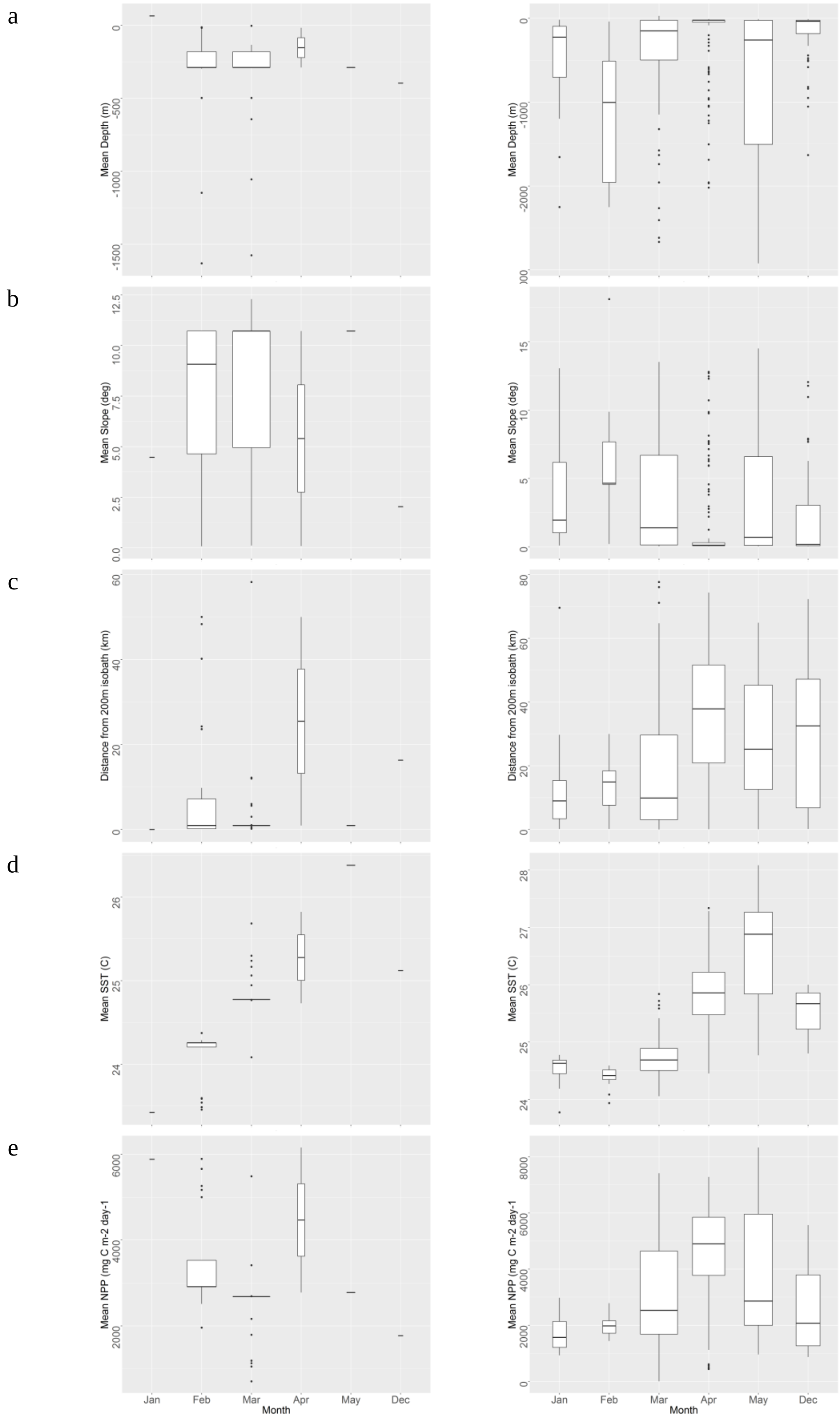

Figure 2. Median and inter-quartile ranges (box plots with outliers) for environmental co-variates extracted from December to May for sightings of humpback whales by vessels (left) and for satellite telemetry (right). Variables include; (a) mean depth (m), (b) mean slope (°), (c) mean distance from the 200m isobath, (d) mean night time sea surface temperature (°C) (e) net primary productivity (mg C m^{-2} day^{-1}). Box plot upper and low hinges set at interquartile ranges (25% and 75%), and box width proportional to the square root of the number of observations in the groups. Whisker plots are set within 1.5* of inter-quartile range and outliers represented by points.

Modelling Framework
We used the EENM approach to identify habitat suitability for ASHWs (Araujo & New, 2007; Rangel & Loyola, 2012). Models used were generalised linear model (GLM), generalised additive model (GAM), generalised boosted model (GBM) and multivariate adaptive regression splines (MARS) run within the Biomod2 package (R Development Core Team, 2008; Thuiller et al., 2013). The environmental suitability index produced as raster outputs of each model run were scored from 0 to 1, with 1 a perfect score indicating greatest habitat suitability, 0.5 areas of typical suitability and 0 and absence. The mean scores of each model type, and mean of means from all runs during an experiment were used for final performance assessment of each experiment.

Models were run using a 10-fold cross validation with a 75/25% random split of location data for calibration and model testing (Pikesley et al. 2013). Three metrics were used for evaluation of model experiments and scaled between 0 to 1. The true skill statistic (TSS) was used to determine the accuracy of the models in predicting the correct category relative to that of random chance. A measure of resolution in discriminating between two alternative events or potential usefulness was performed by the receiver operating characteristic (ROC). Cohen's Kappa coefficient (Heidke skill score; KAPPA) was also used as a measure of agreement occurring by chance and inter-raster agreement for categorical quantitative items.

The importance of environmental variables was also calculated for each model run using a randomisation process (Pikesley et al. 2013), where the correlation between a prediction of all environmental variables and a prediction where the independent variable being assessed was randomly re-ordered. Low correlations were considered important for the model. The relative importance of each variable was calculated from the mean of the correlation coefficient over multiple runs (Thuiller et al. 2009), followed by subtracting these means from 1.

RESULTS

Model evaluation metrics including ROC, KAPPA and TSS scored above 0.9 for both vessel sightings and satellite telemetry data (Table 1). Comparison between environmental suitability index values (Table 2) calculated by the difference raster datasets reveals both sightings and telemetry data to be equal in contribution to the strength of index values (mean= 0.0, SD=0.02).

Variation was found between the influence of the different co-variate data from telemetry and sighting data (Table 3). Distance from the 200m isobath, sea surface temperature and net primary productivity were the primary components from the telemetry data (mean VI= 0.35 (SD 0.04), 0.44 (SD 0.07) and 0.22 (SD 0.09) respectively). Depth, slope and net primary productivity were the primary components of vessel sightings data (mean variable importance= 0.24 (0.06), 0.35 (0.14) and 0.45 (0.12) respectively).

Table 1. Summary of ecological niche modelling evaluation metrics for 10- fold cross validation. Abbreviations: Generalised Linear Model (GLM), Multivariate Adaptive Regression Splines (MARS) and Generalised Boosted Model (GBM).

MODEL EVALUATION METRICS	VESSEL SIGHTINGS			SATELLITE TELEMETRY		
	GLM	MARS	GBM	GLM	MARS	GBM
ROC	0.99	0.96	1.00	0.99	0.99	0.99
KAPPA	0.96	0.91	0.96	0.96	0.97	0.97
TSS	0.96	0.91	0.96	0.96	0.96	0.96

Table 2. Differences between environmental suitability index values from sightings and telemetry models (for min, max and mean values) exported from raster subtraction in ARC GIS v10.3. Positive numbers show dominance of the sightings model, negative models show dominance of the telemetry model.

	MIN	MAX	MEAN	STDDEV
GLM	-0.97	0.99	-0.01	0.18
MARS	-0.98	0.98	0.01	0.15
GBM	-0.99	0.99	-0.01	0.14
MEAN	-0.97	0.92	0.00	0.13
STDEV	0.01	0.01	0.02	

Table 3. Summary of ecological niche modelling variable importance from 10 fold cross validation for vessel surveys and satellite telemetry data.

	ENVIRONMENTAL CO-VARIATE DATA	GLM	MARS	GBM	MEAN OF MEANS	SD OF MEANS	RANK
VESSEL SIGHTINGS	Depth	0.44	0.25	0.02	0.24	0.06	3
	Slope	0.28	0.60	0.18	0.35	0.14	2
	Distance from 200m isobath	0.00	0.00	0.00	0.00	0.00	5
	NSST	0.00	0.18	0.01	0.06	0.07	4
	NPP	0.13	0.85	0.36	0.45	0.12	1
SATELLITE TELEMETRY	Depth	0.03	0.06	0.08	0.06	0.04	4
	Slope	0.04	0.02	0.00	0.02	0.03	5
	Distance from 200m isobath	0.39	0.32	0.34	0.35	0.04	2
	NSST	0.42	0.50	0.41	0.44	0.07	1
	NPP	0.32	0.29	0.23	0.28	0.09	3

Modelling outputs (Figure 3 and Figure 4) revealed similar distribution of environmental suitability from both the sightings and telemetry driven EENMs. Environmental suitability >0.75 is projected from south to north along the continental shelf of Oman. The sightings model predicts similar near-shore habitat along the northern shores of the Arabian Sea spanning between the Straits of Hormuz in the west and Gulf of Kutch in the east. Only the GBM model predicts a similar distribution for telemetry data, with ensemble of the models for this dataset only predicting suitability in the area straddling the EEZ of Pakistan and India. The ensemble telemetry model also indicates suitable habitat north west of the entrance to the Red Sea around the Hanish Archipelago. A narrow margin of high suitability is highlighted along the coastal fringes of Sudan and Saudi Arabia in the central Red Sea by the GLM and MARS telemetry model.

Difference rasters highlighting variation in predictions between the two types of EENMs, one modelled using boat data, the other satellite telemetry highlight the difference between the sightings and telemetry data (Figure 5). The ensemble difference raster reveals dominance of the telemetry outputs for all models along the coast of Oman further from the coast, with both sightings and telemetry datasets performing equally closer to shore. The telemetry model is also dominant in the area offshore from the Gulf of Kutch with the area of high environmental suitability that continues towards the canyon area offshore from the Indus delta. The differences of the models also highlight prediction of sightings model of moderate suitability to the west coast of India and high suitability of an isolated pocket in the Gulf of Rambhat (north west India). Moderate suitability (0.5 - <0.75) is also picked up by this model in the Sea of Oman.

The mean of the ensemble rasters of both datasets (Figure 6) presents a good fit with humpback whale catches by the Soviet whaling fleets between 1962 and 1966 as reported by Mikhalev (2000; IWC Catch Database, extracted 2013). This raster also shows close flanking of whale captures to the area along the Gulf of Kutch, and continues towards the canyon area offshore from the Indus delta. This area is characterised by shelf-brake, with slope and high net primary productivity.

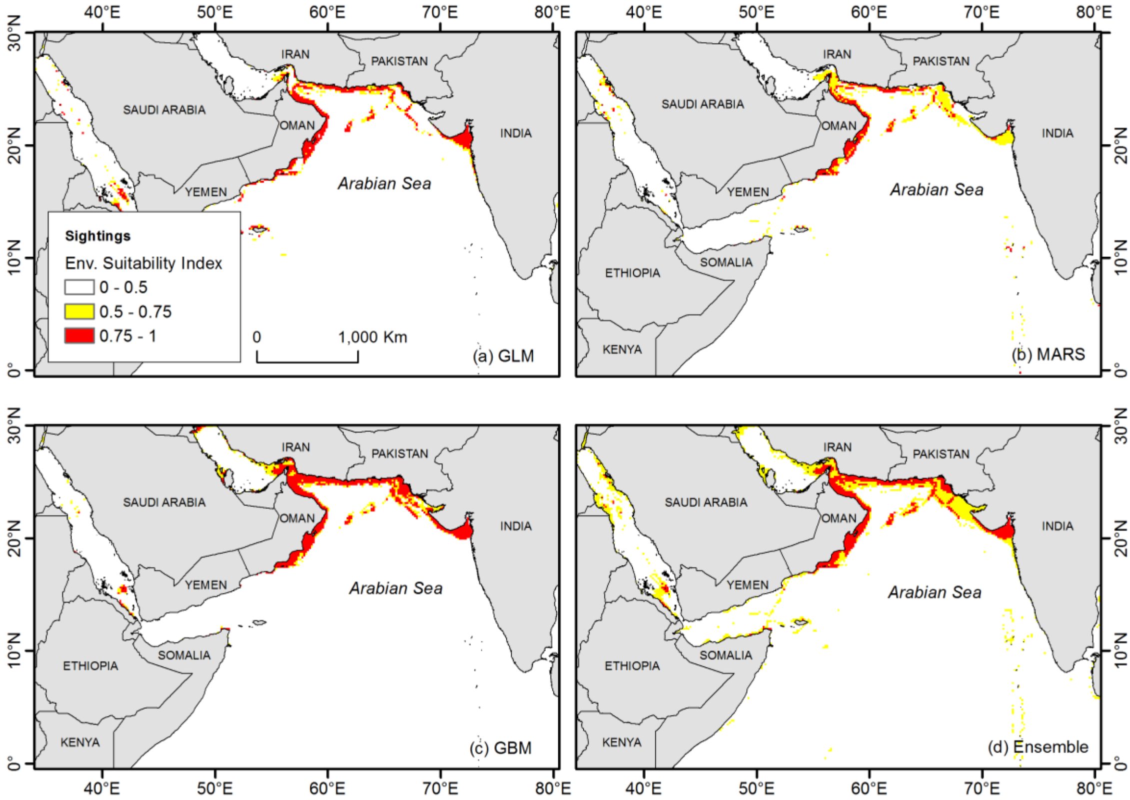

Figure 3 Ecological niche model and ensemble model for Oman humpback whale sightings data (2003-2014, December to May), using (a) Generalised Linear Model (GLM), (b) Multivariate Adaptive Regression Splines (MARS), (c) Generalised Boosted Model and (d) the ensemble modelling algorithms from 'Biomod2' package (R Development Core Team, 2008; R package: biomod2; Thuiller et al., 2013).

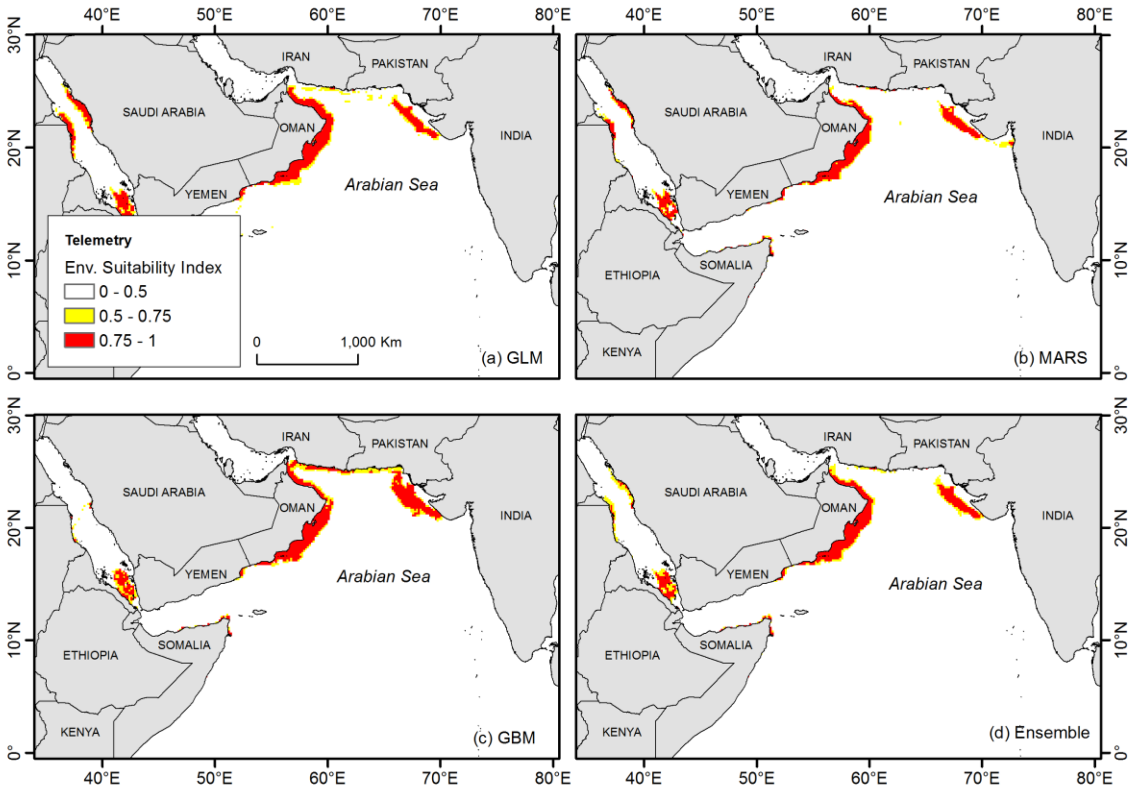

Figure 4 Ecological niche model and ensemble model for Oman satellite telemetry data (2003-2014, December to May), using a) Generalised Linear Model (GLM), b) Multivariate Adaptive Regression Splines (MARS), c) Generalised Boosted Model and d) the ensemble modelling algorithms from 'Biomod2' package (R Development Core Team, 2008; R package: biomod2; Thuller et al., 2013).

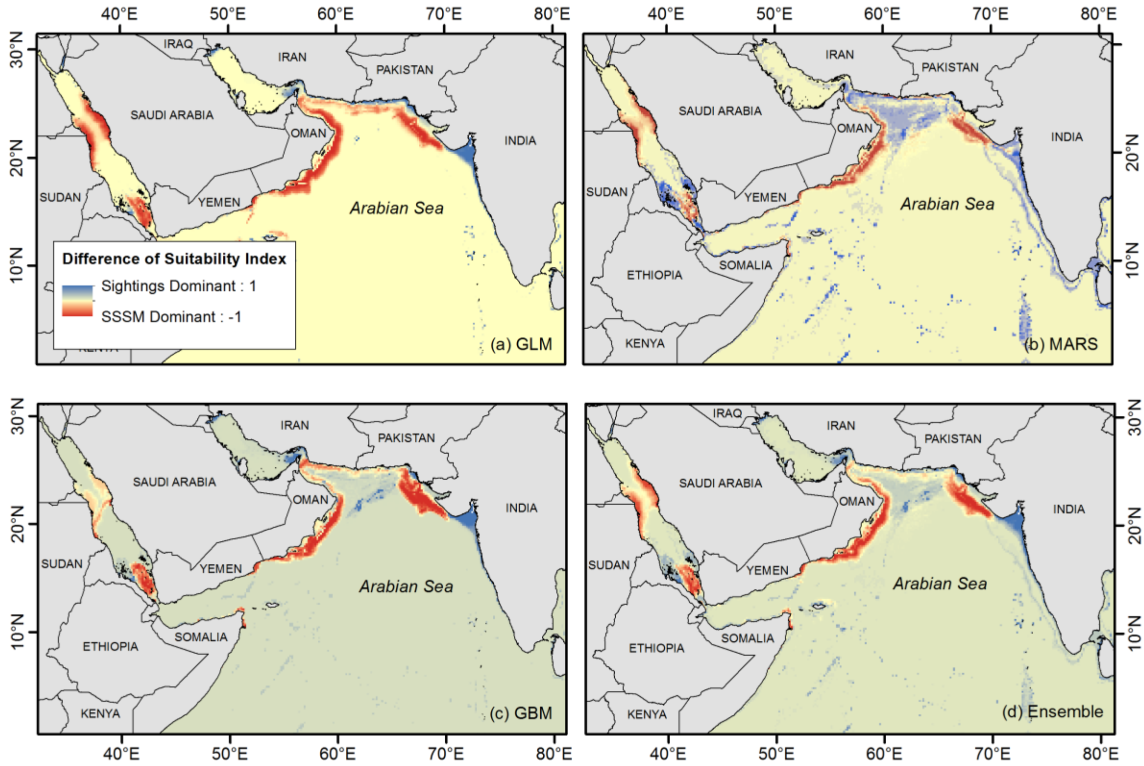

Figure 5 Difference between ecological niche model rasters from Oman whale sightings data and Oman satellite telemetry data focusing a) Generalised Linear Model (GLM), b) Multivariate Adaptive Regression Splines (MARS), c) Generalised Boosted Model and d) the ensemble model.

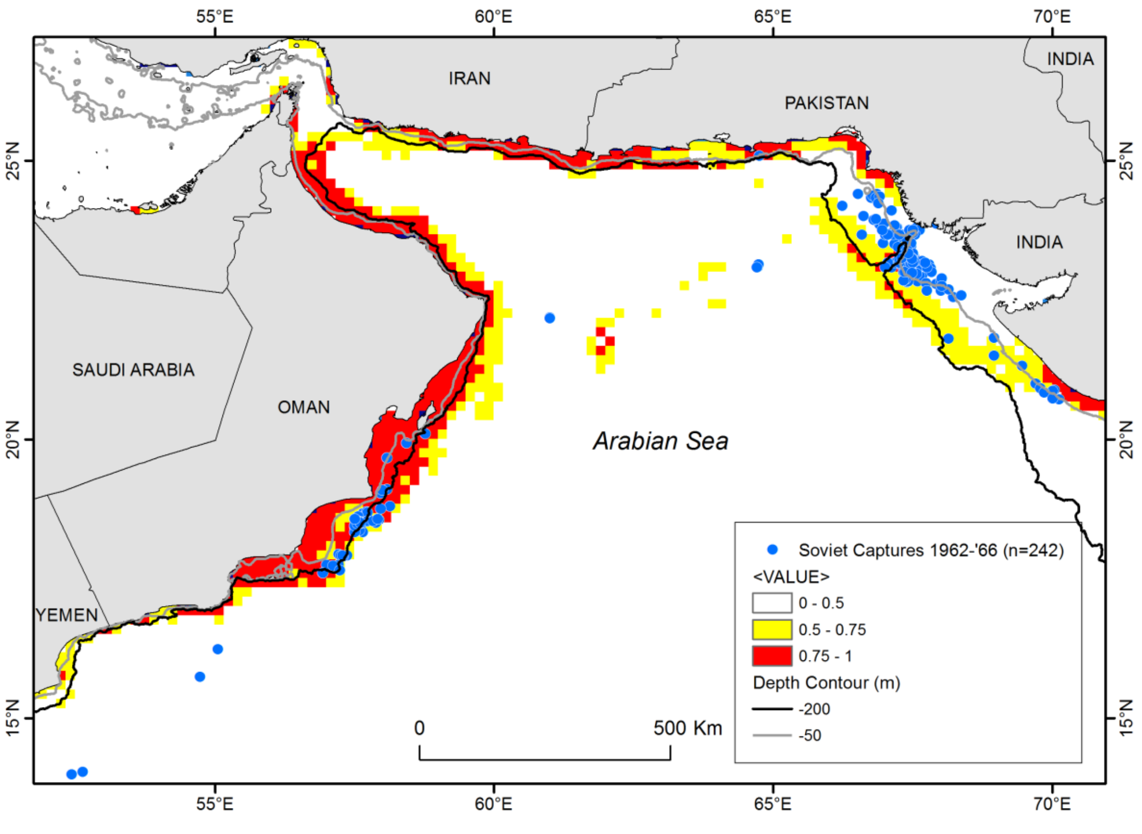

Figure 6 Average of sightings and satellite telemetry ensemble ecological niche models with overlay of historical takes of, humpback whales in the Northern Indian Ocean as documented in Soviet whaling records between 1962 and 1966, (Mikhalev, 2000; IWC Catch Database, extracted 25 October 2013).

DISCUSSION

Knowledge of the distribution and habitat utilisation of ASHWs has been constrained across range states by the access to resources, scientific capacity and seasonally inclement weather conditions that hinder implementation of comprehensive surveys (Willson *et al.* 2016). Our approach to use vessel surveys and telemetry data within a mixed model framework has provided the first range wide predictive models to be developed for the ASHW. The final mean raster of ensemble models addresses the constraints of both data sources to facilitate biological and geographical representation. This output of the study appears to geographically fit well with Soviet whaling capture records, even though the catches were made just prior to the season we performed the modelling for (early to mid-November). If this is a true fit then it is possible the environmental conditions in the Arabian Sea influencing whale distribution are similar in recent years (up to 2016) as they were between 1962 and 1966.

Understanding the co-occurrence of whales with shipping has been noted as a priority by the Arabian Sea Whale Network (Minton et al, 2015), with ecological niche modelling identified as a relevant approach to predict areas of risk. In a preliminary study to understand the association between whales and shipping in the NIO, container shipping was noted to have increased three-fold between 2004 and 2014 (Willson et al. 2016b). Traffic density routes derived from satellite based Automatic Identification System vessel tracks from this previous study show co-occurrence with the predicted areas of high environmental suitability for ASHWs along the coast of southern Oman and northern Arabian Sea between Pakistan and India. A similar overlap with vessel traffic density is also noted by comparison of these same routes with predicted distribution of blue habitat during the northeast and southwest monsoons (Redfern *et al*. 2017) and the Soviet whaling data (Mikhalev, 1997).

Together, models predicting habitat of multiple species should be used to help direct priorities for further investigations of the shipping and whale issue in the Arabian Sea. However, to address management interventions additional field based studies will be required to meet ‘robust risk assessment analysis’ (Silber, 2012) and ‘best available science’ required by the International Maritime Organisation (IMO, 2009). This would necessitate conducting surveys in priority areas to calculate abundance estimates. The latest approaches to shipping and whale management include use of ENM together with abundance estimates to provide a near-real-time predictive model of whale habitat as a tool to advise selection of shipping routes (Hazen *et al.* 2016). Design of future studies in the Arabian Sea should be considered with this approach in mind.

The exercise highlighted interesting differences between the variable importance of environmental parameters between the sightings and telemetry datasets, with only net primary productivity considered important to both. This may be a facet of the limited range and sample size of sightings data. Inputs to the sightings model were limited in number (n=99) and biased in geographic range with 82% of data points sourced from Dhofar across a total range between points of 190km. This data was further constrained by the limitations that small vessel surveys presented to obtain sightings further from shore with individual whales. The telemetry model draws from nine times more location points (n=913) that span 860km between furthest points and is unconstrained by distance of detection from shore a comparison that promotes the utility of this data collection method for ENM. For our defined study periods only one female was instrumented with a telemetry unit out of nine tags (11%) and only 26 females (16%) were identified out of 164 individual encounters from vessel surveys (with sex in 44 encounters unidentified), OMCD,2017). This bias needs to be addressed for future surveys.

CONCLUSION

Although caveats exist, the model outputs are considered consistent, and should be used together with recent blue whale habitat ENM work to guide large whale research and management strategy in the Arabian Sea. The habitat predictions for ASHWs could be immediately strengthened by incorporation of sightings data from other range states. Future work should be expedited to address spatial and temporal gaps in ecological models including the implementation of robust abundance estimates off southern Oman and the northern coasts of the Arabian Sea, and continuing to promote use of remote sensing technology (such as passive acoustics and satellite telemetry) to overcome the issue of detecting whales during the summer monsoon. Planning should be driven by drafting of a comprehensive multi-species approach for large whale research in the Arabian Sea to ensure compatibility of outputs to inform ship strike risk assessments and subsequent mitigation measures.

ACHKNOWLEDGEMENTS

We are grateful to the Ministry of Agriculture and Fisheries Wealth & Ministry of Environment and Climate Affairs, Oman, for participation of staff in field activities and issuing of permits to conduct field research, sampling and analysis. We thank all staff at the Environment Society of Oman (www.eso.org.om) through whom the research project is run and the network of volunteers for continued support of the Renaissance Whale and Dolphin Project in Oman. The team would also like to recognise the outstanding professional support of Darryl (MacGyver) MacDonald, for contribution of his wide-ranging paramedic, photographic, mechanical and whale-finding expertise. Appreciation is expressed for the staff at Five Oceans Environmental Services working in the background for provision of expertise, time and patience in dealing with off the clock whale team related demands. The work is also deeply indebted to the technical expertise, resources and patience of personnel from the Marine Mammal Laboratory of the Alaska Fisheries Center and from Instituto Aqualie, without which the tagging work would not have been possible. Finally, sincere thanks are given for the unwavering faith and financial support provided to this project by Renaissance Services S.A.O.G since 2011.

REFERENCES

Aarts, G., MacKenzie, M., McConnell, B., Fedak, M. and Matthiopoulos, J., 2008. Estimating space-use and habitat preference from wildlife telemetry data. *Ecography*, *31*(1), pp.140-160.

Araújo, M.B. and New, M., 2007. Ensemble forecasting of species distributions. *Trends in ecology & evolution*, *22*(1), pp.42-47.

Baldwin, R.M., Gallagher, M.D. and Van Waerebeek, K. 1999. A review of cetaceans from waters off the Arabian Peninsula. In: Oman's Natural History, eds. Fisher, M., Spalton, A. and Gazanfar, S., Backhuys Publishers, Leiden. Pp. 161-189.

Becker, E.A., Foley, D.G., Forney, K.A., Barlow, J., Redfern, J.V. and Gentemann, C.L., 2012. Forecasting cetacean abundance patterns to enhance management decisions. *Endangered Species Research*, *16*(2), pp.97-112.

Breed, G. A., Jonsen, I. D., Myers, R. A., Bowen, W. D., & Leonard, M. L. (2009). Sex-specific, seasonal foraging tactics of adult grey seals (Halichoerus grypus) revealed by state-space analysis. *Ecology*, *90*(11), 3209-3221.

Breed, G. A., Costa, D. P., Jonsen, I. D., Robinson, P. W., & Mills-Flemming, J. (2012). State-space methods for more completely capturing behavioral dynamics from animal tracks. *Ecological Modelling*, *235*, 49-58.

Breed, G. A., Costa, D. P., Goebel, M. E., & Robinson, P. W. (2011). Electronic tracking tag programming is critical to data collection for behavioral time-series analysis. *Ecosphere*, *2*(1), art10.

Brown SG (1957) Whales observed in the Indian Ocean: notes on their distribution. Mar Obs 27: 157−165

Bruce, J.G., Johnson, D.R. and Kindle, J.C., 1994. *Evidence for eddy formation in the eastern Arabian Sea during the northeast monsoon* (No. NRL/JA7332-93-0010). NAVAL RESEARCH LAB STENNIS SPACE CENTER MS.

Corkeron, P, Minton, G., Collins, T., Findlay, K., Willson, A., and Baldwin., R . 2011. Spatial models of sparse data to inform cetacean conservation planning: an example from Oman. Endangered Species Research Vol. 15:39-52.

GEBCO Digital Atlas published by the British Oceanographic Data Centre on behalf of IOC and IHO, 2003

Guisan, A. and Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecological modelling*, *135*(2), pp.147-186.

Harris, P.T., Macmillan-Lawler, M., Rupp, J. & Baker, E.K. (2014) Geomorphology of the oceans. Marine Geology, 352, 4–24.

Hazen, E.L., Palacios, D.M., Forney, K.A., Howell, E.A., Becker, E., Hoover, A.L., Irvine, L., DeAngelis, M., Bograd, S.J., Mate, B.R. and Bailey, H., 2016. WhaleWatch: a dynamic management tool for predicting blue whale density in the California Current. *Journal of Applied Ecology*.

Kaschner, K., Watson, R., Trites, A.W. and Pauly, D., 2006. Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model. *Marine Ecology Progress Series*, *316*, pp.285-310.

Moore, S.E., Waite, J.M., Friday, N.A. and Honkalehto, T., 2002. Cetacean distribution and relative abundance on the central–eastern and the southeastern Bering Sea shelf with reference to oceanographic domains. *Progress in oceanography*, *55*(1), pp.249-261.

Mikhalev, Y. A. 1997. Humpback whales *Megaptera novaeangliae* in the Arabian Sea. Marine Ecology Progress Series 149:13-21.

Mikhalev, Y.A. 2000. Whaling in the Arabian Sea by the whaling fleets Slava and Sovetskaya Ukraina. pp.141–81. In: Yablokov, A.V., Zemsky, V.A. and Tormosov, D.D. (eds). Soviet Whaling Data (1949– 1979). Centre for Russian Environmental Policy, Moscow. 408pp.

Notarbartolo di Sciara G, Smeenk C, Rudolph P, Addink M and others (2014) Summary review of cetaceans of the Red Sea. Paper presented to the IWC Scientific Committee, May 2014, Cambridge

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., O'Connell, A.F., Miller, P.I. and Louzao, M., 2012. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. *Biological Conservation*, *156*, pp.94-104.

Pikesley, S.K., Maxwell, S.M., Pendoley, K., Costa, D.P., Coyne, M.S., Formia, A., Godley, B.J., Klein, W., Makanga-Bahouna, J., Maruca, S. and Ngouessono, S., 2013. On the front line: integrated habitat mapping for olive ridley sea turtles in the southeast Atlantic. *Diversity and Distributions*, *19*(12), pp.1518-1530.

Piontkovski, S.A., Al-Oufi, H.S. and Al-Jufaili, S., 2014. Seasonal and interannual changes of Indian oil sardine, Sardinella longiceps landings in the governorate of Muscat (the Sea of Oman). *Marine Fisheries Review*, *76*(3), pp.50-60.

Pomilla, C. Amaral, A., Collins, T., Minton, G., Findlay, K., Leslie, M., Ponnampalam, L., Baldwin, R., Rosenbaum, H. 2014. The World's Most Isolated and Distinct Whale Population? Humpback Whales of the Arabian Sea. *PLoS ONE*, 9(12), p.e114162. Available at: http://dx.plos.org/10.1371/journal.pone.0114162.

Randall, J.E. 1995. The Complete Divers' and Fishermen's Guide to Coastal Fishes of Oman. Crawford House Publishing, Bathurst, Australia. 2795. pp 63-68.

Redfern, J.V., Ferguson, M.C., Becker, E.A., Hyrenbach, K.D., Good, C.P., Barlow, J., Kaschner, K., Baumgartner, M.F., Forney, K.A., Ballance, L.T. and Fauchald, P., 2006. Techniques for cetacean–habitat modeling.

Redfern, J.V., McKenna, M.F., Moore, T.J., Calambokidis, J., Deangelis, M.L., Becker, E.A., Barlow, J., Forney, K.A., Fiedler, P.C. and Chivers, S.J., 2013. Assessing the risk of ships striking large whales in marine spatial planning. *Conservation Biology*, *27*(2), pp.292-302.

Slijper, E.J., van Utrecht, W.L. and Naaktgeboren, C. 1964. Remarks on the distribution and migration of whales, based on observations from Netherlands ships. Bijdr. Dierkd. 34: 3–93.

Širović, A. and Hildebrand, J.A., 2011. Using passive acoustics to model blue whale habitat off the Western Antarctic Peninsula. *Deep Sea Research Part II: Topical Studies in Oceanography*, *58*(13), pp.1719-1728.
Smith, J.N., Grantham, H.S., Gales, N., Double, M.C., Noad, M.J. and Paton, D., 2012. Identification of humpback whale breeding and calving habitat in the Great Barrier Reef. *Marine Ecology Progress Series*, *447*, pp.259-272.

Sugimoto, T., and H. Tameishi. 1992. Warmcore rings, streamers and their role on the fi shing ground formation around Japan. Deep-Sea Res. Pt. A 39:S183–S201

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B. (2009) BIOMOD–a platform for ensemble forecasting of species distributions. Ecography 32, 369–373.

Tameishi, H., Y. Naramura, and H. Shinomiya. 1994. Role of warm streamers in the northward migration of Japanese sardine (Sardinops melanostictus) off Sanriku (Japan). Bull. Japan. Soc. Sci. Fish. 60:45-50.
Thompson, P.M., Brookes, K.L. and Cordes, L.S., 2015. Integrating passive acoustic and visual data to model spatial patterns of occurrence in coastal dolphins. *ICES Journal of Marine Science: Journal du Conseil*, *72*(2), pp.651-660.

Willson, A., Baldwin, R., Minton, G., Gray, H., Findlay, K., Collins, T. 2013. Arabian Sea humpback whale research update for 2012/13. Paper SC/65a/SH06 presented to the International Whaling Commission Scientific Committee, Jeju, South Korea, June 2013. 08pp. (Available from the IWC Office).

Willson, A., Collins, T. Baldwin, R., Cerchio, S., Geyer, Y., Godley B., Gray, H., Al-Harthi, S., Minton, Al-Zehlawi, N., M.Witt., Rosenbaum, H., Zerbini, A. 2014. Preliminary results and first insights from satellite tracking studies of male Arabian Sea humpback whales. Paper SC/65b/SH19 presented to the International Whaling Commission Scientific Committee, Slovenia, May 2014. (Available from the IWC Office).

Willson, A., Baldwin, R., Cerchio, S., Collins, T. Findlay, K., Gray, H., Godley B., Gray, H., Al-Harthi, S., Kennedy, A., Minton, Sucunza, F., Zerbini, A., Witt, M. 2016a. Research update on satellite tagging studies of the Arabian Sea humpback whale in the Sultanate of Oman. Paper SC/66b/SH28 presented to the International Whaling Commission Scientific Committee, Slovenia, June 2016. (Available from the IWC Office).

Willson, A., Kowalik, J., Godley, J., Baldwin, R., Struck, A., Struck, L., Nawaz, R. and Witt, M. 2016b. Priorities for addressing whale and ship strike co-occurence off the coast of Oman and the wider North Indian Ocean. Paper SC/66b/HIM/10 presented to the International Whaling Commission Scientific Committee, Slovenia, June 2016. Available from the IWC Office).

Wray, P. and Martin, K.R. 1983. Historical whaling records from the western Indian Ocean. Rep. int. Whal. Commn (special issue)(5): 213–41.

APPENDIX

Table 4. Count of survey effort (days) in Dhofar, Gulf of Masirah and Muscat.

SURVEY AREA	2003	2004	2005	2006	2010	2011	2012	2014	GRAND TOTAL
DHOFAR	16	26				24	26	14	106
GULF OF MASIRAH		16		15	6	7	12		56
MUSCAT	19	3	42	6					70
GRAND TOTAL	**35**	**45**	**42**	**21**	**6**	**31**	**38**	**14**	**274**

Table 5. Count of survey effort (hours) in Dhofar, Gulf of Masirah, and Muscat.

SURVEY AREA	2003	2004	2005	2006	2010	2011	2012	2014	GRAND TOTAL
DHOFAR	177	68				169	194	87	884
GULF OF MASIRAH		164		101	35	36	92		428
MUSCAT	68	18	286	52					424
GRAND TOTAL	**245**	**250**	**286**	**153**	**35**	**205**	**286**	**87**	**2375**

Table 6. Counts of sightings (visual encounters) from dedicated vessel surveys collated by area and year.

AREA	2003	2004	2006	2009	2011	2012	2013	2014	GRAND TOTAL
DHOFAR	16	7			28	7	1	40	99
GULF OF MASIRAH		6	4		5	3	1		19
MUSCAT	1		1						2
OTHER				1					1
GRAND TOTAL	**17**	**13**	**5**	**1**	**33**	**10**	**2**	**40**	**121**

Table 7. Counts of sightings in defined season according to effort level; 1= On effort, 2= Sub-optimal effort and 3= Off effort.

SEASON	1	2	3	GRAND TOTAL
1 - DEC TO MAY	37	17	47	103
2 – JUN TO NOV	7	4	7	18
GRAND TOTAL	**44**	**21**	**54**	**121**

Table 8. Summary of tag deployment details, (sex determined by molecular identification or behaviour).

	DEPLOYMENT DETAILS						**SSSM MODE PERCENTAGES**		
INDIVIDUAL CODE	Perm. Whale ID Code	Deploy date (dd/mm/yy)	Deploy location	Sex (M/F/?)	Social category	Tag longevity (days)	Area Restricted Search (%)	Undefined (%)	Transiting (%)
/	OM02-020	21/02/2014	Hasik	Male	Adult Pair	1	N/A	N/A	N/A
/	OM11-002	23/02/2014	Hasik	Male	Single	1	N/A	N/A	N/A
A	OM14-013	22/11/2015	GoM	Male	Single	18	100	0	0
B	OM15-004	23/11/2015	GoM	Female	Single	23	79	13	7
C	OM01-014	10/03/2015	GoM	Male	Adult Pair	25	96	3	1
D	OM02-019	25/02/2014	Hasik	Male	Adult Pair	41	100	0	0
E	OM00-003	28/02/2014	Hasik	Male	Single	42	71	29	0

F	OM10-001	22/02/2014	Hasik	Male	Single	55	79	16	5
G	OM15-002	21/11/2015	GoM	Male	Adult Pair	62	100	0	0
H	OM02-019	13/03/2015	GoM	Male	Adult Pair	77	100	0	0
I	OM01-006	14/03/2015	GoM	Male	Adult Pair	163	65	26	9
AVERAGE						56	88	10	2
SD						21	12	11	3

Table 9. Summary of Number of Sighting Counts (2003-2014) and Tracking Days (per animal 2014-2015) within each month.

	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	TOTAL
VESSEL SIGHTING COUNTS	1	1	46	52	2	0	0	0	1	0	6	12	121
DAYS OF TELEMETRY DATA	50	10	22	168	94	38	30	31	10	0	0	34	437