Identification of humpback whale breeding and calving habitat in the Great Barrier Reef

Joshua N. Smith[1,5,][*], **Hedley S. Grantham**[2], **Nick Gales**[3], **Michael C. Double**[3],
Michael J. Noad[1], **David Paton**[4]

[1]**Cetacean Ecology and Acoustics Laboratory, School of Veterinary Science, University of Queensland, Gatton, Queensland 4343, Australia**

[2]**School of Biological Sciences, University of Queensland, St. Lucia, Queensland 4072, Australia**

[3]**Australian Marine Mammal Centre, Australian Antarctic Division, Kingston, Tasmania 7050, Australia**

[4]**Blue Planet Marine, PO Box 919, Jamison Centre, Australian Capital Territory 2614, Australia**

[5]*Present address*: **Murdoch University Cetacean Research Unit, Murdoch University, Western Australia 6150, Australia**

ABSTRACT: During the winter months, from June to September, humpback whales *Megaptera novaeangliae* breed and calve in the waters of the Great Barrier Reef (GBR) after migrating north from Antarctic waters. Clearly defined wintering areas for breeding and calving comparable to those identified in other parts of the world have not yet been identified for humpback whales in the GBR Marine Park (GBRMP), mainly because of its large size, which prohibits broad-scale surveys. To identify important wintering areas in the GBRMP, we developed a predictive spatial habitat model using the Maxent modelling method and presence-only sighting data from non-dedicated aerial surveys. The model was further validated using a small independent satellite tag data set of 12 whales migrating north into the GBR. The model identified restricted ranges in water depth (30 to 58 m, highest probability 49 m) and sea surface temperature (21 to 23°C, highest probability 21.8°C) and identified 2 core areas of higher probability of whale occurrence in the GBRMP, which correspond well with the movements of satellite tagged whales. We propose that one of the identified core areas is a potentially important wintering area for humpback whales and the other a migration route. With an estimated increase in port and coastal development and shipping activity in the GBRMP and a rapidly increasing population of whales recovering from whaling off the east Australian coast, the rate of human interactions with whales is likely to increase. Identifying important areas for breeding and calving is essential for the future management of human interactions with breeding humpback whales.

KEY WORDS: Humpback whale · *Megaptera novaeangliae* · Spatial habitat modelling · Great Barrier Reef

INTRODUCTION

Understanding the relationship between a species' distribution and its environment is not only a fundamental requirement of ecology, it is also extremely important for conservation planning (Guisan & Thuiller 2005). For cryptic and highly mobile marine species that range over large areas and are difficult to survey, such as many marine mammals, critical habitat required for successful breeding and foraging is often poorly defined and/or understood (Harwood 2001). The difficulty of surveying marine mammals arises from most of their time being spent underwater and their being difficult to observe visually, often necessitating complementary methods such as acoustic surveys (Evans 2008). Consequently, moni-

toring of marine species in marine ecosystems can be expensive and logistically prohibitive (Redfern et al. 2006). Increased human use of the marine environment has meant that human activities are posing a greater threat to the persistence of marine mammal populations (Evans 2008). Identifying critical habitat is therefore particularly important in areas where human activities are likely to have the greatest impact so that any deleterious human impacts can be mitigated. Predictive modelling of species' distributions and their habitat is one approach that is becoming increasingly popular and proving to be a cost-effective approach when applied to marine mammal populations that are difficult to monitor (Redfern et al. 2006).

Compared with critical habitat for terrestrial species, the identification and designation of critical habitat in the marine environment can be more complex and the boundaries may be less fixed (Evans 2008). The difficulty in identifying critical habitat largely depends on whether the habitat requirements are dependent on stable, geographical environmental parameters such as depth or topography or more fluid oceanographic parameters such as temperature, salinity and ocean current (Evans 2008). For migratory marine species there are often considerable differences in the type of habitat and resources used when feeding compared with breeding. Furthermore, this can be influenced by the species' social organisation (Ersts & Rosenbaum 2003). Humpback whales *Megaptera novaeangliae* are a highly mobile, migratory species that annually undertakes seasonal migrations from temperate or sub-polar feeding areas in high latitudes to tropical breeding and calving habitat in low latitudes (Dawbin 1966). This migration is the longest known for any mammal (Rasmussen et al. 2007, Stevick et al. 2011) and represents a very distinct temporal and geographic separation of activities on an annual cycle, resulting in whales inhabiting very disparate environments over the year. Although their migration can occur over deep oceanic waters, humpback whales are typically a coastal species when feeding and breeding.

Currently, there is limited understanding of the habitat requirements and specific habitat features that breeding humpback whales rely on in wintering areas. That mating has never been observed in humpback whales, despite over 40 yr of research globally, is a contributing factor to this lack of understanding (Clapham 2000). General features of well-known humpback whale wintering areas indicate that many occur close to 20° latitude in both the southern and northern hemispheres (Clapham & Mead 1999, Rasmussen et al. 2007) and that humpback whales generally prefer warm (21.1 to 28.3°C) and shallow waters (15 to 60 m depth) (Whitehead & Moore 1982, Ersts & Rosenbaum 2003, Rasmussen et al. 2007). Depending on the social organization and reproductive status of whales, there also appears to be differential utilisation of habitat (Craig & Herman 2000, Ersts & Rosenbaum 2003). Specifically, females with calves show strong preference for areas of shallow, sheltered water close to the coastline (Ersts & Rosenbaum 2003, Félix & Botero-Acosta 2011). It is possible that factors such as water temperature influence the distribution of humpback whales at a broad scale, whereas the availability of suitable reproductive habitat can influence it at a local scale (Rasmussen et al. 2007).

In Australia, the wintering area for humpback whales off the east coast remains poorly defined, with mating and calving generally considered to occur somewhere in the Great Barrier Reef Marine Park (GBRMP), predominantly south of 21°S (Paterson & Paterson 1984, Simmons & Marsh 1986, Chaloupka & Osmond 1999). As yet, no specific wintering area comparable to those identified in other parts of the world (e.g. Hawaii and the West Indies) (Mattila et al. 1994, Craig & Herman 2000, Johnston et al. 2007) have been identified. The GBRMP is a multiple-use marine park, supporting a wide range of uses, that extends in latitude from approx. 10° to 25°S and has a total area of approx. 344 400 km^2 (GBRMPA 2009). It is prohibitively expensive to survey in its entirety and consequently few studies have been undertaken on the distribution of humpback whales throughout the entire extent of the GBRMP (Paterson & Paterson 1984, Simmons & Marsh 1986, Chaloupka & Osmond 1999). Low numbers of whales in the past as a result of whaling have also hindered attempts to refine aerial surveys to focused areas of interest.

The population of humpback whales off the east Australian coast [IWC substock E(i)] were hunted to near extinction in the 1950s and early 1960s, but since the early 1980s it has rapidly increased in size (Paterson 2001, Noad et al. 2008). The population estimate for 2011 is 14 600 whales, based on a long-term average rate of increase of 10.9% per annum (Noad et al. 2008). The rapid rate of increase of the east Australian population of humpback whales and the now higher numbers of whales could allow for a more definitive assessment of breeding areas, however, recent substantial increases in coastal and port development are among the major management

issues affecting the Great Barrier Reef (GBR). This, along with an associated increase in shipping, will be emerging management issues for the Great Barrier Reef Marine Park Authority (GBRMPA) when managing a rapidly increasing breeding population of humpback whales within the Marine Park (GBRMPA 2009).

While still in its early stages, predictive habitat modelling of cetaceans is proving to be a useful and cost-effective approach for predicting a species' distribution and understanding the ecological processes underlying their distribution (Redfern et al. 2006). Predictive habitat models provide a prediction of occurrence for a species by relating a set of presence localities to relevant environmental variables (Phillips et al. 2006). In the past, the methods for predictive habitat modelling have often required presence/absence or abundance data from systematically designed surveys (Elith & Leathwick 2009). For marine mammals, the collection of absence data can often be difficult and expensive, particularly over large spatial scales. More recent approaches are being applied to modelling presence-only data and one method that performs well in its predictive accuracy is Maxent. Since becoming available in 2004, Maxent has been used for a wide range of ecological applications (Phillips et al. 2006, Elith & Leathwick 2009). This method uses the maximum entropy method, which essentially models a species distribution from a probability distribution of the geographic area of interest (background data) that most closely approximates the empirical average of a target distribution or set of known occurrence points (Phillips et al. 2006, Elith et al. 2010).

A large presence-only data set of whale sightings exists from the Border Protection Command (BPC) aerial surveillance program, which was used in this study. This study aims to analyse existing sighting data to determine the distribution of humpback whales in the GBRMP using the predictive habitat modelling method Maxent to identify potentially important wintering areas for mating and calving. The presence-only Maxent method was chosen because flight paths are not known and there are no effort or environmental condition data associated with the BPC sightings. The outputs from this approach can be used to inform the relationship between humpback whale distribution and the environment in wintering areas as well as for the design of dedicated, systematic aerial surveys aimed at refining our understanding of the distribution of breeding humpback whales in the GBRMP.

MATERIALS AND METHODS

The Great Barrier Reef Marine Park

The GBRMP has a total area of approx. 344 400 km^2 (excluding some islands and internal Queensland waters such as estuaries) and constitutes 99% of the Great Barrier Reef World Heritage Area (GBRWHA; 347 800 km^2). The GBRMP is divided up into 4 management areas and both the GBRMP and GBR-WHA are jointly managed by the Australian Federal (GBRMPA) and Queensland State governments. The GBR extends for approx. 2300 km adjacent to the Queensland coastline and within the GBRMP there are almost 3000 reefs. Coral reefs comprise only 6% of the geographical area of the GBRWHA, whereas the inter-reefal habitats (habitats between the individual coral reefs) comprise 25% of the area, the GBR lagoon (the area between the reefs and the coast) makes up 33% of the area and the remaining 36% of the area is the continental slope that extends past the outer coral reefs down into deep water off the continental shelf (GBRMPA 2009).

BPC Coastwatch sighting data

Humpback whale sighting data from 2003 to 2007 were obtained from the Customs BPC Coastwatch aerial surveillance program, a sub-program of the Australian Customs service. Coastwatch services are available to a range of Commonwealth agencies and the GBRMPA is one key client agency that regularly uses Coastwatch for services that include conducting wildlife monitoring (e.g. recording whale sightings) throughout the entire marine park (ANAO 2000). However, details on flight paths and schedules are confidential, thereby making it difficult to assess whether adequate and unbiased sampling has occurred throughout the marine park and there are no effort or environmental data associated with the sightings data set. Consequently, a major assumption in this study is that there was adequate sampling of the entire GBRMP both spatially and temporally for the time period over which the sighting data were taken. We believe this to be an adequate assumption based firstly on the BPC aerial surveillance program initially having been established (in 1989) according to a stratified random sampling design, which would result in even coverage probability (Pettitt & Haynes 1994). Secondly, graphical examination of the BPC sightings data revealed no bias in sightings with latitude or longitude, demonstrated by whale sightings occurring within the full

range of latitudes and longitudes within any month for all years, indicating consistent coverage of the entire GBRMP. Violation of this assumption would lead to an under-representation of the entire distribution of humpback whales during the breeding season.

The presence-only sighting data included in the model consisted of confirmed species identification and only sightings from the months of July and August. Plots of the sightings data demonstrated there was no bias in latitude in the identification of whales to species (Fig. 1). In total, of 381 sightings of whales during 2003 through 2007, 131 (34 %) were confirmed as humpback whales and 107 of these occurred during July and August. Sightings were restricted to July and August to cover the peak in whale sightings during the breeding season (Chittleborough 1965), in an attempt to minimise the number of potentially migrating whales sighted. During July and August it is assumed that the majority of humpback whales have reached their migratory destination and are using habitat that is important for their breeding activities. Several data sets support this assumption, including whaling data from the mid 20th century from the east coast of Australia. Whaling data show that late July to early August is when calving and the most conceptions occur, and from the end of August, more than half of the population is migrating south away from the breeding grounds (Chittleborough 1965). Furthermore, in the current BPC sighting data, humpback whales were predominantly sighted during May to October, with peak sightings between July and August, indicating that the majority of the population had reached the breeding grounds.

Calves were not identified in the BPC sighting records, so it was not possible to separately model sightings of groups with calves present (i.e. calving areas) and groups without a calf present (i.e. breeding areas). Consequently, patterns in habitat suitability identified within the models are likely to be more general than if habitat models were derived separately from sightings of groups containing and not containing a calf because it incorporates environmental parameters from both. This assumes that habitat preference is potentially different between mating and calving whales, which has been demonstrated for other breeding grounds such as Hawaii (Craig & Herman 2000).

Environmental variables

We selected 4 environmental variables—water depth, sea surface temperature (SST), distance from the coast and seafloor slope (measured in degrees)— based on their importance identified in previously published studies investigating relationships between humpback whale distribution and the environment (Ersts & Rosenbaum 2003, Johnston et al. 2007, Rasmussen et al. 2007, Oviedo & Solís 2008). Two additional environmental variables, distance to nearest reef and seafloor rugosity, were also investigated to assess whether these variables may contribute to humpback whale distribution and habitat preference in the GBR. Seafloor rugosity is a measure of the benthic terrain complexity ('roughness') and is defined as the ratio of surface area to planar area. Rugosity values near 1 represent flat, smooth terrain and higher values reflect increasing terrain roughness. All environmental layers used were raster data at a resolution of 4.8 × 4.8 km (Universal Transverse Mercator [UTM] GDA 1994 Zone 55 projection) that were converted to ASCII files for use in Maxent.

Bathymetry data were obtained from the Geoscience Australia Australian Bathymetry and Topography Grid 2009 (Whiteway 2009). SST data were obtained from the Advanced Very High Resolution Radiometer (AVHRR) Pathfinder version 5 data set at 4.8 km resolution for the years 2003 to 2007 and average SST data for the months of July and August were obtained for each year. The environmental variables distance to coast and distance to nearest reef were calculated in ArcGIS 9.3 using Spatial Analyst Tools and a GBR-MPA base map, whereas seafloor slope was derived from bathymetry data. Seafloor rugosity was also derived from bathymetry data using the Benthic Terrain Modeler (Wright et al. 2005) in ArcGIS 9.3.

Multi-collinearity among environmental variables was tested by calculating Pearson product-moment correlation coefficients and producing scatterplots. Multi-collinearity refers to a linear dependence between 2 variables and the correlation coefficient measures the strength of that linear dependence, ranging from −1 to 1, with 0 suggesting no linear correlation (Sokal & Rohlf 1995). Multi-collinearity among variables can lead to a bias in predictions and the contribution of each variable in the model (Fielding & Bell 1997). A threshold value for correlation coefficients greater than ±0.5 was used to identify correlated variables (Zuur et al. 2009, p. 473), and the variable that contributed the least information to the full model based on a jack-knife test was excluded.

Species distribution modelling

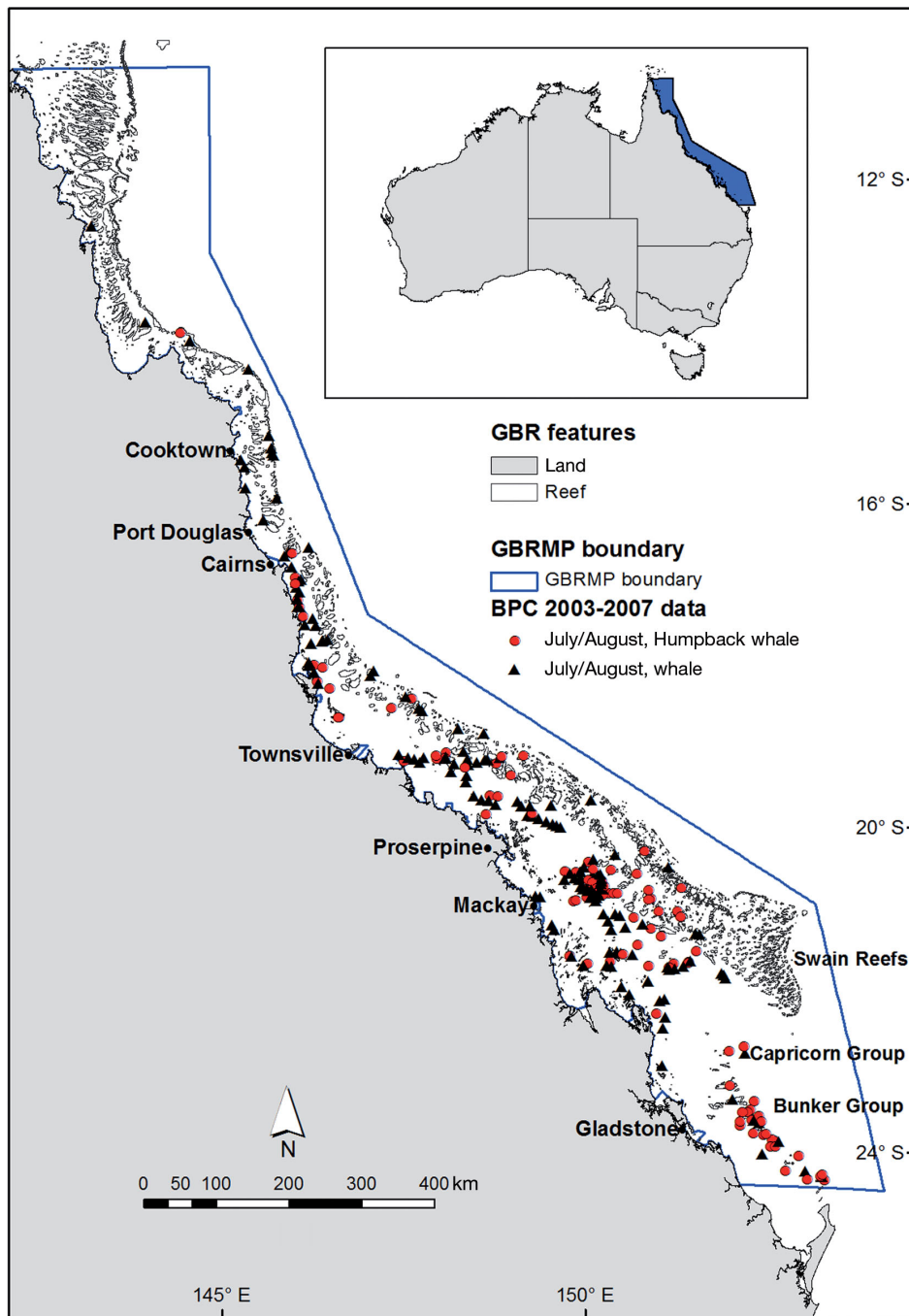Predictive models of humpback whale distribution in the GBRMP were developed using the software

Fig. 1. The Great Barrier Reef Marine Park (GBRMP) off the coast of Queensland, Australia, showing sightings of whales from the Border Protection Command (BPC) aerial surveillance program during July and August 2003 to 2007

Maxent (version 3.3.2). Maxent is based on the maximum entropy method (Phillips et al. 2006, Elith et al. 2010) and has several advantages, including its ability to model presence-only data, perform well with small sample sizes, incorporate interactions among environmental variables, account for spatial bias in the presence data and identify areas that fall beyond the range of environmental conditions used during training when making projections (Phillips et al. 2006).

The underlying theory and assumptions for Maxent have been described in detail elsewhere (Phillips et al. 2006, Dudik et al. 2007, Phillips & Dudík 2008, Elith et al. 2010). The general approach of Maxent is to create a probability distribution for a species by

contrasting occurrence data with background data (pseudo-absences) rather than true absence data. Maxent builds a model of species occurrence starting with a uniform distribution of probability values within each grid over the entire geographic extent (i.e. the background data) and then conducts an optimization routine that iteratively improves model fit measured as gain (i.e. a likelihood or deviance statistic). The gain is closely related to deviance, a measure of goodness of fit used in generalized additive and linear models, and indicates how closely the model is concentrated around the presence samples. The gain is defined as the average log probability of the presence samples, minus a constant that makes the uniform distribution have zero gain. The iterative approach increases the probability value over locations with conditions similar to those of occurrence samples. Among all probability distributions satisfying the constraints (expressed in terms of simple functions of the environmental variables), the one of maximum entropy (i.e. the most unconstrained one) is chosen. The output of Maxent is a probability distribution of environmental suitability for a species, whereby higher values correspond to a prediction of better conditions and higher probability of occurrence (Phillips et al. 2006, Phillips & Dudík 2008).

The final model was developed using 75% of the occurrence data (N = 80) from 2003 to 2007 and tested using the remaining 25% of the data (N = 27), which were randomly drawn from the entire occurrence data set for each of our 50 bootstrap samples. Response curves of the environmental variables were conducted and a jack-knife test was undertaken to evaluate the relative contributions of each environmental variable to the model. Each Maxent predictive model was evaluated using the area under the curve (AUC) of the receiver operator characteristic (ROC), which evaluates how well model predictions discriminate between locations where observations are present and random background data (pseudo-absence points). The AUC is one of the most widely used threshold-independent evaluators of model discriminatory power (Fielding & Bell 1997). The AUC statistic can range from 0 to 1, where an AUC of 0.5 indicates that model performance is equal to that of a random prediction and a value of 1 suggests perfect discrimination between suitable and non-suitable habitat. However, for presence-only and pseudo-absence data, the maximum possible AUC value is less than 1, represented by $1 - a/2$ where a represents a species' true distribution (Wiley et al. 2003, Phillips et al. 2006).

The output of the model was an environmental suitability value for each 4.8×4.8 km cell within the GBRMP. A frequency distribution of environmental suitability values within intervals of 0.1 was constructed and the area of the GBRMP within each interval was calculated. This was done by creating polygons in ArcGIS 9.3 for each environmental suitability value interval and calculating the area for each using the Spatial Analyst Tool Zonal statistics.

Independent data model validation using satellite tag data

In addition to partitioned test data, the spatial habitat model was further validated using an independent data set consisting of humpback whale movements from a satellite tagging program using the Argos satellite system. The tags consisted of a custom-designed, implantable housing that contained Wildlife Computer Spot 5 transmitters. The satellite tagging program was undertaken during 24 to 29 June 2009 off the east coast of Australia at Evans Head, New South Wales, on humpback whales migrating north towards the GBR. During this program, 13 satellite tags were deployed on 8 males, 2 females and 3 whales of unknown sex (Gales et al. 2010). Further details of the tagging procedure and equipment are outlined in Gales et al. (2010).

To validate the habitat model, the tracks from satellite tagged whales were overlayed onto the habitat model in ArcGIS 9.3 and point locations from tag data were used to sample underlying habitat suitability values from the model. The satellite tag data were first filtered to exclude improbable locations based on a conservative maximum swimming speed of 7.2 km h^{-1} using the Speed-Distance-Angle function in the R (R Development Core Team 2008) package 'argosfilter' (Freitas et al. 2008). To limit autocorrelation within the satellite tag data, one daily location of the highest location quality assigned by Argos was used from each whale. Location quality is coded A, B, 0, 1, 2 or 3 in increasing order of position accuracy, with positions 1 to 3 presumably of higher accuracy (location errors <1.0, 0.35 and 0.15 km, respectively) (Argos 1996). Speed of travel between consecutive locations for each whale was then calculated based on a straight path of travel using locations of quality 1 to 3, and data relating to 'migrating' whales were excluded based on a speed of travel greater than 4 km h^{-1}. Swimming speeds of migrating humpback whales have been estimated from several satellite tagging studies which estimate speeds of individuals ranging from 2.6 to 3.8 km h^{-1} (N = 2; Zerbini et al. 2006) and 4.5 to 6.2 km h^{-1} (N = 3; Mate et al. 1998),

with an average migration speed of 4.0 km h^{-1} (N = 4; Lagerquist et al. 2008) and 4.3 km h^{-1} (N = 16; Horton et al. 2011).

A total of 625 satellite tag positions from 13 whales occurred within the GBRMP and, of these, 182 positions from 12 whales (mean ± SD = 11 ± 6 locations per whale) were used to calculate the percentage of locations within areas of varying environmental suitability predicted by the model (Table 1, see Fig. 6). The underlying predictive habitat model was sampled to obtain an environmental suitability value for each grid cell that contained a presence point from the satellite tag data. A frequency distribution of environmental suitability values in intervals of 0.1 was obtained and the percentage of satellite tag locations within each habitat suitability interval was determined. This was compared with the percentage of area in the GBRMP predicted by the habitat model (see Fig. 6).

RESULTS

Model of humpback whale breeding habitat

Based on the Pearson correlation coefficients test for multi-collinearity, the highly correlated variables were seafloor slope and seafloor rugosity, resulting in a correlation coefficient of 0.86. Based on the full model with all variables, the jack-knife test showed that seafloor rugosity contributed the least information in the model compared with seabed slope and was therefore excluded from the final model. The final model used the following 5 environmental variables: SST, water depth, distance to the coast, distance to reef and seabed slope.

The key environmental predictors for humpback whale breeding habitat in the GBRMP were SST, followed by water depth. These 2 environmental variables made higher contributions to the model than the 3 other variables based on the jack-knife test of variable importance. The jack-knife test showed that SST and bathymetry had the most useful information as single variables on training gain (highest gain scores in isolation; Fig. 2A) and the most predictive power (highest AUC in isolation; Fig. 2B) within the model. In particular, SST had the highest gain when used in isolation and therefore appeared to have the most useful information by itself. SST also decreased the gain the most when it was omitted, which suggests that it has the most information that is not present in the other variables.

Response curves characterising the relationship between probability of occurrence and environmental variables demonstrate clear patterns for SST (Fig. 3A), water depth (Fig. 3B) and distance from coast (Fig. 3C). The response curves indicate a preference for waters within SST ranges of 21 to 23°C (highest probability at 21.8°C) and between 30 and 58 m in water depth (highest probability at 49 m). For the distance to coast variable there was a main peak in the suitability curve at a distance between 46 and 109 km from the coast. Both distance to reef and seafloor slope showed no clear patterns in the response curves, indicating no clear preferences for either variables. Seafloor slope is relatively homogenous within the GBRMP and there is a lack of variation in this environmental parameter.

Table 1. *Megaptera novaeangliae*. Sex, deployment dates, duration of tag transmission, number of locations received within the Great Barrier Reef Marine Park (GBRMP), number of positions used in model validation and mean speed of travel of 12 humpback whales tagged off Evans Head, Australia, during 24 to 29 June 2009. Dates in d/mo/yr

| Tag ID	Sex	Tagging date	Last transmission	Tag longevity (d)	No. of locations in GBRMP	Date entered GBRMP	No. locations used	Mean speed of travel in GBRMP (km h^{-1})
88730	Male	28/06/2009	11/08/2009	45	151	3/07/2009	24	1.71
88734	Female	26/06/2009	21/07/2009	26	127	3/07/2009	16	1.76
88736	Male	29/06/2009	13/07/2009	15	74	3/07/2009	5	2.16
88737	Male	29/06/2009	19/07/2009	21	99	3/07/2009	13	2.16
88742	Male	27/06/2009	17/07/2009	21	90	1/07/2009	11	1.84
88747	Unknown	24/06/2009	7/07/2009	14	72	1/07/2009	3	1.74
88748	Unknown	24/06/2009	28/07/2009	35	129	1/07/2009	16	1.65
88751	Male	27/06/2009	18/07/2009	22	86	2/07/2009	10	1.95
88753	Unknown	24/06/2009	22/07/2009	29	145	4/07/2009	2	1.37
88754	Female	27/06/2009	25/08/2009	60	215	2/07/2009	17	1.67
88755	Male	25/06/2009	4/08/2009	41	52	1/07/2009	3	1.10
88756	Male	29/06/2009	24/07/2009	26	75	4/07/2009	8	1.28
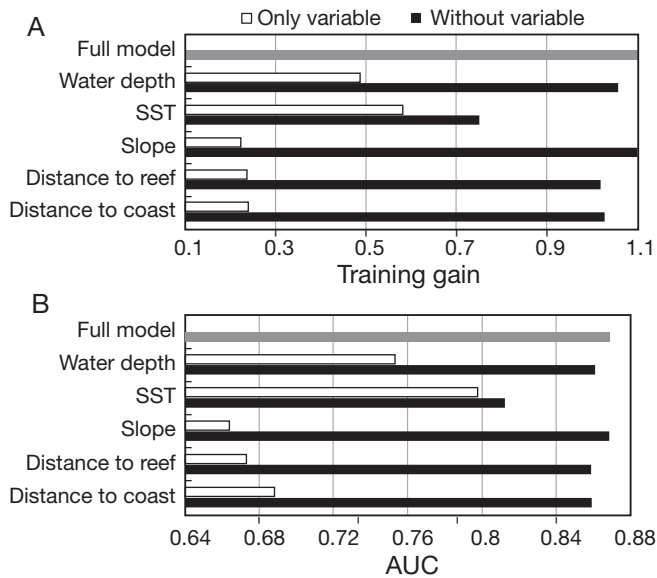
Fig. 2. Variable contributions from the jack-knife test to (A) training gain and (B) area under the curve (AUC) of the model for humpback whales in the Great Barrier Reef Marine Park (GBRMP). Higher values indicate a greater contribution to the model. 'Only variable' indicates the results of the model when a single variable is run in isolation; 'without variable' indicates the effect of removing a single variable from the full model (jack-knife)

The mean test AUC using the Coastwatch data was 0.86 (range 0.80 to 0.92 over 50 bootstrap samples), indicating that the model is reliable at predicting presence sites from random background sites (Elith 2002). The model predicted a range in suitable habitat throughout the GBR lagoon, with areas containing values with a >0.5 or >50% probability of occurrence located predominantly south of 19° S and particularly between 19° and 24° S (Fig. 4). Two core areas of higher habitat suitability, in which there was a >70% probability of occurrence identified in the model, both occurred in the GBR lagoon. They were located in offshore waters, one approx. 100 km east of Proserpine south to Mackay, particularly the inner and outer reefs off Proserpine and outer reefs off Mackay; and the other located at the Capricorn and Bunker Groups of islands and reefs approx. 100 km east of Gladstone.

Independent validation of model using satellite tag data

Based on visual inspection, the tracks of 12 satellite-tagged humpback whales were generally located within areas of high probability of occurrence of the habitat model (Fig. 5). The model predicted less
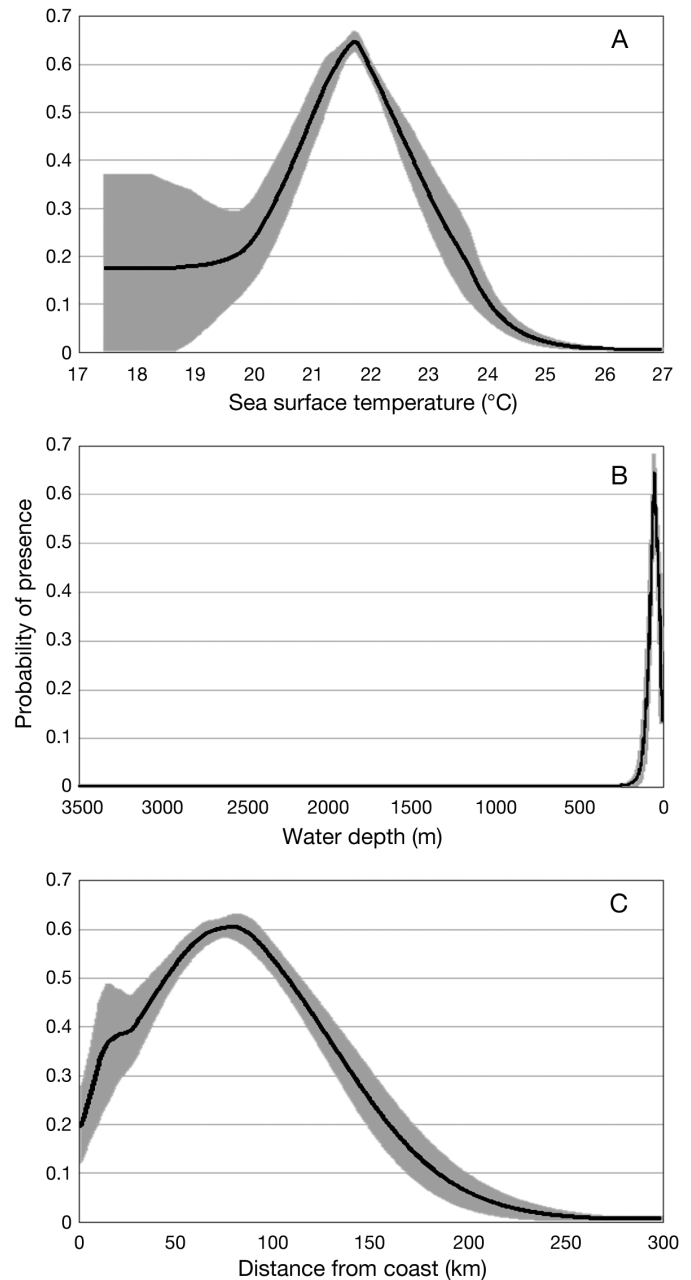

Fig. 3. Response curves of the 3 most informative environmental predictors, illustrating the relationship between probability of occurrence of humpback whales and (A) sea surface temperature, (B) water depth (probability of presence >0.5 = −36 to −66 m) and (C) distance to the coast. These curves show how the shape of the response changes for a particular variable while all other variables are held at their mean sample value. Data are means ± 1SD (grey shading)

suitable habitat in the northern GBR and higher environmental suitability within the southern GBR, corresponding to a higher probability of occurrence in the south, which is where the majority (92%) of tagged whales remained. From the satellite tag data, the movement of 11 of the 12 whales was concentrated
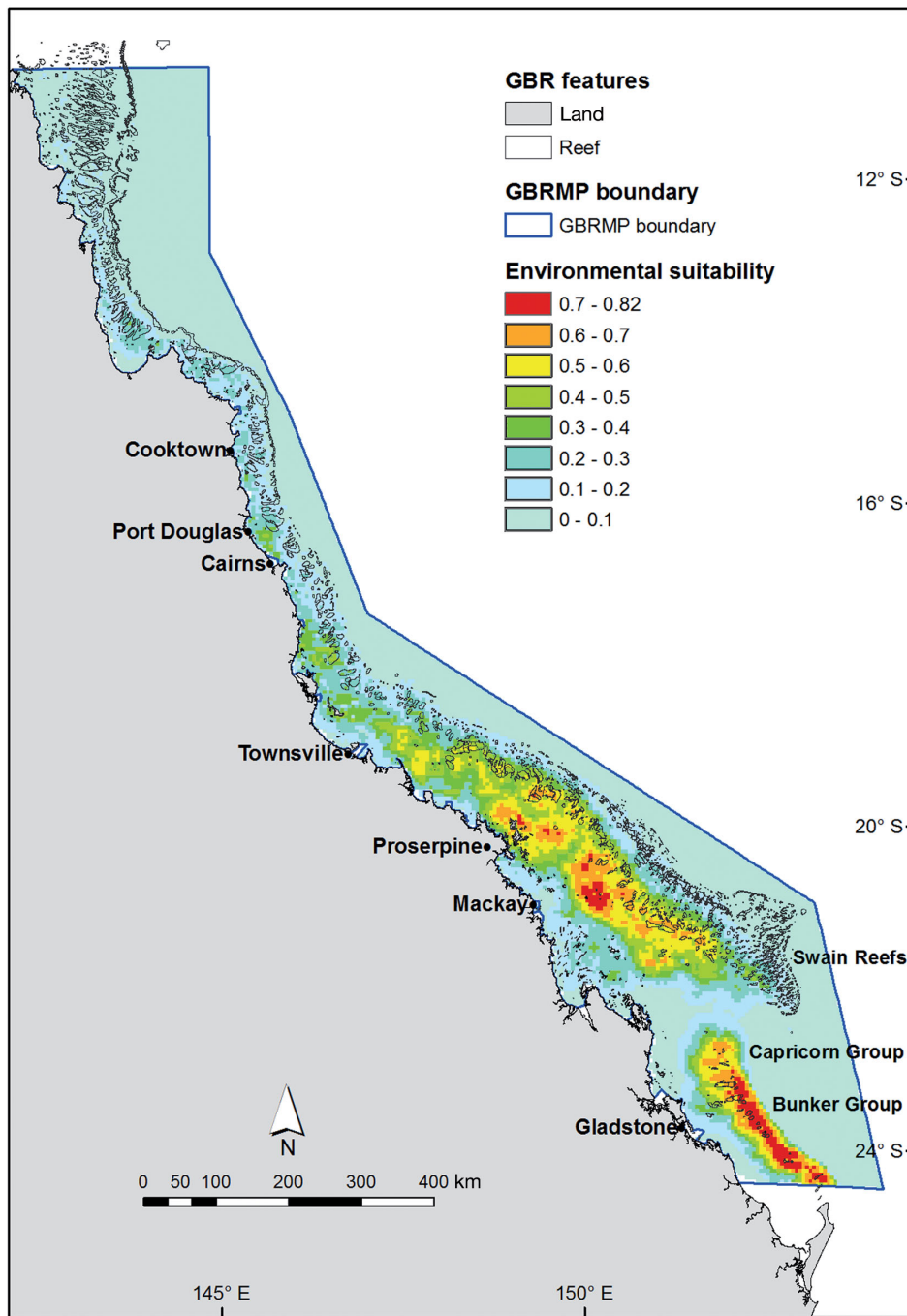
Fig. 4. Model prediction of average environmental suitability (after bootstrapping N = 50) for humpback whales in the Great Barrier Reef Marine Park (GBRMP) for July and August 2003 to 2007. High probability of suitable habitat and occurrence of whales are indicated by dark red

around the area offshore of Mackay, with movements extending north and offshore of Proserpine, which was within the area that the model predicted values of highest habitat suitability. One of the satellite-tagged whales (tag 88730) also travelled through this area and continued north to its most northerly destination south of Cooktown (Fig. 5).

There was a general positive trend between the number of whale satellite tag locations and the probability of humpback whale occurrence based on the model. The model discriminated areas of high habitat suitability within the GBRMP from areas of low habitat suitability, predicting over half of the area of the GBRMP (54%) to contain low
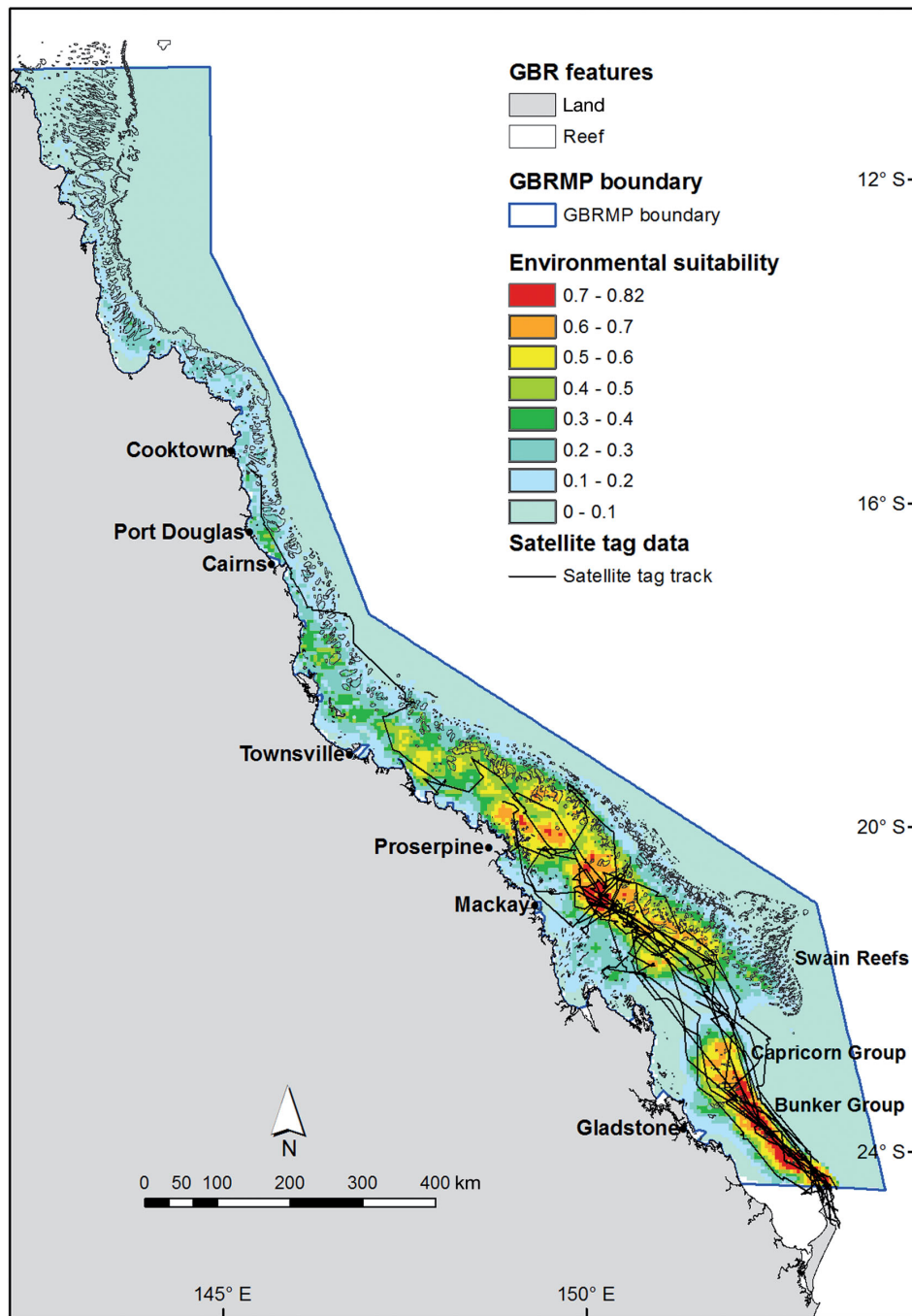
Fig. 5. Model prediction of environmental suitability for humpback whales in the Great Barrier Reef Marine Park (GBRMP) and an overlay of the spatial movement of 12 individual humpback whales (excluding southward migration travel) from satellites tags for the period of July to August 2009

habitat suitability (values between 0 and 0.099) for humpback whale occurrence (Fig. 6). Despite a greater probability of a satellite tag position occurring within lower habitat suitability values, 53% of all satellite tag positions occurred within cells of habitat suitability values >0.5 (therefore non-random

discrimination), which is 10% (33 306 km^2) of the GBRMP. A total of 14% occurred in suitability values >0.7, i.e. 1.2% (4127 km^2) of the GBRMP and no satellite tags occurred in the habitat suitability interval of 0.8 to 0.82, although this represents 0.03% of the GBRMP (93 km^2).
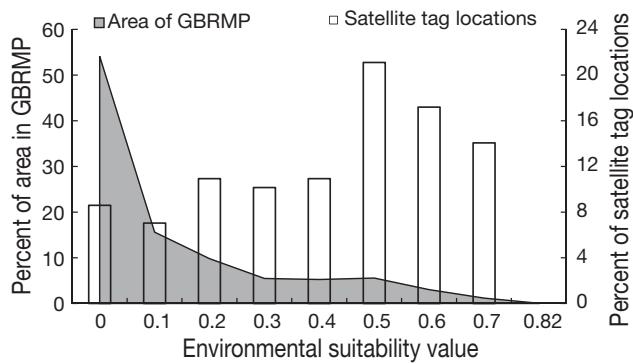
Fig. 6. Percentage of area in the Great Barrier Reef Marine Park (GBRMP) within each environmental suitability interval of 0.1 (e.g. 0.1–0.199) identified by the predictive habitat model and percentage of satellite tag locations from 12 humpback whales (N = 128) within each corresponding habitat suitability value interval

DISCUSSION

Because of the large scale of the GBR, it is logistically difficult and costly to survey the entire area. Few studies have been undertaken on the distribution of humpback whales throughout the entire extent of the GBRMP (Paterson & Paterson 1984, Simmons & Marsh 1986, Chaloupka & Osmond 1999). To date, no specific locations of wintering areas comparable to those identified in other parts of the world had been identified within the GBR. The main wintering area is believed to be in the southern GBR lagoon, from the Whitsundays offshore of Proserpine to south of the Capricorn and Bunker groups (Chaloupka & Osmond 1999). Using presence-only data and habitat modelling, we have refined this knowledge by developing a predictive spatial habitat model that has identified relative core areas in the GBR that could represent an important wintering area and migration route for humpback whales. An independent data set of the movements of satellite tagged humpback whales in the GBR supports the results of the habitat model.

It is important, however, to highlight the limitations of the sightings data used to develop the model. These include the absence of environmental and effort data and the inability to unequivocally determine equal sampling coverage throughout the GBRMP. This has the potential to reduce our ability to accurately identify the true extent of suitable habitat in the GBRMP. Nevertheless, although further validation of the model needs to occur, the independent satellite tag data have demonstrated that the model appears reliable for identifying the distribution of humpback whales during the breeding season and

supports the use of the model to direct further systematic survey work to determine the extent of the breeding grounds. The present study has also shown that the use of presence-only data and development of a spatially explicit habitat model is a valuable and cost-effective approach for informing the relationship between the distribution and environmental features for humpback whales within the GBRMP.

The model identified 2 core areas in which there was a higher probability of humpback whale occurrence (>70% probability) than other areas in the GBRMP: (1) offshore of Proserpine extending south to Mackay within the inner reef lagoon region, and (2) the Capricorn and Bunker groups of islands and reefs approx. 100 km east of Gladstone (Fig. 4). Based on the combination of the habitat model and the independent satellite tag data, we propose that the first core area of high habitat suitability (19.5° to 21.5°S) off Proserpine and Mackay represents an important wintering area, and particularly the area approx. 100 km off Mackay (21°S). We suggest that the second area near the Capricorn and Bunker groups of islands and reefs represents a migration route. The habitat model identifies a discrete area of high probability of occurrence offshore of the Capricorn and Bunker groups of islands and reefs (Fig. 4). Validation with the satellite tag data supports this and shows a restricted distribution (Fig. 5) and highly directed movement of whales through the Capricorn and Bunker groups similar to the directed travel exhibited by migrating whales. Migrating whales are capable of remarkable directional precision and their strong, directional migration is a major characteristic (Mate et al. 1998, Zerbini et al. 2006, Horton et al. 2011). The satellite tag data also show a further movement north past the Capricorn and Bunker groups to the area identified off Mackay, suggesting that the Capricorn and Bunker groups are an important migration route and that the Capricorn group is the terminus of the migration off the east coast of Australia.

Similarly, it is the combined habitat model and satellite tag data, and specifically the absence of consistent directions of travel, reversals of movement patterns and speeds of travel consistent with other wintering areas (Mate et al. 1998, Zerbini et al. 2006, Lagerquist et al. 2008, Horton et al. 2011), that suggest that the other core area offshore of Proserpine and Mackay represents an important wintering area. Within this area of high probability of humpback whale occurrence identified in the habitat model, the satellite tracks of the 12 tagged whales show more erratic directions of travel and speeds of travel typically between 1 and 2 km h^{-1}. The majority of whales satel-

lite tagged (11 of the 12 whales) also did not travel north beyond Proserpine (Fig. 5), indicating that these whales had reached their northernmost point of travel in the GBRMP. For 7 of the 12 whales, the last location was >50 km south of the most northerly location, indicating a southward migrating movement of travel. Although the satellite tag data set is small, it does suggest that this area is an important wintering area, at least for a subset of the population. Interestingly, this core area of high habitat suitability and potential wintering area is situated between latitudes 19.5° and 21.5° S, which are within a latitudinal range similar to that of many of the major known breeding concentrations (typically 20° S) around the world (Clapham & Mead 1999, Rasmussen et al. 2007). The validation process, which compared filtered satellite tag locations with underlying habitat suitability values, demonstrated that although only a small fraction of the GBRMP (1.2 % or 4127 km^2) was predicted to contain habitat suitability values of 0.7 to 0.82, 14 % of satellite tag positions occurred within cells of these habitat suitability values (Fig. 6).

The most important environmental predictors affecting the modelled distribution of humpback whales in the GBRMP were SST and water depth. Although the remaining environmental variables of distance to coast, seafloor slope and distance to reef were also important, they did not provide as much information in the model to explain the humpback whale distribution. The model indicated a strong preference for waters between 30 and 58 m deep (highest probability at 49 m; Fig. 3B); the majority of the GBR lagoon area out to the outer reefs in the GBRMP is ~70 m in depth. The depth range identified by the model is consistent with depth ranges reported for humpback whales in other worldwide breeding grounds, such as Hawaii (depths <183 m; Craig & Herman 2000) and Madagascar (depths of 50 m and less; Ersts & Rosenbaum 2003). It is important to note, however, that it was not possible to separately model mating and calving areas; therefore, the depth ranges presented represent wintering habitat in general for humpback whales during the breeding season. If there were proportionally fewer groups containing a calf in the sighting data, then it is possible that the model has underestimated the use of shallower waters, considering that females with calves prefer shallow waters (Craig & Herman 2000, Ersts & Rosenbaum 2003, Félix & Botero-Acosta 2011).

The model also indicated a discrete range in SST of 21 to 23°C (highest probability at 21.8°C; Fig. 3A), which falls within the lower limits of the temperature range identified for wintering breeding areas world-

wide of 21.1° to 28.3°C (Rasmussen et al. 2007). Rasmussen et al. (2007) suggest that water temperature might influence the distribution of humpback whales during the breeding season at a broad scale, whereas the availability of suitable reproductive habitat can be a major influence at a local scale. They suggested that humpback whales might winter in areas with less than optimal water temperatures if those areas offer suitable shallow, protected conditions. It is possible that the observed patterns in environmental parameters identified within the model are a consequence of the physical structure of the GBR region and, particularly, differences in the structure of the southern GBR compared with the northern GBR. In the northern GBR, the reefs are closer to shore (typically within 80 km of the coast), the lagoon area is shallow but with complex bottom topography and there is a steep continental slope that is closer to the coast. In contrast, the southern section of the GBR consists predominantly of a large lagoonal area of shallow water (<50 m in depth), a gentle continental slope and a bottom topography that varies from smooth in the lagoon to extremely complex towards the outermost reefs that are in deeper waters and hundreds of kilometres from shore (Fig. 1). The available area and structure of the southern GBR might therefore be more suitable as breeding habitat and could explain the distribution patterns within the model. This could account for higher environmental suitability values in areas exhibiting SST below the mean (24.3°C) for worldwide breeding areas.

The southern lagoon area of the GBR does represent habitat with attributes similar to those of breeding grounds in other areas of the world, which consist of warm, shallow, protected waters within a larger available geographic space than the northern GBR. It is also possible, however, that not all of the assumptions regarding the sighting data were adequately met, such as bias in sightings because of sighting conditions or uneven sampling coverage. We know that humpback whales migrate to the northern GBR waters around Port Douglas and Cairns, as indicated by the satellite tag data. If the assumptions were not adequately met, this would result in the distribution of humpback whales in the far northern region to have been under-represented. Further investigation is needed to determine the relative importance of this region to the rest of the GBR for humpback whale mating and calving activities.

The habitat model and satellite tag data suggest that the area located offshore of Mackay is a potentially important wintering area for humpback whales during the breeding season (Fig. 5). This area is also

situated near a major designated shipping area in the GBR and the junction of the inner north–south and east–west shipping routes (GBRMPA 2009). Given that the Australian east coast population of humpback whales is rapidly increasing (Noad et al. 2008), an increase in the interactions between human activities and breeding humpback whales is likely. Specifically, mining and industrial activity along the Queensland coast has fuelled a rapid increase in port development and associated shipping activity and there are proposals for significant expansion of 7 of the 10 major ports along the GBR coast, including the ports off Mackay and Gladstone (GBRMPA 2009). A current and projected increase in population growth along the Queensland coast has also led to an increase in coastal infrastructure and related water runoff and water quality issues (GBRMPA 2009). We could predict from the model that there is likely to be a greater reliance on the coastal waters of Mackay and Gladstone as the season progresses, and mothers with newborn calves utilise these areas more, as indicated by preliminary survey data (M. J. Noad unpubl. data). Consequently, increased port and coastal development and an associated increase in shipping are likely to be emerging management issues for a rapidly increasing breeding population of humpback whales. Increased shipping can lead to increased ship strikes and have a serious effect on the recovery of whale species, which has been well documented in North Atlantic right whales (Ward-Geiger et al. 2005).

The identification of important wintering areas for humpback whales in the GBR is crucial for the proper management of human activities in these areas. There is a high probability of increased rates of human interactions with whales, such as through shipping, and better demarcation of critical habitat for breeding humpback whales in the GBRMP will be important. The current habitat model has been derived from opportunistic sighting data and validated to a small, independent satellite tag data set. This has shown that areas of high probability of occurrence of humpback whales identified by the model closely correspond to areas utilised by a small subset of the Australian east coast humpback whale population that were satellite tagged. However, to better understand the boundaries of breeding areas for humpback whales in the GBRMP, further validation of the model is needed; this could be achieved by undertaking systematic surveys. The current model is reliable for informing the design of systematic surveys and further validation could be achieved by undergoing dedicated aerial surveys that subsample areas in the GBR ranging in probabilities of habitat suitability that include low, medium and high probability of whale occurrence in an effort to better define the breeding grounds and explore environmental relationships. Based on the model, and largely dependent on sampling design, it is possible that only 16% (54 998 km^2) of the GBRMP would need to be surveyed.

LITERATURE CITED

Argos (1996) User's manual. Service Argos, Landover, MD

Australian National Audit Office (2000) Coastwatch: Australian Customs Service, The Audito-General Audit Report No. 38 1999–2000. Performance Audit, Canberra

Chaloupka M, Osmond M (1999) Spatial and seasonal distribution of humpback whales in the Great Barrier Reef region. Am Fish Soc Symp 23:89–106

Chittleborough RG (1965) Dynamics of two populations of the humpback whale, *Megaptera novaeangliae* (Borowski). Aust J Mar Freshw Res 16:33–128

Clapham PJ (2000) The humpback whale: seasonal feeding and breeding in a baleen whale. In: Mann J, Connor RC, Tyack PL, Whitehead H (eds) Cetacean societies: field studies in behavior. University of Chicago Press, Chicago, IL, p 173–196

Clapham PJ, Mead JG (1999) *Megaptera novaeangliae*. Mammal Species 604:1–9

Craig AS, Herman LM (2000) Habitat preferences of female humpback whales *Megaptera novaeangliae* in the Hawaiian Islands are associated with reproductive status. Mar Ecol Prog Ser 193:209–216

Dawbin WH (1966) The seasonal migratory cycle of humpback whales. In: Norris KS (ed) Whales, dolphins and porpoises. University of California Press, Berkeley, CA, p 145–170

Dudik M, Phillips SJ, Schapire RE (2007) Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. J Mach Learn Res 8:1217–1260

Elith J (2002) Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: Ferson S, Burgman M (eds) Quantitative methods for conservation biology. Springer-Verlag, New York, NY, p 39–58

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

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2010) A statistical explanation of MaxEnt for ecologists. Divers Distrib 17:43–57

Ersts PJ, Rosenbaum HC (2003) Habitat preference reflects social organization of humpback whales (*Megaptera novaeangliae*) on a wintering ground. J Zool 260:337–345

Evans PGH (2008) Habitat pressures. In: Perrin WF, Würsig B, Thewissen JGM (eds) Encyclopedia of marine mammals, Vol 1. Academic Press, San Diego, CA, p 513–516

Félix F, Botero-Acosta N (2011) Distribution and behaviour of humpback whale mother–calf pairs during the breeding season off Ecuador. Mar Ecol Prog Ser 426:277–287

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24:38–49

Freitas C, Lydersen C, Fedak MA, Kovacs KM (2008) A simple new algorithm to filter marine mammal Argos locations. Mar Mamm Sci 24:315–325

Gales N, Double MC, Robinson S, Jenner C, Jenner M, King E, Gedamke J, Childerhouse S, Paton D (2010) Satellite tracking of Australian humpback (*Megaptera novaeangliae*) and pygmy blue whales (*Balaenoptera musculus brevicauda*). Paper submitted to the International Whaling Commission Scientific Committee, IWC Morocco, 30 May – 11 June. SC/62/SH21 (Available from IWC Office)

GBRMPA (Great Barrier Reef Marine Park Authority) (2009) Great Barrier Reef outlook report 2009. GBRMPA, Townsville

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8: 993–1009

Harwood J (2001) Marine mammals and their environment in the twenty-first century. J Mammal 82:630–640

Horton TW, Holdaway RN, Zerbini AN, Hauser N, Garrigue C, Andriolo A, Clapham PJ (2011) Straight as an arrow: humpback whales swim constant course tracks during long-distance migration. Biol Lett (in press) doi:10.1098/rsbl.2011.0279

Johnston DW, Chapla ME, Williams LE, Mattila DK (2007) Identification of humpback whale *Megaptera novaeangliae* wintering habitat in the Northwestern Hawaiian Islands using spatial habitat modeling. Endang Species Res 3:249–257

Lagerquist BA, Mate BR, Ortega-Ortiz JG, Winsor M, Urbán-Ramirez J (2008) Migratory movements and surfacing rates of humpback whales (*Megaptera novaeangliae*) satellite tagged at Socorro Island, Mexico. Mar Mamm Sci 24:815–830

Mate BR, Gisiner R, Mobley J (1998) Local and migratory movements of Hawaiian humpback whales tracked by satellite telemetry. Can J Zool 76:863–868

Mattila DK, Clapham PJ, Vasquez O, Bowman RS (1994) Occurrence, population composition, and habitat use of humpback whales in Samana Bay, Dominican Republic. Can J Zool 72:1898–1907

Noad MJ, Dunlop RA, Paton DA, Cato DH (2008) An update of the east Australian humpback whale population (E1) rate of increase. Paper submitted to the International Whaling Commission Scientific Committee, IWC Santiago, 1–13 June. SC/60/SH31 (Available from IWC Office)

Oviedo L, Solís M (2008) Underwater topography determines critical breeding habitat for humpback whales near Osa Peninsula, Costa Rica: implications for marine protected areas. Rev Biol Trop 56:591–602

Paterson RA (2001) Exploitation of humpback whales, *Megaptera novaeangliae*, in the south west Pacific and adjacent Antarctic waters during the 19th and 20th centuries. Mem Queensl Mus 47:421–429

Paterson R, Paterson P (1984) A study of the past and present status of humpback whales in East Australian waters. Biol Conserv 29:321–343

Pettitt AN, Haynes MA (1994) Statistical review of the aerial surveillance database. Summary Report 1: Assessment of the sector/plot aerial surveillance methodology. Consultancy Report prepared for the Great Barrier Reef Marine Park Authority, Office of Commercial Services, Brisbane

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

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190:231–259

R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.r-project.org

Rasmussen K, Palacios DM, Calambokidis J, Saborío MT and others (2007) Southern Hemisphere humpback whales wintering off Central America: insights from water temperature into the longest mammalian migration. Biol Lett 3:302–305

Redfern JV, Ferguson MC, Becker EA, Hyrenbach KD and others (2006) Techniques for cetacean–habitat modeling. Mar Ecol Prog Ser 310:271–295

Simmons ML, Marsh H (1986) Sightings of humpback whales in Great Barrier Reef waters. Sci Rep Whale Res Inst 37: 31–46

Sokal RR, Rohlf FJ (1995) Biometry. W. H. Freeman & Co., New York, NY

Stevick PT, Neves MC, Johansen F, Engel MH, Allen J, Marcondes MCC, Carlson C (2011) A quarter of a world away: female humpback whale moves 10000 km between breeding areas. Biol Lett 7:299–302

Ward-Geiger LI, Silber GK, Baumstark RD, Pulfer TL (2005) Characterization of ship traffic in right whale critical habitat. Coast Manage 33:263–278

Whitehead H, Moore MJ (1982) Distribution and movements of West Indian humpback whales in winter. Can J Zool 60:2203–2211

Whiteway TG (2009) Australian bathymetry and topography grid, June 2009. Record 2009/21 GeoCat #67703, ANZCW0703013116. Geoscience Australia, Canberra

Wiley EO, McNyset KM, Peterson AT, Robins CR, Stewart AM (2003) Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. Oceanography 16:120–127

Wright DJ, Lundblad ER, Larkin EM, Rinehart RW, Murphy J, Cary-Kothera L, Draganov K (2005) ArcGIS Benthic Terrain Modeler, Corvallis, Oregon, Oregon State University, Davey Jones Locker Seafloor Mapping/Marine GIS Laboratory and NOAA Coastal Services Center

Zerbini AN, Andriolo A, Heide-Jørgensen MP, Pizzorno JL and others (2006) Satellite-monitored movements of humpback whales *Megaptera novaeangliae* in the Southwest Atlantic Ocean. Mar Ecol Prog Ser 313:295–304

Zuur AF, Leno EN, Walker N, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R, 1st edn. Springer, New York, NY