UNIVERSIDADE DOS AÇORES
Departamento de Oceanografia e Pescas

SEI WHALE (*BALAENOPTERA BOREALIS*) ECOLOGY AND MANAGEMENT IN THE NORTH ATLANTIC

by

Rui Prieto da Silva

Dissertação apresentada à Universidade dos Açores para obtenção do
Grau de Doutor no Ramo de Ciências do Mar e Especialidade em Ecologia Marinha

Horta, 2014

SEI WHALE (*BALAENOPTERA BOREALIS*) ECOLOGY AND MANAGEMENT IN THE NORTH ATLANTIC

by

Rui Prieto da Silva

Fundação para a Ciência e Tecnologia fellowship under the supervision of

João M. A. Gonçalves

Departamento de Oceanografia e Pescas da Universidade dos Açores

&

Gordon T. Waring

Northeast Fisheries Science Center – Protected Species Branch

National Oceanic and Atmospheric Administration

The work conducent to this dissertation was supported by Fundação para a Ciência e Tecnologia (FCT) doctoral grant SFRH/BD/32520/2006 and was partially developed in the scope of projects CETAMARH (POCTI/BSE/38991/01; FCT), MACETUS (MAC/4.2/M10; EU FP5 Interreg), MARMACII (EU InterregIIIB/FEDER), TRACE (PTDC/MAR/74071/2006; FCT) and MAPCET (M2.1.2/F/012/ 2011; DRCT), co-funded by FCT, Direcção Regional da Ciência, Tecnologia e Comunicações (DRCTC), FEDER, the Competitiveness Factors Operational (COMPETE), QREN European Social Fund, and Proconvergencia Açores Program.

ABSTRACT

The current knowledge and data gaps on the biology and ecology of the sei whale (*Balaenoptera borealis*) in the North Atlantic (NA) are reviewed and quantified. Topics where investment is necessary to ensure the proper management of the species are identified and suggestions are made in how to tackle data deficiencies. State of the art ecological research methods are used to characterize the summer habitat of the sei whale off New England and Nova Scotia, the migration and foraging behaviour of the species in the NA, and to investigate the plausibility of existing stock boundaries. It is shown that research on the species has been stagnant for more than two decades. Essential aspects of the sei whale ecology and biology are still mainly derived from whaling records. Movement patterns and winter distribution are not clear and great uncertainty exists about the subdivision of the NA population in further biological units. Abundance estimates in the NA are fragmentary and restricted to a small part of the known part of the species summering habitat. In order to gain insight in the distribution of the sei whale in the north-western Atlantic, a summer habitat suitability model was fit to presence-only data derived from aerial and shipboard sighting surveys, opportunistic sightings and whaling records. The model predicts high habitat suitability for areas of known sei whale aggregation off New England and Nova Scotia but also indicate that other areas, especially along the shelf edge, hold suitable conditions for the occurrence of sei whales in the summer. Further, the model indicates that the sei whale and the right whale (*Eubalaena glacialis*) habitats partially overlap in the studied region, which can translate in possible competition for prey. The migration of the NA sei whales was investigated using satellite telemetry. Results show a migratory corridor between the Azores islands (Portugal) and the Labrador Sea (LS) during the spring. A Bayesian switching state space model was used to investigate the behaviour of sei whales monitored by satellite tags and it is shown that the LS comprises a foraging ground for the species during spring and summer. The most probable origin of those sei whales is the region south or southeast of the Azores, possibly from wintering grounds thought to exist off northwest Africa. The results of the study uncovered unknown patterns about the distribution, movements and habitat use of the sei whale in the North Atlantic, providing new evidence essential to create plausible hypotheses about the stock structure of the species.

SUMÁRIO

O estado actual da investigação e lacunas de conhecimento sobre a biologia e ecologia da baleia sardinheira (*Balaenoptera borealis*) no Atlântico Norte (AN) são revistos e quantificados. As áreas de investimento necessário para suportar a gestão da espécie são identificadas e são feitas sugestões para resolver deficiências de dados. Metodologias modernas de obtenção e tratamento de dados são utilizadas para caracterizar o habitat de verão da baleia sardinheira no Atlântico noroeste, para estudar a migração e comportamento alimentar da espécie no AN e para investigar a plausibilidade das actuais áreas de gestão populacional. Os resultados demonstram que a investigação acerca da baleia sardinheira estagnou há mais de duas décadas. A maior parte do conhecimento sobre aspectos essenciais da espécie baseia-se em dados da baleação. Os padrões de movimentação e a distribuição durante o inverno são virtualmente desconhecidos e subsiste uma grande incerteza quanto à existência e identidade de unidades populacionais no AN. As estimativas de abundância são dispersas e restringem-se a uma pequena parte do habitat de verão da espécie. Para compreender melhor a distribuição da baleia sardinheira no Atlântico noroeste, um modelo foi ajustado a dados de presença obtidos a partir de avistamentos durante campanhas de amostragem navais e aéreas, observações oportunísticas e capturas nas águas da Nova Inglaterra e Nova Escócia. O modelo prevê a existência de habitat favorável em áreas de conhecida agregação da espécie no verão, mas também indica a existência de outras áreas favoráveis, principalmente ao longo do talude continental. Existe alguma sobreposição do habitat da baleia sardinheira com o da baleia franca (*Eubalaena glacialis*), o que pode resultar em competição alimentar. A migração da baleia sardinheira no AN foi investigada utilizando telemetria por satélite. Os resultados demonstram a existência de um corredor migratório entre os Açores e o Mar do Labrador (ML). Um modelo probabilístico Bayesiano de estados num espaço (*state space model*) foi utilizado para investigar o comportamento das baleias, demonstrando que o ML é uma área de alimentação para a espécie. As baleias vistas nos Açores provavelmente passam o inverno em regiões a sul ou sudeste do arquipélago, possivelmente ao largo da costa ocidental africana. O estudo permitiu identificar padrões desconhecidos sobre a distribuição, movimentos e utilização de habitat da baleia sardinheira no AN, provendo informação nova e actualizada, essencial para a criação de cenários plausíveis sobre a estrutura populacional da espécie.

TABLE OF CONTENTS

FOURTH CHAPTER **111**

FIFTH CHAPTER **141**

SIXTH CHAPTER **165**

APPENDICES **175**

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I am thankful to my supervisors, Gordon Waring and João Gonçalves, for having accepted to oversee this thesis. Their continuous support, encouragement and patience despite my repetitive setbacks and delays were fundamental for reaching the end of this task.

Bob Kenney help with the dataset analysed in the Third Chapter was vital for this work. Bob career and scientific insight were also an inspiration to me.

Ricardo Serrão Santos skilfully steered our research unit for many years. His help and encouragement was also vital to me.

There is no "I" in cetacean research. All the experience design, fieldwork, data analysis, reasoning, writing, artwork and dissemination are the result of teamwork. No part of this dissertation is attributable only to me. Every one of these >61.000 words was supported by the toil of countless people.

Fieldwork was only possible due to the competence and professionalism of our crews. Wasn't it for Norberto Serpa, Paulo Martins, Renato Martins, Víctor Rosa, and other crew members and I would have stayed on land watching cetaceans by a telescope.

When at sea, our eyes are in the cliffs. I am indebted to the help provided by the vigias Antero Soares, Carlos "Vigia", "Chá Preto" (dearly missed), Sr. José, Martins "Vigia", and Sr. Vargas in locating and working with the animals.

I am thankful to all whale watching crews and owners for their support in the field and on land, their knowledge and patience.

All the work we carry out is only possible through the support of a numerous and efficient staff. To all IMAR/DOP staff goes my acknowledgement. I am especially grateful to Sandra & Sandra.

We had too many volunteers to personally tank each one of them. But their work, inquisitiveness and dedication were fundamental for the development of our research. This line is a voucher: if you volunteered for us, present this to collect your personal and heartfelt thankfulness (and a drink) :-«D».

Everyone who at some moment was part of the Cetacean Research Group at DOP/UAç contributed in some way to the construction of a dream: Ana Pinela, Andreia Cruz, Aline Zunino, Cláudia Oliveira, Inês Seabra, Irma Cascão, Janaina Bon de Sousa , João Vaz Medeiros, Margarida Bacelar Santos, Maria João Cruz, Marta Morcillo, Ricardo Antunes, Sara Magalhães, Sónia Santos, Sophie Quérouil, Suzana “Suca” Morais, Tiago Sá, Yves Cuenot.

What to do with so many friends, who kept me alive even when I felt like dying? I cannot list you all, but will kiss you one by one…

Filipa and Amarelo: thanks!

I was supported by a doctoral grant (SFRH/BD/32520/2006) from the Portuguese Foundation for Science and Technology (FCT). Part of the work presented here was supported by projects CETAMARH (POCTI/BSE/38991/01; FCT), MACETUS (MAC/4.2/M10; EU FP5 Interreg), MARMACII (EU InterregIIB/FEDER), TRACE (PTDC/MAR/74071/2006; FCT) and MAPCET (M2.1.2/F/012/ 2011; DRCT).

I thank the International Whaling Commission for access to the individual whale capture database, used in the Second Chapter. Special thanks are owed to Cherry Allison for help with the database.

The Oceanography Branch of the Northeast Fisheries Science Center, NOAA provided data on copepod abundance for the Gulf of Maine. Although the data was not used in the final version of this dissertation, I am grateful to NOAA and especially to John Hare for the access and help with the data.

The sei whale data used in the Third Chapter was graciously provided by the *North Atlantic Right Whale Consortium* (NARWC). I am indebted to the NARWC contributors and personnel.

I also wish to thank everyone at the NOAA Northeast Fisheries Science Center for receiving me, for their insightful suggestions and for their warmth.

I have the rare fortune of having two fathers and two mothers, a big brother and a small sister. I owe you all everything and more. I love you guys.

Throughout my entire career Monica Silva has been an inspiration, a safe harbour, a mentor and above all a faithful friend. If Monica didn’t exist these would be blank pages.

FIRST CHAPTER

General introduction and dissertation organization

GENERAL INTRODUCTION

Whales and Man throughout history

Cetaceans appear in the fossil record during the Eocene, 52.5 million years ago (Mya), and rapidly radiate to attain the peak in diversity during the middle Miocene (approximately 16 Mya) [1]. After the Miocene, cetacean diversity decreased progressively to the present 90 species (of which one, a river dolphin, is possibly extinct) [2]. The extant cetacean are subdivided in two clades, the Mysticeti (the baleen whales, with 14 species) and the Odontoceti (the toothed whales, with 76 species), all fully aquatic.

Due to the physical characteristics of their environment and macroevolutionary factors most cetaceans became large in comparison to land mammals, with whales attaining the largest sizes in the animal Kingdom [3]. Among many adaptations to the aquatic environment, cetaceans developed a thick sub-dermal adipose layer that would eventually be central in the ecological relation that evolved between those animals and Man.

Their large size means that securing a whale carcass guarantees access to large quantities of meat, fat and raw materials. Thus it is not surprising that Man (*Homo sapiens*) and other hominids started exploiting cetaceans as a resource already in pre-historic times, initially opportunistically as result of strandings [4-7] but soon engaging into active hunting and, in a sense, becoming predators of those animals [8-14]. Due to the sheer size of the animals and the fact that they are strictly aquatic, cetacean hunting is logistically and technically challenging and the activity probably played an important role in technological development and seafaring for some cultures [8,10,13,15]. Nevertheless, some whale species are markedly more difficult to capture than others, due to a combination of anatomy and behaviour, dictating that up to the 1860’s only a few species were targeted by most whaling operations. Those were the bowhead whale (*Balaena mysticetus*), the right whales (*Eubalaena glacialis*, *E. australis, E. japonica*) the humpback whale (*Megaptera novaeangliae*), the gray whale (*Eschrichtius robustus*), and the sperm whale (*Physeter macrocephalus*) [16,17].

From the 11th Century onwards, starting in Europe, whaling evolved from a subsistence activity to an extractive industry focused mostly on obtaining oil and baleen that were used as raw materials for other industries [16,18]. With industrialization whaling captures became massive, numbering in the tens of thousands, and started impacting the sustainability of the targeted whale populations [19-22]. The mechanization of vessels and technological improvements in catching and processing equipment (introduced in the mid-19th Century) opened the way to the capture of other species that were until then logistically unavailable, giving rise to what is known as “modern whaling” [16,18]. Modern whaling was characterized by a rapid decline in whale populations, with species being depleted successively according to their relative size, from larger to smaller, a reflex of the sole focus of the industry in oil production (larger species yielding more oil per individual) [18].

Due to the important role of whales as marine predators, collapse of whale populations caused by whaling is thought to have affected the functioning of ecosystems across entire ocean basins and along the water column [23-27]. Most (if not all) whale populations that underwent commercial exploitation are believed to still be under recovery but increasing human and environmental pressures threaten that recovery for some species/populations [28-30]. Thus, there is still much uncertainty about how whale populations will behave in the future.

Independently of the direction of population trends, fluctuations in whale abundance are relevant not only at an ecological level, but also under a socio-economic perspective [31]. For example, substantial increases in whale populations will result in sizeable reductions of their prey and may directly or indirectly affect many species currently targeted by fisheries [23,32].

Yet, despite centuries of exploitation, we advanced little in understanding the ecology and ecological role of most whale species. As it will be shown in Chapter Two of this dissertation, some species (of which the sei whale, *Balaenoptera borealis*, is a paradigmatic example) were essentially neglected by modern ecological science. It is thus clear that new approaches to studying whale ecology and to whale management are of paramount importance. Some authors (e.g. [31]) go to the extreme of suggesting that studies on whale ecology might be more useful in whale and ecosystem management than estimation of whale demography. I disagree, and believe that both approaches are complementary and necessary at this stage.

The sei whale

The sei whale is one of the eight species in the Family *Balaenopteridae*, the rorqual whales [2]. The taxonomy of the sei whale was a complex affair from the very beginning. On February 21, 1819 a whale stranded in the shore of Holstein, Germany, and was transported to Hamburg where it was on paid display during two months, until the stench drove away even the bravest visitors and put an end to the profitable business. It was only then that the eminent scholar from the University of Berlin Karl Asmund Rudolphi acquired the decayed carcass and took it back with him to Berlin to be studied. Rudolphi published his anatomical notes on this specimen in 1822 [33], but erroneously described it as *Balaena rostrata* Fabricius, 1780 – an early synonymous of the common minke whale *Balaenoptera acutorostrata*.

In 1823 Georges Cuvier referred to Rudolphi's specimen as "Rorqual du Nord" but did not use a binomial Latin name, and despite noting several anatomical differences it is unclear if he suspected that it should belong to a different species [34]. It was not until 1828 that Réne Primevère Lesson converted the Cuvier's vernacular designation of "Rorqual du Nord" into Latin as *Balaenoptera borealis*, finally coining the formal binomial designation [35].

Regardless of his mistake, for a long time the English vernacular designation of the species was "Rudolphi's rorqual", honouring the original describer of the holotype. However, the vernacular name most widely used to designate the species is the Norwegian derived "sei whale" (*Sejhval*), in a reference to the arrival of these whales upon the coasts of northern Norway at the same time as the *"seje"* or pollock (*Pollachius virens*).

In the following decades after the description of the species, at least a dozen new similar species and subspecies were described, but eventually all were identified as synonyms to the sei whale. Currently there are two subspecies recognized, although morphological and genetic support for the southern-hemisphere subspecies is weak. The two subspecies are:

- *Balaenoptera borealis borealis* Lesson, 1828 referring to the Northern hemisphere sei whale. The holotype, the specimen collected by Rudolphi, was stored in the Berlin Museum of Natural History but is presumed destroyed as a consequence of bombings during WWII.
- *Balaenoptera borealis schlegellii* Flower, 1865 that refers to all sei whales from the southern hemisphere. The holotype of this subspecies is in the Leiden Museum of Natural History.

Sei whales are around 4.5 m when they are born. As with other baleen whales, adult females grow slightly larger than males with sizes ranging from 13.4 to 19.5 m. Males range between 12.8 to 18.6 m. The larger sizes are attained by the Southern Hemisphere individuals. Adult weights range between 20 to 30 metric tons [36]. They have a slender body and lack complex pigmentation patterns. They are dark grey over most of the body and white on the underside. The body is often covered with round scars presumably from cookie-cutter sharks (*Isistius brasiliensis*).

The dorsal fin is tall and falcate, normally making an angle with the back in excess of 45^{o}. In some individuals the dorsal fin has one or more circular holes near the trailing edge, possibly caused by burying parasites [37]. The rostrum is somewhat pointed and is slightly pitched, having a single ridge in the middle that runs from the splash guard of the blowholes to the tip of the mouth. In the field and at close range, the single ridge on the head is the most reliable feature in distinguishing the sei whale from the very similar Bryde’s whale, which has three ridges [38].

Sei whales have 340 to 350 plates of dark grey baleen with light bristles, measuring between 75 and 80 cm in length. Some individuals can have some white plates near the tip of the rostrum [38]. The ventral grooves (or pleats) that allow the mouth to expand during feeding, end well before the navel, unlike what happens in most other Balaenopteriid whales (except for the minke whale) [36].

Despite their cosmopolitan distribution and having been heavily exploited throughout the world by whaling operations, we are still ignorant of many aspects of the lives of sei whales. Some whale species came under the focus of intense research over the last decades for diverse reasons: for being highly endangered; for having coastal habits which increases study opportunities; for being iconic; for being commercially exploited; or for a combination of those. The sei whale does not fall into any of those categories. Research on the sei whale in the past was mostly fuelled by whaling. When whaling came under intense public scrutiny, the sei whale came out of the research spot-lights as a result of the moratorium on all commercial whaling activity declared by the International Whaling Commission from the 1985–86 pelagic and 1986 coastal seasons [39]. In the last decades, in the North Atlantic (NA), only few localized and short-termed studies exist that focused on or included this species (some examples are presented in the Second Chapter).

Currently any attempt for science driven management of the species at the scale of the North Atlantic is bound to be unsuccessful due to the limited and obsolete nature of the data.

This dissertation examines some aspects of the sei whale ecology in the NA using state of the art data collection and analytical techniques. The motivation of the dissertation is not only to contribute for a better understanding of the ecology of the species in the NA but also to demonstrate that with modern research techniques it is possible to properly inform management decisions, even for intractable species such as whales. By careful use of state of the art techniques I show how it is possible to obtain *much out of little*, in order to improve our understating about the ecology and ecological roles of whales and to aid in the management of these animals and their ecosystems in a cost-effective manner. The results provide much needed information about the ecology of the sei whale in the North Atlantic and show that similar techniques can be invaluable to the study and management of other taxa facing the same problems of data scarcity.

DISSERTATION STRUCTURE

This dissertation is composed by four research chapters, plus the General Introduction and Discussion chapters. The research chapters are intended to stand alone as publishable units and as a result some redundancy in the introduction and methods sections can be found. For the same reason, no cross-references are made between research chapters, but references to published and "In Press" chapters may be found in the text.

By their nature, some chapters made heavy use of references. For the sake of readability references are presented as numbered endnotes to each chapter using the PLoS style.

The Second Chapter, "*The forgotten whale: a bibliometric analysis and literature review of the North Atlantic sei whale (*Balaenoptera borealis*)*" was published in

Mammal Review (DOI: 10.1111/j.1365-2907.2011.00195.x) and is presented as part of this dissertation with acknowledgement to the co-authors, David Janiger, Mónica A. Silva, Gordon T. Waring and João M. A. Gonçalves.

The Fourth Chapter, "*Assessing performance of Bayesian state-space models fit to Argos satellite telemetry locations processed with Kalman filtering*" has been published in PLoS ONE (DOI: 10.1371/journal.pone.0092277) and is presented as part of this dissertation with acknowledgement to the co-authors, Mónica A. Silva, Ian D. Jonsen, Deborah J. F. Russell, Dave Thompson and Mark F. Baumgartner.

The Fifth Chapter, "*Sei whale movements and behaviour in the North Atlantic inferred from satellite telemetry*", has been accepted for publication in Endangered Species Research (DOI:10.3354/esr00630) and is presented as part of this dissertation with acknowledgement to the co-authors, Mónica A. Silva, Gordon T. Waring and João M. A. Gonçalves. Results of this chapter were also included in a working paper presented to the 64th meeting of the International Whaling Commission Scientific Committee (working paper SC/64/RMP6, available from http://iwc.int/iwc64docs), in the scope of the discussion over a proposal by Iceland for the pre-implementation of a management procedure for the eventual resumption of sei whaling by that country.

DISSERTATION OUTLINE

First Chapter: General introduction and dissertation structure

The First Chapter lays down the motivation of the thesis, describes the form of the dissertation and briefly introduces the aims and hypotheses of each of the research chapters.

Second Chapter: The forgotten whale: a bibliometric analysis and literature review of the North Atlantic sei whale (*Balaenoptera borealis*)

In the Second Chapter, I assess the state of the art in North Atlantic sei whale research. This chapter is more than a mere review of published and unpublished information. By using bibliometric techniques I show that research on the sei whale is stagnant in comparison to other similar species. I also clearly quantify the scientific production by subjects and regions of the World, showing where are the greatest deficiencies in knowledge and which are the more relevant issues to resolve for the proper management of the species and the ecosystems it integrates. Research and management funding is limited and understandably funding agencies want to maximize the outcome of their investment and restrict spending to solving essential problems. The approach taken in the Second Chapter is the first step in tackling the data deficiency problem posed by the sei whale. By clearly showing that existing data on sei whale is insufficient for its management and laying down the most appropriate lines of research that should be followed to obtain that data, the Second Chapter makes a better case in defending investment on this species.

Objectives

2.1 To give an in-depth review on the state of the art on the North Atlantic sei whale.

2.2 Provide a clear quantification of the knowledge on the sei whale biology and ecology and research needs to inform management decisions.

2.3 Identify lines of research necessary to resolve the data deficiency.

Hypotheses

2.1 Existing data on the North Atlantic sei whale is insufficient and inappropriate for its proper management.

2.2 Existing level of research on the North Atlantic sei whale is insufficient to solve data deficiencies.

Third Chapter: Habitat suitability of the sei whale (*Balaenoptera borealis*) in New England and Nova Scotia using presence-only modelling

In the Third Chapter, sighting and whaling data are combined to create a model of summer habitat suitability for the sei whale off New England and Nova Scotia. To achieve that goal a Maximum Entropy algorithm was employed to enable the use of presence-only data. Pooling the data from several sources and using low-cost opportunistic data enabled the creation of cost-efficient and credible habitat suitability maps. The study also unveiled overlap of feeding habitats in the study region and potential competition between the sei whale and the endangered North Atlantic right whale (*Eubalaena glacialis*). Finally, the model limitations are discussed and improvements suggested so that it can be properly used when informing management decisions.

Objectives

3.1 To assess the feasibility of using presence-only data derived from distinct sources to develop a habitat suitability model for the sei whale.

3.2 To provide habitat suitability maps of the sei whale to advance scientific knowledge on the species and to inform management decisions.

Hypotheses

3.1 Presence-only data of sei whales is sufficient to produce a habitat suitability model that performs better than random.

3.2 Summer habitat suitability for the sei whale is not homogeneous within the study region.

Fourth Chapter: Assessing performance of Bayesian state-space models fit to Argos satellite telemetry locations processed with Kalman filtering

The Fourth Chapter has a more technical nature, but essential to back the analyses made in the Fifth Chapter. In the Fourth Chapter the reliability of fitting Bayesian state-space models (SSMs) and switching state-space models (SSSMs) to marine mammal satellite tracking data processed with a recently released location algorithm (Kalman Filter; KF) is investigated. This study is relevant since although SSMs and SSSMs have been used to model animal movement and behaviour using an older location algorithm (Least Square; LSq), the implications of using these models with the new algorithm were not known and could lead to biased estimates of movement parameters and behavioural states. The results presented in that chapter show that fitting SSMs and SSSMs to KF derived data offer clear advantages over LSq derived data. Since the KF algorithm can considerably increase the number of received positions and prolong the tracking periods, the results presented in the Fourth Chapter are of the upmost relevance to studies where few messages are received, which applies to many marine and dense forest species.

Objectives

4.1 To assess the spatial accuracy of locations from models fit to data derived from LSq and KF processing algorithms.

4.2 To determine how spatial accuracy varies with observation frequency, temporal resolution and reported precision of Argos locations.

4.3 To investigate how the quality of tracking data affected the similarity of the output from SSSMs fit to LS and KF data.

Hypotheses

4.1 Spatial accuracy of locations from SSM models fit to KF processing algorithm derived data is similar or better than from LSq derived data.

4.2 Data quality has the same effect in the accuracy of the results from SSM models fit to LSq and KF data.

4.3 Parameters of SSSM models fit to LSq and KF data are similar both for location and behavioural estimation.

4.4 Fitting SSM models to KF derived data brings advantages over LSq derived data.

Fifth Chapter: Sei whale movements and behaviour in the North Atlantic inferred from satellite telemetry

In the Fifth Chapter satellite telemetry technology is used to gain insight in the migratory and foraging behaviours of the sei whale in the North Atlantic (NA) and to address a management problem regarding the stock composition of the population. Using a Bayesian switching state-space model (SSSM) whale tracks were reconstructed and movement parameters were investigated to infer behaviour. The method enabled the distinction of two discrete movement phases, corresponding to migratory and foraging behaviours. The study demonstrates that the Labrador Sea is an important feeding ground for the sei whale and also resolved the origin of sei whales seen in that region. More importantly, the results presented in the Fifth Chapter offer unprecedented view about the large scale movements of the sei whale in the North Atlantic, challenging some pre-conceived notions and offering much needed information for the proper management of the species. The Fifth Chapter is also an example of the power of state of the art techniques in obtaining cost-effective, meaningful data that is useful both to gain insight into the ecology of challenging species and to inform management decisions.

Objectives

5.1 To investigate the migratory destinations and behaviour of sei whales seen off the Azores.

5.2 To identify foraging areas for this species.

5.3 To identify wintering areas for this species.

5.4 To assess the pertinence of existing theories on population structure for the species in the North Atlantic.

Hypotheses

5.1 Whales seen in the Azores during the spring are *en-route* to foraging grounds.

5.2 Known feeding grounds of the sei whale in the North Atlantic are discrete.

5.3 Migration of sei whales in the North Atlantic is restricted to movements along a north-south axis, supporting a latitudinal structuring of the population.

5.3 Satellite telemetry can be used to identify plausible hypotheses about movements and stock structure of whales to inform management decisions.

Sixth Chapter: Discussion and Conclusions

The Sixth Chapter reviews the contributions of this thesis to the overall knowledge on the North Atlantic sei whale and makes recommendations on future directions to generate useful data for the proper conservation of the species.

REFERENCES

1. Uhen MD (2010) The origin(s) of whales. Annual Review of Earth and Planetary Sciences 38: 189-219.

2. Committee on Taxonomy (2013) List of marine mammal species and subspecies. www.marinemammalscience.org: Society for Marine Mammalogy.

3. Clauset A (2013) How large should whales be? PLoS ONE 8: e53967.

4. Stringer CB, Finlayson JC, Barton RNE, Fernández-Jalvo Y, Cáceres I, et al. (2008) Neanderthal exploitation of marine mammals in Gibraltar. Proceedings of the National Academy of Sciences 105: 14319-14324.

5. Kandel AW, Conard NJ (2003) Scavenging and processing of whale meat and blubber by later stone age people of the Geelbek Dunes, Western Cape province, South Africa. South African Archaeological Bulletin 58: 91–93.

6. Álvarez-Fernández E, Carriol R-P, Jordá JF, Aura JE, Avezuela B, et al. (2013) Occurrence of whale barnacles in Nerja Cave (Málaga, southern Spain): Indirect evidence of whale consumption by humans in the Upper Magdalenian. Quaternary International.

7. Schindler H (1967-1968) Ein ethnographischer bericht über die Ona aus dem jahre 1765. Wiener Völkerkundliche Mitteilungen 14-15: 33-42.

8. Erlandson JM (2001) The archaeology of aquatic adaptations: paradigms for a new millennium. Journal of Archaeological Research 9: 287-350.

9. Savelle JM (2013) Cumulative bowhead whale (Balaena mysticetus) harvest estimates by prehistoric Thule Inuit in the Canadian Arctic 1200–1500 A.D. : implications for bowhead whale population modeling and Thule demography. Bulletin of the National Museum of Ethnology 34: 593 - 618.

10. Porcasi JF, Fujita H (2000) The dolphin hunters: A specialized prehistoric maritime adaptation in the southern California Channel Islands and Baja California. American Antiquity 65: 543-566.

11. Kroeber AL, Barrett SA (1960) Fishing among the indians of Northwestern California. University of California Anthropological Records 21: 1-210.

12. Lee S-M, Robineau D (2004) Les cétacés des gravures rupestres néolithiques de Bangu-dae (Corée du Sud) et les débuts de la chasse à la baleine dans le Pacifique nord-ouest. L'Anthropologie 108: 137-151.

13. Bockstoce JR (1976) On the development of whaling in the Western Thule culture. Folk 18: 41-46.

14. Bleitz DE (1993) The prehistoric exploitation of marine mammals and birds at San Nicolas Island, California. In: Hochberg FG, editor. 3rd California Islands Symposium, Recent advances in research on the California Islands. Santa Barbara, CA, USA: Santa Barbara Museum of Natural History. pp. 519-536.

15. Betts MW (2007) The Mackenzie Inuit whale bone industry: raw material, tool manufacture, scheduling, and trade Arctic 60: 129-144.

16. Reeves RR, Smith T, D. (2003) A taxonomy of world whaling: operations, eras, and data sources. Woods Hole, MA, USA: National Oceanic and Atmospheric Administration, National Marine Fisheries Service. 28 p.

17. Reeves RR (2002) The origins and character of ‘aboriginal subsistence’ whaling: a global review. Mammal Review 32: 71-106.

18. Tønnessen JN, Johnsen AO (1982) The history of modern whaling. Berkeley and Los Angeles, CA, USA: University of California Press. 798 p.

19. Aguilar A (1986) A review of old Basque whaling and its effect on the right whales (Eubalaena glacialis) of the North Atlantic. Report of the International Whaling Commission Special Issue 10: 191-199.

20. Best PB (1987) Estimates of the landed catch of right (and other whalebone) whales in the American fishery, 1805-1909. Fishery Bulletin 85: 403-418.

21. Scarff JE (2001) Preliminary estimates of whaling-induced mortality in the 19th century North Pacific right whale (Eubalaena japonicus) fishery, adjusting for struck-but-lost whales and non-American whaling. Journal of Cetacean Research and Management Special Issue 2: 261-268.

22. Whitehead H (2002) Estimates of the current global population size and historical trajectory for sperm whales. Marine Ecology Progress Series 204: 295-304.

23. Croll DA, Kudela RM, Tershy BR (2006) Ecosystem impact of the decline of large whales in the North Pacific. In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell Jr RL, editors. Whales, whaling and ocean ecosystems. Berkeley, Los Angeles, London: University of California Press. pp. 202-214.

24. Ballance LT, Pitman RL, Hewitt RL, Siniff DB, Trivelpiece WZ, et al. (2006) The removal of large whales from the Southern Ocean. Evidence of long-term ecosystem effects? In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell Jr RL, editors. Whales, whaling and ocean ecosystems. Berkeley, Los Angeles, London: University of California Press. pp. 215-230.

25. Springer AM, van Vliet GB, Piatt JF, Danner EM (2006) Whales and whaling in the North Pacific Ocean and Bering Sea. Oceanographic insights and ecosystem impacts. In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell Jr RL, editors. Whales, whaling and ocean ecosystems. Berkeley, Los Angeles, London: University of California Press. pp. 245-261.

26. Worm B, Lotze HK, Myers RA (2006) Ecosystem effects of fishing and whaling in the North Pacific and Atlantic Oceans. In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell Jr RL, editors. Whales, whaling and ocean ecosystems. Berkeley, Los Angeles, London: University of California Press. pp. 335-343.

27. Butman CA, Carlton JT, Palumbi SR (1995) Whaling effects on deep-sea biodiversity. Conservation Biology 9: 462-464.

28. Magera AM, Mills Flemming JE, Kaschner K, Christensen LB, Lotze HK (2013) Recovery trends in marine mammal populations. PLoS ONE 8: e77908.

29. Moore MJ (2014) How we all kill whales. ICES Journal of Marine Science: Journal du Conseil doi: 10.1093/icesjms/fsu008.

30. Marino L, Gulland FMD, Parsons ECM (2012) Protecting wild dolphins and whales: current crises, strategies, and future projections. Journal of Marine Biology 2012: e934048.

31. Kareiva P, Yuan-Farrell C, O'Connor C (2006) Whales are big and it matters. In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell Jr RL, editors. Whales, whaling and ocean ecosystems. Berkeley, Los Angeles, London: University of California Press. pp. 379-387.

32. Essington TE (2006) Pelagic ecosystem response to a century of commercial fishing and whaling. In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell Jr RL, editors. Whales, whaling and ocean ecosystems. Berkeley, Los Angeles, London: University of California Press. pp. 38-49.

33. Rudolphi KA (1822) Einige anatomische bemerkungen über Balaena rostrata. Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin 1820-1821: 27-40 + 25 pl.

34. Cuvier G, editor (1823) Recherches sur les ossemens fossiles, où l'on rétablit les caracteres de plusieurs animaux dont les révolutions du globe ont détruit les espéces. nouvelle édition ed. Paris: G. Dufour et E. D'Ocagne. 400 + 427 pl p.

35. Lesson RP (1828) Histoire naturelle générale et particulière des mammifères et des oiseaux découverts depuis 1788 jusqu'a nou jours Lesson RP, editor. Paris: Baudouin Frères, Éditeurs. 443 p.

36. Horwood J, editor (1987) The sei whale: Population biology, ecology & management. London: Croom Helm. 375 p.

37. Gambell R (1966) The dorsal fin of sei and fin whales. Norsk Hvalfangst-Tidende 55: 177-181.

38. Collett R (1886) On the external characteristics of the Rudolphi's rorqual (Balaenoptera borealis). Proceedings of the Zoological Society of London: 243-265.

39. Gambell R (1999) The International Whaling Commission and the contemporary whaling debate. In: Twiss JJR, Reeves RR, editors. Conservation and Management of Marine Mammals. Washington: Smithsonian Institution Press. pp. 179-198.

SECOND CHAPTER

The forgotten whale: a bibliometric analysis and literature review of the North Atlantic sei whale (*Balaenoptera borealis*)

ABSTRACT: A bibliometric analysis of the literature on the sei whale *Balaenoptera borealis* is presented. Research output on the species is quantified and compared with research on four other whale species. The results show a significant increase in research for all species except the sei whale. Research output is characterized chronologically and by oceanic basin. The species' distribution, movements, stock structure, feeding, reproduction, abundance, acoustics, mortality and threats are reviewed for the North Atlantic, and the review is complemented with previously unpublished data. Knowledge on the distribution and movements of the sei whale in the North Atlantic is still mainly derived from whaling records. Movement patterns and winter distribution are not clear. Surveys in some known summering areas show that the species has changed its distribution in parts of its previously known range. With the present information, it is impossible to determine whether or not the North Atlantic sei whale population is subdivided into biological units. Abundance estimates are fragmentary and cover a restricted part of the summering habitat. In the North Atlantic, sei whales seem to be stenophagous, feeding almost exclusively on calanoid copepods and euphausiids. On feeding grounds, they are associated with oceanic frontal systems, but how they find and explore these structures has not been fully investigated.

INTRODUCTION

The sei whale *Balaenoptera borealis* is the third largest member of the *Balaenopteridae* family, after the blue whale *B. musculus* and the fin whale *B. physalus*, and is one of the least known mysticete whales. The species is cosmopolitan, inhabiting the world's temperate to subpolar oceans, with generally antitropical distribution centred in temperate zones. The populations in the North Atlantic, North Pacific and Southern Ocean are almost certainly separated, and may be further subdivided into geographical stocks [1]. Sei whales are thought to undertake seasonal migrations from low-latitude wintering areas to high-latitude summer feeding grounds. The locations of winter calving grounds are unknown [2], and summer distribution on the feeding grounds exhibits great year-to-year variability [3]. Throughout their range, sei whales occur predominately in deep waters. In feeding areas, they are most common over the continental slope or in basins situated between banks [2]. Sei whales feed primarily on calanoid copepods and euphausiids, although small schooling fishes and squid form an important part of their diet in some areas [1].

Due to its smaller size, interest in this species by the whaling industry was low until the 1950s. After the decline of more valuable species such as blue, fin and humpback whales *Megaptera novaeangliae*, the whaling industry started targeting smaller species, particularly sei whales. Subsequently, sei whales were heavily exploited throughout their range and populations decreased drastically [1]. Despite a period of intense commercial hunting, and a subsequent declaration by the International Whaling Commission (IWC) that the management of all whale populations was dependent on comprehensive knowledge of their ecology [4], little information existed on essential aspects of sei whales' life history and ecology.

By the end of the 1970s, it was recognized that the information available on population dynamics of sei whales and Bryde's whales *B. edeni* was insufficient for their proper management, leading to a special meeting on these species in 1977 [5]. Knowledge on the two species increased following that meeting but, subsequently, research interest diminished. With the onset of a moratorium on whaling during the 1985–86 pelagic and 1986 coastal seasons, most of the research effort shifted to other species such as right *Eubalaena glacialis, E. australis,* humpback, blue and fin whales. In addition, more research effort was put into studying minke whales *B. acutorostrata,* the only species still killed in significant numbers for scientific and commercial reasons.

In addition to the low level of scientific and commercial whaling interest in sei whales, the costs of studying cetaceans associated with pelagic habitats hindered research on the sei whale. Sei whales are also difficult to distinguish from Bryde's whales where the two species' ranges overlap, and are often mistaken for fin whales, impacting our understanding of the species' distribution and population size.

Joseph Horwood wrote an influential monograph on the species in 1987, summarizing all the knowledge accumulated on the sei whale to that date, which became the foremost reference on the population biology of the species. Despite the depth of that work, the author stated in the concluding remarks that '[it is] obvious that many aspects [of the population dynamics] can profit by more research' [1], after which he indicated directions for future research. Despite Horwood's advice, and possibly due to the persistence of the reasons mentioned above, sei whale research has advanced little since then.

As in other regions, research on North Atlantic sei whales has been closely linked to whaling. Although commercial whaling ceased in the North Atlantic in 1988, most of the literature available still relies on whaling data, and scant new information has been published since whaling ceased.

This review has three main objectives: (i) to undertake a bibliometric analysis in order to quantify the evolution of research on the sei whale, with emphasis on the North Atlantic; (ii) to summarize the state of knowledge about the species in the North Atlantic, by reviewing historical data, correcting some previous reports based on more recent knowledge and reporting the research produced in the last 20 years, including new unpublished data; and (iii) to identify knowledge gaps and further research needed for the proper management of this species.

MATERIAL AND METHODS

Bibliometric analysis

The Marine Mammal Database and Digital Library of the Natural History Museum of Los Angeles County (hereafter designated as MMD) was used to search for publications. This database was chosen because it is one of the most comprehensive literature databases specialized in marine mammals, with 52039 entries at the time of accession (28 July 2010). Records in this database date back to 1665, but there are great temporal gaps between entries until the mid-19th century. To avoid bias due to a poor representation of older references, the search was restricted to records published after 1860.

To enable comparison of the results with those from similar species, search included publications on blue, fin and humpback whales as well as on the two forms of minke whales *B. acutorostrata* and *B. bonaerensis*. The results for the two species of minke whales were combined due to the difficulty of separating them in works published prior to the late 1990s. The database was filtered by the known vernacular English names as well as the current scientific denominations and synonymia for each species in the Title, Keywords and Abstract fields. Even though the method may miss some entries, it was considered that these false negatives are negligible in the scope of the total number of references detected. Moreover, concerning the comparison among species, there is no reason to think that the method favours the detection of any specific species in the literature over the others.

The database includes several types of contributions, but only peer-reviewed publications and scientific books or monographs were considered. Conference proceedings, book reviews, non-peer reviewed technical reports, unpublished documents, unpublished theses, popular articles and similar were rejected. Duplicates and false positives (e.g. under ‘sei’ results such as ‘seismic’, ‘seining’, ‘*hosei*’, etc.) were eliminated from the resulting bibliographic listings. Known mistaken species attributions (e.g. the Bonin Islands Bryde’s whales that were wrongly reported as sei whales for a long time) were also eliminated from the listings.

Each publication on sei whales was then assigned to a unique oceanic basin, whenever that information was available (North Atlantic, South Atlantic, North Pacific, South Pacific, Mediterranean, Indian Ocean and Southern Ocean). Publications dealing with more than one of the above ocean basins or with global significance (e.g. some physiological studies) were assigned to a ‘world’ category, and those with no information on oceanic basin were classified as ‘unknown’. Subsequently, descriptors were assigned to each publication according to the main focus of the publication. The descriptors used were as follows: phylogeny, anatomy, physiology, feeding ecology, reproduction/social ecology, distribution/movements, stock structure, acoustics, population assessment/management, pathology/natural mortality/ threats, molecular biology/genetics, age determination/growth. As a result of the overlapping nature of some of the descriptors and depending on the focus of the publication, more than one descriptor could be assigned to the same publication.

All publications were considered as having equal weight since the investigation regarded research activity, not the influence of individual publications over the subjects investigated [6].

The main bibliometric indicator of production used was Price’s Law on scientific literature growth. This law is the most widely used indicator for the analysis of productivity of a specific scientific field, and states that scientific production follows an exponential growth [6]. Failure in the fulfilment of this postulate indicates either that (i) a field has reached a saturation point in which the growth changes from exponential to linear and can ultimately decrease steadily towards zero or that (ii) artificial boundaries preclude the assignment of adequate resources to the progress of the subject under investigation.

A second indicator used in the present analysis is the national participation index (PI) for overall scientific production. The PI was calculated as the ratio of the number of publications per million people generated by country and the total number of publications on the subject for the period. Only the lead author was considered for country assignment of each publication since in many publications the affiliations of the co-authors were not mentioned. To calculate the number of publications per million people, the mid-decadal country populations were estimated using the average of the fifth and sixth year of each decade. All demographics were obtained from the U. S. Census Bureau except for the years 1880–89 that were derived from the Great Britain General Register Office [7].

Literature review

The literature review included reports from published material and ‘grey literature’ from the MMD and from other sources, including the author´s personal and other scientific libraries. In some cases specialists were contacted for relevant unpublished information on North Atlantic sei whales. Whenever possible, species confusion between sei and Bryde’s whales was investigated. Problematic identifications are discussed in the text, especially with regard to whaling statistics.

RESULTS

Bibliometric analysis

At a global level, a total of 261 publications were detected on sei whales, 332 on blue whales, 515 on fin whales, 867 on minke whales and 954 on humpback whales.

Research output was negligible prior to the 1940s and has increased more for the other species than for the sei whale in the last 20 years (Figure 2-1). In order to assess whether the scientific output for each of the species follows Price’s Law of exponential literature growth, linear and exponential models were fitted to the data. Only data from 1940 onwards were used to coincide with the initial growth of scientific literature on these species.

Except for the sei whale, there was a significant growth in research output for all species, and the exponential model best explained that growth (blue r^2: 0.52; fin r^2: 0.63; humpback r^2: 0.75; minke r^2: 0.75; $P < 0.05$ for all species). In some cases, the exponential and linear models both had a close fit towards the last 20 years, indicating a change from exponential to linear growth as predicted by Price’s Law.

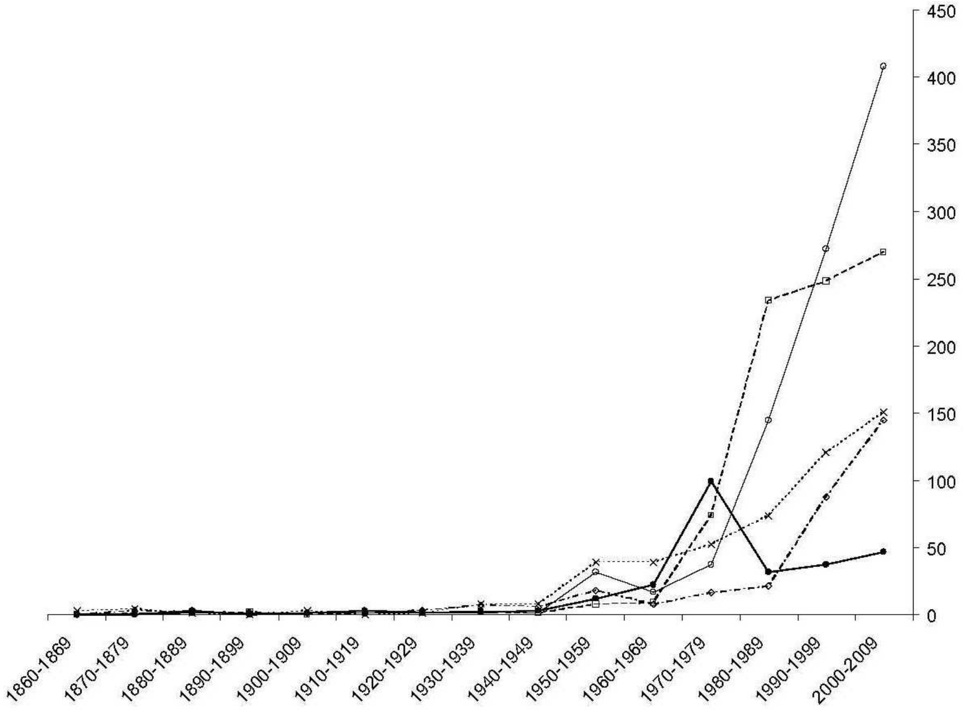

Figure 2-1: Numbers of peer-reviewed publications on sei, blue, fin, humpback and minke whales in the Marine Mammal Database and Digital Library of the Natural History Museum of Los Angeles County, by decade, between 1860 and 2009. Closed circles: sei whale; open diamonds: blue whale; crosses: fin whale; open squares: Minke whale; open circles: humpback whale.

For the sei whale, research output is almost stagnant, and the Price's Law postulate is not fulfilled (Linear model: $r^2 = 0.014$; $P > 0.1$. Exponential model: $r^2 = 0.08$; $P < 0.01$; Figure 2-2). The data points from 1977 to 1978 (corresponding to the outputs of the IWC special meeting on sei and Bryde's whales) were identified as highly influential on the fitting of the models by an analysis of residuals, and new models were fitted without those data. In the new models, the values for r^2 were still low, indicating that most of the variance remained unexplained (Linear model $r^2 = 0.09$; $P < 0.01$. Exponential model $r^2 = 0.13$; $P < 0.01$).

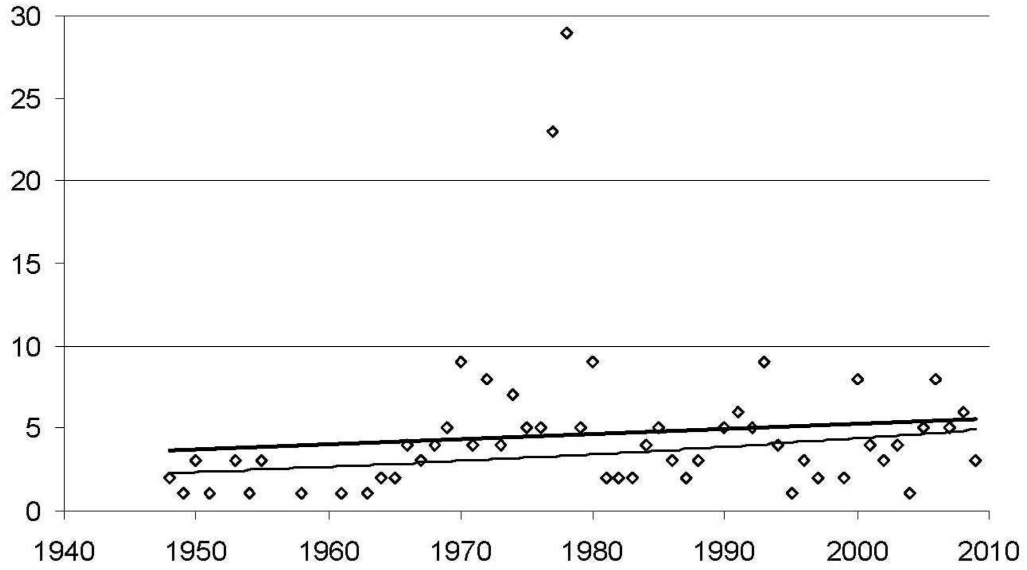

Figure 2-2: Numbers of peer-reviewed papers on sei whales published per year. Linear and exponential models were fitted to the data to check for compliance with Price's Law on scientific literature growth. Linear model (thick line): y = 0.0318x-58.298 (r^2 = 0.014; $P > 0.1$). Exponential model (narrow line): y = 1-11e0.0125x (r^2 = 0.08; $P < 0.01$).

The distribution of publications by ocean basin is presented in Figure 3. Nearly one-third of the literature refers to the Southern Ocean, which can be explained by the importance of the modern whaling activity which developed there [8]. Research concerning the North Atlantic sei whale is documented in 56 publications (21% of the total number). Half of those publications (28) were published after 1986 (when the whaling moratorium established by the IWC came into effect), but 50% (14) of those either used whaling data or were focused on management for whaling, showing that only half of the recent research on the North Atlantic sei whale is independent of whaling.

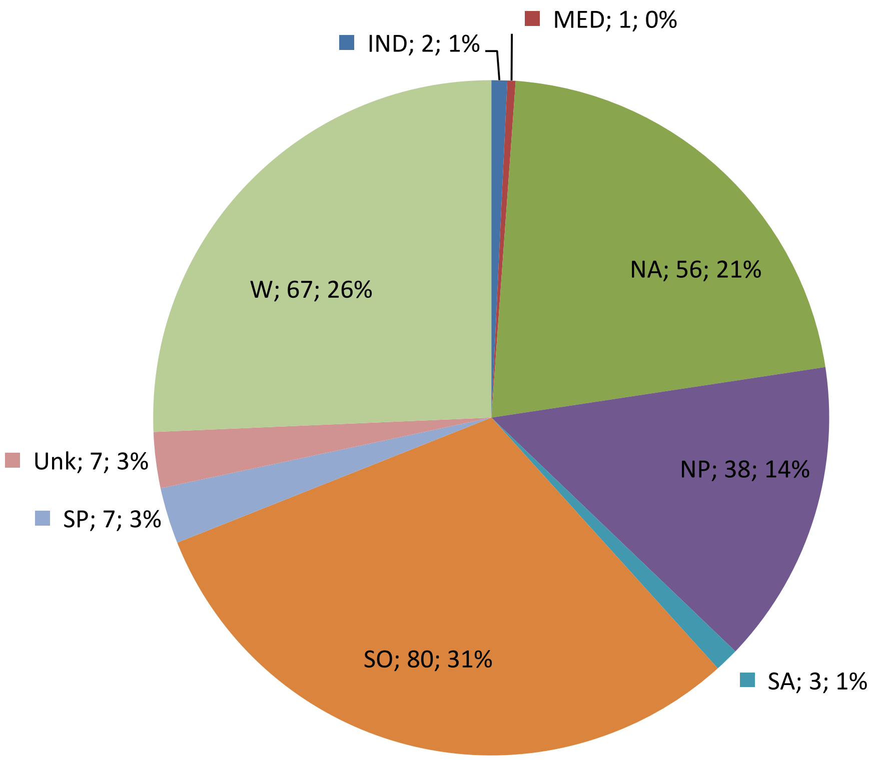

Figure 2-3: Peer-reviewed publications on sei whales in the Marine Mammal Database and Digital Library of the Natural History Museum of Los Angeles County, USA, by ocean basin (Ocean; *n*; % of total). Ocean basins: IND, Indian Ocean; MED, Mediterranean; NA, North Atlantic; NP, North Pacific; SA, South Atlantic; SP, South Pacific; SO, Southern Ocean; W, World; Unk, unknown.

The research output in subject categories by ocean basin, and for North Atlantic sei whales by decade, are presented in Table 2-1 and Table 2-2 respectively. The subjects that have received the most attention are related to whaling (distribution and population assessment/management). Almost one-third of the publications on anatomy (30%, 13 publications) and almost half of the publications on distribution and movements (45%, 31) relate to the North Atlantic but, in contrast, relatively few publications exist on vital parameters such as age determination and growth, reproduction and natural mortality of the North Atlantic population.

Distribution/movements is the only subject for which there has been a clear increase over the years, in investigations on the North Atlantic sei whale (Table 2-2). In fact, it is the only subject that appears to have grown exponentially in accordance with Price's Law since the 1940s, although that result was based on a reduced sample size (exponential curve equation $y = 0.5771e^{0.4055x}$, $r^2 = 0.77$, $P < 0.01$, $n = 31$)

In the North Atlantic, although researchers from the United Kingdom and the United States produced most of the publications on the sei whale during the period 1880–2009 (17 and 13, respectively), Iceland and Norway have higher PI values (1.97 and 0.07, respectively; Table 2-3).

In the last 20 years, the distribution of publications has tended to be more balanced between nations, although the PI values vary widely due to different population sizes. It is not surprising that the nations listed in Table 2-3 were all involved in commercial sei whaling in the North Atlantic at some point in time, with the exception of the Netherlands (represented by a single publication on a stranding). Conversely, all the nations involved in commercial sei whaling in the North Atlantic are represented in that table, except for Ireland, since whaling was conducted under the British flag before the Anglo–Irish treaty and ceased after that treaty.

Table 2-1: Scientific literature production (number of published papers) on sei whales, by subject and by ocean basin, for the period 1860–2009

	Phylogeny	Anatomy	Physiology	Feeding ecology	Reproduction/ social ecology	Distribution/ movements	Stock structure	Acoustics	Population assessment/ management	Pathology/ natural mortality/ threats	Molecular biology/ genetics	Age determination/ growth
Indian Ocean						2				2		
Mediterranean						1				1		
North Atlantic	3	13	3	11	6	31	8	2	11	3	3	2
North Pacific	2	6	4	10	4	8	6	3	6	3	1	
South Atlantic				1		1				1		
South Pacific				1	1	5				1		
Southern Ocean		6	5	10	17	17	7	2	35	5		7
World	5	18	30	6	6	5	4		7	5	7	
Unknown				2	1	2	2		4	1		
Total	**10**	**43**	**42**	**41**	**35**	**69**	**27**	**7**	**63**	**19**	**11**	**9**

Table 2-2: Scientific literature production (number of published papers) on North Atlantic sei whales, by subject and by decade

	Phylogeny	Anatomy	Physiology	Feeding ecology	Reproduction/ social ecology	Distribution/ movements	Stock Structure	Acoustics	Population assessment/ management	Pathology/ natural mortality/ threats	Molecular biology/ genetics	Age determination/ growth
1880–89		3				2						
1890–99												
1900–09						1						
1910–19												
1920–29		1				2						
1930–39												
1940–49												
1950–59						1						
1960–69		1		1								
1970–79		2	1	2	2	6	4		4			1
1980–89		5		1	2	3	1		2			1
1990–99	2	1	2	2		6	2		4	1	2	
2000–09	1			5	2	10	1	2	1	2	1	
Total	**3**	**13**	**3**	**11**	**6**	**31**	**8**	**2**	**11**	**3**	**3**	**2**

Table 2-3: Distribution of publications on the North Atlantic sei whale among countries, by decade

	Canada	Denmark and Greenland	Iceland	Netherlands	Norway	Portugal	Spain	United Kingdom	United States
1880–89								3 (0.028)	
1890–99									
1900–09									1 (0.012)
1910–19									
1920–29									2 (0.009)
1930–39									
1940–49									
1950–59									1 (0.006)
1960–69									1 (0.005)
1970–79	3 (0.012)	1 (0.018)			1 (0.023)		1 (0.003)	2 (0.003)	3 (0.001)
1980–89	2 (0.007)	1 (0.018)	1 (0.379)				1 (0.002)	6 (0.010)	
1990–99	1 (0.003)		4 (1.362)		1 (0.021)		2 (0.005)	2 (0.003)	1 (0.0003)
2000–09		2 (0.024)	1 (0.226)	1 (0.004)	2 (0.029)	1 (0.006)		4 (0.004)	4 (0.001)
Total	**6 (0.022)**	**4 (0.060)**	**6 (1.966)**	**1 (0.004)**	**4 (0.073)**	**1 (0.006)**	**4 (0.010)**	**17 (0.048)**	**13 (0.035)**

Values in parentheses correspond to the national participation index, calculated as the ratio of the number of publications per million people generated by country and the total number of publications on the subject for the period.

Literature review

Taxonomy

In the last three decades, our understanding of cetacean origins has rapidly advanced with data from diverse fields such as palaeontology, anatomy, stratigraphy and, more recently, molecular biology [9]. Immunological and amino acid comparisons made between the 1950s and the 1990s suggested a close relationship between the clades Cetacea and the Artiodactila [10], supporting early hypotheses about the association between these two groups based on morphological data [11,12]. Traditionally, the Cetacea clade has been grouped as the Order Cetacea, but most of the post-1994 molecular studies have placed the cetaceans nested within the artiodactyls, leading to the fusion of these two clades into the Order Cetartiodactyla [9,13]. Despite the molecular evidence, reference to the Order Cetacea is still widespread among marine mammalogists both for convenience and in the interest of taxonomic stability.

Classification within the Cetartiodactyla is still not fully resolved [13-15] and the clades Mysticeti (baleen whales) and Odontoceti (toothed whales), usually referred to as suborders by marine mammalogists, have not yet been ranked [9].

Despite the intense recent activity in cetacean phylogeny research, few researchers have dealt specifically with the sei whale. Two subspecies were proposed by Flower, the northern *B. borealis borealis* and the southern *B. borealis schlegelli* forms, but both morphological and genetic support for a Southern Hemisphere subspecies are weak [16]. In contrast, a maximum parsimony phylogeny of baleen whales revealed a strongly supported (bootstrap support > 95%) clade for North Atlantic sei whales [17].

A prioritized list of species in need of further taxonomic research was produced at a cetacean taxonomy workshop held in April 2004 [18]. The list was based on both taxonomic uncertainty and conservation status, and taxa were ranked in three categories: high, medium and low priority (for taxonomic research). The two forms of sei whale are listed under medium priority, both due to taxonomic uncertainty and conservation status, indicating that the taxonomic status of the species is still not completely resolved and that taxonomic research is needed.

Distribution and movements

In most contemporary references, the main source of information on the distribution and movements of the sei whale are records from the whaling industry [1,2,19]. These records, in the majority, reflect the preferred whaling areas, which were chosen based on several factors including a preference for known grounds, the presence of other favoured species, proximity to land stations and prevailing weather conditions. As such, these data do not reflect the entire distribution of the species and, in some cases, may no longer be accurate.

The whaling records, however, document the oceanwide distribution of the species. Sei whales seem to prefer offshore, temperate waters and seldom venture into the colder waters of the polar seas ([1,2]), although larger individuals enter colder waters [20]. Influxes into inshore waters may occur apparently in response to prey distribution [3,21,22].

Sei whales undertake seasonal migrations from low-latitude breeding grounds to high-latitude feeding areas in both hemispheres [1]. While feeding areas are fairly well defined in all oceans, the locations of breeding areas remain unknown.

Catch data from the Southern Hemisphere and North Pacific suggest that migration is structured by sex and age class, and that pregnant females lead the migration both to and from the feeding grounds [23-28].

Historical distribution

Eastern North Atlantic

In the northeastern North Atlantic, sei whales were first captured with the onset of modern whaling in Finmark, in northern Norway [29,30]. Thereafter, modern whaling spread to other areas, and land stations were created in northern and western Norway, the Scottish islands (Orkney, Shetland and the Hebrides), the Faroe Islands, Ireland, Iceland, East Greenland, the Iberian Peninsula and northern Morocco [3,31-34]. Descriptions of sei whale distribution and movements were largely based on the records from these stations.

Ingebrigtsen [35] is still widely cited to describe the apparent spring–summer migratory sequence of sei whales in the northeastern North Atlantic (e.g. by [1,2,26,36]) despite the fact that his assumptions remain unverified. According to Ingebrigtsen [35], sei whales wintered in the southern part of the North Atlantic, possibly off southern Spain and Portugal and northwest Africa. They then migrated north up the mid-Atlantic and off the western continental slope of Europe to the areas around Iceland, Scotland, western Norway, and Finmark, and as far north as Bear Island and Spitsbergen.

The timing and extent of sei whale movements through western European waters has always displayed great seasonal and inter-annual fluctuation. Reports of years in which unusually large numbers of sei whales occur in an area, often followed by their disappearance from that area for several years, have been termed ‘invasion years’ [3].

Sei whales were captured along the shelf edges around the Scottish islands, from April to October with a peak in June [1,31,33]. In Norway, sei whales first arrived at the waters off Møre, in the west coast, as soon as late April or early May (but more commonly in mid-May). Sei whales were captured closer to shore in the beginning of the season, between May and June, apparently moving offshore towards the end of the season [3]. In Finmark, they were captured mostly from the western coasts (although they were also captured to the east in smaller numbers) from mid-May to August, with the largest catches in June [29,30,37-40]. In 1885, sei whales seem to have made an incursion in great numbers east of the North Cape – over 700 were captured by Norwegian and Russian whalers [41]. In Iceland, sei whales first appeared in late May or early June; the peak catches occurred in August or September [1,3,42].

Off the Iberian Peninsula, sightings of sei and Bryde’s whales are reported from logbooks of sperm whaling expeditions in the North Atlantic in the 19th century, in the Atlantic side of the Straits of Gibraltar but, at the time, they were not taken [43]. Sanpera and Aguilar [34] reviewed whaling off the Iberian Peninsula during the 20th century and described three major whaling grounds: off northwest Spain, off central Portugal, and in the Atlantic side of the Straits of Gibraltar. Sei whales were reported from the northwestern Spain and Straits of Gibraltar grounds, although some confusion over the species identification still persists. In the 1920s, sei whales were declared in the catch from the Straits of Gibraltar in small numbers, but the distinction between sei and Bryde’s whales may be unreliable. In northeastern Spain, sei whales may have been confused with fin whales in the same way [34].

Horwood [1] after Tønnessen and Johnsen [8] reported that 66 sei whales were caught off the west coast of Portugal in 1925–27, but probably Tønnessen and Johnsen [8] were referring to Spanish catches. Otherwise, the figures do not add up for Portugal, and catches are above 100% (749 sei whales are reported, and catches were 93% fin + 68/749 = 9% sei and blue + an undetermined number of sperm whales). Nevertheless, those figures do not agree with more recent reviews of Iberian whaling, which report no sei whales [34] or just one sei whale [44] captured off central Portugal, indicating that sei whales were never common in the grounds off central Portugal, at least during the whaling season.

From 1930 to 1945, whaling off the Iberian Peninsula was drastically reduced, resuming only in the late 1940s. After that period, there are no reported catches of sei whales from the central Portugal grounds, but relatively high catches of sei whales were declared in the Straits of Gibraltar grounds from 1950 to 1953. That is in contrast with the figures for the 1920s catches, suggesting that there could be confusion in the identification of sei and Bryde's whales [34,45]. Off northwest Spain, catches were resumed only after the 1950s, and no subsequent catches of sei whales were declared [34].

Sei whales were reported, sometimes in large numbers in the beginning of the 20th century, from the area from Madeira Islands, past the Canary Islands towards the Cape Verde Islands [3]. In the light of present knowledge, it is possible that those sightings were primarily of Bryde's whales (see discussion on current distribution).

From 1976 to 1979, the unregulated catcher boat *Sierra* and associated vessels operated in the eastern North Atlantic, spending the summer off the Iberian Peninsula and the winter off north-western Africa [46]. These vessels used to stay in the region 18–22°N, 20–25°W from November to May, reportedly in search of sei whales. The catch statistics for this operation are fragmented. Regarding sei whales, the only figures available are the catch of one whale in 1976 and 110 whales for the year 1978 [46]. Although some confusion with Bryde's whales cannot be ruled out, the fact that a distinction is made between Bryde's and sei whales in the catch for the year 1976 indicates that some care was taken in differentiating the two species. Up to the present, this is the best indication of a well-defined sei whale wintering area in the Northeast Atlantic.

Western North Atlantic

Sei whales were seldom taken in the Davis Strait and western Greenland. Only eight whales were reported from whaling boats based in western Greenland in the periods between 1924–39 and 1946–50. Sei whales were considered rare, occurring only in years with warm water intrusions [47]. Around the south-eastern Labrador Sea, sei whales were uncommon in inshore waters; however, data from the Fisheries Research Board of Canada, collected between 1966 and 1973, indicated the occurrence of sei whales in parts of the offshore waters of the Labrador Sea [48]. Whaling off Newfoundland was episodic, with long periods of inactivity. Sei whales were taken in low numbers since the late 1890s [49]. From 1966, an operation from Blandford, Nova Scotia, took a substantial catch of sei whales until 1972, when commercial whaling was terminated in Canada [49].

In the Labrador Sea, sei whales occurred from the first week of June [49]. Off Nova Scotia, Mitchell [48] described two ‘runs’ of sei whales taken by the Blandford station. One run began in early June, peaking in June–July, and another began in late August and early September, peaking in September–October. Mitchell [48] hypothesized that the whales migrated from south of Cape Cod along the continental slope in June and July and then returned south in September–October.

From stranding records along the western US coast, northern Gulf of Mexico and in the Greater Antilles, Mead [50] hypothesized that sei whales from Nova Scotia migrate south along the Atlantic coast of the US, to winter in the waters of the Gulf of Mexico.

Current distribution

In recent years, most of the data on distribution come from sighting surveys, either dedicated or opportunistic, and there are a few accounts from strandings and incidental mortality. Much of the survey information has not been published in peer-reviewed journals and is available only through technical reports, sometimes making it difficult to assess and interpret.

Sei whales seem to be present at low densities over most of the areas covered by dedicated sighting surveys in the North Atlantic. However, that could be explained in part by the fact that most surveys cover areas over the continental shelves, and sei whales are known to prefer the deeper waters off the continental slope [1].

Mediterranean

In the Mediterranean, sei whales are thought to be occasional visitors – rare sightings and strandings are reported from Spain, Gibraltar, France and possibly Tunisia [51]. Recent dedicated sighting surveys failed to detect sei whales [52,53], supporting the idea that the species is an occasional visitor.

Eastern North Atlantic

Over the easternmost part of the North Atlantic, most sei whale sightings during the summer are concentrated in the deep waters of the Irminger Sea area between east Greenland and west Iceland and, in a lesser extent, in the Faroe–Shetland Channel [36,54-58].

Sighting rates in the Norwegian and adjacent waters are extremely low; just a few sightings are reported for the Norwegian Sea and near Jan Mayen [59].

In the Bay of Biscay, sei whales are observed in late summer and early winter along the shelf edge, although with much lower sighting rates than those of most other species [60,61].

A concentration of sei whales was reported in early summer near the frontal area just north and southwest of the Charlie Gibbs Fracture Zone (53°N) during the Census of Marine Life survey conducted over the Mid-Atlantic Ridge from 4 June 2004 to 2 July 2004 [62].

During a survey conducted in July 2007 along the continental shelves of Britain, Ireland, France and northern Spain, sei whales were only detected in small numbers off north-western Spain in the areas of Castro Terrace and the Charcot Seamounts [63].

Only two records of sei whales along the Portuguese continental shore could be found. In September 1978, a sei whale was stranded in Carrapateira, south-western Portugal, in an advanced stage of decomposition which implies it could had been drifting for several days [64]. A single sei whale with an estimated length of 9m was sighted 8 nautical miles off Sagres (south-western coast of Portugal) in October 2010 by a whale watching enterprise (Sara Magalhães pers. comm.). Pictures of the head and body were kindly provided, enabling confirmation of species identity. The estimated size of this individual is similar to the size estimated by Lockyer [65] for weaning in the Antarctic (8m long).

In the Azores Archipelago, the species was first recorded in 1989 [66]. Since then, it has been shown that the species is a regular visitor to these islands during the spring migration; sightings span from early spring to late summer, and relative abundance peaks in April and May [67-69].

The first confirmed record of sei whale presence in the Madeira Islands is from 2002 [70]. The fact that the species occurs simultaneously with the much more abundant Bryde’s whale (that is present from late spring to early autumn) makes it difficult to characterize its seasonality in that archipelago (Luís Freitas pers. comm.).

In the Canary Islands, sei whales are seen only in low numbers in autumn and winter, apparently during migration to lower latitudes [71,72].

The consistent presence of Bryde’s whales during the summer months both in the Madeira and Canary archipelagos (López-Jurado et al. [72], Luís Freitas pers. comm.) may have in the past created some confusion due to the difficulty in distinguishing both species. The discrepancy between the reports of large numbers of sei whales in that region in the beginning of the 20th century mentioned by Jonsgård and Darling [3] and the more recent data mentioned above, indicating a less pronounced presence, may in part be because Bryde’s whales in the region were misidentified as sei whales in the past.

Along the western coast of Africa, information on sei whales is very limited. During a dedicated sighting survey in April 2005, the Atlantic waters contiguous to the Mediterranean were surveyed between the latitudes of 20°N and 37°N [53]. During this survey, only one sei whale was detected in off-shelf waters. Another dedicated sighting survey off southern Morocco between 24°N and 22°N in the Cintra Bay area, undertaken from January to February 1996, did not detect any sei whales [73].

Off Mauritania, a sighting of a single sei whale was reported in March 2003, during a seabird and cetacean survey in offshore waters near the Chinguetti oilfield [74]. However, the picture reproduced in the report shows a very falcate dorsal fin more characteristic of Bryde’s whale. It is unclear whether this sighting was positively differentiated from Bryde’s whale.

In mid April 2011, at least two sei whales were sighted in waters off Mauritania by naturalists on a touristic expedition vessel (Richard White pers. comm.). Pictures were kindly provided and enabled confirmation of species identity. These sightings occurred in the vicinity of the region where the unregulated whaling boat *Sierra* reportedly caught sei whales during winter (see discussion on historical distribution above).

Other records along the west coast of Africa come only from incomplete specimens and strandings, most of which have not been positively identified [75-77]. Confirmed records include a stranding of an 11-m long individual in Mauritania in 1981, and a series of nine baleen plates recovered from the stomach of a tiger shark *Galeocerdo cuvier* captured off Senegal [76]. The Senegal specimen presents a problem in interpretation since it is impossible to know where the shark swallowed the baleen. Tiger sharks are known to move over great distances [78], and baleen probably has very low digestibility and, thus, has long permanency times in the stomach.

Hazevoet and Wenzel [75] reviewed a stranding of a baleen whale calf in Cape Verde that occurred in November 1983 and concluded that it was probably a sei whale, although they could not rule out Bryde’s whale that also occurs in the archipelago.

Western North Atlantic

In western Greenland, sei whale sighting rates are low and seem to be concentrated in two specific areas: the banks at 66°N at Lille Hellefiske and off the southern tip of Greenland [57,79,80]. Data from autonomous acoustic recording units deployed in the Davis Strait in 2007 showed that sei whales occurred in the area from late July to early October (when the equipment was collected) with two peaks: one in late August and a second, less pronounced, in late September (Kate Stafford pers. comm.).

During a July–August 2007 sighting survey in Canadian continental shelf waters along the Labrador and Newfoundland coasts, three sei whales were detected [81]. One was off southern Newfoundland, and two isolated individuals were recorded over the Nova Scotia shelf.

Off the north-eastern US coast, sei whales concentrate during spring in the deeper waters of the margins of eastern and southern Georges Bank, and make episodic incursions into shallower waters, including the Great South Channel, Stellwagen Bank and the southern Gulf of Maine [21,22,82]. Recently, sei whales have also been detected with an array of autonomous ocean gliders deployed in the south-western Gulf of Maine during summer [83,84].

Over 20 cetacean species have been observed during aerial and shipboard surveys conducted during all seasons in the oceanic waters of the northern Gulf of Mexico, but the only baleen whale reported was Bryde’s whale [85-88]. Records of sei whales in the area are restricted to four strandings in eastern Louisiana [86].

Several sighting surveys undertaken in the Bahamas and Caribbean in late winter and spring did not detect sei whales either in coastal or offshore waters [89-91]. Mignucci-Giannoni [92] reported two sightings of sei whales in waters off Puerto Rico and the Virgin Islands, but it is not clear in which circumstances these sightings were made and whether the whales were positively differentiated from Bryde’s whale.

In the oceanic waters of the Central North Atlantic, the only information available comes from a pilot study using yachtsmen sailing from the Caribbean to the Azores in 1998 [93]. No cetaceans were sighted to the west of a longitude of 43°W or south of a latitude of 32°N, but a possible sei whale was reported on June 16, northeast of the Azores at 43°50′N, 21°31′W. However, the sample was small (only two out of 50 boats reported their results), and the use of untrained observers might explain the low sighting rates and uncertainty in identification.

Recent surveys in the North Atlantic show that sei whales still use some of the areas identified from whaling data, and support the notion that after the 1960s sei whales may have drastically changed their distribution in some parts of their range [36], having apparently either abandoned or been extirpated from areas where they were previously recorded with some frequency, such as the Norwegian Sea.

Information from tropical waters on both sides of the North Atlantic is fragmented and imprecise, and the location of the breeding grounds is still unknown. In the offshore pelagic habitats of the Central North Atlantic, apart from localized surveys near the Atlantic islands, no survey has been conducted.

Distribution data are both scarce and sparse over temporal and spatial scales, and are clearly insufficient to understand the present distribution and movement pat-terns of the species over the North Atlantic, particularly with regards to winter distribution and the location of breeding grounds.

Stock identification

Although the term ‘stock’ represents the fundamental population unit for most management purposes, its formal definition varies broadly in scientific literature and management schemes. A typological definition may be unrealistic since it is dependent on management objectives [94]. The IWC, for some species and areas, does not clearly define ‘stocks’ or ‘management units’, leading to the designation of management areas that are not well founded as defining biologically isolated units [95].

Little attention was paid to the definition of stocks in the North Atlantic in the early years of the IWC. It was only after 1977 that the current stock boundaries for fin, sei and minke whales in the North Atlantic were established [95].

Eight concentrations of sei whales were identified for the North Atlantic by the IWC (southwest Norway, north Norway, Faeroe, Denmark Strait, Labrador Sea, Nova Scotia, Gulf of Mexico–Caribbean Sea and southeast North Atlantic). Although Horwood [1] interpreted these as a first attempt at stock definition, in fact, the IWC did not suggest these areas as necessarily representing separate stocks [95]. Mitchell and Chapman [49] presented some evidence of distributional differentiation between the Labrador and Nova Scotia sei whales and proposed that these units constituted two different stocks; both were considered separate from eastern Atlantic sei whales by the IWC [95].

The current IWC boundaries for the management areas of sei whales in the North Atlantic (Figure 2-4) were adopted in 1977 based chiefly on statistical convenience and historic catch data rather than on scientific evidence of stock structure [95].

Horwood [1] suggested a separation of whales taken in Icelandic and Canadian operations based on reproductive data (see discussion on reproduction below).

Daniélsdóttir et al. [96] studied electrophoretic variation in sei whales from Icelandic catches from 1985 to 1988 and did not find significant differences in allele frequencies between samples from different years, indicating that those samples originated from a homogeneous population. No further information is available regarding genetic differentiation of sei whales in the North Atlantic.

Recently, movements of sei whales between the Azores and the Labrador Sea have been recorded with the aid of satellite telemetry [97,98]. These movements show that the migration patterns of this species may be more complex than previously assumed, comprising not only the expected large latitudinal movements but also wide longitudinal displacements. These findings have clear implications for the interpretation of the current management areas for sei whales in the North Atlantic.

Given the weak genotypic information on stock identity, the virtual lack of knowledge regarding the location of breeding grounds, and sparse evidence for stock separation on the feeding grounds, it is presently impossible to ascertain if the North Atlantic sei whale population is subdivided into biological units or if it represents a panmictic population.

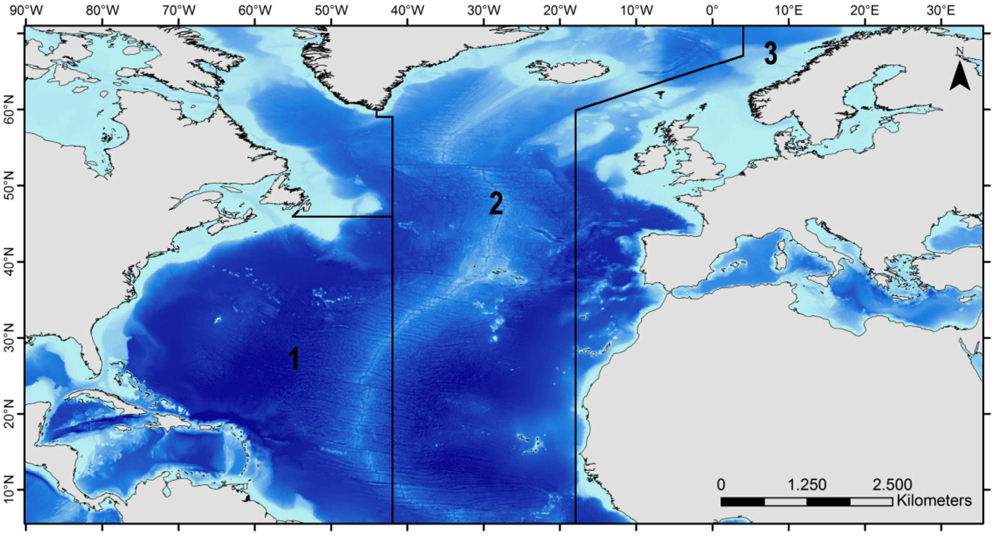

Figure 2-4: International Whaling Commission stock boundaries for North Atlantic sei whale stocks. (1) Nova Scotia, (2) Iceland–Denmark Strait, (3) Eastern.

Feeding ecology

The sei whale is the only rorqual species that seems to have evolved the ability to capture prey both by engulfment, as do the other rorquals, or by skimming on relatively low prey concentrations, as do the right and bowhead whales *Balaena mysticetus* [1,22,99,100]. The ability to switch between feeding strategies results from anatomical adaptations of the internal baleen fringe, which is much finer than that of other rorquals [30], and of the mouth cross section, which has some features reminiscent of right whales [101].

This adaptation is reflected in the variety of prey recorded for the species. The sei whale has the most extensive list of prey of any baleen whale, including copepods, euphausiids, amphipods, decapods, cephalopods and fish, but the prey preferences are highly dependent on ocean basin and swarming characteristics of the prey, as shown in reviews by Budylenko [102] and Horwood [1].

Notwithstanding, in the North Atlantic, sei whales seem to be nearly stenophagous, feeding almost exclusively on the copepod *Calanus finmarchicus* and, to a lesser extent, on the euphausiids *Meganyctiphanes norvegica* and *Thysanoessa inermis* [1,30,103,104]. Off Iceland, this pattern seems to be inverted, and euphausiids are reported as the main prey, followed by copepods [42,104]. In samples from whales taken in the Labrador Sea in 1971 and 1972, copepods were predominant in the stomach contents during the early whaling season (June–September); later in the season (October–November), they were replaced by euphausiids [105]. This pattern probably reflected seasonal changes in the zooplankton assemblage and abundance. Care should be taken in extrapolating that information to the present time, since copepods show substantial differences in abundance and distribution linked to the phase of the North Atlantic Oscillation, which has shifted from predominantly negative in the 1960s to predominantly positive since the mid-1980s [106,107].

In this ocean basin, even when apparently suitable schooling fish are available, sei whales demonstrate a preference for zooplankton [100]. Nevertheless, 2% of Icelandic whales examined consumed either sandeel *Ammodytes tobianus*, lumpfish *Cyclopterus lumpus* or capelin *Mallotus villosus* [42].

Sei whales feeding on copepods seem to prefer the late copepodite stages [22,108]. After the nauplius phase, copepods grow through five copepodite stages (CI-CV) and develop an oil sac filled with wax esters so that the latter stages have high caloric content [109]. In the same manner as for right whales [110], by preying on the late-stage copepodites, sei whales probably maximize their energy intake.

The stenophagy of sei whales found in the North Atlantic contrasts with the more euryphagic nature of the species found in other areas. In the North Pacific, for example, sei whales seem to switch between prey according to availability [108,111]. In most cases, an individual focuses on one prey species during single feeding bouts, and only rarely do individuals have more than one prey species in their stomachs [108].

Their ability to switch between feeding strategies theoretically allows sei whales to feed in the less productive wintering areas (perhaps with lower energetic gain) as has been reported for other areas [1,112]. Unfortunately, in part due to the scarcity of information about the winter distribution of sei whales in the North Atlantic, no information about their feeding ecology in lower latitudes could be found.

After the last catches of sei whales in the North Atlantic by Iceland in 1988 under a special IWC permit for scientific research, very little research has been carried out on the feeding preferences of the species. No data on stomach contents from stranding records was found. This is not surprising since sei whale strandings are relatively rare due to the pelagic habits of the species.

Studies on the diet of North Atlantic sei whales using alternative methodologies, such as analysis of fatty acids, stable isotopes or faeces, are also scarce. The only information on sei whale feeding in the North Atlantic from non-lethal methods comes from an analysis of faeces collected during an episodic influx into the southern Gulf of Maine in 1986 that revealed only mandibles of *C. finmarchicus* [22].

Some synoptic studies of the occurrence of potential prey and sei whales are available and support the diet preferences of the species in the North Atlantic. In western and southern Greenland, sightings of sei whales in September 2005 were in areas with the highest densities of krill *M. norvegica* and *Thysanoessa* sp. [80].

A concentration of sei whales reported in early summer 2004 near the frontal area of the Charlie Gibbs Fracture Zone, over the Mid-Atlantic Ridge, overlapped with the only area where concentrations of *C. finmarchicus* were found in the upper 100m of the water column [113]. Similarly, off New England, the preference for *C. finmarchicus* is supported by the consistent synoptic occurrence of the whales with that copepod [21,22,83].

When preying on zooplanktonic organisms, sei whales show feeding periodicity, which seems to be associated with the diel vertical migration of the prey, and take advantage of prey concentrations near the water surface between dusk and dawn [1,83]. This may be a strategy to maximize foraging efficiency [83].

On feeding grounds, baleen whales (and sei whales in particular) tend to associate with oceanic frontal systems such as fronts, eddies and upwelling systems [114-118]. In the Pacific, sei whales tend to follow and be found in the warmer waters of the major mixing zones and tend to be associated with eddies broken from the fronts [114]. Similarly, aggregations of sei whales have recently been found to be associated with fine scale frontal processes in the North Atlantic [113].

These associations with oceanic frontal systems probably enhance foraging efficiency by association with large concentrations of prey because of advection processes or due to increased prey biomass related to primary production [113,118]. Nevertheless, the ways in which sei whales detect these systems and subsequently exploit them has not yet been investigated in detail. Since climactic changes may bring changes in the distribution and intensity of frontal systems, it is important to understand better the role that these oceanographic features play in sei whale feeding behaviour and how predicted changes in climate may impact feeding ecology.

An affinity with submarine canyons is suggested by a sighting of more than 40 sei whales in a multispecies assemblage over and near Hydrographer Canyon in 1980 [119].

Reproduction

Horwood [1] reviewed information on sei whale reproduction and, since then, no new information has become available for the North Atlantic. The most recent reproductive research in the North Atlantic was based on data collected in western Norway between 1951 and 1957 [3], from Iceland in 1967, 1969, 1972–75 and 1977–81 [120], and from the north-western Atlantic from 1966 to 1972 [105]. All reproduction studies on sei whales have been based on histological and morphological examination of material from catches.

Males

Male sexual maturity in cetaceans is normally assessed by examination of testes weight, spermatogenesis and histological examination of the testes tubules; animals are classified as immature, pubertal, active and resting [121]. The most commonly used techniques to calculate the body length and age at sexual maturity are not comparable for most purposes, and some care should be taken when comparing results from different studies [122].

Horwood [1] used the body length when 50% of males were sexually mature as to define the length at sexual maturity. He estimated lengths at 50% maturity of 12.0m from Canada, 12.7m from Iceland and 12.8m from Norway. The values from Canada are lower, but he noted that better agreement would be achieved if a different combined testes weight at maturity, comparable with that used in the other studies, was used in the

Canadian study. The median age at sexual maturity was not calculated in the same manner since it was considered prone to being highly underestimated due to selection for large whales. Lockyer [123] pooled all the available information to build a table of baleen whale reproductive parameters in which the mean length at sexual maturity estimated for North Atlantic male sei whales was 12.9m, and the mean age at sexual maturity, estimated by ear plug growth layer analysis was 8 years.

The evidence for seasonal cycles in spermatogenesis is equivocal for sei whales. Data from the northwest Atlantic suggest a seasonal cycle with increased activity towards the end of the summer [105]. The recent use of hormonal techniques in samples originating from Icelandic whaling showed a rise in hormonal activity during July–September, supporting a seasonal breeding cycle [124]. Results for the Southern Hemisphere do not support that hypothesis since no seasonal cycles in spermatogenesis were found [23,28]. That could be due to a low sensitivity of anatomical and histological methods to these cycles but could also reflect physiological differences between populations.

Females

Several methods and datasets have been used to estimate dates of conception of sei whales. Results indicate that conception occurs over a range of months, with a peak in June and July in the Southern Hemisphere, in November and December in the North Pacific and in December and January in the North Atlantic [1]. Based on the average length at birth (4.5m), the gestation period in the North Atlantic was calculated to be 10.7 months [120].

Lockyer and Martin [120] calculated a pregnancy rate of 0.40–0.44 and an ovulation rate of 0.59 for Icelandic sei whales, which is in disagreement with the ovulation rate of 0.26 calculated for the northwest Atlantic by Mitchell and Kozicki [105]. This discrepancy might be due to the method of calculation used by Mitchell and Kozicki [105], in which the zero *corpora* classes are included in the regression of *corpus* number at age, thus rendering that estimate not directly comparable with others [123]. An ovulation rate of 0.26 and low pregnancy rate of 0.30 were interpreted by the authors as indications of a 3- to 4-year cycle, while all other studies from the North Atlantic and other oceans indicate a 2-year cycle [1,123]. However, it is noteworthy that after accounting for the bias against lactating females, Rice [125] estimated a true pregnancy rate of 0.36 suggesting a 3-year pregnancy cycle for the great majority of female sei whales. Using hormonal methods, Kjeld [124] found an apparent pregnancy rate in mature females of between 0.37 and 0.38, but when indecisive results were accounted for, that value rose to 0.41, closely agreeing with the results from Lockyer and Martin [120].

Two studies from the North Pacific indicated the existence of age-specific fecundity within the populations studied [24,125]. Lockyer and Martin [120] did not detect evidence for age-specific ovulation in Icelandic females up to 45 years old.

Using the same approach as described above for males, Horwood [1] calculated the body length at 50% mature for females to be 13.1m for Iceland, 13.1–13.4m for Canada and 13.4 m for Norway. As with males, he did not carry out the same analysis for age. Lockyer [123] estimated the mean length at sexual maturity for North Atlantic female sei whales to be 13.3m and the mean age at sexual maturation (t_m) to be 8 years.

Several studies on ear plug transition phase reported a decline in t_m over time for sei, fin and minke whales. Although the calculations may be affected by several types of biases, at least for some cases, there is evidence for a real decline in the age at sexual maturation [123].

For the Icelandic catches, Lockyer and Martin [120] analysed the data for mean and confidence limits separately in groups of cohorts to avoid biases associated with using linear regression to describe the trends (see Lockyer [123] for discussion). They detected a significant reduction in the t_m between the periods pre- and post-1940, in the order of 1.7 years.

It has been hypothesized that the decline in whale populations, due to intensive exploitation, may have led to a lower intra- and interspecific competition for prey that, in turn, may have led to faster growth rates inducing precocious puberty [126,127]. Although the trends of decreasing age at sexual maturation seem, at least partially, to reflect the effect of intensive exploitation, the implications of that decline are still uncertain [123].

After the cessation of sei whaling in the North Atlantic, few further studies on reproduction were undertaken, and those that were undertaken were based on whaling-derived samples [124]. If, as it seems, there was an effect of exploitation on the reproductive parameters of the population, cessation of whaling may have reversed it, at least partially. However, other factors may have also influenced those parameters, such as a rearrangement of the North Atlantic ecosystem owing to changes in whaling, fishing (including overfishing) and hydrometeorology [128-130]. Consequently, it is unlikely that reproductive ecology of the sei whale in the North Atlantic remains unchanged. Since ecological and management models normally need reproductive components, it is important that new studies using recent samples are conducted to provide more up-to-date figures.

Recently, gonadal steroid hormones have been used to determine pregnancy in marine mammals [131]. The dramatic differences in the concentration of these hormones in the blubber [132] and hormonal metabolites in faecal samples [133] of pregnant and non-pregnant cetaceans opened the possibility of assessing the pregnancy status of large numbers of free ranging cetaceans without the need for capture. There are some complications with the use of these techniques, since sometimes covariates are required to discriminate pregnancy fully [134], but in general results are reliable. The work by Kjeld *et al.* [135] using blood samples from sei whales caught in Icelandic whaling operations was important in establishing ground values for steroid hormones of North Atlantic sei whales backed by anatomical and histological methods. That work proved that hormonal methods not only complement anatomical and histological methods but can in some instances actually be more sensitive. Nonetheless, the authors also pointed out disagreements between serum oestradiol values between sei and fin whales, advising further study.

No hormonal studies on the sei whale using tissues other than blood have been published, but based on studies of other species there is no reason to think that data on reproduction could not be obtained from non-lethal techniques such as biopsy sampling or faecal collection. A hormonal study using blubber and skin from sei whales is necessary in order to test whether, and to what extent, the reproductive rate of the species can be assessed using these techniques. It must be acknowledged that recalculation of reproductive parameters using only hormonal techniques would be both technically and logistically challenging and may not be practical presently. Nevertheless, in the absence of other data sources, the feasibility of using hormonal methods should at least be investigated.

Abundance

There is no information on the pre-exploitation abundance of sei whales in the North Atlantic. The total catch of sei whales by modern whalers in the late 19th and 20th centuries has been estimated to be 16.933 [136]. That figure is, however, probably affected by some confusion in the identification of sei and Bryde's whales and does not include any proportion of the 12.322 unspecified species included in the catch [136].

As discussed above, sei whales appear to occur at low densities over most of the areas covered by systematic sighting surveys in the North Atlantic either because the surveys encompassed areas not frequented by sei whales or were carried out during a season when they are scarce. Here, only the most recent estimates based on cetacean surveys specifically designed for abundance estimation are presented. Previous estimates based on tagging from mark-recapture programmes [49] are considered inadequate given that, for such programmes to be effective, large numbers of animals have to be tagged and recaptured in each sampling period [137]. Similarly, estimates from early line transect surveys are likely to be negatively biased due to poor sampling design and failure to correct for submerged animals and ship avoidance [137].

No population estimate for the entire North Atlantic is available. Some estimates have been made for restricted areas of the eastern and western North Atlantic (Table 2-4). These abundance estimates cover only parts of the known summering habitat of North Atlantic sei whales, and where data are available, it is uncertain what fraction of the population was surveyed. If separate ecological units exist in the North Atlantic, it is possible that in some areas they overlap both spatially and temporally and have been treated as a single unit in abundance estimates.

A series of synoptic multinational sighting surveys in the north-eastern and central North Atlantic (NASS) were conducted in 1987, 1989, 1995 and 2001, from late June to early August, with some differences in area coverage between years [58,138]. Abundance estimates for sei whales were produced from the NASS surveys for the years 1987 and 1989, but the areas and survey periods differed between these 2 years. For the 1987 survey, a first estimate of 1243 sei whales was calculated by Gunnlaugsson and Sigurjónsson [139] but, due to timing and coverage area, the estimate was considered to represent only part of the population previously targeted by Icelandic whalers. During 1989, the survey was undertaken later in the season (mid-July to mid-August) and extended to blocks farther south from the blocks surveyed in previous years. An abundance estimate of 10300 sei whales (coefficient of variation; CV = 0.268) was calculated for that survey; the data from 1987 were reanalysed using the same methodology yielding an estimate of 1293 sei whales (CV = 0.603). The difference between the two estimates was interpreted as being caused by the substantial survey effort made in 1989 in the area south of the area surveyed in 1987. The estimate for 1989 for a comparable area to the 1987 survey was 1590 whales (CV = 0.4) similar to the estimate for 1987 [36].

During the multinational Cetacean Offshore Distribution and Abundance in the European Atlantic survey (CODA), conducted in July 2007 along the British, Irish, French and northern Spanish continental shelves, sei whales were only detected in one of the four survey blocks off northwestern Spain; the sei whale abundance estimate for that block was 366 whales (CV = 0.33; Anonymous [57]).

Three recent abundance estimates are available for parts of the Canadian and US Atlantic Exclusive Economic Zone during the summers of 2002, 2004 and 2006. The 2004 estimate (386; CV = 0.85) is considered the best estimate for the Nova Scotia stock, although it is considered conservative because of uncertainties about population structure and whale movements between surveyed and unsurveyed areas [82].

Acoustics

Very few accounts of the acoustic characteristics and behaviour of the sei whale exist. Baumgartner et al. [84] reviewed the scarce literature available and added some information from their own research, resulting in only five published reports describing sei whale calls. The current knowledge about sei whale call characteristics is summarized in Table 2-5. Recordings made in different areas and seasons present much dissimilarity that could be a result of geographic separation and population-specific acoustic characteristics, or because different types of calls have distinct ecological functions [84,141].

Table 2-4: Recent abundance estimates and CVs for the North Atlantic sei whale

Area	Period	Abundance	CV	Reference
Southern Gulf of Maine to Maine (United States)	August 2002	71	1.01	Waring et al. [82]
Gulf of Maine to lower Bay of Fundy (United States)	June–July 2004	386	0.85	Waring et al. [82]
Southern Gulf of Maine to upper Bay of Fundy to Gulf of St. Lawrence (United States, Canada)	August 2006	207	0.62	Waring et al. [82]
Iceland–Denmark Strait (excluding southernmost blocks)	June–July 1987	1293	0.60	Cattanach et al. [36]
Iceland–Denmark Strait (including southernmost blocks)	July–August 1989	10300	0.27	Cattanach et al. [36]
CODA block 3 (northwestern Spain)	July 2007	366	0.33	Anonymous [57]

CV, coefficient of variation; CODA, Cetacean Offshore Distribution and Abundance.

Table 2-5: Summary of acoustic characteristics of known sei whale calls

Type of call	Minimum frequency (Hz)	Maximum frequency (Hz)	Average duration (s)	Phrase composition (parts)	Area	Reference
FM down-sweep	34.0	82.3	1.38	1–3	New England, United States	Baumgartner et al. [84]
FM up-sweep	–	3k	0.7	7–10	Nova Scotia-Newfoundland	Thompson et al. (1979) *in* Baumgartner et al. [84]
FM up-sweep	1.5k	3.5k	0.03–0.04	10–20	Nova Scotia	Knowlton et al. [140]
FM down-sweep	44.6	100.3	1.2	–	Hawaii	Rankin and Barlow [141]
FM down-sweep	21.0	39.4	1.2	–	Hawaii	Rankin and Barlow [141]
FM down-, up-sweeps	200	600	1.1	–	Antarctic Peninsula	McDonald et al. [142]
Tonal call	100	950	0.45	1–5	Antarctic Peninsula	McDonald et al. [142]
Broadband signals	100	600	1.5	–	Antarctic Peninsula	McDonald et al. [142]

FM, frequency modulated.

Mortality and threats

Natural mortality and threats

Killer whales *Orcinus orca* have been reported to prey on all great whales and are considered their only significant natural predator [143,144]. The contribution of large cetaceans to the diet of mammal-eating killer whales is difficult to infer [145]. Reciprocally, the scale and ecological significance that predation by killer whales have on large whale populations is still a matter of debate in the scientific literature [146].

There are relatively fewer reports of killer whale predation on sei whales than on other species of large whales [143,145], but that does not necessarily mean a lower incidence of predation since predation could be occurring farther offshore and be more difficult to detect.

Parasites have been known to cause major health problems in marine mammals [147] and are considered one of the most important sources of natural mortality for sei whales [1]. Helminth parasites can cause severe complications, especially when infecting the liver, urinary and respiratory systems and brains [147-149].

From a small sample from Iceland (*n* = 24) Lambertsen [150] found a high incidence of the Acantocephalan *Bolbossoma* spp. in the colon. In the Antarctic, sei whales presented a high incidence of infection with several species of helminth parasites; most infections were restricted to the genitalia, stomach and intestine [151]. Nevertheless, the implications of these infections on sei whales' health and mortality are unknown.

Marine mammal viral diseases have started to be investigated only recently but can induce high mortality rates and be persistent at a population level [147,152]. Very little information exists on the occurrence of viral diseases in sei and other baleen whales, and there are no records of epizootics in baleen whales [153]. The only evidence for viral diseases in North Atlantic sei whales comes from Lambertsen [150], who found inflammation consistent with a viral pathogen in the lungs of 14% of the sei whales examined in Iceland, but could not detect the causative agent.

An unknown disease with an incidence of 7% in sei whales off California caused the shedding of baleen plates, which impaired their ability to feed [154]. There is no evidence for the same condition in North Atlantic sei whales.

Estimates of sei whale natural mortality rates have been impeded by several sources of bias. The estimates have been based on catch at age techniques that are likely to underestimate rates and may not be representative of the populations studied [1]. A wide range of mortality rate estimates, between 0.05 and 0.1, have been presented with large confidence limits.

Lockyer [155] estimated mortality rates for Durban and South Georgia in the Southern ocean from age-frequency distributions and calculated that natural mortality in Durban was in the range of 0.047–0.066. For South Georgia, she argued that, though the amount of pelagic whaling for the area was difficult to assess, the natural mortality rate varied between 0.054 and 0.068.

For the eastern North Pacific, Rice [125] used a similar approach to calculate mortality rates for animals with at least eleven earplug layers (the mean number at sexual maturity estimated for that population) and estimated a total (natural plus human-induced) mortality rate of 0.088 for females and 0.103 for males, but did not try to calculate the natural mortality rate.

Based on age-frequency distributions, Lockyer and Martin [120] calculated instantaneous total mortality rates from 0.092 to 0.096 (males) and from 0.103 to 0.104 (females) for sei whales from Iceland but cautioned that these rates suffered from some sources of bias, including the irregular catching history of the species and a preference for larger (and hence older) animals in the catch. Natural mortality was not calculated in this study, and the authors only allude to an estimate of 0.06 from the Southern Hemisphere. The same authors also suggested that the age distributions within the catch might not be representative of the whole population due to geographic segregation of feeding grounds [120].

Human-induced mortality and threats

In addition to direct exploitation, expanding anthropogenic activities in the marine environment probably impact sei whales, such as fisheries, shipping, pollution, military activities, and deep-water hydrocarbon extraction.

Fisheries. Entanglement in fishing gear is a major source of mortality for many species of small and large cetaceans, particularly great whales occupying coastal habitats in heavily fished regions [156]. Sei whales primarily occupy pelagic habitats and are, therefore, less likely to encounter static fishing gear than coastal species such as humpback and minke whales. Serious injury and mortality of sei whales due to entanglement in fishing gear off the northeast United States and New Brunswick, Canada, have been reported though at a much lower rate than for right, humpback, fin and minke whales [157].

Fisheries may also have ecological impacts on marine mammals through the large-scale removal of animal biomass from marine ecosystems, which can induce direct and indirect competition for resources between fisheries and marine mammals [158]. The ways in which fishery-induced changes to the marine ecosystem may affect marine mammals are complex and have proved to be difficult to test [158]. Nevertheless, in theory, the rapid removal of biomass from marine ecosystems could eventually lead to a collapse or change in the community structure that can reflect at several ecological levels with unpredictable consequences including to the great whales [153]. The feeding strategy of sei whales should provide them with flexibility to switch between planktonic and fish prey. However, in the North Atlantic, the sei whale seems to be stenophagous, so the species' adaptability to rapid changes in prey assemblages in this ocean is questionable.

Ship collisions. Enhanced monitoring in recent years has resulted in an increasing number of ship collisions with great whales being documented, which have been attributed to increases in shipping traffic as well as to the increasing size and speed of ships [159-161]. Between 1970 and 1990, it is estimated that between 13% and 20% of the strandings of great whales in the United States, Italy, France and South America were caused by collisions with ships [159]. Sei whales are also affected by this increasing phenomenon [159,161,162]. Although reports of collisions with sei whales are comparatively less common than reports of collisions with other, more coastal, species [159], that could be due to the low probability of detecting an event when it occurs offshore. The ocean-wide impact of ship strike rates on sei whale mortality is poorly understood but may be significant due to increases in merchant shipping traffic.

Noise. Human-generated noise in the oceans has been recognized as a potential threat to large whales since the early 1970s, but it was not until recently that the problem started receiving more attention [163].

Ocean noise levels have shown a continued increase in association with increasing human activities such as commercial shipping, seismic surveys, military activities, scientific research, etc. It has been estimated that in many parts of the world, between 1950 and 2000, there was an overall increase in low-frequency noise from shipping of in the order of 16dB [164].

Direct mortality caused by anthropogenic noise is a controversial issue, but the potential risk seems to be higher in deep diving cetaceans and less pronounced for baleen whales [165-168].

Acoustic interference, or masking, can have detrimental effects on cetaceans by interfering with conspecific communication as well as by reducing their ability to detect other biologically important sounds and abiotic sound cues [163]. Effects may be short- or long-term and may contribute to reduced survivorship and reproductive success as a result of persistent degradation of acoustic habitat over an animal's lifetime [163].

The acoustic repertoire of sei whales is poorly described (Baumgartner et al. [84], this review), and the way they use their acoustic environment is virtually unknown. These parameters require further research to evaluate the impact of human-induced noise on sei whale ecology.

Contaminants. Currently, there is no evidence that organochlorine and heavy metal contaminants pose a major threat to baleen whales. Existing data on mysticetes support the view that feeding at lower trophic levels results in smaller contamination burdens than found in most odontocetes that feed at higher trophic levels [169]. Studies comparing contamination by organochlorine compounds in odontocetes and mysticetes from the same areas show levels of contamination at least one order of magnitude higher in the odontocetes [170-173]. That holds true for the North Atlantic sei whale, as has been shown by a study of Polychlorinated biphenyls (PCB) and Dichlorodiphenyltrichloroethane family compounds (DDTs) using samples from sei, fin, sperm *Physeter macrocephalus* and pilot *Globicephala melas* whales, white-sided dolphins *Lagenorhynchus acutus* and harbour porpoises *Phocoena phocoena* from the north-eastern North Atlantic, where sei whale total DDTs and PCB concentrations (mg kg^{-1}) were from 14 to 136 times and from 23 to 146 times, respectively, lower than those of the odontocete species [170]. Heavy metal contamination also tends to be lower in mysticetes and in the sei whale in particular; differences from odontocetes vary from non-existent to two orders of magnitude [172,173].

Climate change. Global warming is unequivocal [174] and is likely to have detrimental impacts on marine mammal populations [175]. Climate change is predicted to affect cetacean distribution, timing and range of migrations, abundance, mortality, reproductive success and prey resources [176,177]).

Rates of temperature change in marine ecosystems have not been evenly distributed globally. Some of the more drastic increases in sea surface temperature have been reported for regions of the North Atlantic, including the Newfoundland–Labrador, west Greenland, Scottish and Icelandic shelves as well as the Faroe Plateau [178]. These regions include much of the known sei whale summer habitat.

Marine ecosystems have a critical thermal boundary, where a small increase in temperature triggers abrupt ecosystem shifts at multiple trophic levels [179]; that process may already be causing ecosystem reorganization in marine copepod biodiversity in the North Atlantic [130,179]. In the north-eastern North Atlantic and adjacent seas, there has been a shift in the copepod assemblage: warmer water species are moving poleward more than 10° in latitude, and there is an associated decrease in the number of subarctic and arctic species. West of the Mid-Atlantic Ridge, especially in the Labrador Sea, the trend is opposite: an increase and expansion to the south of the cold water species is taking place [130]. Those changes may have special implications for sei whales given that copepods seem to constitute their main prey in this ocean basin. Nevertheless, it is difficult to predict if sei whales will suffer or profit from these changes due to uncertainties about the ecosystem reorganization and about sei whales' ability to cope with prey assemblage changes in the North Atlantic.

Resumption of whaling. Commercial sei whaling in the North Atlantic ceased in 1986 after the moratorium in commercial whaling implemented by the IWC. Under special scientific permits, however, Icelandic whalers captured 70 sei whales between 1986 and 1988. Aboriginal subsistence hunters in Greenland were allowed to take two whales in 1989 and one in 2006 [136].

The resumption of commercial whaling in the North Atlantic is still a contentious issue within the IWC, and the possibility of operations targeting sei whales cannot be ruled out in the near future.

DISCUSSION AND CONCLUSIONS

In his seminal publication on the sei whale, Horwood [1] identified research projects required to fill critical knowledge gaps, which in his opinion could be conducted using non-lethal techniques and existing data. Topics were as follows:

(i) stock identity and genetics; (ii) distribution and migration; (iii) reproduction and mortality; and (iv) abundance estimation. To date, little progress has been made on these recommendations.

The bibliometric analysis undertaken in the present work shows that research on the sei whale has been almost stagnant for the last 20 years. There was a clear departure from the growth expected from Price's Law on scientific output. According to Price's Law, if scientific output on a subject does not follow exponential growth, either the field has reached saturation, which is clearly not the case, or not enough resources are being assigned to research.

The literature on virtually every aspect of the population biology of the North Atlantic sei whale is meagre and in most cases outdated. Furthermore, much of the research carried out over the last two decades relies on old whaling data, and little effort has been directed to applying current methodologies to understanding sei whale population biology.

Meanwhile, some paradigms on population management and conservation have changed. The impacts of human activities are no longer restricted to single species and locations but have become global and transverse entire ecological communities. For example, as fishing pressure on several marine species increases, the lack of sustainable fishery management is a threat not only to the species being harvested but also to the equilibrium of the communities. Detrimental effects occur at all trophic levels [180].

As a result of these changes, it has been argued that marine research has shown a major shift in orientation to face an ever increasing environmental deterioration, and is now driven by the urgency to understand and ameliorate disturbed marine ecosystems, with implications for research priorities and the allocation of funds [181]. An example of that shift is a change in the focus of cetacean research from basic biological and ecological issues to conservation-related topics [182]. One problem with this new paradigm is that it works against species or ecosystems for which there is little or no ongoing research. Since it is not possible to detect if they are facing conservation problems, investment is channelled away to other pressing issues, even if in reality they are in trouble. These species or ecosystems may end up being ‘forgotten’ by science, the general public and decision makers. That can create a paradox in which our ignorance about a resource puts it at risk by diverting investment to conserve other resources.

To avoid that paradox, it is important to focus efforts on gaining more information on those species and ecosystems for which our understanding is severely limited as is the case with cryptic species such as the sei whale.

The current research needs for the North Atlantic sei whale still include those identified by Horwood more than 20 years ago. To them are added the need to understand (i) the role of this species in the ecosystems it integrates into, and (ii) how changes in this ocean’s marine environment will affect the species. The existing scientific knowledge on sei whales, largely derived from historical whaling data, is insufficient for making sound scientific and management decisions. This information is critically important given current human impacts on the marine environment and the likelihood of large-scale ecosystem changes due to climate change.

Research on the sei whale has been hindered in the past in part by its pelagic nature, which makes the study of the species both logistically difficult and expensive, compared with other more coastal species. With the aid of the most recent techniques to derive and analyse data, such as satellite telemetry and passive acoustic monitoring, it is now possible to conduct cetacean research efficiently even in pelagic habitats, and some of the research goals mentioned above can be achieved with fairly low human and financial investment. Other goals demand more investment but are essential to fill the data gaps. Among the most important are population assessments using synoptic multinational surveys that cover a wide latitudinal and longitudinal range and the clear definition of stock identity within the North Atlantic.

Overall, it is essential to bring research on North Atlantic sei whales out of its stagnancy and into the 21st century to ensure proper management of the species and the ecosystems it integrates with in the North Atlantic Ocean.

REFERENCES

1. Horwood J, editor (1987) The sei whale: Population biology, ecology & management. London: Croom Helm. 375 p.

2. Perry S, DeMaster DP, Silber GK (1999) The sei whale (*Balaenoptera borealis*). Marine Fisheries Review 61: 52-58.

3. Jonsgård A, Darling K (1977) On the biology of the eastern North Atlantic sei whale, *Balaenoptera borealis* Lesson. Reports of the International Whaling Commission Special issue 1: 124-129.

4. Donovan GPe (1989) The comprehensive assessment of whale stocks: The early years. Reports of the International Whaling Commission.

5. International Whaling Commission (1977) Report of the special meeting of the scientific committee on sei and Bryde's whales. Cambridge, UK: International Whaling Commission. 62 p.

6. Price DJ (1951) Quantitative measures of the development of science. Archives Internationales d'Histoire des Sciences 14: 85-93.

7. Great Britain General Register Office (1891) Preliminary report of the census of the United Kingdom. Journal of the Royal Statistical Society 54: 457-491.

8. Tønnessen JN, Johnsen AO (1982) The history of modern whaling. Berkeley and Los Angeles, CA, USA: University of California Press. 798 p.

9. Uhen MD (2010) The origin(s) of whales. Annual Review of Earth and Planetary Sciences 38: 189-219.

10. Gatesy J, O'Leary MA (2001) Deciphering whale origins with molecules and fossils. Trends in Ecology and Evolution 16: 562-570.

11. Flower WH (1883) On whales past and present, and their probable origin. Nature 28: 199-202.

12. Flower WH (1883) On whales past and present, and their probable origin II. Nature 28: 226-230.

13. Price SA, Bininda-Emonds ORP, Gittleman JL (2005) A complete phylogeny of the whales, dolphins and even-toed hoofed mammals (Cetartiodactyla). Biological Reviews 80: 445-473.

14. Geisler JH, Uhen MD (2005) Phylogenetic relationships of extinct cetartiodactyls: Results of simultaneous analyses of molecular, morphological, and stratigraphic data. Journal of Mammalian Evolution 12: 145-160.

15. Spaulding M, O'Leary MA, Gatesy J (2009) Relationships of Cetacea (Artiodactyla) among mammals: increased taxon sampling alters interpretations of key fossils and character evolution. PLoS ONE 4: e7062.

16. Perrin WF, Mead JG, Brownell Jr. RL (2009) Review of the evidence used in the description of currently recognized cetacean subspecies. http://noaa.ntis.gov/view.php?pid=NOAA:ocn540583513: US Department of Commerce, National Oceanic and Atmospheric Administration. NOAA-TM-NMFS-SWFSC-450 NOAA-TM-NMFS-SWFSC-450. 35 p.

17. Baker CS, Dalebout ML, Funahashi N, Ma YU, Steel D, et al. (2004) Market surveys of whales, dolphins and porpoises in Japan and Korea, 2003-2004, with reference to stock identity of sei whales. pp. 8.

18. Taylor B (2005) Identifying units to conserve. In: Reinolds JEI, Perrin WF, Reeves RR, Montgomery S, Ragen TJ, editors. Marine mammal research - Conservation beyond crisis. Baltimore, ML, USA: John Hopkins University Press. pp. 149-162.

19. Horwood J (2009) Sei whale *Balaenoptera borealis*. In: Perrin WF, Würsig B, Thewissen JGM, editors. Encyclopedia of Marine Mammals. 2nd ed. London & New York: Academic Press. pp. 1001-1003.

20. Horwood J (1986) The lengths and distribution of Antarctic sei whales. Scientific Reports of the Whales Research Institute 37: 47-60.

21. Payne PM, Wiley DN, Young SB, Pittman S, Clapham PJ, et al. (1990) Recent fluctuations in the abundance of baleen whales in the southern Gulf of Maine in relation to changes in selected prey. Fishery Bulletin 88: 687-696.

22. Schilling MR, Seipt I, Weinrich MT, Kuhlberg AE, lapham PJC (1992) Behavior of individually-identified sei whales, *Balaenoptera borealis*, during an episodic influx into the southern Gulf of Maine in 1986. Fishery Bulletin 90: 749-755.

23. Gambell R (1968) Seasonal cycles and reproduction in sei whales of the Southern Hemisphere. Discovery Reports 35: 31-134.

24. Masaki Y (1976) Biological studies on the North Pacific sei whale. Bulletin of the Far Seas Fisheries Research Laboratory 14: 1-104 + 101 pl.

25. Lockyer CH (1977) Some possible factors affecting age distribution of the catch of sei whales in the Antarctic. Reports of the International Whaling Commission Special issue 1: 63-70.

26. Gregr EJ, Nichol L, Ford JKB, Ellis G, Trites AW (2000) Migration and population structure of northern pacific whales off coastal british columbia: an analysis of commercial whaling records. Marine Mammal Science 16: 699-727.

27. Best PB, Lockyer CH (2002) Reproduction, growth and migrations of sei whales *Balaenoptera borealis* off the west coast of South Africa in the 1960s. South African Journal of Marine Sciences 24: 111-133.

28. Matthews LH (1938) The sei whale, *Balaenoptera borealis*. Discovery Reports 17: 183-290, +182 pls.

29. Cocks AH (1885) Additional notes on the finwhale fishery on the North European coast. The Zoologist - Third series 9: 134-143.

30. Collett R (1886) On the external characteristics of the Rudolphi's rorqual (Balaenoptera borealis). Proceedings of the Zoological Society of London: 243-265.

31. Haldane RC (1904) Whaling in Shetland. The Annals of Scottish Natural History 0: 74-77.

32. Southwell T (1904) On the whale fishery from Scotland, with some account of the changes in that industry and of the species hunted. The Annals of Scottish Natural History 50: 77-90.

33. Thompson DaW (1919) On whales landed at the Scottish whaling stations, especially during the years 1908-1914 - parts IV to VI. The bottlenose, humpback and finner whales. The Scottish Naturalist: 1-46.

34. Sanpera C, Aguilar A (1992) Modern whaling off the Iberian Peninsula during the 20th century. Report of the International Whaling Commission 42: 723-730.

35. Ingebrigtsen A (1929) Whales captured in the North Atlantic and other seas. Rapports et Proces-Verbaux des Reunions Conseil International pour l'Exploration de la Mer 56: 1-26.

36. Cattanach KL, Sigurjónsson J, Buckland ST, Gunnlaugsson T (1993) Sei whale abundance in the North Atlantic, estimated from NASS-87 and NASS-89 data. Report of the International Whaling Commission 43: 315-321.

37. Cocks AH (1887) The finwhale fishery of 1886 on the Lapland coast. The Zoologist - Third series 11: 207-222.

38. Cocks AH (1888) The finwhale fishery off the Lapland coast. The Zoologist - Third series 12: 201-208.

39. Cocks AH (1889) The finwhale fishery off the Lapland coast in 1888. The Zoologist - Third series 13: 281-290.

40. Andrews RC (1916) Whale Hunting with Gun and Camera. New York: D. Appleton and Company.

41. Cocks AH (1886) The finwhale fishery of 1885 on the North European coast. The Zoologist - Third series 10: 121-136.

42. Sigurjónsson J, Víkingsson GA (1997) Seasonal abundance of and estimated food consumption by cetaceans in Icelandic and adjacent waters. Journal of Northwest Atlantic Fishery Science 22: 271-287.

43. Aguilar A, Borrell A (2007) Open-boat whaling on the Straits of Gibraltar ground and adjacent waters. Marine Mammal Science 23: 322-342.

44. Brito C (2008) Assessment of catch statistics during the land-based whaling in Portugal. Marine Biodiversity Records 1, e92 doi:10.1017/S175526720700930X.

45. Aguilar A (1985) Further information on the movements of the sperm whale (*Physter macrocephalus*) in the North Atlantic. Mammalia 49: 421-424.

46. Best PB (1992) Catches of fin whales in the North Atlantic by the M.V. Sierra (and associated vessels). Report of the International Whaling Commission 42: 697-700.

47. Kapel FO (1979) Exploitation of large whales in West Greenland in the twentieth century. Reports of the International Whaling Commission 29: 197-214.

48. Mitchell ED (1975) Preliminary report on Nova Scotia fishery for sei whales (*Balaenoptera borealis*). Reports of the International Whaling Commission 25: 218-225.

49. Mitchell ED, Chapman DG (1977) Preliminary assessment of stocks of northwest Atlantic sei whales (*Balaenoptera borealis*). Reports of the International Whaling Commission Special issue 1: 117-120.

50. Mead JG (1977) Records of sei and Bryde's whales from the Atlantic coast of the United States, the Gulf of Mexico, and the Caribbean. Reports of the International Whaling Commission Special issue 1: 113-116.

51. Bompar J-M (2000) Les cétacés de Méditerranée. La Calade, Aix-en-Provence: Édisud. 186 p.

52. Notarbartolo Di Sciara G, Venturino MC, Zanardelli M, Bearzi G, Borsani JF, et al. (1993) Cetaceans in the central Mediterranean Sea: Distribution and sighting frequencies. Bolletin Zoológica 60: 131-138.

53. Boisseau O, Lacey C, Lewis T, Moscrop A, Danbolt M, et al. (2010) Encounter rates of cetaceans in the Mediterranean Sea and contiguous Atlantic area. Journal of the Marine Biological Association of the United Kingdom.

54. Weir CR, Pollock C, Cronin C, Taylor S (2001) Cetaceans of the Atlantic Frontier, north and west of Scotland. Continental Shelf Research 21: 1047-1071.

55. MacLeod K, Simmonds MP, Murray E (2003) Summer distribution and relative abundance of cetacean populations off north-west Scotland. Journal of the Marine Biological Association of the United Kingdom 83: 1187-1192.

56. MacLeod K, Simmonds MP, Murray E (2006) Abundance of fin (*Balaenoptera physalus*) and sei whales (*B. borealis*) amid oil exploration and development off northwest Scotland. Journal of Cetacean Research and Management 8: 247-254.

57. Anonymous (2009) NAMMCO Annual Report 2007-2008. Tromsø, Norway: North Atlantic Marine Mammal Commission. 399 p.

58. Sigurjónsson J, Gunnlaugsson T, Payne M (1989) NASS-87: shipboard sighting surveys in Icelandic and adjacent waters Jun-July 1987. Report of the International Whaling Commission 39: 395-409.

59. Øien N (2009) Distribution and abundance of large whales in Norwegian and adjacent waters based on ship surveys 1995-2001. NAMMCO Scientific Publications 7: 31-40.

60. Cresswell G, Walker D, editors (2001) Whales & Dolphins of the European Atlantic - The Bay of Biscay and the English Channel. Hampshire - UK: Wild Guides.

61. Carlisle A, Coles P, Cresswell G, Diamond J, Gordman L, et al. (2001) A report on the whales, dolphins and porpoises of Bay of Biscay and English Channel 1999. Orca 1: 5-53.

62. Waring GT, Nottestad L, Olsen E, Skov H, Vikingsson GA (2008) Distribution and density estimates of cetaceans along the mid-Atlantic Ridge during summer 2004. Journal of Cetacean Research and Management 10: 137-146.

63. Anonymous (2009) Cetacean offshore distribution and abundance in the European Atlantic (CODA). Unpublished report to the European Commission. http://biology.st-andrews.ac.uk/coda/documents/CODA_Final_Report_11-2-09.pdf. 42 p. + Annexes p.

64. Sequeira M, Inácio A, Reiner F, editors (1992) Arrojamentos de mamíferos marinhos na costa continental portuguesa entre 1978 e 1988: Serviço Nacional de Parques, Reservas e Conservação da Natureza.

65. Lockyer CH (1977) Some estimates of growth in the sei whale, *Balaenoptera borealis*. Reports of the International Whaling Commission Special issue 1: 58-62.

66. Gordon J, Steiner L, Goncalves JM (1990) Sei whales (*Balaenoptera borealis*) encountered in the Azores: A new record for the region. *Arquipelago* - Life and Earth Sciences 8: 97-100.

67. Silva MA, Prieto R, Magalhães SM, Cabecinhas R, Cruz A, et al. (2003) Occurrence and distribution of cetaceans in waters around the Azores (Portugal), Summer and Autumn 1999-2000. Aquatic Mammals 29: 77-83.

68. Santos MBO (2008) Distribuição e padrão de residência das baleias-de-barbas (Família *Balaenopteridae*) no Arquipélago dos Açores [p.]. Ponta Delgada: University of the Azores. 35 p.

69. Prieto R, Silva MA (2010) Marine Mammals. In: Borges PAV, Costa A, R. C, Gabriel R, Gonçalves V et al., editors. A list of the terrestrial and marine biota from the Azores. Cascais, Portugal: Princípia. pp. 344-345.

70. Freitas L, Dinis A, Alves F, Nóbrega F, editors (2004) Cetáceos no Arquipélago da Madeira. Machico, Portugal: Museu da Baleia do Machico.

71. Carrillo M, Pérez-Vallazza C, Álvarez-Vázquez R (2010) Cetacean diversity and distribution off Tenerife (Canary Islands). Marine Biodiversity Records 3: e97.

72. López-Jurado LF, André M, Martin V (1996) Inventário de los cetáceos de las aguas Canárias: aplicación de la Directiva 92/43/CEE. http://www.mma.es/portal/secciones/aguas_marinas_litoral/prot_medio_marino/biodiversidad/bm_bbdd_inventario.htm: Environment, Rural and Marine Ministry of the Spanish Government. pp. 64.

73. Notarbartolo-di-Sciara G, Politi E, Bayed A, Beaubrun P-C, Knowlton AR (1998) A winter cetacean survey off southern Morocco, with a special emphasis on right whales. Reports of the International Whaling Commission 48: 547-551.

74. Anonymous (2003) Seabird and cetacean surveys in the vicinity of the Chinguetti Oil Field, offshore Mauritania. Unpublished report prepared as part of an environmental impact assessment for the Chinguetti development project. http://www.miga.org/documents/ Mauritania_tr1.pdf. pp. 44 p. + appendixes.

75. Hazevoet CJ, Wenzel FW (2000) Whales and dolphins (Mammalia, Cetacea) of the Cape Verde Islands, with special reference to the Humpback Whale *Megaptera novaeangliae* (Borowski, 1781). Contributions to Zoology 69: 197-211.

76. Van Waerebeek K, Ndiaye E, Djiba A, Diallo M, Murphy P, et al. (2000) A survey of the conservation status of cetaceans in Senegal, The Gambia and Guinea-Bissau. Bonn, Germany: UNEP/CMS Secretariat. 80 p.

77. Weir CR (2010) A review of cetacean occurrence in West African waters from the Gulf of Guinea to Angola. Mammal Review 40: 2-39.

78. Heithaus M, Wirsing A, Dill L, Heithaus L (2007) Long-term movements of tiger sharks satellite-tagged in Shark Bay, Western Australia. Marine Biology 151: 1455-1461.

79. Compton R, Banks A, Goodwin L, Hooker SK (2007) Pilot cetacean survey of the sub-Arctic North Atlantic utilizing a cruise-ship platform. Journal of the Marine Biological Association of the United Kingdom 87: 321-325.

80. Laidre KL, Heide-Jørgensen M-P, Heagerty P, Cossio A, Bergstrom B, et al. (2010) Spatial associations between large baleen whales and their prey in West Greenland. Marine Ecology Progress Series 402: 269-284.

81. Lawson JW, Gosselin JF (2009) Distribution and preliminary abundance estimates for cetaceans seen during Canada’s marine megafauna survey - A component of the 2007 TNASS. Department of Fisheries and Oceans, Canadian Science Advisory Secretariat. Research Document 2009/031 Research Document 2009/031. vi + 28 p.

82. Waring GT, Josephson E, Maze-Foley K, Rosel PE (2009) U.S. Atlantic and Gulf of Mexico Marine Mammal Stock Assessments – 2009. http://www.nefsc.noaa.gov/publications/tm/tm213/: US Department of Commerce, National Oceanic and Atmospheric Administration 528 p.

83. Baumgartner MF, Fratantoni DM (2008) Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders. Limnology and Oceanography 53: 2197–2209.

84. Baumgartner MF, Van Parijs SM, Wenzel FW, Tremblay CJ, Esch HC, et al. (2008) Low frequency vocalizations attributed to sei whales (*Balaenoptera borealis*). Journal of the Acoustical Society of America 124: 1339-1349.

85. Davis RW, Fragion GS (1996) US Department of Interior ,Minerals Management Service, Gulf of Mexico OCS Region. New Orleans, LA. OCS Study MMS 96-0026 OCS Study MMS 96-0026. 29 p.

86. Jefferson TA, Schiro AJ (1997) Distribution of cetaceans in the offshore Gulf of Mexico. Mammal Review 27: 27-50.

87. Davis RW, Evans WE, Würsig B (2000) Cetaceans, sea turtles and seabirds in the northern Gulf of Mexico: distribution, abundance and habitat associations. New Orleans, LA: U.S. Department of the Interior, Geological Survey, Biological Resources Division, USGS/BRD/CR-1999-0006 and Minerals Management Service, Gulf of Mexico OCS Region. 1445-CTO9-96-0004 and 1445-1A09-96-0009 1445-CTO9-96-0004 and 1445-1A09-96-0009. 27 p.

88. Baumgartner MF, Mullin KD, May LN, Leming TD (2001) Cetacean habitats in the northern Gulf of Mexico. Fishery Bulletin 99: 219-239.

89. Roden CL, Mullin KD (2000) Sightings of cetaceans in the northern Caribbean Sea and adjacent waters, winter 1995. Caribbean Journal of Science 36: 280-288.

90. Swartz SL, Martinez A, Stamates J, Burks C, Mignucci-Giannoni AA (2001) Acoustic and visual survey of cetaceans in the waters of Puerto Rico and the Virgin Islands: February–March 2001.

http://www.sefsc.noaa.gov/docs/Cetaceans-survey-PuertoRico&VI2001.pdf: US Department of Commerce, National Oceanic and Atmospheric Administration. NOAA Technical Memorandum NMFS-SEFSC- 463 NOAA Technical Memorandum NMFS-SEFSC- 463.

91. Yoshida H, Compton J, Punnett S, Lovell T, Draper K, et al. (2010) Cetacean sightings in the eastern Caribbean and adjacent waters, spring 2004. Aquatic Mammals 36: 154-161.

92. Mignucci-Giannoni AA (1998) Zoogeography of cetaceans off Puerto Rico and the Virgin Islands. Caribbean Journal of Science 34: 173-190.

93. Boisseau O, Prieto R, Nicholson D, Gordon J (1999) An Atlantic cetacean survey using yachtsmen: a pilot study. Arquipelago - *Life and Marine Sciences* 17A: 101-107.

94. Dizon AE, Lockyer CH, Perrin WF, DeMaster DP, Sisson J (1992) Rethinking the stock concept: a phylogeographic approach. Conservation Biology 6: 24-36.

95. Donovan GP (1991) A review of IWC stock boundaries. Report of the International Whaling Commission special issue 13: 39-68.

96. Daniélsdóttir AK, Duke EJ, Joyce P (1991) Preliminary studies on genetic variation at enzyme loci in fin whales (*Balaenoptera physalus*) and sei whales (*Balaenoptera borealis*) form the north Atlantic. Reports of the International Whaling Commission Special Issue 13: 115-124.

97. Olsen E, Budgell P, Head E, Kleivane L, Nottestad L, et al. (2009) First satellite-tracked long-distance movement of a sei whale (*Balaenoptera borealis)* in the North Atlantic. Aquatic Mammals 35: 313-318.

98. Prieto R, Silva MA, Berube M, Palsbøll PJ (2012) Migratory destinations and sex composition of sei whales (*Balaenoptera borealis*) transiting through the Azores. IWC/64/Rep1rev1 SC/64/RMP6. http://iwc.int/iwc64docs: International Whaling Commission. pp. 7.

99. Gill CD, Hughes SE (1971) A sei whale, *Balaenoptera borealis*, feeding in Pacific saury, *Cololabis saira*. California Fish and Game 57: 218-219.

100. Watkins WA, Schevill WE (1979) Aerial observation of feeding behavior in four baleen whales: *Eubalaena glacialis*, *Balaenoptera borealis*, *Megaptera novaeangliae*, *Balaenoptera physalus*. Journal of Mammalogy 60: 155-163.

101. Brodie PF, Víkingsson G (2009) On the feeding mechanisms of the sei whale (*Balaenoptera borealis*). Journal of Northwest Atlantic Fishery Science 42: 49-54.

102. Budylenko GA (1978) On sei whale feeding in the Southern Ocean. Reports of the International Whaling Commission 28: 379-385.

103. Andrews RC (1916) Monographs of the Pacific Cetacea. II. The sei whale (*Balaenoptera borealis* Lesson). 1. History, habits, external anatomy, osteology, and relationship. Memoirs of the American Museum of Natural History New Series 1: 289-388 +214 pls.

104. Sigurjónsson J (1995) On the life history and autecology of North Atlantic rorquals. In: Blix AS, Walloe L, Ulltang Ø, editors. Whales, Seals, Fish and Man. Amsterdam: Elsevier Science. pp. 720.

105. Mitchell ED, Kozicki VM (1974) The sei whale (*Balaenoptera borealis*) in the northwest Atlantic Ocean. pp. 48+ 48 fig.

106. Fromentin J-M, Planque B (1996) *Calanus* and environment in the eastern North Atlantic. 2. Influence of the North Atlantic Oscillation on *C. finmarchicus* and *C. helgolandicus*. Marine Ecology Progress Series 134: 111-118.

107. Head EJH, Harris LR, Yashayaev I (2003) Distributions of *Calanus* spp. and other mesozooplankton in the Labrador Sea in relation to hydrography in spring and summer (1995–2000). Progress in Oceanography 59: 1-30.

108. Konishi K, Tamura T, Isoda T, Okamoto R, Hakamada T, et al. (2009) Feeding strategies and prey consumption of three baleen whale species within the Kuroshio-Current extension. Journal of Northwest Atlantic Fishery Science 42: 27-40.

109. Miller CB, Crain JA, Morgan CA (2000) Oil storage variability in *Calanus finmarchicus*. ICES Journal of Marine Science 57: 1786-1799.

110. Baumgartner MF, Mayo CA, Kenney RD (2007) Enormous carnivores, microscopic food, and a restaurant that's hard to find. In: Kraus SD, Rolland RM, editors. The Urban Whale: North Atlantic Right Whales at the Crossroads. Cambridge, MA, USA; London, England: Harvard University Press. pp. 138-171.

111. Flinn RD, Trites AW, Gregr EJ, Perry RI (2002) Diets of fin, sei, and sperm whales in British Columbia: An analysis of commercial whaling records, 1963-1967. Marine Mammal Science 18: 663-679.

112. Kawamura A (1973) Food and feeding of sei whale caught in the waters south of 40°N in the North Pacific. Scientific Reports of the Whales Research Institute of Tokyo 25: 219-236.

113. Skov H, Gunnlaugsson T, Budgell WP, Horne J, Nottestad L, et al. (2008) Small-scale spatial variability of sperm and sei whales in relation to oceanographic and topographic features along the Mid-Atlantic Ridge. Deep-Sea Research Part II 55: 254-268.

114. Nasu K (1966) Fishery oceanographic study on the baleen whaling grounds. Scientific Reports of the Whales Research Institute of Tokyo 20: 157-210.

115. Moore SE, Waite JM, Friday NA, 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: 249-261.

116. Tynan CT, Ainley DG, Barth JA, Cowles TJ, Pierce SD, et al. (2005) Cetacean distributions relative to ocean processes in the northern California Current System. Deep-Sea Research Part II 52: 145-167.

117. Doniol-Valcroze T, Berteaux D, Larouche P, Sears R (2007) Influence of thermal fronts on habitat selection by four rorqual whale species in the Gulf of St. Lawrence. Marine Ecology Progress Series 335: 207-216.

118. Bost CA, Cotté C, Bailleul F, Cherel Y, Charrassin JB, et al. (2009) The importance of oceanographic fronts to marine birds and mammals of the southern oceans. Journal of Marine Systems 78: 363-376.

119. Kenney RD, Winn HE (1987) Cetacean biomass near submarine canyons compared to adjacent shelf/slope areas. Continental Shelf Research 7: 107-114.

120. Lockyer CH, Martin AR (1983) The sei whale off western Iceland. II. Age, growth and reproduction. Report of the International Whaling Commission 33: 465-476.

121. Perrin WF, Donovan GP (1984) Report of the Workshop on reproduction on whales, dolphins and porpoises. Report of the International Whaling Commission Special Issue 6: 1-17.

122. DeMaster DP (1984) Review of techniques used to estimate the average age at attainement of sexual maturity in marine mammals. Report of the International Whaling Commission Special Issue 6: 175-179.

123. Lockyer CH (1984) Review of baleen whale (Mysticeti) reproduction and implications for management. In: Perrin WF, Brownell Jr RL, DeMaster DP, editors. Reproduction in whales, dolphins and porpoises. Cambridge, UK: International Whaling Commission. pp. 27-50.

124. Kjeld M (2003) Salt and water balance of modern baleen whales: rate of urine production and food intake. Canadian Journal of Zoology 81: 606-616.

125. Rice DW (1977) Synopsis of biological data on the sei whale and Bryde's whale in the eastern North Pacific. Reports of the International Whaling Commission Special issue 1: 92-97.

126. Lockyer CH (1974) Investigation of the ear plug of the southern sei whale, Balaenoptera borealis, as a valid means of determining age. Journal du Conseil International Pour L'Exploration de la Mer 36: 71-81.

127. Laws RM (1977) Seals and whales of the Southern Ocean. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 279: 81-96.

128. Merrick RL (1997) Current and historical roles of apex predators in the Bering Sea ecossystem. Journal of Northwest Atlantic Fishery Science 22: 343-355.

129. Beaugrand G (2004) The North Sea regime shift: Evidence, causes, mechanisms and consequences. Progress in Oceanography 60: 245-262.

130. Beaugrand G, Reid PC, Ibañez F, Lindley JA, Edwards M (2002) Reorganization of North Atlantic marine copepod biodiversity and climate. Science 296: 1692-1694.

131. Baker JD, Westgate A, Eguchi T (2010) Vital rates and population dynamics. In: Boyd IL, Bowen WD, Iverson SJ, editors. Marine mammal ecology and conservation, a handbook of techniques. Oxford, UK: Oxford University Press. pp. 119-143.

132. Kellar WM, Trego ML, Marks CI, Dizon AE (2006) Determining pregnancy from blubber in three species of delphinids. Marine Mammal Science 22: 1-16.

133. Rolland RM, Hunt KE, Kraus SD, Wasser SK (2005) Assessing reproductive status of right whales (Eubalaena glacialis) using fecal hormone metabolites. General and Comparative Endocrinology 142: 308-317.

134. Gardiner KJ, Boyd IL, Racey PA, Reijnders PJH, Thompson PM (1996) Plasma progesterone concentrations measured using an enzyme-linked immunosorbent assay useful for diagnosing pregnancy in harbor seals (*Phoca vitulina*). Marine Mammal Science 12: 265-273.

135. Kjeld M, Víkingsson GA, Alfredsson A, Olafsson O, Arnason A (2003) Sex hormone concentrations in the blood of sei whales (*Balaenoptera borealis*) off Iceland. Journal of Cetacean Research and Management 5: 233-240.

136. Allison C (2008) IWC Individual Catch Database Version 3.5. The Red House, 135 Station Road, Cambridge, Cambridgeshire CB24 9NP, UK: Available from International Whaling Commission Secretariat.

137. Buckland ST, Anderson DR, Burnham KP, Laake JL (1993) Distance Samppling: estimating abundance of biological populations. Lonkon, UK: Chapman and Hall. 446 p.

138. Pike DG, Gunnlaugsson T, Vikingsson GA, Desportes G, Bloch D (2009) Estimates of the abundance of minke whales (*Balaenoptera acutorostrata*) from Faroese and Icelandic NASS shipboard surveys. NAMMCO Scientific Publications 7: 81-94.

139. Gunnlaugsson T, Sigurjónsson J (1990) A note on the problem of false positives in the use of natural marking data for abundance estimation (SC/A88/ID46). Report of the International Whaling Commission: 143-145.

140. Knowlton AR, Clark CW, Kraus SD. Sounds recorded in the presence of sei whales *Balaenoptera borealis*; 1991 1991; Chicago. Society of Marine Mammalogy. pp. 40.

141. Rankin S, Barlow J (2007) Vocalizations of the Sei Whale *Balaenoptera borealis* off the Hawaiian islands. Bioacoustics: the international journal of animal sound and its recording 16: 137-145.

142. McDonald MA, Hildebrand JA, Wiggins SM, Thiele D, Glasgow D, et al. (2005) Sei whale sounds recorded in the Antarctic. Journal of the Acoustical Society of America 118: 3941-3945.

143. Jefferson TA, Stacey PJ, Baird RW (1991) A review of killer whale interactions with other marine mammals: Predation to co-existence. Mammal Review 21: 151-180.

144. Springer AM, Van Vliet GB, Piatt JF, Danner EM (2006) Whales and whaling in the North Pacific Ocean and Bering Sea. In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell Jr. RL, editors. Whales, whaling and ocean ecosystems. Berkeley, CA, USA: University of California. pp. 245-261.

145. Branch TA, Williams TM (2006) Legacy of industrial whaling: could killer whales be responsible for declines in Southern Hemisphere sea lions, elephant seals and minke whales? In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownwell Jr. RT, editors. Whales, whaling and ocean ecosystems. Berkeley, CA, USA: University of California. pp. 262-278.

146. Reeves RR, Berger J, Clapham PJ (2006) Killer whales as predators of large baleen whales and sperm whales. In: Estes JA, DeMaster DP, Doak DF, Williams TM, Brownell Jr. RL, editors. Whales, whaling and ocean ecosystems. Berkeley, CA, USA: University of California. pp. 174-187.

147. Dailey MD (2001) Parasitic diseases. In: Dierauf LA, Gulland FMD, editors. CRC handbook of marine mammal medicine. Boca Raton, FL, USA: CRC Press. pp. 357-379.

148. Lambertsen RH (1986) Disease of the common fin whale (*Balaenoptera physalus*): crassicaudiosis of the urinary system. Journal of Mammalogy 67: 353-366.

149. Lambertsen RH, Birnir B, Bauer JE (1986) Serum chemistry and evidence of renal railure in the North Atlantic fin whale population Journal of Wildlife Diseases 22: 389-396.

150. Lambertsen RH (1990) Disease biomarkers in large whale populations of the North Atlantic and other oceans. In: McCarthy JE, Shugart LR, editors. Biomarkers of environmental contamination. Boca Raton, FL, USA: Lewis Publishers. pp. 395-417.

151. Dailey MD, Vogelbein WK (1991) Parasite fauna of three species of Antarctic whales with reference to their use as potential stock indicators. Fishery Bulletin 89: 355-365.

152. Van Bressem MF, Van Waerebeek K, Raga JA (1999) A review of virus infections of cetaceans and the potential impact of morbiliviruses, poxviruses and papillomaviruses on host population dynamics. Diseases of Aquatic Organisms 38: 53-65.

153. DeMaster DP (2001) Predation and competition: the impact of fisheries on marine-mammal populations over the next one hundred years. Journal of Mammalogy 82: 641-651.

154. Mizroch SA, Rice DW, Breiwick JM (1984) The sei whale, *Balaenoptera borealis*. Marine Fisheries Review 46: 25-29.

155. Lockyer CH (1977) Mortality rates for mature southern sei whales. Reports of the International Whaling Commission Special issue 1: 53-57.

156. Clapham PJ, Young SB, Brownell RL, Jr. (1999) Baleen whales: conservation issues and the status of the most endangered populations. Mammal Review 29: 35-60.

157. Glass AH, Cole TVN, Garron M (2010) Mortality and serious injury determinations for baleen whale stocks along the United States eastern seaboard and adjacent Canadian Maritimes, 2004-2008. Woods Hole, USA: US Deptarment of Commerce, Northeast Fish Science Center. Northeast Fish Sci Cent Ref Doc. 09-04 Northeast Fish Sci Cent Ref Doc. 09-04. 19 p.

158. Northridge SP (2009) Fishing industry, effects of. In: Perrin WF, Würsig B, Thewissen JGM, editors. Encyclopedia of marine mammals. 2nd ed. London & New York: Academic Press. pp. 443-447.

159. Laist DW, Knowlton AR, Mead JG, Collet AS, Podesta M (2001) Collisions between ships and whales. Marine Mammal Science 17: 35-75.

160. Hazel J, Lawler IR, Marsh H, Robson S (2007) Vessel speed increases collision risk for the green turtle *Chelonia mydas*. Endangered Species Research 3: 105-113.

161. Van Waerebeek K, Baker AN, Félix F, Gedamke J, Iñinguez M, et al. (2007) Vessel collisions with small cetaceans worldwide and with large whales in the Southern Hemisphere, an initial assessment. LAJAM 6: 43-69.

162. Glass AH, Cole TVN, Garron M (2009) Mortality and serious injury determinations for baleen whale stocks along the United States eastern seaboard and adjacent Canadian Maritimes, 2003-2007. Woods Hole, USA: US Deptarment of Commerce, Northeast Fish Science Center. Northeast Fish Sci Cent Ref Doc. 09-04 Northeast Fish Sci Cent Ref Doc. 09-04. 19 p.

163. Clark CW, Ellison WT, Southall BL, Hatch LT, Van Parijs SM, et al. (2009) Acoustic masking in marine ecosystems: intuitions, analysis, and implication. Marine Ecology Progress Series 395: 201-222.

164. Committee on Potential Impacts of Ambient Noise in the Ocean on Marine Mammals (2003) Ocean Noise and Marine Mammals. Washington, D.C., USA: The National Academies Press. 208 p.

165. Frantzis A (1998) Does acoustic testing strand whales? Nature 392: 29.

166. Jepson PD, Arbelo M, Deaville R, Patterson IAP, Castro P, et al. (2003) Gas-bubble lesions in stranded cetaceans. Nature 425: 575-576.

167. Fernández A, Edwards JF, Rodriguez F, De Los Monteros AE, Herraez P, et al. (2005) Gas and fat embolic syndrome" involving a mass stranding of beaked whales (Family *Ziphiidae*) exposed to anthropogenic sonar signals. Veterinary Pathology 42: 446-457.

168. Cox TM, Ragen TJ, Read AJ, Vos E, Baird RW, et al. (2006) Understanding the impacts of anthropogenic sound on beaked whales. Journal of Cetacean Research and Management 7: 177-187.

169. O'Shea TJ, Brownell RL, Jr. (1994) Organochlorine and metal contaminants in baleen whales: A review and evaluation of conservation implications. Science of the Total Environment 154: 179-200.

170. Borrell A (1993) PCB and DDTs in blubber of cetaceans from the northeastern North Atlantic. Marine Pollution Bulletin 26: 146-151.

171. Marsili L, Casini C, Marini L, Regoli A, Focardi S (1997) Age, growth and organochlorines (HCB, DDTs and PCBs) in Mediterranean striped dolphins Stenella coeruleoalba stranded in 1988-1994 on the coasts of Italy. Marine Ecology Progress Series 151: 273-282.

172. Haraguchi K, Endo T, Sakata M, Masuda Y (2000) Contamination survey of heavy metals and organochlorine compounds in cetacean products purchased in Japan. Food Hygiene and Safety Science (Shokuhin Eiseigaku Zasshi) 41: 287-296.

173. Focardi S, Marsili L, Leonzio C, Zanardelli M, Nortarbartolo di Sciara G (1992) Organochlorines and trace elements in skin biopsies of fin whale *Balaenoptera physalus* and striped dolphin *Stenella coeruleoalba*. European Research on Cetaceans 6: 230-233.

174. IPCC Core Writing Team (2007) Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland: Intergovernmental Panel on Climate Change. 104 p.

175. Simmonds MP, Eliott WJ (2009) Climate change and cetaceans: concerns and recent developments. Journal of the Marine Biological Association of the United Kingdom 89: 203-210.

176. Simmonds MP, Isaac SJ (2007) The impacts of climate change on marine mammals: early signs of significant problems. Oryx 41: 19-26.

177. Learmonth JA, MacLeod CD, Santos MB, Pierce GJ, Crick HQP, et al. (2006) Potential effects of climate change on marine mammals. Oceanography and Marine Biology: An Annual Review 44: 431-464.

178. Belkin IM (2009) Rapid warming of large marine ecosystems. Progress in Oceanography 81: 207-213.

179. Beaugrand G, Edwards M, Brander K, Luckzak C, Ibanez F (2008) Causes and projections of abrupt climate-driven ecosystem shifts in the North Atlantic. Ecology Letters 11: 1157-1168.

180. Sherman K, Aquarone MC, Adams S (2009) Sustaining the World's large marine ecosystems. Gland, Switzerland: IUCN. 142 p.

181. Stachowitsch M (2003) Research on intact marine ecosystems: a lost era. Marine Pollution Bulletin 46: 801-805.

182. Rose NA, Janiger D, Parsons ECM, Stachowitsch M (2011) Shifting baselines in scientific publications: A case study using cetacean research. Marine Policy 35: 477-482.

THIRD CHAPTER

Habitat suitability of the sei whale (*Balaenoptera borealis*) in New England and Nova Scotia using presence-only modelling

ABSTRACT: The waters off New England and Nova Scotia hold an important summering habitat for the sei whale (*Balaenoptera borealis*). This region is also one of the busiest shipping areas in the world and holds important fishing grounds, resulting in enhanced pressure over marine fauna, including whales and other marine mammals. Despite those threats, until now there has been no characterization of the habitat suitability for the sei whale throughout the region. Sighting surveys for the highly endangered North Atlantic right whale (*Eubalaena glacialis*) have collected accessory data on other species, including the sei whale, but have a deficient coverage of the region. Other complementary sources of occurrence data exist but do not contain absence data. In order to characterize the habitat suitability for the sei whale in the region, data from several sources, including aerial and shipboard sighting surveys, opportunistic sightings and whaling records, were pooled together to develop an habitat suitability model for the species during the summer. A Maximum Entropy algorithm was used in order to allow the use of all available data. Results of the model show that not all the region comprises suitable habitat for the sei whale enabling a more efficient allocation of management resources. The results further show some overlap between summer habitats of the sei and right whales, and attention is drawn to the potential competition of the two species. The limitations of the model are discussed and suggestions for future improvements are provided.

INTRODUCTION

Understanding the dynamics of organisms distributions is one of the pillars of ecological science [1]. However, for most species, information on distribution is difficult and expensive to obtain and in the majority of the cases only a fraction of the real habitat has been surveyed.

Additionally to the limited information, the environment is changing rapidly with reflexes in the distribution of organisms. Natural spaces are progressively being permeated and fragmented as result of virtually all human activities, including (but not exclusively) land clearing, urbanization, opening of new terrestrial and marine communication routes, intensification of traffic over land, water and in the air, manipulation of water resources, exploration and exploitation of live and mineral resources both on land and at sea [2]. The rate of change in distribution and population fragmentation has increased over the last few decades for many terrestrial and marine habitats, in response to climactic changes (regardless of their causes) [3-5]. Marine mammals are also expected to experience changes in distribution due to direct effects of climactic changes and also in response to the reorganization of their ecosystems as result of changes in the thermal habitat [6-14].

With the increasing human influence in natural ecosystems at a global scale and the realization of the intrinsic economic and social value of those ecosystems to human welfare, detailed and reliable information about the spatial distribution of organisms becomes essential for creating robust resource management frameworks [15-18]. In face of the lack of data in sites that have not been surveyed this critical information can be provided through the fitting of species distribution models (SDM) that relate biological records to relevant independent spatial environmental variables [19-21].

There are several approaches to fitting SDMs (also commonly referred in the literature as ecological niche models: ENM [22]) with different levels of complexity, requirements and assumptions, making the choice of the appropriate method for each situation a non-trivial matter [20,21,23]. The most widely used approaches in SDM fitting fall into five broad categories (Table 3-1):

1) Heuristic models (also known as envelope models), that characterise sites that are located within the climactic niche occupied by a species regardless of other biophysical and biotic variables, of which BIOCLIM is the most eminent;
2) Statistical, including generalised linear models (GLM), generalised linear mixed models (GLMM) and generalised additive models (GAM), multivariate adaptive regression splines (MARS), among others, that are extensively used for the strong statistical foundation and proven ability to realistically model ecological relationships;
3) Bayesian approach, based on inductive learning process carried out using Bayes theorem, such as the BAYES algorithm;
4) Machine learning techniques stemming from informatics science and including artificial neural networks (ANN), boosted regression trees (BRT), random forests (RF), and maximum entropy (MaxEnt) [note: “MaxEnt” is used throughout to denote the modelling approach while “MAXENT” is used to denote the software];
5) Evolutionary computation, such as the genetic algorithm for rule-set prediction (GARP) modelling environment that uses machine learning approaches but is particular for using a genetic algorithm.

Table 3-1: Main species distribution models (SDM) categories and examples. See text for explanation of acronyms.

Model family	Examples	References
Heuristic	BIOCLIM*	[24,25]
Statistical	GAM; GLM; GLMM; MARS	[19,26-28]
Bayesian	BAYES	[29]
Machine learning	RF; BRT; ANN; MAXENT*	[23,30-35]
Evolutionary computation	GARP*	[36,37]

*Methods specifically designed to be used with presence-only data.

With the multiplicity of approaches to developing SDMs it is not surprising that there is much discussion and confusion in the literature about the strengths and weaknesses, as well as the utility, of each approach (e.g. [23,38-43]). Only a handful review and synthesis articles have tried to put some order in the apparent confusion and clear misconceptions [19,20,44-46].

One fundamental difference between some of the above approaches regards to the goals of the SDMs [20,47,48]. For some applications with strong ecological focus SDMs may be used to gain insight into the causal drivers of species distributions. However, in many studies with focus on conservation and space management the main goal is accurate prediction of distributions, regardless of the ecological processes involved. For the former case, approaches in the “Statistical” family in Table 3-1 are the most appropriate but they tend to be data hungry and difficult to implement. On the other hand, when the goal is to produce accurate predictions of distributions, machine learning and Bayesian approaches have been shown to have a very good predictive ability even in face of scarce or non-systematic data [20,23,48]. Although being first received with some scepticism among classical ecologists and still fuelling some discussion for being perceived as “black boxes” (*sensu* Fraser [49]), they have steadily gained acceptance and are increasingly found in the ecologist’s toolbox [20,35,38,48,50].

Another aspect to take into account when considering SDM approaches is the nature of the dependent variable. Traditionally SDMs were created using only presence-absence data, describing known occurrences (presence) and known absences, collected under specifically designed surveys. The advantage of presence-absence data is enabling analyses of biases regarding the surveyed locations and calculating the probability of detection of vagile species at any given moment, i.e. the prevalence of the species [51]. Although being widely accepted as the best practice in SDM, some discussion exists about the potential that presence data has of introducing confounding information in some situations (e.g. [52,53]). Notwithstanding, in many cases records consist of presence-only data (telemetry data, museum/historical records, platforms of opportunity, citizen science). In fact, for some rare or cryptic species, or species in remote areas and extreme environments, presence-only records may be the sole data available. Notwithstanding, modelling of presence-only data is still a challenge. Formerly this problem was only addressed by envelopes and distance measures [24,25]. More recently the use of pseudo-absences (the artificial creation of absences, using randomly selected sites where the species of interest has not been detected) became widespread, despite the potential for bias introduction. In an effort to reduce bias, new methods for the choice of optimal pseudo-absences are constantly being proposed [23,36,51,52,54-56]. Few approaches to SDM were created with the main purpose of handling presence-only data (Table 3-1), the most prominent being the maximum entropy method (MaxEnt) proposed by Phillips et al. [34].

Numerous marine species, including marine mammals, present some of the challenges that are addressed by presence-only methods such as the MaxEnt approach. Many have low detectability, occur in remote sites difficult to survey (e.g. at depth, amid sea ice, in remote ocean areas), are rare or cryptic. Undertaking systematic surveys for marine species involves high operational costs which can make obtaining absence data prohibitively expensive. Moreover, obtaining reliable absence data for marine taxa is difficult, if not impossible, for many species. Nevertheless, for many of those species presence-only data exist or can be obtained at much lower cost. That is the case for the sei whale (*Balaenoptera borealis*). Sei whales occur mainly in pelagic habitats and are seldom recorded by dedicated surveys [57]. Additionally, at sea sei whales are difficult to distinguish from fin (*B. physalus*) and Bryde’s (*B. edeni*) whales where their ranges overlap, invalidating a great number of potential sightings. Thus the sei whale is a clear candidate for the use of approaches such as MaxEnt which allow using data obtained from non-systematic surveys and opportunistically. MaxEnt has been tested more extensively with land species but a growing body of literature demonstrates that the approach can be successfully applied to marine species as well (e.g. [58-61]) and specifically to cetaceans (e.g. [62-65]).

The presence of the sei whale in western North Atlantic was first reported in 1902 [66] and since then it has been established that the species uses the waters off New England and Nova Scotia as a feeding habitat during the summer [57].

The same region is also the summering habitat for the endangered North Atlantic right whale (*Eubalaena glacialis*; hereby referred only as right whale). The region harbours some of the busiest commercial ports in the world resulting in an intense ship traffic [67]. In addition, it includes very important fishing grounds and marine development sites. The development of those anthropogenic activities has been shown to be a major threat to the conservation of the right whales using the area, which elicited the implementation of several management measures including the investment in dedicated sighting surveys to obtain reliable demographic and distribution data [68,69]. Those data enabled the development of spatially explicit habitat models for the right whale, used in developing spatial/temporal management actions aimed to decrease threats posed to the whales by the human activities [64,68].

Data on the occurrence of other species were also obtained during those surveys. However, since surveys were optimized to cover the right whale habitat they may be sub-optimal for other species and need to be complemented with other sources of data to be representative of the true habitat of those non-target species.

Here a SDM is developed to characterize the summer habitat of the sei whale off New England and Nova Scotia. Data combines sightings made during aerial and shipboard surveys specifically designed for the right whale, sightings from opportunistic records and capture positions made by whaling operations. The combination of these data cover a much wider area than any of the sources by itself and probably yield a better representation of the true summer habitat of the sei whale in the region (Figure 3-1-A). Since not all sources contain absence and effort data, the MaxEnt approach was used. In addition, by using that approach it was also possible to use data obtained opportunistically during the sighting surveys, when no effort was being recorded (designated off-effort data) consequently increasing the data available for modelling.

MATERIAL AND METHODS

Maximum entropy concepts and application to species distribution models

It is well known that the concept of entropy (“*S*”) is intrinsically connected to the second law of thermodynamics (hence this being also known as *entropy law*).

The term was coined by Rudolph Clausius, who built upon the work of Sadi Carnot (after whom the Carnot cycle was named) and Émile Claperion on the conversion of heat energy to mechanical work [70,71]. In a series of memoirs and lectures posteriorly compiled in one volume [72], Clausius introduced and developed the concept of entropy, which was later considered his most important single contribution to science [73].

The concept of entropy stemmed from the investigations of Clausius on the Carnot cycle. It was first derived as a state function (or *magnitude* in the words of Clausius) involved in the calculation of the energy flow under a thermodynamic process. The classical definition, derived by Clausius, is in the form of equation (1) below:

$$dS = \frac{\delta Q}{T} \tag{1}$$

where Q denotes the *heat content* (later termed *internal heat energy* in molecular theory) and T the temperature of the system.

In Clausius’ own words, “We might call S the *transformational content* of the body (…) [but] I propose to call the magnitude S the *entropy* of the body, from the Greek word *τροπή, transformation.*” [74].

Clausius proved that in isolated systems (systems that do not receive energy from the exterior) S tends to be spontaneously maximized or, in other words, that these systems tend to evolve towards thermodynamic equilibrium. Since there is no evidence that the universe can be an open system, this proof becomes a universal law which was highlighted by Clausius at the end of his *Mechanical Theory of Heat* [74]:

> “(…) *we may express in the following manner the fundamental laws of the universe which correspond to the two fundamental theorems of the mechanical theory of heat.*
>
> *The energy of the universe is constant.*
>
> *The entropy of the universe tends to a maximum.*"

For certain applications entropy may be interpreted as a measure of the states that a system may take, which can be viewed as a measure of order (or disorder). In that view, the maximization of entropy leads to a maximum number of states, which most of the time translate into a minimal organization (but see [75,76] for examples on the contrary). That has led to the common misconception about entropy being the same as disorder, a qualitative entity. Styer [75] illustrates how trying to define the qualitative character of entropy may prove full of pitfalls and shows how a single definition may not be possible. That is in great part because entropy is a quantitative rather than qualitative measure.

From its onset in the thermodynamics theory, entropy rapidly found its way into other fields becoming fundamental in the development of physical chemistry, statistical mechanics (which was started by the works of James C. Maxwell, Ludwig Boltzmann and Josiah Gibbs as the molecular theory of gases) and quantum mechanics.

As part of his work on the molecular theory of gases, Ludwig Boltzmann was responsible for creating the means to compute the entropy based on counting the number of available microscopic states of particles (microstates). That became known as the Boltzmann principle [77] and is translated by the following equation[1]:

$$S = k \ln W \tag{2}$$

Here k is the thermodynamic unit of measurement of entropy, or Boltzmann constant ($1.33 \mathrm{X} 10^{-16}$erg/°C). Nevertheless, as it will be seen below, k is an arbitrary constant which determines the units in which entropy is measured and can assume other values.

---

[1] As a side note, Boltzmann’s tombstone is peculiar by having his entropy equation engraved at the top. At the time of his death, in 1906, the atomic theory was still under debate and it was not before 1909 that the experimental proof of the Brownian motion by Jean Perrin [78] settled the matter. Boltzmann was fiercely antagonized by many of his peers during his career due to his ideas and died before his statistical theories on mechanics and thermodynamics could be validated experimentally. The inscription of the entropy equation on his tombstone acts both as a posthumous vindication against his detractors and as a testimony to his significant contribution to science.

$\mathcal{W}$ is the total number of microstates of the particles in a system (the kinetic energy of each molecule) that are compatible with the macroscopic state of that system.

The mathematical principles developed in statistical and quantum mechanics had a profound impact in the applications of entropy theory, since they are applicable in any statistical system outside the realm of physics and chemistry.

The applicability of those principles in other fields was perceived by Claude Shannon who used them in developing his theory of communication in 1948 [79].

The drive behind Shannon's work was the creation of a framework for encoding and maximizing the information transmitted by a communications system without loss of fidelity. By adapting the thermodynamics principles to information theory, he showed how fidelity in message transmission could be represented as a probability function of the distributions of the possible messages generated at the source and the messages recovered at the receiving point.

Shannon perceived the equivalence of the quantity S in statistical mechanics with the uncertainty over the true content of a message.

It can be shown [80] that Equation (2) under certain axiomatic conditions becomes:

$$S = -k_B \sum p_i \ln p_i \tag{3}$$

where p_i is the probability that particle i will be in a given microstate. That equation, known as the Boltzmann-Gibbs entropy, has the exact formulation as Shannon's *uncertainty* (not surprisingly also represented as S), with the only difference being that in the latter p_i stands for the probability for the value of a given information bit [79].

Building on Shannon's work, in 1957 Edwin T. Jaynes published two key articles that would revolutionize information theory and, later, many other fields. In his two companion articles *Information theory and statistical mechanics, I; II* [81,82] Jaynes reformulated and simplified the mathematics behind the statistical mechanics principle of maximum entropy, leading to extensions of the theory and allowing for its use in inference based on partial information.

The basic concept behind Jaynes' Maximum Entropy Principle (MEP) is simple to understand and aims to single out the least biased distribution that better approaches the “true” distribution: “Given a model of probability distributions, choose the distribution with highest entropy” [83].

Jaynes intuited that the problem in drawing inferences from partial information was one of finding a probability assignment that avoids bias and at the same time includes all the existing information. He elegantly structured the solution to that problem in the first of his two seminal articles of 1957 [81].

He started by formulating the problem which is, in fact, a classic problem in theory of probabilities:

> *Problem:* A quantity is capable of assuming discrete values x_i $(i = 1, 2, \ldots, n)$; we are ignorant of the corresponding probabilities p_i, but know the expectation value of the function $f(x)$:
>
> $$\langle f(x) \rangle = \sum_{i=1}^{n} p_i \, f(x_i) \tag{4}$$
>
> What is the expectation value of the function $g(x)$?

This is a known problem of specification of probabilities with limited or no information and it seems unsolvable since there is no information to define the probabilities p_i, because Equation (4) and the only other known condition, the normalization condition

$$\sum p_i = 1 \tag{5}$$

have to be supplemented by (n – 2) additional conditions before $\langle g(x) \rangle$ can be found.

Jaynes proceeds by recalling Shannon’s information theory and its *uncertainty* magnitude, expressed by Equation (3) (which Jaynes synonymises to *entropy* for clarity) and showing, by way of logical reasoning, that using the information given by the entropy of the probability distribution is the solution for making inferences on the basis of partial information.

The reasoning is that by using the probability distribution which has maximum entropy within the constraints of existing information, one can ensure that no arbitrary assumptions of information are made when choosing among concurring solutions. That follows from Shannon’s proof [79] that S as formulated in Equation (3) is a quantity that is positive, which increases with increasing uncertainty and is additive for independent sources of uncertainty. The distribution with maximum entropy is the only one that makes no assumptions from the existing data, but gives weight to all the information that is present.

It remains to show how to maximize the entropy in order to single out the unbiased distribution. The basic process of maximizing the entropy proposed by Jaynes is sketched below.

In order to maximize (3) subject to the constraints (4) and (5), Jaynes resourced to the Langrangian multipliers λ, μ, obtaining

$$p_i = e^{-\lambda-\mu f(x_i)} \tag{6}$$

Substituting the constants λ, μ in (4) and (5), one obtains

$$\langle f(x)\rangle = -\frac{\partial}{\partial_\mu}\ln Z(\mu) \tag{7}$$

$$\lambda = \ln Z(\mu) \tag{8}$$

$$Z(\mu) = \sum_i e^{-\mu f(x_i)} \tag{9}$$

Jaynes then shows that this can be generalized to any number of functions $f(x)$, which yields the maximum entropy probability distribution as

$$p_i = exp\{-[\lambda_1 f_1\,(x_i) + \cdots + \lambda_m f_m\,(x_i)]\} \tag{10}$$

in which the constants are determined from

$$\langle f_r(x) \rangle = -\frac{\partial}{\partial_{\lambda_r}} \ln Z \tag{11}$$

$$\lambda_0 = \ln Z \tag{12}$$

It follows that the maximum entropy of the distribution reduces to

$$S_{max} = \lambda_0 + \lambda_1 \langle f_1(x) \rangle + \cdots + \lambda_m \langle f_m(x) \rangle \tag{13}$$

where the constant k in (3) is set to 1.

Summarizing, the MEP offers a useful method to draw inferences from partial information while ensuring that no unconscious arbitrary assumptions are made in the process. That has obvious impact in scientific inference, as it helps overcoming the arbitrariness involved in decision making under the *Principle of Insufficient Reason*, or *Principle of Indifference*, (i.e.: if under severe uncertainty, all the states should be regarded as equally likely; [81,84]).

In the years that followed its publication, the principle became well established and found numerous applications not only in information theory but also in physics, biological systems, ecology, demography, economy, geosciences, astronomy, medical sciences, among others [83,85-87]. If the principle is easy to understand, methods for entropy maximization and its applications comprise a complex and rich field of investigation, even deserving a dedicated, indexed, peer-reviewed monthly Journal (*Entropy*; published by MDPI; ISSN 1099-4300, CODEN: ENTRFG). Alternative measures of information entropy and ways of maximizing it are constantly being developed, dependent on diverse axiomatic foundations. An excellent and fairly recent review is given in Beck [86].

Recently Phillips et al. [34] and Dudík et al. [88] devised a way of applying the MEP to SDMs using presence-only data, implemented through a software called MAXENT (freely available through http://www.cs.princeton.edu/~schapire/maxent). In their method the unknown probability distribution p_i in equation (4) is denoted by π and is drawn over a finite set X of points (that for practical purposes correspond to pixels in the study region). The distribution π assigns a non-negative probability $\pi(x)$ to each point x in X, totalling 1. The approximation to π is given by another probability distribution $\hat{\pi}$, which entropy's is given by

$$S(\hat{\pi}) = -\sum_{x \in X} \hat{\pi}(x) \ln \hat{\pi}(x) \tag{14}$$

The probability distribution $\hat{\pi}$ respects a set of constraints derived from the species occurrence data, expressed as a set of functions (transformations) of the explanatory variables (*features* in the MAXENT terminology) [89]. This is achieved by making the mean of each feature to closely match the empirical average over the presence sites. From the MEP it follows that among all probability distributions that satisfy those constraints, the one with the maximum entropy will be the most unconstrained or, in other words, that will produce the least under-specified model [81,88]. The full methods of implementation are described in Phillips et al. [34] and Dudík et al. [88] and further developed in Dudík et al. [90] and Phillips and Dudík [89]. At this point it is relevant to point out that the MaxEnt approach also applies to presence-absence data, in which case the method provides estimates of a quantity that is monotonously related to the probability of presence of the modelled species, conditioned on the environmental predictors [87].

It is clear that this approach tends to closely match the training data, an effect known as "overfitting" that limits the generalization ability of the model [91,92]. To handle that effect, MAXENT allows the setting of an error bound around the samples empirical feature means, denoted as γ_j in equation (15)[2]:

$$\gamma_j = \gamma \sqrt{\frac{s^2[h_j]}{m}} \tag{15}$$

This *regularization parameter* γ_j acts over feature h_j in equation (15). The feature h_j has variance $s^2[h_j]$ over the m presence sites. Since γ_j corresponds to the width of the confidence interval it is computed as the standard error (thus the radical) multiplied by

---

[2] In the notation used by Dudík et al. [92] the regularization multiplier is denoted by the Greek letter lambda (λ). However, to avoid confusion with the notation used before in this work, here the letter gamma (γ) was used instead.

the desired confidence level set by γ that will vary according to the type and complexity of feature classes that MAXENT tries to fit.

The gammas in equation (15) allow smoothing the distribution (making it more regular, thus the term *regularization parameter*) and comprise a specific type of regularization called L1, also known as the *Lasso* or *elastic-net* penalty, that is a common approach to model improvement [92-94]. In MAXENT it can be set manually. However, in order to simplify model fitting and to avoid bias inherent to the data a set of feature class-specific tuned parameters is pre-loaded in the program based on results from modelling a large multitaxa dataset using presence-absence data covering 226 species, 6 regions and sample sizes from 2 to 5822 [23,89].

Since complexity of the functions describing the environmental variables comes at the expense of an increased sample size, depending on the size of the dataset the feature classes considered are limited by default in the program (for smaller datasets only simpler functions are used). Nevertheless, that can also be overridden by the user.

The program MAXENT became publicly available in 2004 and since then it has become one of the preferred methods of modelling species distributions with presence-only data. It has been shown that it not only successfully models species distributions as it is among the highest performing methods even when compared with more established methods including those using presence-absence data [23,95,96]. It has been used with diverse aims, including finding correlates of species occurrences, mapping current distributions, and predicting to new times and places [33].

However, the MaxEnt approach has not been free of criticism. Being in the machine learning model family (Table 3-1) the MaxEnt approach has often been classified as a “black box” (e.g. [41,97]). The term “black box” was first used to designate models that failed to describe internal sources of variation [49] and is often applied to refer to statistical models that do not quantify the influence of the independent variables in the model and a confidence measure of that contribution. That hardly applies to the MaxEnt method since quantitative information about the relative contributions of the explanatory variables can be calculated in several ways [98]. Even MaxEnt not being the best option when it comes to understand the habitat preferences of a species [95], outputs of a model ran using the MAXENT software can include up to four variable contribution measures depending on the specifications of the model.

The term “black box” is also widely used when referring to methods that give little or no option for model manipulation. In that regard it must be noted that the MAXENT software does offer a series of tuning options that allow optimization of model specification. Nevertheless, a recent review has shown that most users tend to ignore that and opt to use the default values in the program [98]. Critics to the MaxEnt approach have also referred the reduced availability of inferential and model checking tools in the MAXENT program [41], which was partially acknowledged by the creators of the software [34]. Nevertheless, several approaches for internal model performance assessment and for model evaluation have been suggested and can be used in conjunction with MAXENT (e.g. [99-104]).

Recently it has been suggested that the MaxEnt approach is mathematically equivalent to a point process model (PPM), often implemented as GLMs [98,105,106]. If that equivalence is confirmed, there are important implications: 1) It cannot be argued that MaxEnt performs worse or better than a PPM since as long as the methods are applied in the same manner the results should be identical; 2) MaxEnt cannot be dubbed as a “black box” method, since it can be reexpressed as a PPM; 3) model checking can be made using traditional tools available for GLM, GAM and other well established approaches to SDM [98,107].

Dependent variable

Occurrences of sei whales were obtained from the North Atlantic Right Whale Consortium database [108,109]. The North Atlantic Right Whale Consortium database (NARWCd) is a centralized archive of survey and sighting data held by the University of Rhode Island that receives data from every organization and agency conducting surveys for right whales, from Florida to Atlantic Canada. The NARWCd also includes historical records of captures by whaling activities and opportunistic sightings as well as data on other marine mammals and sea turtles. All sightings are georreferenced and coded according to origin of the record and reliability of the identification. Identification reliability is divided in three categories, increasing from 1 (unsure) to 3 (definitive) with an additional value of 9 when that information is unknown or has not been recorded.

For sei whales four sources of records were available: aerial surveys, shipboard surveys, whaling and opportunistic sightings. All sources were combined and only sightings with an identification level of 3 were used, resulting in 1133 records (Figure 3-1-A), spanning from 1966 to 2008. Aerial surveys contributed with 56.13% (636) of the records, whaling with 25.68% (291), shipboard surveys with 16.24% (184) and opportunistic sightings with 1.94% (22). From sighting surveys (aerial and shipboard), both on- and off-effort sightings were considered, with off-effort records accounting for 39.51% (324) of the sightings. Multiple occurrences within the same environmental grid cells were removed using ENMTools [103], reducing the number of occurrences to 855 records. One record was outside the modelled area and was discarded, resulting in a final number of occurrences of 854 (Figure 3-1-B) that were used through the modelling process described below.

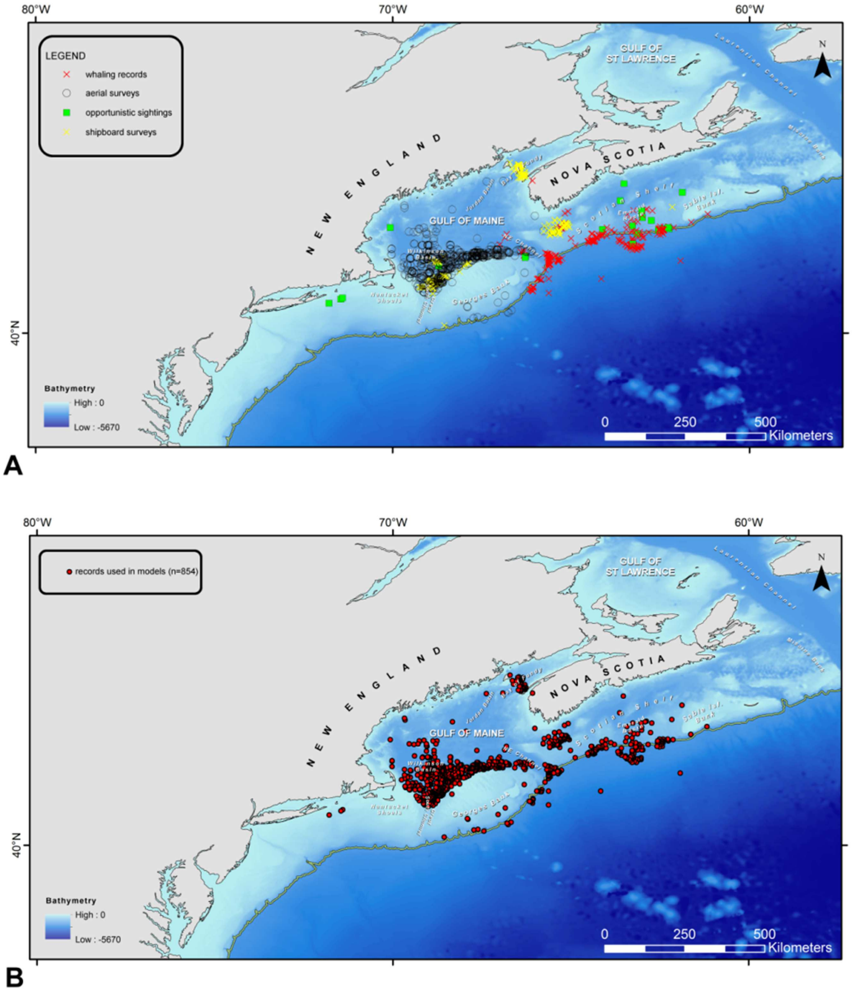

Figure 3-1: A – Sei whale records in the study area during summer from 1966 to 2008 (n=1133), according to type of sources. B – Records used in MAXENT models after data clean-up (n=854). The main topographic features mentioned in the text are shown and bathymetry is represented as shades of blue. The 500 m isobath, corresponding to the approximate position of the shelf-break, is shown as a continuous yellow line.

Explanatory variables

Several candidate predictor variables were chosen as potentially having explanatory power for the distribution of whales. Choice of the candidate variables was done by a combination of revision of the literature to identify the variables that are believed to drive whale's movements [57], identification of proxies for prey distribution and data availability.

Candidate variables originated from different sources with varying resolutions (Table 3-2).

Variables were treated in a Geographic Information System (ArcGIS 10.1; ESRI, Inc.; hereby referred as ArcGIS) as raster datasets. Prior to processing all rasters were projected to an Equidistant Cylindrical projection with datum WGS84. Some variables were derived from other and suffered some pre-processing, explained below. All rasters were then resampled to adjust cell size for the chosen resolution (0.01666667 arc-minutes) and to achieve correct alignment. Finally all variable layers were clipped to the same extent and exported to ascii file format for use in MAXENT. In order to facilitate this process, a set of custom tools was developed using the Model Builder facility in ArcGIS (Figure A-I.1; Appendix I).

The relevance of using each of the chosen variables as predictors of whale distribution and specific processing steps for each raster are briefly explained in the following sections.

Table 3-2: Explanatory variables used in the construction of the MAXENT models.

Variable	Acronym	Units	Resolution	References
Depth	*bath*	meters	1 arc-minute**	[110]
Aspect (northness)*	*northness*	radians	1 arc-minute**	[110]
Aspect (eastness)*	*eastness*	radians	1 arc-minute**	[110]
Slope*	*slope*	degrees	1 arc-minute**	[110]
Rugosity*	*rugosity*	unitless	1 arc-minute**	[110]
Distance to shore*	*dist_shore*	kilometers	1 arc-minute**	[110]
Chlorophyl-a concentration	*lnchl_smmr*	$\ln(mg/m^3)$	0.01667 arc-minute	[111]
Summer sea surface temperature	*sst_smmr*	°Centigrade	0.01667 arc-minute	[111]

*Derived from Depth; **Downsampled to 0.01667 arc-minute

Geomorphological

Geomorphological data were all derived from the ETOPO-1 digital elevation model (DEM; [110]). The geomorphological data layers used were *depth*, *aspect* (decomposed in two variables: *northness* and *eastness*), *slope*, *rugosity* and *distance to shore*. Apart from *depth* that was measured directly from the DEM, all other geomorphological data had to suffer some transformation prior to being processed for use in MAXENT.

Aspect measures the geographic orientation of terrain surfaces and is relevant in ecology since habitat quality may vary depending on terrain orientation. Exposure to sunlight or predominant currents, water temperature and type of substrate are some of the characteristics that can be influenced by *aspect* in the marine environment and that can, in turn, drive the distribution of organisms. *Slope* measures the inclination of the terrain in respect to a reference surface (in this case the geoid representing the Earth). Similarly to terrain *aspect,* terrain *slope* can also affect the quality of the habitat for different organisms. It is well known that terrain *slope* affects a plethora of oceanographic phenomena, including vertical and horizontal water and particle transport, creating phenomena such as up- and down-welling with reflexes on primary and secondary productivity.

Aspect and *slope* rasters were created using the *DEM Surface Tools* extension in ArcGIS [112,113]. Although ArcGIS offers tools to calculate *slope* and *aspect*, there is no option to choose the algorithm on which these calculations are based. Jones [114] compared several algorithms that are used in the computation of *slope* and *aspect* and found that the Fleming and Hoffer's method with the derivation by Zevenbergen and Thorne [115] using four neighbouring cell values was consistently the most accurate. Unlike the standard ArcGIS tools, the *DEM Surface Tools* extension offers the option to choose between different algorithms, including Fleming and Hoffer's method using a four cell neighbourhood, which was the method chosen to create the *slope* and *aspect* rasters used here. In this work slope was calculated in degrees. Slope can also be (and is often) expressed as a percentage of change in steepness. Nevertheless while the change in steepness expressed in degrees is linear, percent slope will increase non-linearly, being 0% at 0^0 and rapidly approaching infinity near 90^0 (100% change is attained at 45^0). *Aspect* presents another type of problem, by being a circular measure (unlike with what happens with *slope* which can vary only between 0^0 and 90^0, *aspect* can take any value between 0^0 and 360^0). Two approaches can be used to cope with that problem: 1) to categorize the values, or 2) to calculate the "northness" and "eastness" components of the orientation values, which will vary between -1 and 1. Here the second approach was used. To create the *northness* and *eastness* raster datasets, *aspect* values were first converted to absolute radians (in ArcGIS, flat areas, with no clear orientation are represented as -1 making this transformation necessary). *Northness* and *eastness* were calculated by computing respectively the cosine and sine of each grid-cell value of the resulting absolute radians raster dataset.

Terrain *rugosity* translates terrain complexity, which can drive organisms' distribution in several ways, for example by offering sheltered areas and hiding places, creating local micro-climates, or affecting ease of movement. In aquatic environments it can further create turbulence, advecting sediment and other particles into the water which will in turn influence turbidity and productivity conditions and may create niches to some organisms [116]. Most traditional methods of computing *rugosity* are indices directly derived from terrain *slope* (e.g. [117]; [118]). More recently Sappington et al. [119] proposed a *vector rugosity measure* (VRM) of terrain that computes *rugosity* as the variation in three-dimensional orientation of grid cells within a neighbourhood, incorporating both the *aspect* and *gradient* components of the terrain. This is a more inclusive way of measuring the heterogeneity of the terrain and has the further advantage of being more independent of *slope* than traditional methods, enabling the simultaneous utilization of both measures in modelling.

The final geomorphological variable derived from the ETOPO-1 DEM was distance to shore. Cetacean distribution can be often linked to the distance they are from shore (e.g. [120-122]). As a result, that variable is often included in habitat modelling for those animals. Nevertheless, the ecological significance of the variable must not be taken at face value. For species that in some way depend on land, distance to shore may have a clear ecological significance. For instance, the at sea distribution of polar bears, pinnipeds and marine birds may be influenced by the energetic cost involved in round trips from haul-out sites and colonies to foraging grounds (e.g. [123-125] and references therein). Cetaceans, on the other hand, live an exclusively aquatic life and the ecological meaning of distance to shore is not so clear. The variable can in many cases represent only a proxy to other relevant variables, such as depth, position of the shelf break, turbidity, prevailing currents, or other (e.g. [126-129]). In fact, the distribution according to distance to shore for the same species in different regions sometimes varies widely [130], indicating that the variable may be dissociated from the ecological mechanism(s) driving the distribution. Here the variable was included in the candidate explanatory variables for being a traditional variable in this type of studies.

The raster dataset representing distance to shore was created using the *Spatial Analyst* tools in ArcGIS. First, a shoreline was derived from the ETOPO-1 DEM using the *Countour* tool in *Spatial Analyst* to create a contour line for bathymetry 0. Then, a constant raster was created having the same boundaries and dimensions of the remaining datasets, and the shortest distance of the centroid of each grid cell to the shoreline was calculated using the *Euclidean Distance* tool in *Spatial Analyst*.

Remote sensed

The distribution of the sei whale is known to be in some way related to sea surface temperature [57]. Remotely sensed sea surface temperature (SST; °C) was derived from standard mapped images (level 3, seasonal summer composite 1998-2008, 4.6 km resolution) collected by the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard NASA’s Aqua satellite and obtained from the Ocean Color Discipline Processing System [111].

An attempt to use *in-situ* data from one of the chief preys of the sei whale in the North Atlantic, the copepod *Calanus finmarchicus*, was made using a kriging approach. Nevertheless, the resulting map did not cover the entire study area, or even all the sites with known occurrences of the sei whale. Thus that map was not used in modelling the sei whale distribution since it would introduce undesired bias in the models. Instead the natural logarithm (ln) of Chlorophyll-a concentration was used as a proxy for primary productivity. Remotely sensed near-surface primary productivity indicated by Chlorophyll-a concentration (Chl-a; mg/m^3) data was also derived from data collected by Acqua MODIS, with the same resolution and for the same period as SST.

Model fitting

All models were fitted using MAXENT 3.3.3k. Unless otherwise stated, settings were left at their default values. During test trials it was verified that the default 500 iterations were not sufficient to guarantee model convergence. Thus the number of maximum iterations was set to 5.000 for all models. To speed-up data processing, MAXENT was set to use two processing threads and was allocated 2 GB of RAM memory. Due to the large number of models, MAXENT was run using scripts and not the graphical user interface. When no independent data exist to test model performance, it is common practice to withhold a part of the presence data to test model performance [33].

Nevertheless, since model performance was tested using the methodology proposed by Raes and ter Steege [101] (see below), all sei whale sightings were used to train the models.

The choice of number and type of explanatory variables is fundamental to model performance. As a rule of thumb, the best model is the one which combines simplicity (in terms of the number of model parameters) with predictive ability, i.e., the most parsimonious model [87,131]. A very complex model in terms of number of explanatory variables and parameters is said to be saturated and tends to overfit the data. Thus, an important part of the modelling process is to investigate which, among those available, is the set of explanatory variables that will produce the most parsimonious model. This includes eliminating highly correlated variables. Ecologically, two or more highly correlated explanatory variables are bound to be analogous representations of the same phenomenon acting over the modelled organism. Therefore, including correlated variables goes against the parsimony principle. Additionally, the correlation nature may differ in other scenarios (other region or time), which will reduce the transferability of the model [132].

As referred previously, MAXENT includes a regularization parameter that balances model fit and complexity. In MAXENT, different default regularization values are included, according to the type of features used in calculating the model parameters. These were based in a large dataset [89] but may not be appropriate for all systems and should be tuned to achieve better model performance [104]. Tuning can be done individually or through a regularization multiplier that acts over all default regularization values, which is the preferred option for most applications [104].

Tuning of model complexity followed the methodology proposed by Warren and Seifert [104] and Warren et al. [132]. As a first step, a set of 76 models was built with all 8 variables, varying the regularization multiplier from 0 to 15 in steps of 0.2. Output format was set to raw and models were compared using the sample corrected Akaike information criterion (AICc) scores [107,133] calculated using ENMTools. Only the model with the lowest AICc score was retained (hereafter the starting model).

As a second step, the contribution scores of each environmental predictor variable in the starting model were inferred by the permutation importance of the variables, assigned by MAXENT. The permutation importance is calculated by randomly permuting the values of each variable on training and background data; the model is then re-evaluated on the permuted data and the metric represents the resulting drop in training AUC normalized as percentage. These scores were used in conjunction with the spatial correlation to select the final set of environmental variables. Testing for correlation between predictor variables was performed by calculating the Kendal's tau correlation coefficient in R (R Development Core Team 2012). The Pearson's correlation coefficient is commonly used as an indicator of correlation in spatial data analysis literature. Nevertheless, that coefficient is sensible to outliers and non-normality [134]. Rank-based coefficients are more appropriate in these cases, as they are distribution-free. Spearman's rho is widely used as a non-parametric measure of correlation but Xu et al. [135] have argued in favour of the use of the Kendall's method over Spearman's when in presence of large sample sizes. High collinearity between any two predictors was assumed to be present when the absolute value of the correlation index exceeded a value of 0.7 [132,136].

Variables with a contribution score <5 in the starting model were discarded. The variable with the highest score was retained and any other variables correlated with that were also discarded. This process was followed with the remaining variables until all were either discarded or included in the final set of variables (hereby called the reduced set).

Finally, a new set of models was created with the reduced set of variables to determine the optimal value of the regularization multiplier. Although that was performed earlier as the first step to determine the ideal set of variables, and even though MaxEnt does not estimate interaction terms for variables, there may be interactions between those variables in how they affect regularization (Warren, pers. comm.). Thus, eliminating one pair of a correlated set of variables may require that the function for the other variable be re-estimated, with a different number of parameters. Thus, in order to determine the optimal regularization multiplier, the procedure performed in the first step was repeated with the reduced set of variables. The resulting model with the lowest AICc score was chosen as the final SDM (SDM_f).

Model evaluation

After its introduction to the field of ecological modelling, in 1997 [137], receiver operating characteristic (ROC) curve analysis became one of the most widely used methods to assess SDM performance. The value of the area under the ROC curve

(AUC) is the metric used to estimate the predictive accuracy of SDMs, and is implemented in MAXENT as a measure of model performance [33,89].

With presence-absence data, AUC values range from 0 to 1, with a value of 0.5 indicating model accuracy not better than random and a value of 1.0 indicating perfect model fit. However, some authors have questioned the way the AUC metric is used to measure model performance, especially with presence-only data (e.g. [138]). One of the major problems is that when using presence-only data, pseudo-absences must be used to calculate the AUC. Given a sufficient number of pseudo-absences in relation to the size of the study area, a reasonable representation of the environmental variation of the study area is achieved and the AUC is still considered a useful metric for model evaluation [34,36,139]. However, when using pseudo-absences the maximum achievable AUC (AUC_{max}) is less than 1. In fact, in that case AUC_{max} is given by $1-a/2$ where "a" is the fraction of the geographical area covered by the species' true (and unknown) distribution [34]. This invalidates the use of standard thresholds to indicate good SDM performance, such as the often used value of |0.7| [140,141].

To overcome that problem, Raes and ter Steege [101] suggested a methodology to assess whether the SDM prediction differs from what would be expected by chance alone. This is achieved by testing the AUC value calculated for the SDM against a null distribution of expected AUC values based on a random dataset, the null hypothesis being that it does not differ from a prediction obtained by chance [101].

To create the null AUC distribution, locations randomly distributed in the study area, in the same number of the occurrences used to fit the SDM (n=854), were created without replacement. For the present study, 500 sets of those random points were created and used to train an equal number of null models, using the exact same settings in MAXENT as used to fit the SDM_f. A one-sided 99% confidence interval (CI) was then calculated as in conventional statistical analysis, to test if the SDM_f performs significantly better than expected by chance. The AUC values of the 500 null models were ranked and the 495th value (0.99 X 500) was set as the 99% CI upper limit. AUC values higher than that upper limit are representative of an SDM with accuracy significantly better than expected by chance alone ($p<0.01$).

RESULTS

Model fitting

Comparison of the AICc scores of the 76 competing models showed that the best model was that with a regularization multiplier of 0.6 (Table A-II.1; Appendix II), making this the starting model. The permutation importance of each variable in the starting model is shown in

Table 3-3 and correlation between variables in Figure 3-2. The variables *distance to shore*, *northness*, *eastness*, *rugosity* and *slope* were discarded for having contribution scores <5 (Table 3-3). The remaining variables were not highly correlated to each other (Figure 3-2) and were used to build the next set of models to investigate the best value of the regularization multiplier.

Of these new 76 models using the reduced set of variables, the model with the best AICc score was that with a regularization value of 0.4 (Table A-II.2; Appendix II) and was, thus, chosen as the final SDM (SDM_f).

Table 3-3: Permutation importance of variable in the starting model. Variables with a permutation value above 5 were retained to fit the final SDM.

Variable	Permutation importance
sst_smmr	49.7
bath	36.7
lnchl_smmr	6.6
dist_shore	3
northness	1.3
rugosity	1.2
slope	1.5
eastness	0

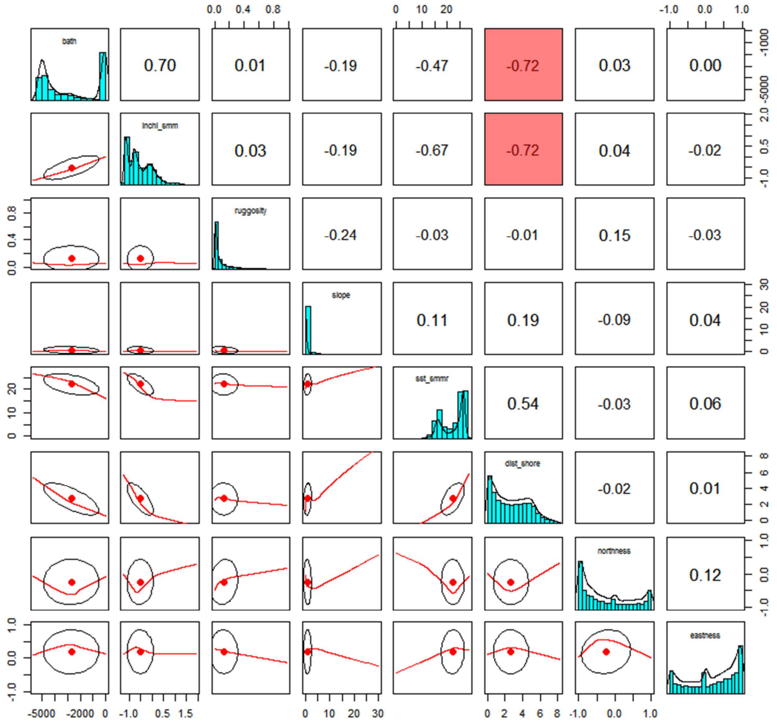

Figure 3-2: Frequency distributions of explanatory variables considered for modelling (central diagonal) with ellipse plots (lower left; e*llipses* indicate density envelopes that contain the specific mass of points, as determined by the 95% probability) and associated Kendall's correlation coefficients (upper right) between pairs of explanatory variables. General relationships between the pairs of explanatory variables are shown as red trend lines in the ellipse plots and magnitude of correlation is represented by increasing text size of the Kendall's coefficients. Red squares represent correlation values above/below the absolute cut-off threshold of 0.7. Refer to text for abbreviations.

Model evaluation

The AUC scores of the 500 null models ranged between 0.612 and 0.662 (mean = 0.633), close to the theoretical expected value for a random model (0.5), while that of the SDM_f was 0.932 (Figure 3-3). Based on the distribution of the AUC values of the null models, the 99% CI upper limit for a model that performs better than random was set at AUC = 0.651, showing that the SDM_f performs significantly better than expected by chance ($p<0.01$).

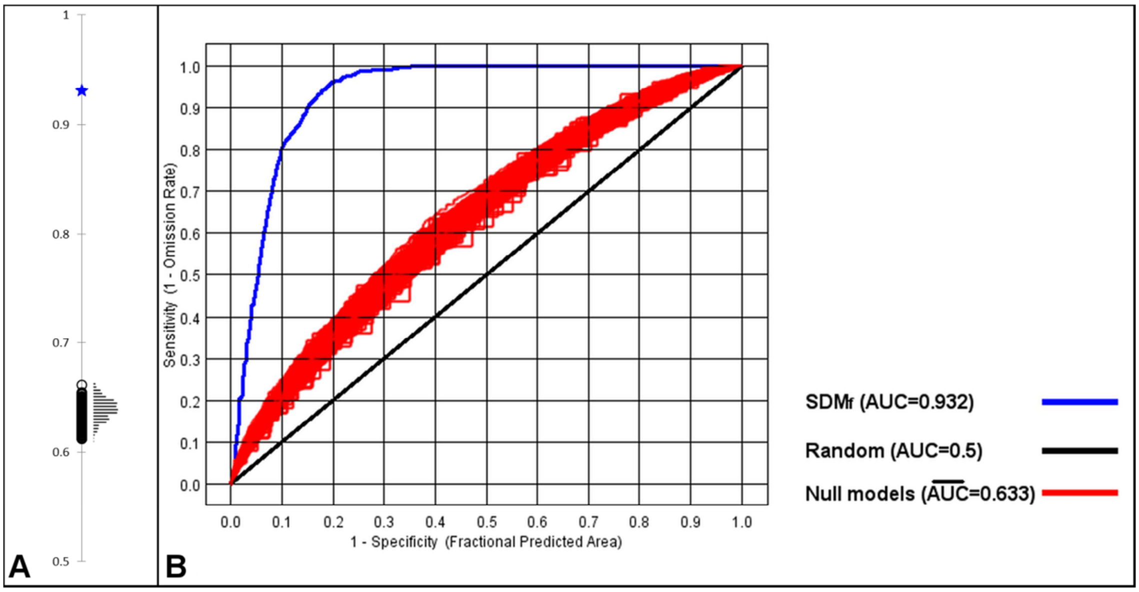

Figure 3-3: A - Training AUC values of the SDMf model (star) and the 500 null models (circles; frequency histogram shown on the right side) using randomly drawn cells from the study area. B - The receiver operating characteristic (ROC) curves for the SDM_f model (blue) and 500 null models (red).

Habitat suitability

Since the SDM_f performance based on the ROC analysis was much better than expected by chance, it was considered a valuable model for characterization of the sei whale summer habitat in the study region. A new model was produced with the exact same settings as SDM_f except regarding the output format that was changed from “raw” to “logistic”. The logistic output enables the calculation of habitat suitability probabilities, conditioned on the environmental variables [89]. These values were used to create two types of habitat suitability maps.

In Figure 3-4 habitat suitability values are represented by a colour gradient, giving a representation of the relative suitability of the habitat throughout the study area. A highly suitable region is clearly seen inside the Gulf of Maine, in the area corresponding to the Wilkinson Basin region. The model also gives high suitability scores to the area along the shelf break. Areas of lower but still noticeable suitability are placed inside the Bay of Fundy, over the Emerald Basin (Scotian Shelf) and over the margin of the Gulf of St. Lawrence along the Laurentian channel. In contrast suitability clearly decreases over shallow banks, which can clearly be seen over Georges, Sable Island and Misaine Banks.

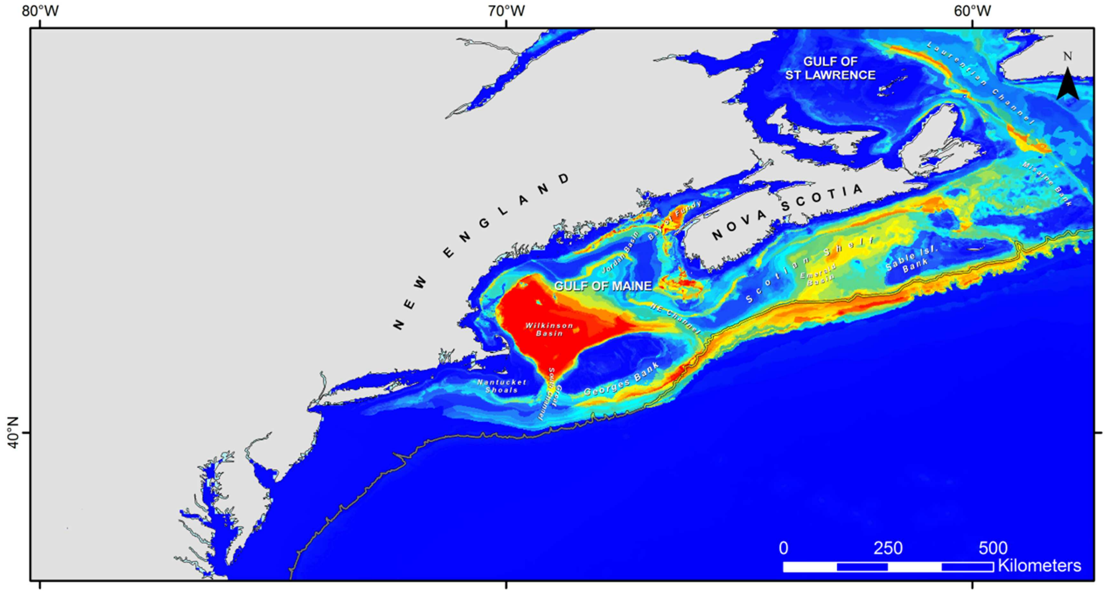

Figure 3-4: Summer habitat suitability for the sei whale in the study region. Warmer colours (red) represent areas that are more suitable, whereas cooler colours (blue) are less suitable. The isobath of 500 m, corresponding to the approximate position of the shelf-break, is shown as a continuous yellow line.

For some applications a state-structured model may be more appropriate. The 10% minimum probability of suitable habitat calculated by MAXENT was used to produce the map in Figure 3-5. Assuming that some errors may exist in the data, this threshold defines the suitable habitat by including the best scoring 90% of the data used in training the model and thus eliminating most minor errors in the data. The resulting threshold corresponds to a minimum suitability probability of 0.2783. Other thresholds could be used, depending on management goals and confidence in the data quality.

In Figure 3-5 is clear that over the continental shelf, suitable habitat is found in some of the deepest areas. Nevertheless, not all deep areas are included in the suitable habitat defined by the model, which is clear in the case of the Jordan Basin in the Gulf of Maine and the Laurentian Channel northeast of the Gulf of St. Lawrence.

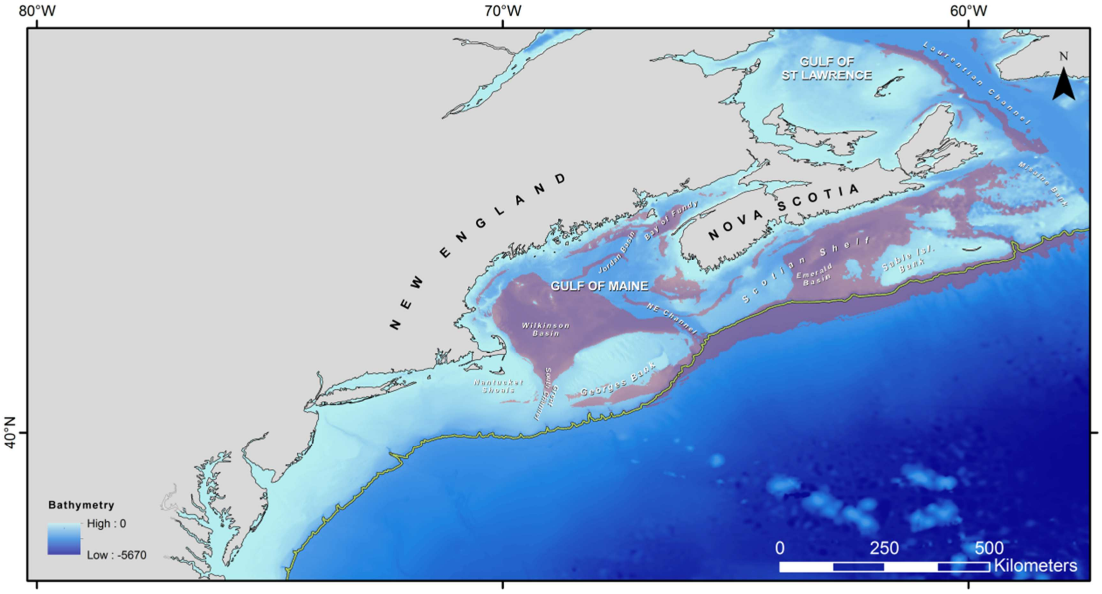

Figure 3-5: Habitat suitability for the sei whale using a threshold of 0.2783 suitability probability (transparent red shade). See text for further explanation of threshold calculation. Bathymetry is represented as shades of blue. The isobath of 500 m, corresponding to the approximate position of the shelf-break, is shown as a continuous yellow line.

DISCUSSION AND CONCLUSIONS

Overall model performance and issues

The goal of this study was to produce a spatial model of the sei whale summering habitat off New England and Nova Scotia that could be used to aid in conservation management. Since the available systematic survey data clearly did not cover the entire habitat used by the species, it was necessary to use a presence-only modelling approach. By carefully choosing ecologically relevant variables and tuning the model, the resulting SDM was significantly better than expected by chance and can be a valuable tool for the management of this and other whale species in that region.

No model is perfect and there is always space for refinement. Inclusion of prey abundance as a predictor would be one of the ways of achieving that in the present case. The inclusion of prey abundance as a predictor improved model accuracy of habitat suitability maps for the right whale using the MaxEnt approach [64]. As mentioned above, a model of prey abundance was also developed for the present work but the available data was insufficient to cover the entire study region. Within a different or smaller study area, using prey abundance could be an option and would probably improve model accuracy. But that improvement would come at the expense of transferability of the ENM since prey data is lacking for most areas. Thus, for the goals of the present work, the inclusion of prey abundance in the model was considered counter-productive and was abandoned. Instead, variables that probably have low ecological significance to the sei whale but are believed to influence prey distribution and abundance (such as Chl-a concentration and some topographic variables) were included in the model. As a result, the variables retained in the ENM_f probably reflect not only those important to the whales but also those that effect prey distribution and as a result are influential in shaping the whales' feeding habitat.

Models are based in underlying assumptions that must be carefully considered for their plausibility. Probably the strongest assumption in the model presented here is that the 10 year climatologies for SST and Chl-a concentration are representative of the entire period covered by the sei whale sightings data used to train the model. Since SST had the most contribution to the model among the retained variables, major changes in this variable over the study period may influence the results. Large Marine Ecosystems (LMEs) thermal characteristics are changing, but changes are not uniform and strong regional variations are noticeable [142]. The area for which the SDM was developed in this study encompasses two LMEs: the Scotian shelf and the NE US shelf. The NE US shelf is warming slowly but changes in the Scotian shelf are considered fast with a net increase of 0.89°C between 1982 and 2006 [142]. It is important to understand how this could have affected the resulting SDM. Of the 854 presence points used to train the model only 192 (22.5%) were positioned over the Scotian shelf. Thus any effects of the warming of the Scotian shelf over the predictions of the model are restricted. Moreover, the net increase of 0.89°C is much lower than the inter-annual variability that whales (and prey) have to experience in the region [142]. Consequently this increase is probably within the variation tolerated by the animals. The option to deal with this temporal mismatch between the climatologies and part of the sighting data would be to use only sightings from the same period as the climatologies. However, the resulting sightings dataset would comprise mostly observations from surveys that were shown to be spatially biased (Figure 3-1-A). Taken all into consideration, the spatial bias was regarded to be more adverse than the bias from ocean warming.

The fact that an area is deemed suitable by an SDM does not imply that it is in fact used by the observed species. Model predictions are no more than the geographic projection of the environmental conditions most closely associated with the occurrence data, identified during model training [19]. Even if an area is deemed suitable there can be factors that preclude its use by the observed species such as physical barriers, competition with other species, inability of occupying all the available habitat (by small or depleted populations, for example), among other [1]. A clear example is evident in the present work by the stretch along the margin of the Gulf of St. Lawrence identified by the SDM_f as suitable habitat for the sei whale. Evidence of sei whale presence in that area could not be found in the literature.

In a way, this constitutes a Type I (*commission*) error since it can be interpreted as the model predicting the presence of the species where it in fact is not (but it should be stressed that the model is predicting potential and not realized habitat). A commission error is the converse of Type II (*omission*) error, when a model predicts the absence of the species where it is in fact present [137].

The implications of Type I and Type II errors for conservation planning must be understood and weighted when deciding on management procedures. If only part of the suitable habitat identified by the SDM is to be selected for management purposes, commission errors may lead to the selection of areas that are irrelevant in protecting the observed species. Moreover, creating management measures in areas where the species is absent has direct and indirect costs and should be avoided. On the other hand, omission errors may lead to neglecting the selection of sites that may include essential habitat for the species. Ideally SDMs would be devoid of those errors but in reality perfect truth cannot be attained by a model [19,143]. Thus, critically interpreting model results and deciding on how to balance model credibility (based on a desired, subjective, confidence level) and applicability (regarding to which extent model results are useful) is an important part of conservation management practices [144,145].

In the present case the fact that there are no known records of the sei whale in the Gulf of St. Lawrence is a strong indicator that we are in the presence of a commission error. In many cases, however, this type of error would not be as easily detectable. In fact, one of the goals in creating an SDM may be identifying new areas where a species has not been detected for lack of survey effort (e.g. [146]). In such cases, other measures can be undertaken. As mentioned before, the suitability values calculated by the model can be used in setting a threshold to down-weight marginal habitats [144]. Another approach is to survey areas identified as potential habitat to confirm the presence of the target species. The latter option obviously has costs but since the potential habitat areas identified by a model are normally much smaller than the whole study area, creating an SDM can be an efficient way of allocating limited funds.

The model developed here was designed to give a wide scale picture, both at spatial and temporal dimensions. The areas identified as suitable habitat are probably oversized at any given moment, as whale distribution in the feeding areas tends to be patchy as reflex of prey distribution. Depending on the management goals, models can be developed at much higher resolution and/or at smaller spatial and temporal scales. For instance, Pendleton et al. [64] developed a predictive model of habitat suitability for the right whale at 8 day resolution that can aid in creating near real-time adaptive management measures. In principle, that can also be achieved with the model created in the present work. After the model has been trained it can be used to forecast changes in habitat suitability (*projecting* in the MAXENT terminology) for new environmental configurations, as long as the predictors used to train the model are maintained [89]. Thus the model predictions could be updated at any desired time scale by projecting predictions into different environments resulting from changes of highly dynamic predictors such as SST and Chl-a concentration.

Ecological interpretation and consequences of model results

The resulting suitability scores from the SDM_f clearly show distinct areas of high and low habitat suitability, demonstrating that whales do not use the whole study area in the same manner. Thus more efficient management can be attained by focusing efforts and resources in specific areas inside the study region. These results can in fact be used to create cost-effective survey designs that can improve data quality for future ecological and population studies of the sei whale in the region.

The MaxEnt approach is more focused in creating accurate SDMs than in unveiling the form of the relationship of the dependent and predictor variables. For that goal other approaches are more appropriate [95]. Nevertheless, the ranking of the permutation importance scores of the variables calculated by MAXENT can be viewed as proxy for habitat preferences, giving some insight into what is relevant to the animals [64,89]. In the final model, SST had the highest permutation importance score, followed by bathymetry. These two variables are known to influence the sei whale distribution [57] and it is natural that they contributed the most to the model. Despite bathymetry not being the most important predictor, it still seems to be important in defining the preferred habitat of sei whale. According to the SDM_f the shallower areas do not comprise suitable habitat (Figure 3-5), which is in agreement with what is known about the habitat preferences of the species [57]. Nevertheless not all deep areas (see for example the Jordan Basin) seem to comprise suitable habitat, which is natural given the contribution of the other variables.

The only other variable retained in the SDM_f was *lnChl-a*. Chlorophill-a concentration is a proxy for primary production and it probably has low direct ecological relevance to the sei whale. However, as mentioned above, primary production drives the distribution and abundance of grazing zooplankton such as the *C. finmarchicus,* which in turn drives whale distribution. The retention of the variable in the model shows that it still has some explanatory power but its low permutation importance is probably due to the indirect influence in the habitat preferences of the sei whale.

Some of the most important areas identified in the model were expected to be important for the sei whale either by the sheer number of sightings in a given area or by ecological relevance. The Wilkinson basin came up as a highly suitable habitat for the sei whale. Judging by the clustering and number of sightings in the area this was an expected outcome. The shelf break also seems to comprise important summering habitat for the sei whale. Strong shelf-slope fronts develop in the areas southeast of Georges bank and off the Scotian shelf [147] and may explain this apparent affinity of the whales for the shelf break. The association of marine predators with fronts is well documented throughout the literature, namely for pelagic fishes (e.g. [148,149]), sea turtles (e.g. [150,151]), seabirds (e.g. [152-154]) and marine mammals, including the sei whale (e.g. [127,150,152,155-157]). It is believed that these associations with oceanic fronts enhance foraging efficiency by concentration of prey through convergence or by increased prey biomass due to elevated primary production [152].

It comes as an interesting outcome that the variable *slope*, presumably a proxy for the presence of the shelf break, was not retained in the SDM_f. The fact that the model was able to identify the shelf break as suitable habitat even after the discarding of that variable indicates that the oceanographic processes are more important to the animals than the topographic configuration of the bottom.

Part of the habitat characterized as suitable for sei whales in this study overlaps with the feeding habitat of the highly endangered right whale. The southernmost part of the

Wilkinson basin and the entrance to the Great South Channel have been identified as important feeding habitat for the right whale due to the dense and predictable aggregations of the copepod *C. finmarchicus* [64,158]. This copepod is known to be the main prey of the sei whale in the Gulf of Maine [159-161] and the coincidence of feeding habitats unveiled by this study raises the question about how much competition for prey exist between the sei and right whales in the study region.

The North Atlantic right whale is one of the most endangered cetacean species, with population effectives numbering on few hundred animals and with an exceedingly slow recovery rate from past whaling [67]. Calf production in the right whale is influenced by the abundance of *C. finmarchicus* and competition between this whale species with other predators such as zooplanktivorous fish and sei whales may be among several factors slowing down the recovery of the population [160-162].

Debate about the potential competition between sei and right whales for the same resources is not new, but Kenney et al. [162] suggested that in the shelf waters of the U.S. this should not be an issue due to little distributional overlap between the two species. That reasoning is in contrast with the findings from this work. The present work focuses on a much wider scale than that alluded by Kenney et al. [162] and may miss some detail at lower scale. On the other hand, compared to what was available to Kenney et al. [162], this work benefited from a longer sampling period and larger sightings dataset, as well as a state of the art habitat modelling tool. It is possible that at scales in the order of few kilometres some segregation between the two species do exist, which in practical terms would mean that the whales are not competing for the same prey patches. Nevertheless, the results reported here indicate that the feeding habitats of the two species partially overlap at a broader scale. Even if whales are not competing for the same patches of food, they are probably using the same prey population. Interspecific competition between the predators of *C. finmarchicus* in the region are complex and difficult to quantify, but should be more noticeable in years of lower copepod abundance [160]. In fact it seems that competition may be more detrimental for individual sei whales than for right whales since the latter can feed both at depth and at the surface, while sei whales seem to feed only near the surface [159,160,163]. Nevertheless, due to its endangered status, the right whale probably is the species of more concern regarding effects at a population level from competition with other predators. In face of the results from this study, further investigation should be carried on the potential effects of competition between sei and right whales, as well as with other *C. finmarchicus* predators in the region.

Conclusions

Between climactic changes and human activities, the threats to whales are increasing and tools for their proper management are invaluable [164,165]. It is critical that we understand how these animals use their habitat, how we are impacting that habitat and how environmental changes may affect them. This is especially relevant for insufficiently studied species such as the sei whale [57]. In that sense the development of species distribution models such as the one presented here are of paramount importance for their contribution in the development of proper management measures.

The utility of the SDM that resulted from the present work seems clear. By knowing which are the most relevant areas for the animals, direct and indirect threats can be investigated and acted upon if necessary. For example, with the results of the SDM_f it will be easier to access if sei whales are facing habitat degradation or loss in the region or if (and where) they face direct threats such as ship strikes or becoming entangled in fishing gear [165].

The SDM created in the present work improves our understanding not only about the sei whale but also the relationship of this whale with another important species in the study region, such as the right whale. This result shows how coupling SDMs for different species is beneficial not only to the focal species but can also contribute for a more comprehensive management of the ecosystem.

Some of the limitations of the model and strategies for its improvement have been identified and depending on data availability more refined models can be created using the same procedures laid here. Even at its present state, the model could be easily used to predict habitat suitability at different spatial and temporal scales including for near real-time management approaches such as that proposed by Pendleton et al. [64]. The inclusion of prey as an explanatory variable would probably improve model accuracy and efforts should be made on that direction in the future.

REFERENCES

1. Krebs CJ (1994) Ecology: the experimental analysis of distribution and abundance. New York, USA: Harper Collins College Publishers. 801 p.

2. Goudie A (2013) The human impact on the natural environment. Multiple cities: Wiley-Blackwell. 410 p.

3. Chen I-C, Hill JK, Ohlemüller R, Roy DB, Thomas CD (2011) Rapid range shifts of species associated with high levels of climate warming. Science 333: 1024-1026.

4. Beaugrand G (2009) Decadal changes in climate and ecosystems in the North Atlantic Ocean and adjacent seas. Deep Sea Research Part II: Topical Studies in Oceanography 56: 656-673.

5. Pitois SG, Fox CJ (2006) Long-term changes in zooplankton biomass concentration and mean size over the Northwest European shelf inferred from Continuous Plankton Recorder data. ICES Journal of Marine Science 63: 785-798.

6. Kovacs KM, Lydersen C (2008) Climate change impacts on seals and whales in the North Atlantic Arctic and adjacent shelf seas. Science Progress 91: 117-150.

7. Moore SE, Huntington HP (2008) Arctic marine mammals and climate change: impacts and resilience. Ecological Applications 18: S157-S165.

8. Salvadeo CJ, Lluch-Belda D, Gomez-Gallardo A, Urban-Ramirez J, MacLeod CD (2010) Climate change and a poleward shift in the distribution of the Pacific white-sided dolphin in the northeastern Pacific. Endangered Species Research 11: 13-19.

9. Baker JD, Harting AL, Wurth TA, Johanos TC (2011) Dramatic shifts in Hawaiian monk seal distribution predicted from divergent regional trends. Marine Mammal Science 27: 78-93.

10. Johannessen OM, Miles MW (2011) Critical vulnerabilities of marine and sea ice–based ecosystems in the high Arctic. Regional Environmental Change 11: S239-S248.

11. Kovacs KM, Lyndersen C, Overland JE, Moore SE (2011) Impacts of changing sea-ice conditions on Arctic marine mammals. Marine Biodiversity 41: 181-194.

12. Kirkman SP, Yemane D, Oosthuizen WH, Meÿer MA, Kotze PGH, et al. (2013) Spatio-temporal shifts of the dynamic Cape fur seal population in southern Africa, based on aerial censuses (1972–2009). Marine Mammal Science 29: 497-524.

13. Torres LG, Smith TD, Sutton P, MacDiarmid A, Bannister J, et al. (2013) From exploitation to conservation: habitat models using whaling data predict distribution patterns and threat exposure of an endangered whale. Diversity and Distributions: n/a-n/a.

14. Williams N (2007) Polar bears shift from thinning ice. Current Biology 17: R571-R572.

15. World Bank, editor (2006) Where is the wealth of nations? Measuring capital for the 21st century. Washington: World Bank Publications. 188 p.

16. Grumbine RE (1994) What is ecosystem management? Conservation Biology 8: 27-38.

17. Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, et al. (2013) Predicting species distributions for conservation decisions. Ecology Letters 16: 1424-1435.

18. Costanza R, d'Arge R, de Groot R, Farber S, Grasso M, et al. (1997) The value of the world's ecosystem services and natural capital. Nature 387: 253-260.

19. Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecological Modelling 135: 147-186.

20. Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40: 677-697.

21. Austin M (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecological Modelling 200: 1-19.

22. Warren DL (2012) In defense of niche modeling Trends in ecology & evolution 27: 497-500.

23. Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129-151.

24. Beaumont LJ, Hughes L, Poulsen M (2005) Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. Ecological Modelling 186: 251-270.

25. Nix HA (1986) A biogeographic analysis of Australian Elaphid snakes. Atlas of Elapid snakes of Australia: Australian Government Publishing Service. pp. 4-15.

26. Nelder JA, Wedderburn RWM (1972) Generalized Linear Models. Journal of the Royal Statistical Society Series A (General) 135: 370-384.

27. Hastie T, Tibshirani R (1986) Generalized Additive Models. Statistical Science 1: 297-318.

28. Guisan A, Edwards Jr TC, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling 157: 89-100.

29. Aspinall R (1992) An inductive modelling procedure based on Bayes' theorem for analysis of pattern in spatial data. International Journal of Geographical Information Systems 6: 105-121.

30. Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. Journal of Animal Ecology 77: 802-813.

31. Breiman L (2001) Random forests. Machine Learning 45: 5-32.

32. Lawler JJ, White D, Neilson RP, Blaustein AR (2006) Predicting climate-induced range shifts: model differences and model reliability. Global Change Biology 12: 1568-1584.

33. Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, et al. (2011) A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17: 43-57.

34. Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259.

35. Olden JD (2008) Machine learning withouth tears: a primer for ecologists. The Quarterly Review of Biology 83: 171-193.

36. Stockwell D, Peters D (1999) The GARP modelling system: problems and solutions to automated spatial prediction. International Journal of Geographical Information Science 13: 143-158.

37. Stockwell DRB, Noble IR (1992) Induction of sets of rules from animal distribution data: A robust and informative method of data analysis. Mathematics and Computers in Simulation 33: 385-390.

38. Uriarte M, Yackulic CB (2009) Preaching to the unconverted: Ecological Applications. Ecological Applications 19: 592-596.

39. Monk J, Ierodiaconou D, Harvey E, Rattray A, Versace VL (2012) Are we predicting the actual or apparent distribution of temperate marine fishes? PLoS ONE 7: e34558.

40. Pueyo S (2012) Comparison of two views of maximum entropy in biodiversity: Frank (2011) and Pueyo et al. (2007). Ecology and Evolution 2: 988-993.

41. Yackulic CB, Chandler R, Zipkin EF, Royle JA, Nichols JD, et al. (2012) Presence-only modelling using MAXENT: when can we trust the inferences? Methods in Ecology and Evolution: n/a-n/a.

42. Fitzpatrick MC, Gotelli NJ, Ellison AM (2013) MaxEnt versus MaxLike: empirical comparisons with ant species distributions. Ecosphere 4: art55.

43. Sinclair SJ, White MD, Newell GR (2010) How useful are species distribution models for managing biodiversity under future climates? Ecology and Society 15: 8p. [online].

44. Schröder B (2008) Challenges of species distribution modeling belowground. Journal of Plant Nutrition and Soil Science 171: 325-337.

45. Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters 8: 993-1009.

46. Richards CL, Carstens BC, Lacey Knowles L (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. Journal of Biogeography 34: 1833-1845.

47. Mac Nally R (2000) Regression and model-building in conservation biology, biogeography and ecology: The distinction between – and reconciliation of – ‘predictive’ and ‘explanatory’ models. Biodiversity & Conservation 9: 655-671.

48. Breiman L (2001) Statistical modeling: the two cultures. Statistical Science 16: 199-231.

49. Fraser DAS (1968) A Black Box or a comprehensive model. Technometrics 10: 219-229.

50. Clark JS, Gelfand AE (2006) A future for models and data in environmental science. Trends in Ecology & Evolution 21: 375-380.

51. Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, et al. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19: 181-197.

52. Jiménez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. Diversity and Distributions 14: 885-890.

53. Soberón J (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics 2: 1-10.

54. Warton DI, Shepherd LC (2010) Poisson point process models solve the "pseudo-absence problem" for presence-only data in ecology. Annals of Applied Statistics 4: 1388-1402.

55. Mateo RG, Croat TB, Felicísimo ÁM, Muñoz J (2010) Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. Diversity and Distributions 16: 84-94.

56. VanDerWal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? Ecological Modelling 220: 589-594.

57. Prieto R, Janiger D, Silva MA, Waring GT, Gonçalves JM (2012) The forgotten whale: a bibliometric analysis and literature review of the North Atlantic sei whale *Balaenoptera borealis*. Mammal Review 42: 235-272.

58. Monk J, Ierodiaconou D, Versace VL, Bellgrove A, HArvey E, et al. (2010) Habitat suitability for marine fishes using presence-only modelling and multibeam sonar. Marine Ecology Progress Series 420: 157-174.

59. Pittman SJ, Brown KA (2011) Multi-scale approach for predicting fish species distributions across coral reef seascapes. PLoS ONE 6: e20583.

60. Martín-García L, González-Lorenzo G, Brito-Izquierdo IT, Barquín-Diez J (2013) Use of topographic predictors for macrobenthic community mapping in the Marine Reserve of La Palma (Canary Islands, Spain). Ecological Modelling 263: 19-31.

61. Oppel S, Meirinho A, Ramírez I, Gardner B, O'Connell AF, et al. (2012) Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biological Conservation 156: 94-104.

62. Gregr EJ (2011) Insights into North Pacific right whale *Eubalaena japonica* habitat from historic whaling records. Endangered Species Research 15: 223-239.

63. Moura AE, Sillero N, Rodrigues A (2012) Common dolphin (*Delphinus delphis*) habitat preferences using data from two platforms of opportunity. Acta Oecologica 38: 24-32.

64. Pendleton DE, Sullivan PJ, Brown MW, Cole TVN, Good CP, et al. (2012) Weekly predictions of North Atlantic right whale Eubalaena glacialis habitat reveal influence of prey abundance and seasonality of habitat preferences. Endangered Species Research 18: 147-161.

65. Smith JN, Grantham HS, Gales N, Double MC, Noad MJ, et al. (2012) Identification of humpback whale breeding and calving habitat in the Great Barrier Reef. Marine Ecology Progress Series 447: 259-272.

66. True FW (1903) First record of the pollack whale (Balaenoptera borealis) in the western North Atlantic. Science 17: 150.

67. Kraus SD, Rolland RM (2007) The urban whale. North Atlantic right whales at the crossroads. Cambridge, MA, USA & London, UK: Harvard University Press. 543 p.

68. Mullen KA, Peterson ML, Todd SK (2013) Has designating and protecting critical habitat had an impact on endangered North Atlantic right whale ship strike mortality? Marine Policy 42: 293-304.

69. Garrison LP (2007) The big picture: modeling right whales in space and time. In: Kraus SD, Rolland RM, editors. The urban whale North Atlantic right whales at the crossroads. Cambridge, MA, USA & London, UK: Harvard University Press. pp. 460-487.

70. Carnot S (1824) Reflexions sur la puissance motrice du feu et suer les machines propres a développer cette puissance. Paris: Bachelier. 102 p.

71. Clapeyron É (1834) Mémoire sur la pussaince motrice de la chaleur. Journal de l'École Polytechnique XXIIIe Cahier: 153–190.

72. Clausius RJE (1867) The mechanical theory of heat, with its appplications to the steam-engine and to the physical properties of bodies. London: John Van Voorst. 374 p.

73. Magie WF (1899) The second law of thermodynamics. Memoirs by Carnot, Clausius and Thomson; Ames JS, editor. New York and London: Harper & Brothers Publishers. 151 p.

74. Clausius RJE (1899) On the motive power of heat, and on the laws which can be deduced from it for the theory of heat. In: Magie WF, editor. The second law of thermodynamics Memoirs by Carnot, Clausius and Thomson. New York and London: Harper & Brothers Publishers. pp. 65-106.

75. Styer DF (2000) Insight into entropy. American Journal of Physics 68: 1090-1096.

76. Goldman T, Sharp DH (2012) Order from disorder in closed systems via time-reversal violation. EPL (Europhysics Letters) 97: 61003.

77. Campisi M, Kobe DH (2010) Derivation of the Boltzmann principle. American Journal of Physics 78: 608-615.

78. Perrin J (1909) Mouvement brownien et realité moléculaire. Annales de Chimie et de Physique 8: 5-114.

79. Shannon C (1948) A mathematical theory of communication. Bell System Technical Journal 27: 379-423, 623-656.

80. Chakrabarti CG, De K (2000) Boltzmann-Gibbs entropy: axiomaticc characterization and application. International Journal of Mathematics and Mathematical Sciences 23: 243-251.

81. Jaynes ET (1957) Information theory and statistical mechanics. Physical Review 106: 620-630.

82. Jaynes ET (1957) Information theory and statistical mechanics. II. Physical Review 108: 171-190.

83. Harremoës P, Topsøe F (2001) Maximum Entropy fundamentals. Entropy 3: 191-226.

84. Keynes JM (1921) A treatise on probability. London: Macmillan and Co., Ltd. 466 p.

85. Banavar JR, Maritan A, Volkov I (2010) Applications of the principle of maximum entropy: from physics to ecology. Journal of Physics: Condensed Matter 22: 063101.

86. Beck C (2009) Generalized information and entropy measures in physics. Contemporary Physics 50: 495-510.

87. Harte J (2011) Maximum entropy and ecology. A theory of abundance, distribution and energetics; May R, Harvey P, Godfray C, Dunne J, editors. Oxford, UK: Oxford University Press. 280 p.

88. Dudík M, Phillips SJ, Schapire RE (2004) Performance guarantees for regularized maximum entropy density estimation. In: Shawe-Taylor J, Singer Y, editors. Learning Theory: Springer Berlin Heidelberg. pp. 472-486.

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

90. Dudík M, Phillips SJ, Schapire RE (2005) Correcting sample selection bias in maximum entropy density estimation. Cambridge, Massachusetts, USA: MIT Press.

91. Schaffer C (1993) Overfitting avoidance as bias. Machine Learning 10: 153-178.

92. Dudík M, Phillips SJ, Schapire RE (2007) Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. Journal of Machine Learning Research 8: 1217-1260.

93. Tibshirani R (2011) Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73: 273-282.

94. Tibshirani R (1996) Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society Series B-Methodological 58: 267-288.

95. Aguirre-Gutiérrez J, Carvalheiro LG, Polce C, van Loon EE, Raes N, et al. (2013) Fit-for-Purpose: species distribution model performance depends on evaluation criteria – Dutch hoverflies as a case study. PLoS ONE 8: e63708.

96. Elith J, Graham CH (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32: 66-77.

97. Royle JA, Chandler RB, Yackulic C, Nichols JD (2012) Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. Methods in Ecology and Evolution 3: 545-554.

98. Halvorsen R (2013) A strict maximum likelihood explanation of MaxEnt, and some implications for distribution modelling. Sommerfeltia 36: 1-132.

99. Phillips SJ, Elith J (2010) POC plots: calibrating species distribution models with presence-only data. Ecology 91: 2476-2484.

100. Radosavljevic A, Anderson RP (2013) Making better Maxent models of species distributions: complexity, overfitting and evaluation. Journal of Biogeography: n/a-n/a.

101. Raes N, ter Steege H (2007) A null-model for significance testing of presence-only species distribution models. Ecography 30: 727-736.

102. Shipley B (2010) Inferential permutation tests for maximum entropy models in ecology. Ecology 91: 2794-2805.

103. Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33: 607-611.

104. Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecological Applications 21: 335-342.

105. Fithian W, Hastie T (2013) Finite-sample equivalence in statistical models for presence-only data. Annals of Applied Statistics 7: 1917-1939.

106. Renner IW, Warton DI (2013) Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. Biometrics 69: 274-281.

107. Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. Sociological Methods & Research 33: 261-304.

108. Kenney RD (2001) The North Atlantic Right Whale Consortium databases. Maritimes 43: 3-5.

109. Kenney RD (2008) The North Atlantic Right Whale Consortium Database: a guide for users and contributor. Narragansett, Rhode Island, USA: North Atlantic Right Whale Consortium. 119 p.

110. Amante C, Eakins BW (2008) ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. Boulder, Colorado, USA: NOAA. 19 p.

111. Campbell JW, Blaisdell JM, Darzi M (1995) Level-3 SeaWiFS data products: spatial and temporal binning algorithms. In: Hooker SB, Firestone ER, Acker JG, editors. NASA Tech Memo 104566, vol 32. Greenbelt, MD, USA: NASA Goddard Space Flight Center.

112. Jenness JS (2004) Calculating landscape surface area from digital elevation models. Wildlife Society Bulletin 32: 829-839.

113. Jenness JS (2009) Tools for Graphics and Shapes: Extension for ArcGIS. Jenness Enterprises.

114. Jones KH (1998) A comparison of algorithms used to compute hill slope as a property of the DEM. Computers & Geosciences 24: 315-323.

115. Zevenbergen LW, Thorne CR (1987) Quantitative analysis of land surface topography. Earth Surface Processes and Landforms 12: 47-56.

116. Portner RA, Dickinson JA, Daczko NR (2011) Interaction of gravity flows with a rugged mid-ocean-ridge seafloor: an outcrop example from Macquarie Island. Journal of Sedimentary Research 81: 355-375.

117. Beasom SL, Wiggers EP, Giardino JR (1982) A technique for assessing land surface ruggedness. The Journal of Wildlife Management 47: 1163-1166.

118. Riley SJ, DeGloria SD, Elliot R (1999) A terrain ruggedness index that quantifies topographic heterogeneity. Intermountain Journal of Sciences 5: 23-27.

119. Sappington JM, Longshore KM, Thompson DB (2007) Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave desert. Journal of Wildlife Management 71: 1419-1426.

120. Best PB, Lockyer CH (2002) Reproduction, growth and migrations of sei whales *Balaenoptera borealis* off the west coast of South Africa in the 1960s. South African Journal of Marine Sciences 24: 111-133.

121. Silber GK, Newcomer MW, Silber PC, Pérez-Cortés M H, Ellis GM (1994) Cetaceans of the northern Gulf of California: distribution, occurrence, and relative abundance. Marine Mammal Science 10: 283-298.

122. Pittman SJ, Costa B (2010) Linking cetaceans to their environment: spatial data acquisition, digital processing and predictive modeling for marine spatial planning in the Northwest Atlantic. In: Cushman SA, Huettmann F, editors. Spatial Complexity, Informatics, and Wildlife Conservation: Springer Japan. pp. 387-408.

123. Durner GM, Whiteman JP, Harlow HJ, Amstrup SC, Regehr EV, et al. (2011) Consequences of long-distance swimming and travel over deep-water pack ice for a female polar bear during a year of extreme sea ice retreat. Polar Biology 34: 975-984.

124. Magalhães MC, Santos RS, Hamer KC (2008) Dual-foraging of Cory's shearwaters in the Azores: feeding locations, behaviour at sea and implications for food provisioning of chicks. Marine Ecology Progress Series 359: 283-293.

125. Sharples RJ, Moss SE, Patterson TA, Hammond PS (2012) Spatial variation in foraging behaviour of a marine top predator (*Phoca vitulina*) determined by a large-scale satellite tagging program. PLoS ONE 7: e37216.

126. Smultea MA (1994) Segregation by humpback whale (Megaptera novaeangliae) cows with a calf in coastal habitat near the island of Hawaii. Canadian Journal of Zoology 72: 805-811.

127. Doniol-Valcroze T, Berteaux D, Larouche P, Sears R (2007) Influence of thermal fronts on habitat selection by four rorqual whale species in the Gulf of St. Lawrence. Marine Ecology Progress Series 335: 207-216.

128. Gill PC, Morrice MG, Page B, Pirzl R, Levings AH, et al. (2011) Blue whale habitat selection and within-season distribution in a regional upwelling system off southern Australia. Marine Ecology Progress Series 421: 243-263.

129. Gannier A, Petiau E (2006) Environmental variables affecting the residence of spinner dolphins (*Stenella longirostris*) in a bay of Tahiti (French Polynesia). Aquatic Mammals 32: 202-211.

130. Forcada J (2002) Distribution. In: Perrin WF, Würsig B, Thewissen JGM, editors. Encyclopedia of marine mammals. London, UK: Academic Press. pp. 316-321.

131. Clark JS (2007) Models for ecological data: an introduction. Princeton, N.J.: Princeton University Press. xiii, 617 p. p.

132. Warren DL, Wright AN, Seifert SN, Shaffer HB (2013) Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. Diversity and Distributions early view: n/a-n/a.

133. Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Berlin, Germany: Springer-Verlag. XXVI, 488 p.

134. Kowalski CJ (1972) On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. Journal of the Royal Statistical Society Series C (Applied Statistics) 21: 1-12.

135. Xu W, Hou Y, Hung YS, Zou Y (2013) A comparative analysis of Spearman's rho and Kendall's tau in normal and contaminated normal models. Signal Processing 93: 261-276.

136. Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, et al. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36: 027-046.

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

138. Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17: 145-151.

139. Ferrier S, Watson G, Pearce J, Drielsma M (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. Biodiversity & Conservation 11: 2275-2307.

140. Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. Ecological Modelling 133: 225-245.

141. Manel S, Williams HC, Ormerod SJ (2001) Evaluating presence–absence models in ecology: the need to account for prevalence. Journal of Applied Ecology 38: 921-931.

142. Belkin IM (2009) Rapid warming of large marine ecosystems. Progress in Oceanography 81: 207-213.

143. Oreskes N, Shrader-Frechette K, Belitz K (1994) Verification, validation, and confirmation of numerical models in the Earth sciences. Science 263: 641-646.

144. Loiselle BA, Howell CA, Graham CH, Goerck JM, Brooks T, et al. (2003) Avoiding pitfalls of using species distribution models in conservation planning. Conservation Biology 17: 1591-1600.

145. Rykiel Jr EJ (1996) Testing ecological models: the meaning of validation. Ecological Modelling 90: 229-244.

146. Tittensor DP, Baco AR, Brewin PE, Clark MR, Consalvey M, et al. (2009) Predicting global habitat suitability for stony corals on seamounts. Journal of Biogeography 36: 1111-1128.

147. Belkin IM, Cornillon PC, Sherman K (2009) Fronts in large marine ecosystems. Progress in Oceanography 81: 223-236.

148. Santos AMP, Fiúza AFG, Laurs RM (2006) Influence of SST on catches of swordfish and tuna in the Portuguese domestic longline fishery. International Journal of Remote Sensing 27: 3131-3152.

149. Worm B, Sandow M, Oschlies A, Lotze HK, Myers RA (2005) Global patterns of predator diversity in the open oceans. Science 309: 1365-1369.

150. Etnoyer P, Canny D, Mate BR, Morgan LE, Ortega-Ortiz JG, et al. (2006) Sea-surface temperature gradients across blue whale and sea turtle foraging trajectories off the Baja California Peninsula, Mexico. Deep-Sea Research Part II 53: 340-358.

151. Seminoff JA, Zárate P, Coyne M, Foley DG, Parker D, et al. (2008) Post-nesting migrations of Galápagos green turtles *Chelonia mydas* in relation to oceanographic conditions: integrating satellite telemetry with remotely sensed ocean data. Endangered Species Research 4: 57-72.

152. Bost CA, Cotté C, Bailleul F, Cherel Y, Charrassin JB, et al. (2009) The importance of oceanographic fronts to marine birds and mammals of the southern oceans. Journal of Marine Systems 78: 363-376.

153. Charrassin J-B, Bost C-A (2001) Utilization of the oceanic habitat by king penguins over the annual cycle. Marine Ecology Progress Series 221.

154. Kai ET, Rossi V, Sudre J, Weimerskirch H, Lopez C, et al. (2009) Top marine predators track Lagrangian coherent structures. PNAS 106: 8245-8250.

155. Bailleul F, Charrassin J-B, Monestiez P, Roquet F, Biuw M, et al. (2007) Successful foraging zones of southern elephant seals from the Kerguelen Islands in relation to oceanographic conditions. Philosophical Transactions of the Royal Society of London B 362: 2169-2181.

156. Baumgartner MF, Cole TVN, Campbell RG, Teegarden GJ, Durbin EG (2003) Associations between North Atlantic right whales and their prey, *Calanus finmarchicus*, over diel and tidal time scales. Marine Ecology Progress Series 264: 155-166.

157. Laptikhovsky V (2009) Oceanographic factors influencing the distribution of South American fur seal, Arctocephalus australis around the Falkland Islands before the breeding season. Journal of the Marine Biological Association of the United Kingdom 89: 1597-1600.

158. Baumgartner MF, Mate BR (2005) Summer and fall habitat of North Atlantic right whales (*Eubalaena glacialis*) inferred from satellite telemetry. Canadian Journal of Fisheries and Aquatic Sciences 62: 527-543.

159. Baumgartner MF, Fratantoni DM (2008) Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders. Limnology and Oceanography 53: 2197–2209.

160. Baumgartner MF, Lysiak NSJ, Schuman C, Urban-Rich J, Wenzel FW (2011) Diel vertical migration behavior of *Calanus finmarchicus* and its influence on right and sei whale occurrence. Marine Ecology Progress Series 423: 167-184.

161. Payne PM, Wiley DN, Young SB, Pittman S, Clapham PJ, et al. (1990) Recent fluctuations in the abundance of baleen whales in the southern Gulf of Maine in relation to changes in selected prey. Fishery Bulletin 88: 687-696.

162. Kenney RD, Hyman MAM, Owen RE, Scott GP, Winn HE (1986) Estimation of prey densities required by western north Atlantic right whales. Marine Mammal Science 2: 1-13.

163. Baumgartner MF, Mate BR (2003) Summertime foraging ecology of North Atlantic right whales. Marine Ecology Progress Series 264: 123-135.

164. Van Der Hoop JM, Moore MJ, Barco SG, Cole TVN, Daoust P-Y, et al. (2013) Assessment of management to mitigate anthropogenic effects on large whales. Conservation Biology 27: 121-133.

165. Moore MJ (2014) How we all kill whales. ICES Journal of Marine Science: Journal du Conseil.

FOURTH CHAPTER

Assessing performance of Bayesian state-space models fit to Argos satellite telemetry locations processed with Kalman filtering

ABSTRACT: Argos recently implemented a new algorithm to calculate locations of satellite-tracked animals that uses a Kalman filter (KF). The KF algorithm is reported to increase the number and accuracy of estimated positions over the traditional Least Squares (LS) algorithm, with potential advantages to the application of state-space methods to model animal movement data. The performance of two Bayesian state-space models (SSMs) fitted to satellite tracking data processed with KF algorithm were tested. Tracks from 7 harbour seals (*Phoca vitulina*) tagged with ARGOS satellite transmitters equipped with Fastloc GPS loggers were used to calculate the error of locations estimated from SSMs fitted to KF and LS data, by comparing those to "true" GPS locations. Data on 6 fin whales (*Balaenoptera physalus*) were used to investigate consistency in movement parameters, location and behavioural states estimated by switching state-space models (SSSM) fitted to data derived from KF and LS methods. The model fit to KF locations improved the accuracy of seal trips by 27% over the LS model. 82% of locations predicted from the KF model and 73% of locations from the LS model were <5 km from the corresponding interpolated GPS position. Uncertainty in KF model estimates (5.6±5.6 km) was nearly half that of LS estimates (11.6±8.4 km). Accuracy of KF and LS modelled locations was sensitive to precision but not to observation frequency or temporal resolution of raw Argos data. On average, 88% of whale locations estimated by KF models fell within the 95% probability ellipse of paired locations from LS models. Precision of KF locations for whales was generally higher. Whales' behavioural mode inferred by KF models matched the classification from LS models in 94% of the cases. State-space models fit to KF data can improve spatial accuracy of location estimates over LS models and produce equally reliable behavioural estimates.

INTRODUCTION

The collection of individual animal movement data has become widely utilized by ecologists in the last decade due to the improvement of the underlying technologies and reduction of operational costs involved in animal telemetry. Of the several technologies available, one of the most popular is that based on satellite tags (platform transmitter terminals, PTTs) using the Argos system [1]. However, most satellite tags record observations at irregular intervals and with considerable error, meaning that movements are observed neither continuously nor with complete accuracy. The Argos service provider assigns a quality index, or location class (LC), to each position based on its estimated precision. The radius of error (assumed to include 68% of positions) for each LC is: LC 3<250m, LC 2 250-500m, LC 1 500-1500m, LC 0 >1500m, and LC A, B and Z for which no estimate of error is provided [2]. However, attempts to measure spatial error of Argos locations using either stationary tests or double-tagging experiments with free-ranging animals consistently reported larger errors than those indicated by Argos.

Many of these studies also provided error estimates for location classes A and B, showing these could be in the range of tens to hundreds of kilometres (reviewed in [3]).

Varying accuracy and precision, and unevenness in space and time of telemetry data can affect the determination of distribution, habitat use and behavioural patterns of animals and severely bias the calculation of movement metrics [3-5]. Therefore, advanced statistical methods are necessary to account for spatial error and temporal irregularity in the data and to understand the movement behaviour of the tracked animals.

Jonsen et al. [6] proposed a state-space framework for analysis of movement data that was further developed in Jonsen et al. [7-9], in order to deal with the biological and statistical complexities associated with animal tracking data. State-space models (SSMs) offer a powerful way to infer latent movement from imperfect estimates of animal locations by allowing uncertainty in both the observations and in the movement dynamics to be accounted for separately in the estimation process. Additionally, movement models can include behavioural or environmental effects, enabling a better understanding of the interaction between an animal's behaviour and its environment [8-11]. SSMs have been used widely among ecologists and are currently one of the tools of choice for analysing tracking data of several taxa and across environments [5,12-17].

Geolocation of animals tracked with Argos systems is based on the Doppler shift of the tag's fixed transmission frequency; i.e. the frequency shift of the tag's signal received at the orbiting satellite as it approaches and moves away from the tag [18]. The system estimates two possible positions, which are symmetrical on each side of the satellite ground track. Until recently, Argos used a non-linear Least Squares (LS) algorithm to refine the tag's position estimates and to select the one with the minimal residual error. However, the LS positioning algorithm presented a number of limitations and affected the quality of the tracks obtained. For instance, when the LS algorithm could not complete the refinement routine or check the validity of the most plausible location estimate, no position was provided. In addition, the process required at least two transmissions (also called messages) during a single satellite pass to compute a position and at least four messages to produce an error estimate.

In May 2011, Argos implemented a new algorithm that accounts for movement dynamics and uses a Kalman filter (KF) to estimate positions [1,19]. The algorithm uses a correlated random walk model to predict the next position and its estimated error based on the previous positions and estimated error. It then uses the Doppler frequency-shift measurements acquired during a satellite pass to update the position predicted by the model and return a final position. Compared to the LS method, the Kalman filtering estimator is reported to improve the accuracy of estimated positions and to increase the number of positions up to 13% [19]. Such improvements may have a significant impact in studies where relatively few messages are received with each satellite pass, which is the case for many marine and dense forest species.

Although the new processing algorithm may bring significant advantages, it may also introduce changes in the autocorrelation structure of the Argos satellite data. Given that many published SSM applications for animal tracking data do not currently account for the potential autocorrelation in location errors introduced by the new KF algorithm, models fit to datasets with differing degrees of autocorrelated errors could lead to biased estimates of movement parameters, behavioural states, and their uncertainties. Several studies have examined the validity of SSMs applied to data obtained with the LS positioning algorithm and quantified the precision of predicted locations (e.g. [20-22]), none examined how changes introduced by the KF algorithm might affect the application of these models.

The aim of the present work is to assess the performance of Bayesian SSMs fit to satellite tracking data processed with the new KF positioning algorithm introduced by Argos. This work used two real datasets from marine taxa that differ greatly in their movement ranges– harbour seals (*Phoca vitulina*) and fin whales (*Balaenoptera physalus*) – as SSMs are known to be sensitive to the scale of movement [20]. Using data from 7 harbour seals instrumented with ARGOS satellite transmitters equipped with Fastloc GPS loggers (hereafter GPS/Argos tags), estimated locations from a hierarchical SSM (hSSM) fit to data processed with KF and LS algorithms were compared to the GPS positions obtained from the same tag to (1) assess spatial accuracy of locations from models fit to data derived from each algorithm; and (2) determine how spatial accuracy varies with observation frequency, temporal resolution and reported precision of Argos locations. Models fit to fin whale tracks could not be evaluated through comparison with GPS data because whales were instrumented with Argos-only transmitters. Satellite tracks of 6 fin whales were used to compare location and behavioural states estimated from a switching state-space model (SSSM) fit to the KF data to those from models fit to the classical LS algorithm. Whale tracks were analysed with different temporal resolutions to test whether and how the quality of tracking data affected the similarity of the output from SSSMs fit to LS and KF data.

MATERIAL AND METHODS

Ethics Statement

All seal handling and tagging procedures were carried out under license number 60/4009 issued by the UK Home Office under the Animals (Scientific Procedures) Act 1986. Fieldwork and whale tagging in the Azores were developed under research permits by the administrative authorities of the Autonomous Region of the Azores. All procedures in whales followed the guidelines of the American Society of Mammalogists [23].

Data collection and processing

In the interest of clarity the following terminology will be used throughout this work: i) LS locations/data and KF locations/data refer to the locations/data provided by Argos that were derived from the application of the LS and KF algorithms, respectively; and ii) LS or KF model refer to the state-space models fit to data derived from the application of either the LS or KF algorithm. As explained below, the same models were fit to LS and KF datasets.

Harbour seal data

GPS/Argos tags were deployed on harbour seals in the Eden Estuary, south-east Scotland and around Eday, Orkney between May and July 2012. Animals were caught on or close to haul-out sites using hand, seine or tangle nets and subsequently anesthetised with Zoletil® as detailed in Sharples et al. [24]. Tags were attached to the fur at the back of the neck using Loctite® 422 Instant Adhesive. Tag duration ranged from 25 to 65 days (median 41 days).

The Fastloc GPS data used in this study were transmitted via the Argos system, providing high resolution at sea locations. The Argos transmissions also generated a concurrent series of standard Argos locations. In order to conduct this study, messages from the satellite transmitters were processed by the Argos service provider (CLS, Ramonville Saint-Agne, France) using both the LS and KF processing algorithms.

Fastloc GPS positions are more accurate and precise than Argos locations and in the present study were assumed to represent the seals' "true" position. However, GPS accuracy is known to decrease when Fastloc calculations are based on fewer satellites [25,26] and when residual error is high [27]. GPS data were therefore cleaned according to the Sea Mammal Research Unit protocol where locations estimated with <5 satellites and with residual errors=0 or >25 were removed [27]. Tests on land showed that over 95% of the cleaned locations had an error of <50 m [27].

As central-place foragers harbour seals haul-out on land between foraging trips. Thus, it was necessary to remove haul-out locations from the data before fitting any models. Although the GPS/Argos tags have a wet/dry sensor which records haul-out events, only a subset of these records are received via the Argos system. These animals often range in near shore waters and the large measurement error in Argos observations means such observations could not be used to define whether a location fell on land. Thus, the Fastloc GPS positions were used to define the precise time seals departed and returned to land. Positions within 200 m from all shorelines were also considered as haul-out to buffer against errors in GPS positions and because harbour seals haul-out on intertidal sandbanks. This procedure may have excluded valid parts of a few foraging trips but this shouldn't affect algorithm comparison in anyway. Consecutive at-sea locations between haul-out events thus formed an individual foraging trip. A series of trips were defined within each seal GPS track and, for each trip, all LS and KF locations obtained between 5 minutes prior to and 5 minutes after the trip were selected. Only trips with ≥30 LS and KF locations were subsequently used for model fitting. The seal dataset analysed in the next sections consisted of 1174 GPS, 1339 Argos LS and 2083 Argos KF positions obtained during 31 foraging trips of 7 different seals (Table A-II.3; Appendix II).

Fin whale data

The data consisted of Argos-derived surface positions obtained from PTTs (model SPOT5-implantable, Wildlife Computers, Redmond, Washington, USA) attached to the flanks of 6 fin whales. Whales were tagged off Faial and Pico islands (38ºN 28ºW), Archipelago of the Azores (Portugal), in September 2009, April and May 2010. All tags were programmed to transmit on a daily basis, every hour of the day up to a maximum of 500 messages per day. Details about the tagging methodology, movements and inferred behaviours of these whales are described in Silva et al.[28]; here the focus is on the analyses of model fitting and performance. Like in the case of the harbour seal data, location data was requested to be processed with both the LS and KF algorithms.

The KF algorithm consistently yielded more positions per individual whale than the LS algorithm (Table A-II.4; Appendix II). To compare the regular, estimated locations from the LS model with those from the KF model for each whale dataset, only the positions from the KF data that were within 2 minutes of a LS position were selected (hereafter called the KF reduced dataset). A second model was fitted to all KF locations to investigate how the tracks from a model fitted to the full KF dataset compared to those from a LS model.

State-space models

State-space models couple two stochastic models: a process model (transition equation) that estimates the current state (e.g. location and behavioural state) of an animal given its previous state, and an observation model that relates the unobserved location states estimated by the process model to the observed data (locations obtained from Argos).

The SSM described in Jonsen et al. [8] uses a first-difference correlated random walk (DCRW) as the process model to describe movement dynamics. The SSSM also uses a DCRW as the process model but allows movement parameters to change between two discrete behavioural states – for example, transiting versus area-restricted search (ARS; [29]) – by including a different DCRW model for each [9].

Model fit to harbour seal data

Initially an attempt to fit a SSSM to the harbour seal data was made but the same problems noted by Breed et al. [20] using simulated tracks were encountered. These authors showed that when the scale of movement is small relative to observation error and frequency, the models are unable to accurately estimate location and behavioural states. Even though the temporal resolution of the seal data was reasonably high (Table A-II.3; Appendix II), the SSSM provided a poor fit, resulting in unreliable location and behavioural estimates, irrespective of the algorithm used (although models fitted to KF data behaved slightly better). It is possible that movements of harbour seals are best analysed with different models (e.g. [30]) but this evaluation is beyond the scope of this paper.

Thus a SSM [8] was fit to the harbour seals' satellite locations derived from the LS and KF algorithms. The SSM was fit as a single hierarchical model (hSSM) [5] to all trips of all seals simultaneously, as this significantly improved parameter estimation, especially for data-sparse trips.

By letting k index each individual harbour seal trip, the transition equation of the SSM formulated within a hierarchical framework becomes:

$$\mathrm{d}_{t,k} \sim N_2\left(\gamma_{t,k}\mathrm{T}\left(\theta_{t,k}\right)\mathrm{d}_{t-1,\mathrm{k}}, \Sigma\right)$$

where $\mathbf{d}_{t\text{-}1}$ is the displacement between unobserved locations $\mathbf{x}_{t\text{-}1}$ and $\mathbf{x}_{t\text{-}2,}$ and $\mathbf{d}_t$ is the displacement between unobserved locations $\mathbf{x}_t$ and $\mathbf{x}_{t\text{-}1}$. $\mathbf{T}(\theta)$ is a transition matrix that provides the rotation required to move from $\mathbf{d}_{t\text{-}1}$ to $\mathbf{d}_t$, where θ is the mean turning angle. γ is the move persistence coefficient (i.e. combined autocorrelation in direction and speed). N_2 is a bivariate Gaussian distribution with covariance matrix $\sum$ and represents the randomness in animal movement.

The observation equation accounts for the irregularity and variable errors in the observed Argos locations. Errors in latitude and longitude are modelled with a t-distribution using independent parameter estimates derived for each Argos location class [8,31]. The same observation equation was fitted to data processed with LS and KF algorithms. Further details about the SSM are provided in Jonsen et al. [5,8].

Model fit to fin whale data

The Bayesian switching state-space model (SSSM) described in Jonsen et al. [9] was fitted to the Argos satellite-based location estimates of fin whales derived from the LS and KF algorithms. The transition equation for the SSSM is similar to that of a SSM:

$$\mathrm{d}_t \sim N_2\left(\gamma_{b_t}\mathrm{T}\left(\theta_{b_t}\right)\mathrm{d}_{t-1}, \Sigma\right)$$

but in this case the movement parameters θ and γ are indexed by behavioural state b. At each displacement t, the estimated behavioural state b corresponds to the set of parameters θ and γ that provide the best model fit.

The observation equation used to model the irregularly observed LS and KF fin whale locations was that same used for the SSM.

Model implementation

Models were fit using R (R Development Core Team 2012) code provided in the supplement to Jonsen et al. [5]. The code implements hSSM and SSSM using Markov Chain Monte Carlo (MCMC) methods using the program Just Another Gibbs Sampler (JAGS).

The hSSM was fitted separately to the harbour seals' location data (excluding Z class locations) obtained from each algorithm using a time step of 2 hours, corresponding to the average temporal resolution of the LS data. For the hSSM fit to the KF and LS satellite datasets, two MCMC chains for 60.000 iterations were run, dropping the first 50.000 samples as a burn-in and retaining every 10th sample from the remaining 10.000 assumed post-converge samples from each chain to reduced sample autocorrelation. Thus, model parameters and estimates of seals' locations were calculated using a total of 2.000 MCMC samples.

The SSSM was fitted separately to the fin whales' data obtained from each algorithm (after removing Z class positions from both datasets) using a time step of 3 hours, corresponding to the average temporal resolution of the LS data. For each SSSM two MCMC chains for 45.000 iterations were run, discarding the first 40.000 samples and retaining every 5th from the remaining 5.000 samples from each chain. A total of 2.000 MCMC samples were used to calculate model parameters and estimates of whales' locations and behaviours.

hSSM and SSSM convergence and sample autocorrelation were assessed by visually inspecting trace and autocorrelation plots and using the Gelman and Rubin scale reduction factor (R-hat) diagnostic available in R package boa.

Data analysis

The Argos locations per seal trip greatly exceeded those of Fastloc GPS, and the latter were also more irregular in time (Table A-II.3; Appendix II). Therefore, in order to estimate the accuracy of locations predicted by LS and KF models, only those locations within 30 min of a GPS position were selected at first. Then the "true" position of the seal at the time of those modelled locations was estimated by linear interpolation between two consecutive GPS positions [32]. Finally, the linear error and absolute latitudinal and longitudinal errors between each modelled location and the corresponding interpolated GPS position were calculated.

To investigate if and how the quality of Argos telemetry data affects spatial accuracy of LS and KF models, location errors from seal trips with different temporal resolutions, spatial precisions and frequency of observations were compared. Linear mixed-effects models were used, with seal and individual trip as random effects to account for behavioural differences among seals and unequal sample sizes across trips. Errors were log transformed to ensure linearity with continuous predictors. Algorithm (LS vs. KF) was included in the model as a categorical predictor and continuous predictors were number of Argos locations used to fit the model, average length of time between locations (hereafter time step), and proportion of positions of LC 0, A and B (hereafter LC 0-B). Values of these continuous predictors for each seal trip are given in Table A-II.3 of Appendix II. A model with interactions between algorithm and all continuous predictors was fitted to allow investigating if the effect of data quality was consistent among the LS and KF models.

In the case of the SSSM fit to the whale data, it was only possible to determine how well the KF models performed in relation to models fit to the LS algorithm. For each whale, the medians, inter-quartiles and 95% credible limits (95%CL) of parameter estimates of LS and KF models were compared. The longitudinal and latitudinal differences between pairs of location estimates from the LS and KF models for each whale were also calculated. For each location predicted by the LS model a probability ellipse was estimated, determined by the 95%CL obtained from the model. The proportion of location estimates from the reduced KF model that fell within the 95% probability ellipse of the corresponding LS position was then estimated.

To understand if the KF algorithm introduced significant changes in the ability of the SSSM to resolve behavioural state, the percentage of agreement in behavioural classification between the LS and KF models was calculated. Whale behaviour at each 3-h location was inferred from the output of the SSSM. Because behaviour is treated as a binary variable, MCMC samples can only assume the values 1 (inferred as transiting) or 2 (inferred as ARS), *b* at each location was estimated as the mean value of the MCMC samples. The same cut off points as Jonsen et al. [9] were used: locations with mean estimates of *b*<1.25 were assumed to represent transiting, *b*>1.75 ARS, and between these values were considered “uncertain”.

Finally, the whale tracks from a model fitted to the full KF dataset were compared to those from the models applied to LS data. The SSSM was fitted to the full KF data using the same time step as above. For each whale the distance (in km) from locations estimated by the full KF model to the track estimated by the LS model were calculated. Only data from days when both methods delivered satellite locations were used in that comparison.

Means are presented ± standard deviation (SD) throughout. All distances were calculated using a great-circle route. Statistical analyses were performed in R software using packages *nlme* and *MASS*.

RESULTS

Accuracy and precision of LS and KF models fit to harbour seal data

The KF algorithm provided 2083 locations, 1.5 times more than the LS algorithm and 1.8 times more than the GPS transmitted via Argos (Table A-II.3; Appendix II). The increase in the number of locations per trip in relation to the LS data ranged from 12 to 137% with an average of 56%. A total of 368 LS and 375 KF model locations were within 30 min of a GPS position and were used to compare spatial accuracy of locations derived from each hSSM (Table 4-1).

Errors in locations estimated from LS and KF models showed the same elliptical distribution in relation to interpolated GPS positions, with a clear directional bias in the longitudinal error component (Figure 4-1). Average longitudinal errors ranged between -0.20º–0.20º (mean=-0.003) for LS models and between -0.36º–0.17º (mean=0.001) for KF models. Latitudinal errors ranged between -0.10º–0.08º (mean=-0.002) for LS models and between -0.10º– 0.09º (mean=0.001) for KF models. Overall, the mean distance of KF model locations to interpolated GPS positions was lower (2.9±2.9km) than that of LS model locations (3.5±3.0 km) (Table 4-1). About 31% of all locations predicted from the KF model were within 1km from the interpolated GPS position and 82% were less than 5 km. For locations predicted from the LS model, 24% and 73% were respectively within 1 km and 5 km from the corresponding interpolated GPS position. The KF model produced smaller mean errors for 27out of 31 trips (Table 4-1).

Predicted trips from the KF model were 27% (range: 1–57%) more accurate than trips derived from the LS model. However, standard deviations of KF errors were sometimes higher suggesting that location accuracy varied considerably within the same trip (Table 4-1). Average errors of trips increased as the average distance between locations (step length) increased. For trips with an average step length ≥6 km, the average error of KF modelled locations was 4.6±0.4 km, and of LS modelled locations was 5.9±1.2 km.

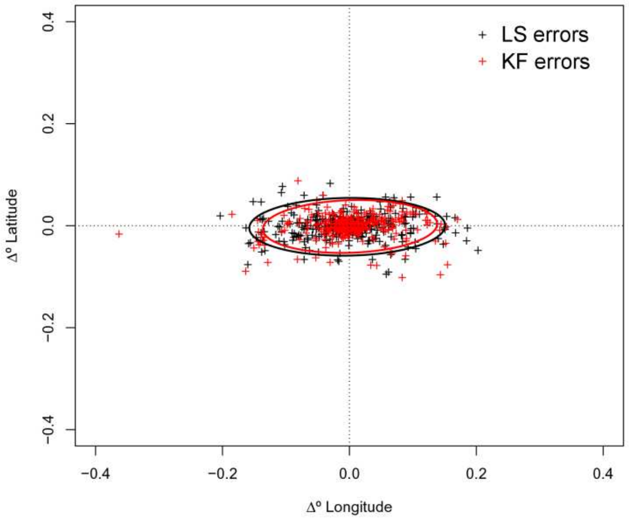

Figure 4-1 Errors in locations estimated from LS and KF models. Errors in harbour seal locations estimated from state-space models fit to Least Squares (LS) (black) and Kalman filtered (KF) (red) data are plotted as offsets from "true" GPS positions. Standard ellipses were fitted to 95% of LS (black line) and KF (red line) error points.

Two representative tracks of foraging trips reconstructed using GPS positions, and LS and KF modelled locations are shown in Figure 4-2. In general, modelled tracks closely matched the GPS tracks, especially during periods of directed movement. Yet, tracks predicted by the LS model occasionally diverged greatly from the GPS track and tended to extend over a wider area in periods of torturous movements.

Table 4-1: Errors in locations estimated from models fit to Least Squares (LS) and Kalman filtered (KF) data for all harbour seal trips.

Seal	Trip	LS model			KF model			Variation in mean errors (%)**
		N*	Mean error (km)	Error range (km)	N*	Mean error (km)	Error range (km)	
1545	11	8	2.6±1.7	0.4–5.1	8	1.7±1.4	0.1–3.9	-34
	21	9	1.9±2.0	0.1–5.6	8	0.8±0.4	0.4–1.5	-57
	23	5	2.2±1.3	1.2–4.3	6	0.9±0.3	0.6–1.4	-56
	27	14	1.9±2.3	0.1–8.6	15	2.2±2.4	0.1–7.9	16
	28	13	1.5±2.1	0.1–7.6	13	1.5±1.5	0.1–5.1	-2
28503	11	12	2.2±1.9	0.2–6.9	14	2.0±2.3	0.2–8.5	-10
	18	12	3.9±2.4	1.7–9.4	11	3.1±2.7	0.3–9.3	-23
	19	14	4.5±3.1	0.2–10.5	14	4.4±3.4	0.8–13.9	-3
	23	8	4.8±1.8	1.6–6.8	8	4.2±2.9	0.9–8.9	-13
	42	5	2.9±1.7	1.0–5.5	6	2.4±1.1	1.1–3.8	-19
43844	4	15	2.5±2.6	0.1–10.0	14	1.6±1.6	0.1–5.7	-36
	8	8	1.7±1.5	0.2–4.6	8	1.2±0.6	0.2–1.9	-29
	14	10	1.7±2.4	0.1–8.0	12	1.4±1.7	0.1–6.4	-23
	16	13	1.5±1.5	0.2–6.3	11	1.0±0.7	0.2–2.3	-32
	22	14	4.9±3.7	0.9–13.6	14	2.5±2.7	0.2–9.0	-48
43871	7	24	3.6±3.5	0.3–12.7	23	3.4±3.4	0.3–11.7	-5
	8	17	5.9±3.6	0.8–11.8	17	4.3±3.2	0.7–10.1	-27
	13	21	4.9±3.7	0.3–10.8	20	3.0±2.5	0.1–9.4	-39
	19	20	5.9±2.9	0.6–11.5	20	4.5±2.8	0.2–10.6	-24
120346	24	13	1.9±1.4	0.5–5.6	13	1.2±0.7	0.3–2.6	-37
	25	15	2.2±1.7	0.4–5.4	15	1.6±1.3	0.2–4.1	-31
	26	5	4.3±2.9	1.3–8.4	5	3.0±3.2	0.9–8.7	-30
	30	13	2.1±1.4	0.4–4.9	13	1.1±1.0	0.1–3.2	-48
	32	7	1.7±1.4	0.6–4.1	7	1.7±1.9	0.2–5.9	-1
120349	3	8	4.2±2.3	1.2–7.1	8	4.5±3.5	0.6–10.8	6
	4	8	3.5±2.8	0.4–7.9	8	4.2±3.9	0.9–12.2	20
	5	5	3.7±1.7	2.0–6.1	5	1.9±1.7	0.8–5.0	-48
	6	10	7.7±3.5	1.1–11.1	10	5.0±3.0	0.5–8.8	-35
120350	3	13	3.7±3.1	0.5–9.4	13	3.1±2.3	0.1–3.9	-16
	4	14	3.2±3.1	0.3–12.5	13	5.0±4.5	0.4–1.5	56
	5	15	4.2±2.4	0.7–7.5	23	4.9±4.3	0.6–1.4	17
Total		**368**	**3.5±3.0**		**375**	**2.9±2.9**		

*N: Number of locations used to calculate errors in locations estimated from LS and KF models.**Variation in mean errors was calculated for each trip as the difference in the mean error estimated for KF and LS models divided by the mean error of the LS model.

Uncertainty in KF model estimates, as indicated by the width of the 95%CL (measured in km), was significantly lower than that of LS model estimates (KF model: 5.6±5.6 km; LS model: 11.6±8.4 km; t-test=-11.41, df=741, *P*<0.001).

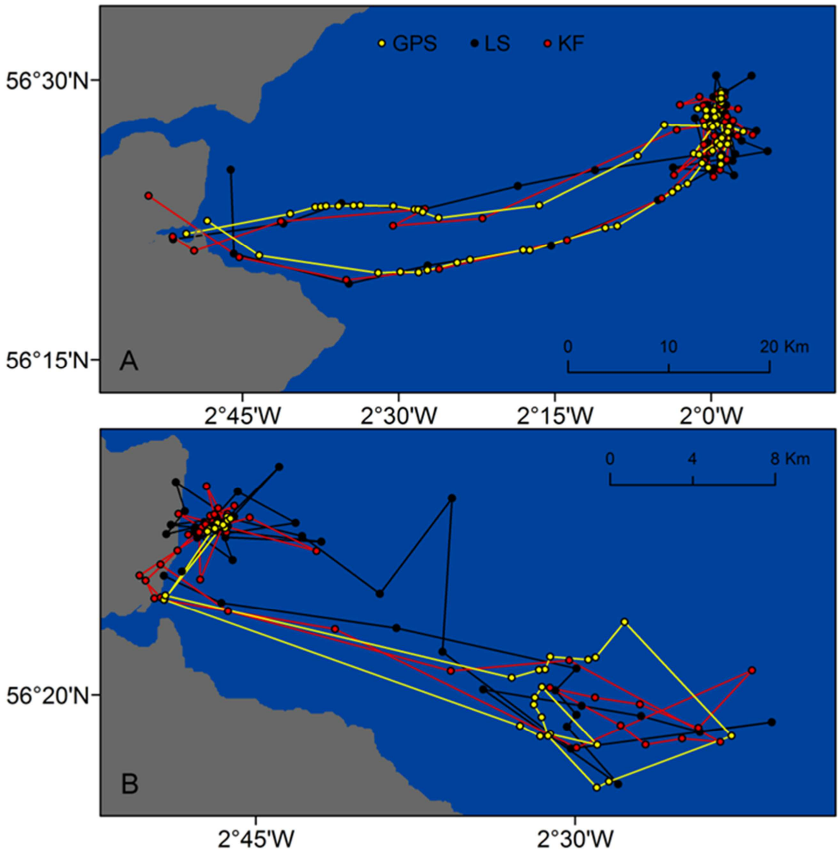

Figure 4-2 Harbour seal tracks obtained from GPS (yellow), LS (black) and KF modelled (red) locations. Estimated locations (circles) and tracks (lines) of harbour seals obtained from fitting state-space models to Least Squares (LS) (black) and Kalman filtered (KF) (red) data, in relation to the “true” GPS positions and track (yellow). A. Example of a trip with higher quality of Argos data: trip 7 of harbour seal #43871. B. Example of a trip with lower quality of Argos data: trip 22 of harbour seal #43844.

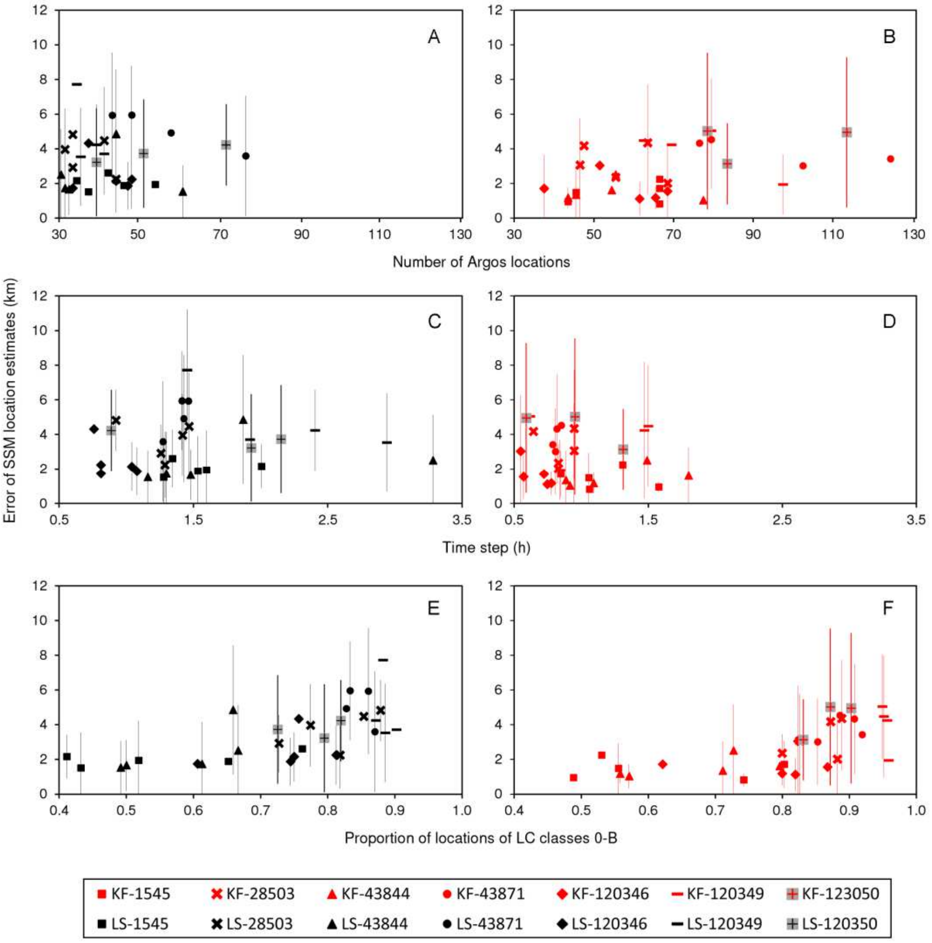

Figure 4-3. Trip-averaged error in locations estimated from LS and KF models relative to Argos data quality. Relationship between mean errors (±SD shown as vertical bars) in locations estimated from state-space models fit to Least Squares (LS) (black) and Kalman filtered (KF) (red) data per harbour seal trip and quality of Argos telemetry data used to fit the models: A-B. Number of locations. C-D. Time step (h) between locations. E-F. Proportion of locations of LC 0-B. Different trips from the same seal have the same symbol.

Effect of data quality on accuracy of LS and KF models fit to harbour seal data

Observation frequency, temporal resolution and spatial precision of Argos data used to fit the SSMs varied among seals and trips and between the LS and KF models (Figure 4-3, Table 4-1). Expectedly, the increase in number of locations that resulted from the application of the KF algorithm improved the temporal resolution of the KF data for all trips. However, it also increased the proportion of locations of lower spatial precision (Argos LC 0, A and B) in each trip. With few exceptions, trips from the same seal tended to have similar number of locations, time steps, and proportion of LC 0-B, suggesting an individual effect in the quality of Argos data. This could be due to tag (e.g. battery power), instrumentation (e.g. tag placement) or behavioural-specific (e.g. surface behaviour and diving time) differences among seals or to a combination of all these factors.

Mean errors (±SD) of LS and KF modelled trips were plotted in relation to the Argos quality parameters described above (Figure 4-3). Accuracy of modelled trips did not seem to improve with the observation frequency or temporal resolution of Argos data, but mean errors (and respective SD) in LS and KF estimated locations appeared to increase with increasing proportions of locations LC 0-B.

A linear mixed-effects model was fitted to examine the effects of type of algorithm and of Argos quality parameters (spatial precision, observation frequency and time step) on estimated errors of modelled locations. The interactions between algorithm and the continuous predictors were the first to be dropped from the linear mixed-effects model based on AIC results, suggesting that quality of Argos data influenced the accuracy of LS and KF models in a similar way. The best fitting model indicated that observation frequency and time step of Argos data had no effect on the errors of locations estimated from the models, and only algorithm and proportion of locations of LC 0-B were significant (Table A-II.5; Appendix II). Contrary to what was expected, there was little variability among different seals in addition to the trip-to-trip variability and both the AIC and the likelihood ratio test indicated that individual seal could be dropped from the model (L=3.95×10^{-7}, P=0.499), leaving trip as the only random effect. The best fitting model predicted larger errors for locations estimated from LS models compared to locations from KF models (Figure 4-4, Table A-II.5; Appendix II). On average, LS models will estimate locations that are 1.6 km farther from the true seal position relative to KF locations. Also, errors (on a logarithmic scale) are expected to increase as proportion of Argos locations with lower precision increases, and this relationship was similar for LS and KF models (Figure 4-4, Table A-II.5; Appendix II).

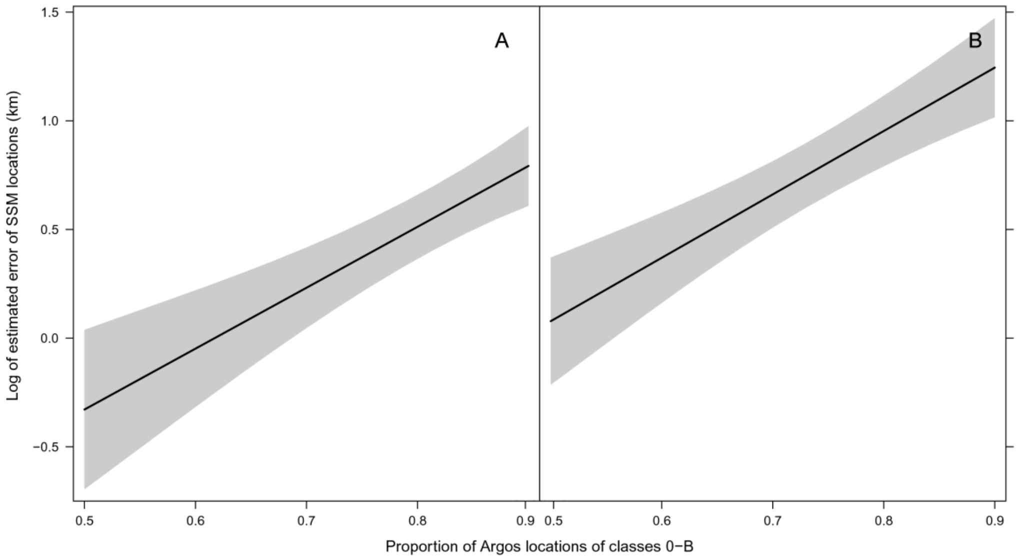

Figure 4-4. Predicted error in locations estimated from LS and KF models. Predicted error in harbour seal locations according to the best fitting linear mixed-effects model for A. State-space models fit to Kalman filtered (KF) data. B. State-space models fit to Least Squares (LS) data.

Comparison of LS and KF models fit to fin whale data

Medians and 95%CL of estimated model parameters of the reduced dataset were similar across whales and between the LS and KF algorithms. Both the LS and KF models distinguished well between the two behavioural modes (transiting and ARS), as indicated by the parameter estimates that aggregated into two non-overlapping groups.

The estimated locations inferred from the KF model applied to the reduced dataset differed little from the locations output by the LS model. Differences in latitude and longitude between paired KF-LS locations were centred around zero but the latter showed a wider range of values (range for latitude: -1.1–0.7º; range for longitude: -1.2–2.0º) (Figure 4-5). Differences in paired KF-LS locations were considerably higher for whale #80716. Removing data from this whale resulted in a considerable reduction in the range of latitudinal (-0.1–0.2º) and longitudinal (-0.5–0.4º) distances between KF and LS locations. Differences in latitude and longitude between paired locations showed no obvious trend with latitude, longitude, date, number of positions per track, or behavioural mode (not shown).

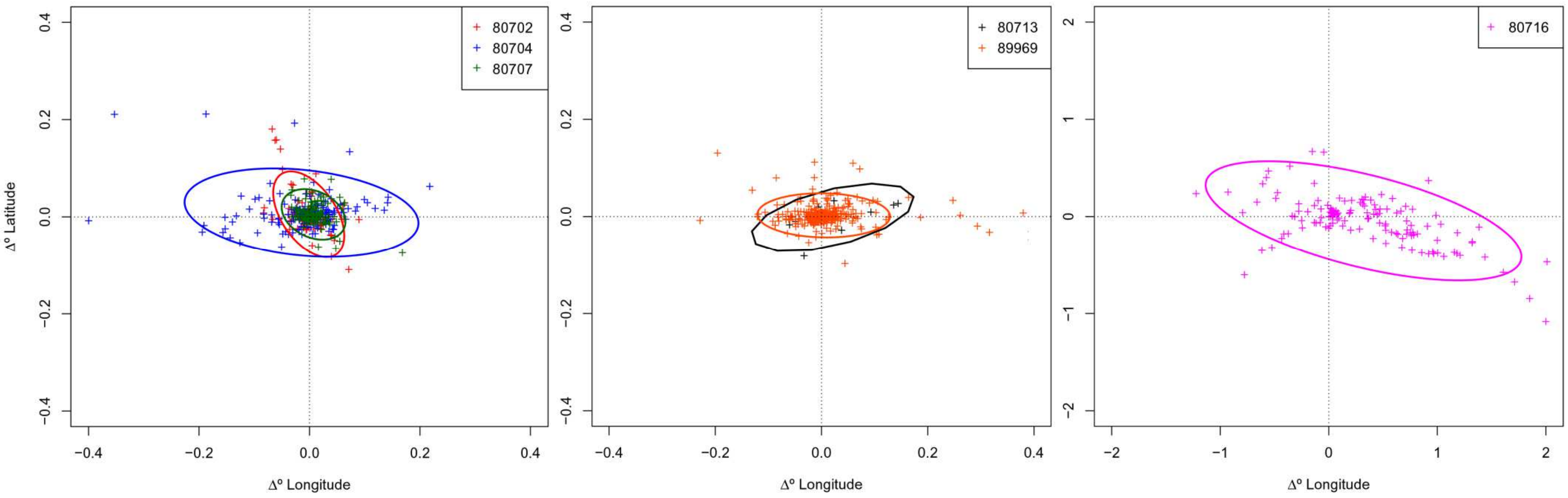

Figure 4-5. Differences in locations estimated from KF and LS models for all fin whales. Differences in locations estimated from switching state-space models fit to Kalman filtered (KF) (red dots) data are plotted as offsets from locations calculated from the same models fit to Least Squares (LS) data. Standard ellipses were fitted to 95% of KF data points. A. Fin whales #80702 (red), #80704 (blue) and #80707 (green). B. Fin whales #80713 (black), #89969 (orange). C. Fin whale #80716 (pink).

The proportion of estimated locations from the SSSM applied to the reduced KF data that fell within the 95% probability ellipse of locations inferred by the LS model varied between whales but was very high, ranging from 69 to 100% (mean=88%). Differences in the width (measured in km) of the 95%CL between pairs of locations estimated from the model fit to the reduced KF data and the LS data were also compared. For five whales, the reduced KF model resulted in lower average widths of 95%CL (paired t-test: $P<0.05$ for all whales), although differences were generally small (mean difference: -2.2±3.9 km). For whale #80716, however, the 95%CL of the reduced KF model were significantly wider than that of LS data (paired t-test: $t=-11.15$, $P<0.001$; mean difference: 76.2±80.3 km).

In 94% of the cases, the behavioural mode inferred by the KF model matched the classification from the model fit to the LS data (Table 4-2). Agreement was highest for locations inferred as transiting (98%), followed by ARS (93%). Changes in behavioural classification between the two models were from transiting or ARS to “uncertain” and vice-versa, but never from transiting to ARS or vice-versa.

Table 4-2 Agreement between fin whale behavioural modes inferred by the models fit to Least Squares (LS) and Kalman filtered (KF) data.

		KF model		
		Transit	ARS*	uncertain
LS model	Transit	353	0	6
	ARS	0	524	40
	uncertain	5	6	83

The matrix shows the number of fin whale locations classified in each behavioural mode by the LS model that were assigned to each of the behavioural modes by the KF model. *ARS: Area-restricted search.

As expected, the KF processing algorithm yielded more positions and improved the temporal resolution of the 6 whale tracks. The increase in number of locations per track ranged from 18 to 272% with an average of 75%. The average number of daily locations per whale track varied between 6.0–38.6 for the full KF data, compared to 1.6–30.8 for the LS data (Table A-II.4; Appendix II). There was also an increase in track duration (3 and 11 days) for two whales but this came at the expense of a few gaps (maximum of 3 days) in those tracks (Table A-II.4; Appendix II). In contrast, the KF algorithm provided several positions within a 5-day gap in the LS tracking data of whale #80716.

The width of the 95%CL of locations estimated by the full KF model (47.3±76.9 km) was significantly lower than the width of 95%CL of locations estimated from the LS model (57.2±113.0 km) (t=2.38, *P*=0.017). Still, locations from the full KF model fitted well the paths inferred from the LS data, except when gaps in the LS data exceeded 1 day (Figure 4-6, Figure 4-7). Combining data from all whales, over 49% of locations estimated by the full KF model were <1 km away from the tracks derived from the LS model and 77% were <5 km.

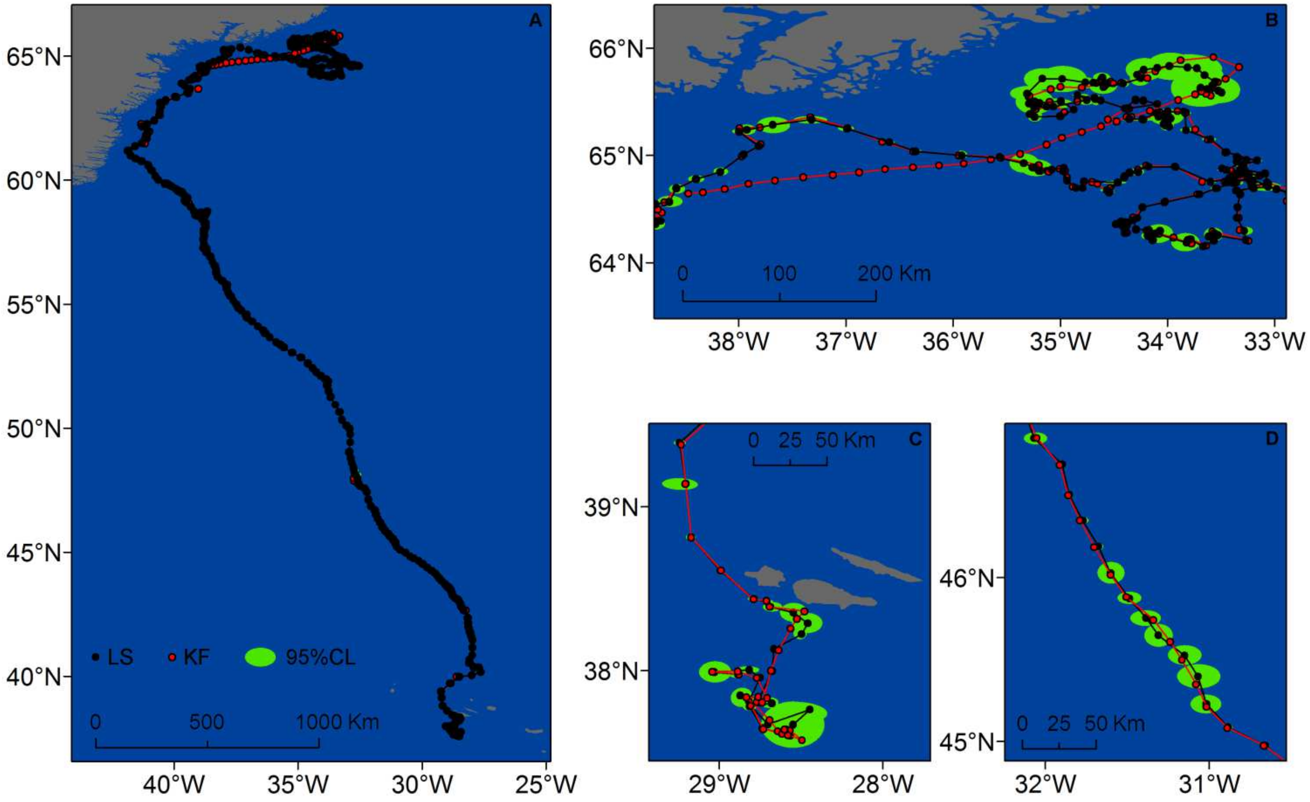

Figure 4-6. Fin whale #89969 tracks obtained from LS (black) and KF modelled (red) locations. Estimated locations (circles) and tracks (lines) of fin whale #89969 obtained from fitting a switching state-space model to Least Squares (LS) (black) and the full Kalman filtered (KF) (red) data. The 95% probability ellipses of locations derived from the LS-based model are shown in green. A. Complete tracks showing the increase in track length resulting from the application of the KF algorithm (red). B, C, D. Detail of the tracks showing the majority of KF locations within the 95% probability ellipses of LS locations.

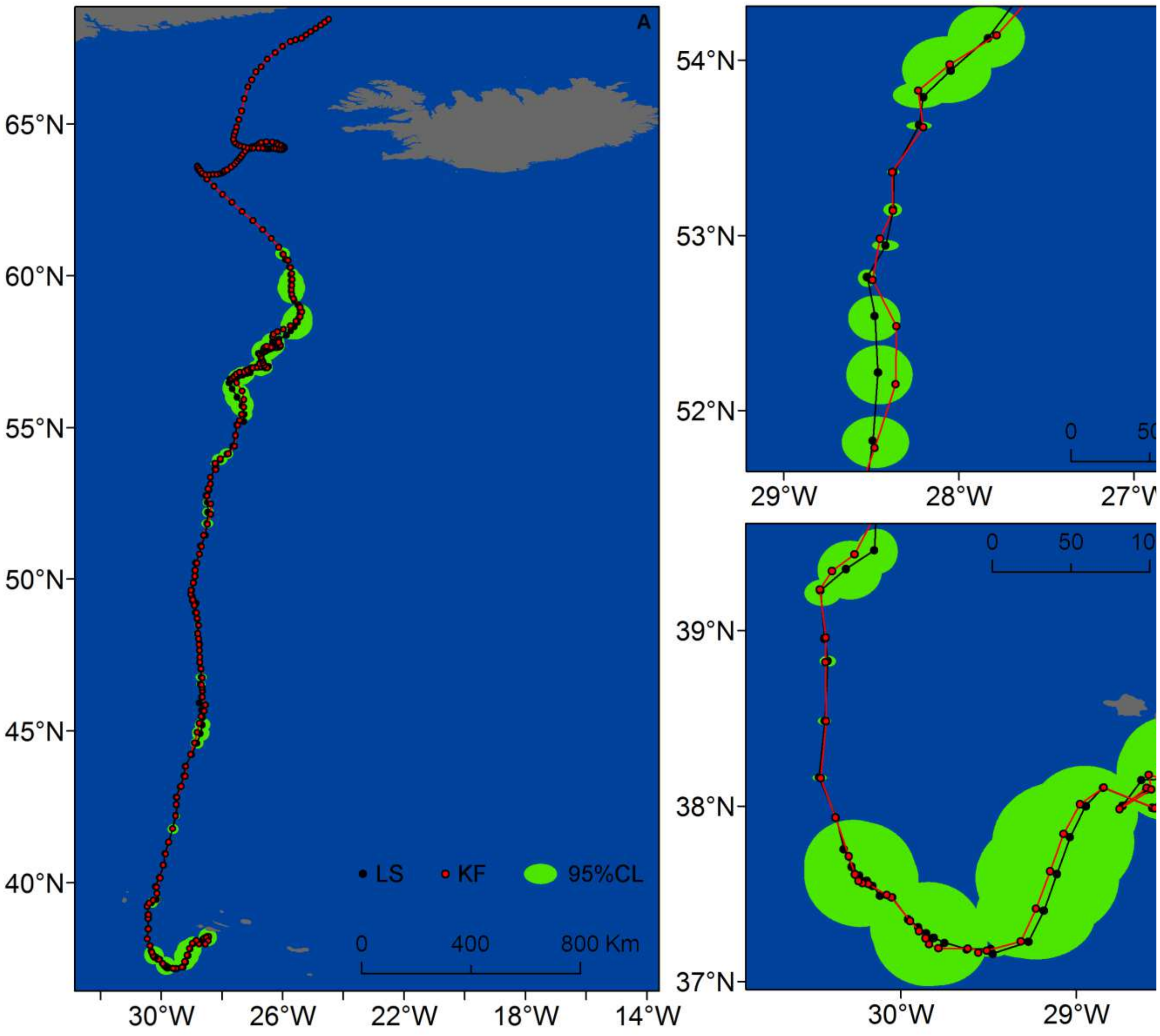

Figure 4-7. Fin whale #80704 tracks obtained from LS (black) and KF modelled (red) locations. Estimated locations (circles) and tracks (lines) of fin whale #89969 obtained from fitting a switching state-space model to Least Squares (LS) (black) and full Kalman filtered (KF) (red) data. The 95% probability ellipses of locations derived from the LS-based model are shown in green. A. Complete tracks showing the increase in track length resulting from the application of the KF algorithm (red). B, C. Detail of the tracks showing the majority of KF locations within the 95% probability ellipses of LS locations.

DISCUSSION AND CONCLUSIONS

Since the recent introduction of the Kalman filtering (KF) algorithm for the processing of satellite tracking data by the Argos system, the service providers have made this the default processing method for new transmitters (PTTs), giving the user the option to choose the Least Squares (LS) algorithm in alternative. The data processing of old PTTs that were already being processed with the LS algorithm remains unchanged, unless KF processing is requested, and stored data from 2008 onwards can be reprocessed using either method (albeit with additional processing costs). Processing of data with the new KF algorithm is bound to become more common as old PTTs end their life, and data processed with this algorithm will soon become the standard for Argos-based tracking.

State-space modelling approaches provide the statistical rigor needed in analysing animal movement data, but SSMs are not simple and require considerable care in their use [5]. Understanding the implications of using data processed with the new KF algorithm is essential when interpreting modelling results. This is the first time that performance of SSMs applied to KF tracking data has been directly validated with known locations of free-ranging animals. This was achieved by fitting the same model to Argos satellite locations obtained on 7 harbour seals processed with LS and KF algorithms and by comparing locations derived from each model against the "true" interpolated positions of the seals obtained by Fastloc GPS technology. In addition, the results of fitting the Bayesian switching state-space model (SSSM) to KF data were compared to those of LS models, using tracking data from 6 fin whales. Although in the latter case it not being possible to assess the accuracy of model-derived locations, it enabled evaluating how SSSMs fit to KF data performed in relation to SSSMs fit to data processed with the LS algorithm, which until recently was the standard processing algorithm used to deliver satellite locations.

This study shows that Kalman filtering consistently provided more estimated locations per animal track than the LS algorithm, supporting previous claims by the Argos service [19]. The increment in estimated locations was substantial for both species (fin whales: 75%; harbour seals: 56%). Compared to the findings of the present work, Boyd and Brightsmith [33] reported only a modest 28% increase in locations computed with the KF algorithm. However, their estimate is based on data obtained from static platforms, while the estimates in the present work and those from Argos come from free-ranging tagged animals. Stationary land tests are closer to the "ideal" conditions for satellite communications and are unlikely to adequately represent most of the problems known to affect the transmission of signals from satellite tags and/or the reception of messages at Argos satellites, especially for marine taxa. Understandably, the potential benefit of the KF method should be higher under circumstances (e.g. areas with limited satellite coverage) and for species more prone to transmission difficulties, and for which the frequency of uplinks is usually low. Not surprisingly, the major increase in estimated locations was for fin whales that typically have shorter surface intervals than harbour seals, and can be more adversely affected by wave wash due to improper antenna orientation and poor environmental conditions.

Like Boyd and Brightsmith [33], this study also found that the majority of additional locations in KF data came from fixes with only 1 message (Argos LC B) (fin whales: 29%; harbour seals: 33%) with a very slight increase in the proportion of locations with 4 or more messages (LC 2 and 3) observed only for fin whales. If, as a result of KF processing, tracks acquire a disproportionate number of locations with low spatial precision, this may impact the analysis and interpretation of animal movement data, particularly when this analysis is based on the raw satellite positions and doesn't take into account variability in measurement errors. Implications could be even more severe if the gain in 1-message LC B locations is not homogeneously distributed along the track and depends, for instance, on the geographic location or behaviour of the animals, therefore being more prevalent in certain areas or during specific activities occurring in preferred habitats.

The results of this work demonstrate that the Jonsen et al. [8] SSM provided a good fit to the data processed with the KF dataset, despite the potentially increased autocorrelation in the location errors imposed by the KF algorithm. The greater spatial accuracy and precision of locations estimated from the KF model compared to those from the LS model was likely due to a combination of increased accuracy in KF-estimated locations and the higher temporal resolution of the KF data.

Although the overall difference in mean errors between the two algorithms appeared small (mean error in LS models was 3.5±3.0 compared to 2.9±2.9 in KF model) the model fit to KF data improved the accuracy of seal trips by 27% over the LS model. The linear mixed-effects model indicated that, despite significant variations in trip accuracy, errors in locations predicted for LS trips were significantly larger than those predicted for KF trips. For both models the largest deviances from true locations occurred along the east/west axis. This is not unexpected since Argos location errors are strongly biased towards the longitudinal component, regardless of the processing algorithm [1,3,31-33], and the SSM does not explicitly account for this directional bias. However, there was no evidence of the non-uniform distribution of extreme errors documented in other studies [3] suggesting that the model was able to handle this problem.

Tracks reconstructed from the models applied to KF and LS data provided faithful representations of the true seal trajectories measured with Fastloc GPS. However, the LS track tended to deviate more from the true track when seals were making short displacements and frequently changing direction. This is likely due to the correlated random walk model employed in the KF algorithm which would tend to smooth out uncommonly large changes in direction and/or displacement. As a result, LS locations tended to spread over a wider area compared to the KF. This was a common feature to several LS modelled tracks that can have major implications if these data are used to calculate sizes of home ranges or ARS patches.

The SSMs were fit as hierarchical models to the LS and KF data, meaning that data from all seal trips were combined to estimate model parameters, leading to improved location estimates. Probably larger errors would be obtained if models were fitted separately to each trip. Yet, there is no reason to expect that the hierarchical formulation behaved differently when applied to LS and KF data, so there is no reason to think that the comparison between algorithms is not valid.

The same observation equation was fitted to data processed with LS and KF methods, thus assuming that the new algorithm did not change substantially the distribution or magnitude of the errors. A recent study demonstrated that both LS and KF location errors are better described by a long-tailed lognormal distribution [33]. In the present work, errors were modelled with generalized t-distributions which are known to be robust to extreme values [8]. Boyd and Brightsmith [33] also compared mean errors in KF and LS processed locations showing these did not differ significantly for most location classes, except for LC 2, for which LS errors were about half the KF errors, and LC B, for which LS errors were nearly 4 times greater than KF errors. In contrast, Argos reported better accuracies with the KF method for locations computed with ≥4 (LC 2 and 3) and 2-3 messages (LC A and B) [19]. In any case, given the predominance of LC classes A and B in both datasets, it is likely that fitting the same observation equation to LS and KF data might have resulted in an overestimation of KF errors relative to LS errors, and not the other way around.

Regardless of which processing method is used, this study showed that accuracy of modelled tracks was sensitive to precision of the raw input data. As the proportion of locations with poor precision increased, the ability of the SSMs to recover accurate locations was significantly worse. This is consistent with findings from other researchers that showed that high measurement error not only impacts accuracy and precision of locations estimated from state-space methods [20,34] but can also affect our ability to discern behavioural patterns and quantify habitat use patterns [4,26,32,35,36].

On the other hand, there was no evidence that observation frequency and temporal resolution of Argos data influenced the magnitude of SSM errors, in contrast to a recent study that suggested that frequency and regularity of raw data may be as important as spatial precision for obtaining accurate estimates of locations from state-space methods [20]. There are two main reasons for the different results between the analysis performed here and that of Breed et al. [20]. First, Breed's analysis of model accuracy was based on a reduced number of simulated tracks to which were imposed different observation frequencies and temporal gaps spanning a much larger range than the number of Argos locations and time steps observed in the seal data (see Table A-II.3; Appendix II). Second, in Breed's study a separate SSM was fit to each simulated track while here a hierarchical approach was adopted. By combining information from all trips to estimate model parameters, potential effects of between-trip data quality likely were lessened and more accurate location estimates were obtained for all trips.

The results presented here strongly suggest that application of SSSM to the whale tracking data processed with the KF algorithm was appropriate and that models fitted well. Estimated parameters from KF models were very similar across all tracks and to parameters from the LS model despite the fact that models were fit separately to each whale LS/KF-processed dataset.

Paths inferred from both models were also similar, with most of the locations from the reduced KF model falling within the 95% probability ellipses of locations estimated from the LS model, and the majority of locations from the full KF model being close to the whale tracks inferred by the LS model. Similar to what was observed for the seal data, the longitudinal bias in Argos errors caused the reduced KF locations to differ more from their paired LS positions in the east/west than in the north/south axis.

The estimated precision of locations inferred from the SSSM fit to the reduced KF data was higher for 5 out of 6 whale tracks, as indicated by the lower average width of the credible limits. However, the KF model behaved significantly worse than the LS model in the case of the whale track (#80716) for which less than 2 satellite positions were received per day. This cannot be accounted for by variations in Argos location classes because 28 of 29 positions were assigned the same class in both datasets. A close inspection of the raw KF and LS data indicates that the poorer performance of the reduced KF model was likely associated with the highly tortuous whale path evident in the KF data (and not in the LS data) and caused by the way the data regularization approach used in the SSSM's observation model dealt with this tortuosity. Because the interval between raw satellite positions was considerably longer than the 3-hourly interval at which the SSSM positions were being estimated, raw positions have more weight on model estimates as the model "forces" derived locations to exactly match raw satellite positions. Such an effect tends to be more pronounced with decreasing linearity of the tracks [37], explaining why uncertainty in the model estimates was greater for the more sinuous KF path and the higher discrepancy in relation to the LS path.

It should be stressed that the application of the KF algorithm increased the total number of locations in this whale track from 29 to 108 (see Table A-II.4; Appendix II), resulting in a remarkable decrease in the uncertainty of SSSM location estimates (average 95%CL width: 86.0±69.5 km) when compared to the LS model. Differences in the remaining tracks were less pronounced but the KF processing algorithm produced an overall increase in number of locations obtained and a decrease in the uncertainty of SSSM estimates.

Estimates of behavioural mode from the KF model agreed well with inferences from the LS model – with 94% of whale locations being assigned the same behavioural category in both models – indicating that the KF algorithm did not introduce appreciable changes in the ability of the SSSM to recover latent behaviours from satellite positions.

These results lead us to conclude that application of widely-used Bayesian state-space models [5] to Argos satellite locations processed with a KF method is appropriate and, as was the case of the SSM fit to harbour seal data, can produce more reliable location estimates than when LS data are used to fit the same models. Also, behavioural modes could be equally well detected from SSSM fit to whale tracking data processed with KF and LS methods. Since the KF algorithm generally yields more positions and longer tracks, there may be clear advantages in using the KF model over the LS model. This is especially true in telemetry studies of species that spend prolonged periods underwater or under dense vegetation cover, for which the number of daily fixes is generally low, precluding examination of movement and behaviour of animals in more detail. However, as seen here, the KF algorithm can increase the number of positions of lower precision (LC B) by nearly 30%, which in turn can degrade accuracy of modelled tracks. Even with LC B positions estimated by the KF method being several times more accurate than LS locations of equal class [19,33], when accuracy and precision are critical for the analysis, researchers may consider removing 1-message positions before fitting state-space models.

REFERENCES

1. Douglas DC, Weinzierl R, C. Davidson S, Kays R, Wikelski M, et al. (2012) Moderating Argos location errors in animal tracking data. Methods in Ecology and Evolution 3: 999-1007.

2. CLS (2011) Argos User's Manual. http://www.argos-system.org/web/en/76-user-s-manual.php: CLS. 62 p.

3. Costa DP, Robinson PW, Arnould JPY, Harrison A-L, Simmons SE, et al. (2010) Accuracy of ARGOS locations of pinnipeds at-sea estimated using fastloc GPS. PLoS ONE 5: e8677.

4. Bradshaw CJA, Sims DW, Hays GC (2007) Measurement error causes scale-dependent threshold erosion of biological signals in animal movement data. Ecological Applications 17: 628-638.

5. Jonsen ID, Basson M, Bestley S, Bravington MV, Patterson TA, et al. (2013) State-space models for bio-loggers: A methodological road map. Deep Sea Research Part II: Topical Studies in Oceanography 88–89: 34-46.

6. Jonsen ID, Myers RA, Flemming JM (2003) Meta-analysis of animal movement using state-space models. Ecology 84: 3055-3063.

7. Jonsen ID (2006) Robust hierarchical state–space models reveal diel variation in travel rates of migrating leatherback turtles. Journal of Animal Ecology 75: 1046-1057.

8. Jonsen ID, Flemming JM, Myers RA (2005) Robust state-space modeling of animal movement data. Ecology 86: 2874-2880.

9. Jonsen ID, Myers RA, James MC (2007) Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. Marine Ecology Progress Series 337: 255-264.

10. Bestley S, Jonsen ID, Hindell MA, Guinet C, Charrassin J-B (2013) Integrative modelling of animal movement: incorporating in situ habitat and behavioural information for a migratory marine predator. Proceedings of the Royal Society B: Biological Sciences 280.

11. Breed GA, Costa DP, Jonsen ID, Robinson PW, Mills-Flemming J (2012) State-space methods for more completely capturing behavioral dynamics from animal tracks. Ecological Modelling 235-236: 49-58.

12. Bailey H, Fossette S, Bograd SJ, Shillinger GL, Swithenbank AM, et al. (2012) Movement patterns for a critically endangered species, the leatherback turtle (*Dermochelys coriacea*), linked to foraging success and population status. PLoS ONE 7: e36401.

13. Block BA, Jonsen ID, Jorgensen SJ, Winship AJ, Shaffer SA, et al. (2011) Tracking apex marine predator movements in a dynamic ocean. Nature 475: 86-90.

14. Flemming JM, Jonsen ID, Myers RA, Field CA (2010) Hierarchical state-space estimation of leatherback turtle navigation ability. PLoS ONE 5: e14245.

15. Hammerschlag N, Gallagher AJ, Lazarre DM (2011) A review of shark satellite tagging studies. Journal of Experimental Marine Biology and Ecology 398: 1-8.

16. Patterson TA, Thomas L, Wilcox C, Ovaskainen O, Matthiopoulos J (2008) State–space models of individual animal movement. Trends in Ecology and Evolution 23: 87-94.

17. Schick RS, Loarie SR, Colchero F, Best BD, Boustany A, et al. (2008) Understanding movement data and movement processes: current and emerging directions. Ecology Letters 11: 1338-1350.

18. Fancy SG, Pank LF, Douglas DC, Curby CH, Garner GW, et al. (1988) Satellite Telemetry: a new tool for wildlife research and management. Washington D.C.: United States Department of the Interior. 74 p.

19. Lopez R, Malardé JP (2011) Improving ARGOS Doppler location using Kalman filtering. CLS-DT-MEMO-11-65. 14 p p.

20. Breed GA, Costa DP, Goebel ME, Robinson PW (2011) Electronic tracking tag programming is critical to data collection for behavioral time-series analysis. Ecosphere 2: art10.

21. Hoenner X, Whiting SD, Hindell MA, McMahon CR (2012) Enhancing the use of Argos satellite data for home range and long distance migration studies of marine animals. PLoS ONE 7: e40713.

22. Winship AJ, Jorgensen SJ, Shaffer SA, Jonsen ID, Robinson PW, et al. (2012) State-space framework for estimating measurement error from double-tagging telemetry experiments. Methods in Ecology and Evolution 3: 291-302.

23. Gannon WL, Sikes RS, The Animal Care and Use Committee of the American Society of Mammalogists (2007) Guidelines of the American Society of Mammalogists for the use of wild mammals in research. Journal of Mammalogy 88: 809-823.

24. Sharples RJ, Moss SE, Patterson TA, Hammond PS (2012) Spatial variation in foraging behaviour of a marine top predator (*Phoca vitulina*) determined by a large-scale satellite tagging program. PLoS ONE 7: e37216.

25. Bryant E (2007) Location Accuracy Statistics for Fastloc® Cores Running Firmware Versions 2.2 & 2.3. Leeds, UK: Wildtrack Telemetry Systems Ltd. Technical Report Fastloc/TR01 Technical Report Fastloc/TR01. 5 p.

26. Hazel J (2009) Evaluation of fast-acquisition GPS in stationary tests and fine-scale tracking of green turtles. Journal of Experimental Marine Biology and Ecology 374: 58-68.

27. Russel DJF, Matthiopoulus J, McConnell BJ (2011) SMRU seal telemetry quality control process. SCOS Briefing paper 11/17. Available from Sea Mammal Research Unit, University of St Andrews. St Andrews, UK.

28. Silva MA, Prieto R, Jonsen I, Baumgartner MF, Santos RS (2013) North Atlantic blue and fin whales suspend their spring migration to forage in middle latitudes: building up energy reserves for the journey? PLoS ONE 8: e76507.

29. Tinbergen N, Impekoven M, Franck D (1967) An experiment on spacing-out as a defence against predation. Behaviour 28: 307-320.

30. McClintock BT, Russell DJF, Matthiopoulos J, King R (2012) Combining individual animal movement and ancillary biotelemetry data to investigate population-level activity budgets. Ecology 94: 838-849.

31. Vincent C, McConnell BJ, Ridoux V, Fedak MA (2002) Assessment of Argos location accuracy from satellite tags deployed on captive gray seals. Marine Mammal Science 18: 156-166.

32. Witt MJ, Åkesson S, Broderick AC, Coyne MS, Ellick J, et al. (2010) Assessing accuracy and utility of satellite-tracking data using Argos-linked Fastloc-GPS. Animal Behaviour 80: 571-581.

33. Boyd JD, Brightsmith DJ (2013) Error properties of Argos satellite telemetry locations using Least Squares and Kalman filtering. PLoS ONE 8: e63051.

34. Kuhn CE, Johnson DS, Ream RR, Gelatt TS (2009) Advances in the tracking of marine species: using GPS locations to evaluate satellite track data and a continuous-time movement model. Marine Ecology Progress Series 393: 97-109.

35. Dragon A-C, Bar-Hen A, Monestiez P, Guinet C (2012) Comparative analysis of methods for inferring successful foraging areas from Argos and GPS tracking data. Marine Ecology Progress Series. pp. 253-267.

36. Morehouse AT, Boyce MS (2013) Deviance from truth: Telemetry location errors erode both precision and accuracy of habitat-selection models. Wildlife Society Bulletin 37: 596-602.

37. Bailey H, Shillinger GL, Palacios DM, Bograd SJ, Spotila JR, et al. (2008) Identifying and comparing phases of movement by leatherback turtles using state-space models. Journal of Experimental Marine Biology and Ecology 356: 128-135.

FIFTH CHAPTER

Sei whale movements and behaviour in the North Atlantic inferred from satellite telemetry

ABSTRACT: The stock structure of the sei whale, *Balaenoptera borealis*, in the North Atlantic (NA) is presently unknown, despite continued whaling from the mid-ninetieth century until the moratorium on commercial whaling implemented by the International Whaling Commission in 1986. The proper management of the species in this ocean basin is dependent on the creation of plausible hypothesis about the stock structure and movements of animals, but available data is both outdated and insufficient to achieve that goal. Between 2008 and 2009 the tracks of eight sei whales were obtained with the aid of satellite telemetry, seven during the spring migration and one in late September. Using a hierarchical switching state space model the behaviour of the whales and the role of distinct areas in the life history of those animals were investigated. Two distinct phases corresponding to migratory and foraging movements were identified. A migratory corridor between the Azores and the Labrador Sea (LS) is clearly identifiable from the data. Inside the LS behaviour consistent with foraging was observed frequently, lasting from eight to 132 hours. The data challenge the notion that migration of this species in the NA follows a north-south axis at either side of the basin and show that whales seen in the Azores make transverse migrations, possibly linking the feeding grounds in the LS to breeding grounds off northwest Africa. Those findings raise new hypotheses about the stock structure of the NA sei whale and contribute to filling the data gaps on the natural history of the species.

INTRODUCTION

The sei whale (*Balaenoptera borealis*) can be found in all oceans, predominantly in deep waters or near the continental slopes of temperate to subpolar waters. Commercial harvesting in the North Atlantic (NA) began sometime in the second half of the nineteenth century with the debut of modern whaling. Similar to most other baleen whales, sei whales were heavily hunted worldwide until the implementation of the International Whaling Commission (IWC) moratorium on commercial whaling, in effect from the beginning of the 1986 whaling season [1,2]. In the NA, nearly 17.000 animals were taken during commercial whaling and despite pre-exploitation population levels being unknown, catches are assumed to have led to a substantial population decline with subsequent effects on distribution and significant changes in the size and age composition of the population [3]. There are no recent population estimates for the entire population. Estimates from restricted areas indicate that current population may be over 10.000 animals [4].

After the introduction of the moratorium on commercial whaling, the Scientific Committee of the IWC started a process to revise its harvest control rules in order to overcome the problems with earlier approaches and improve the management of whale resources. This new approach, the Revised Management Procedure (RMP), aims to be a robust management tool even under the uncertainties about historical and present population dynamics of whale stocks. The conceptual basis and form of application of this management-procedure are fully described by Punt & Donovan [5].

Under the RMP, uncertainties regarding stock structure are examined through population simulation under various scenarios [5]. Prior to the adoption of the RMP a set of simulation trials must be run, focusing particularly on the uncertainties about how stocks mix spatially and temporally. These simulations need accurate and up to date information to create plausible alternative hypotheses about stock structure, mixing and other uncertainties such as non-natural mortality [5].

Currently, stock boundaries are already defined by the IWC for several whale populations worldwide. Unfortunately, data used to establish stock identity in the past (catch distribution, length distributions, mark-recapture, and biological parameters) is not always appropriate to define biological stocks [6]. Thus the plausibility of existing stock boundaries has to be reassessed prior to implementing the RMP.

Based largely on catch data, the IWC identified eight concentrations of sei whales in the NA: southwest Norway, north Norway, Faeroe, Denmark Strait, Labrador Sea, Nova Scotia, Gulf of Mexico–Caribbean Sea and southeast NA [6,7].

The limited evidence of stock identity for sei whales in the NA was reviewed by Horwood [7] and Mitchell and Chapman [8]. Based on seasonality and mark-recapture data, Mitchell and Chapman [8] suggested that two stocks occurred simultaneously in summertime off the Canadian Atlantic, termed the “Labrador stock” and the “Nova Scotia stock”. They further suggested that whales from the Nova Scotia stock had their distributional range along the North-American continental slope from Nova Scotia to somewhere off or south of Florida (Fig. 1 in Mitchell and Chapman [8]). A possible link between sei whales in the Labrador Sea and the Denmark Strait was held “less likely but entirely possible” [8]. Horwood [7] reviewed the calculations of length at sexual maturity reported for whales taken in Norway, Canada (off Nova Scotia; Mitchell & Kozicki 1974) and Iceland. He noted that although the values were similar for females, those reported for males from Canada were lower than for Norway and Iceland (12.0, 12.8 and 12.7 m, respectively). However, he also remarked that better agreement would have been achieved if the methodologies used in the Canadian study were similar to those used in the other two studies. Despite acknowledging that the reproductive data was limited, he suggested a segregation between whales caught off Iceland from whales caught off Canada and the United States.

Regardless of the meagre evidence for stock structure and the admission that the identification of stocks for this species in the NA was very difficult, three stocks were adopted by the IWC in 1977 (Figure 5-1 A). These were apparently designed to accommodate existing whaling operations and to match boundaries of ICES fishing areas instead of trying to reflect actual population structure [6].

Migration behaviour of the species in the NA is also poorly understood and existing theories, discussed by Prieto et al. [4], are highly speculative and based in what can be called the “rook paradigm”: migratory movements of whales are believed to be developed at either side of the ocean basin in a north-south axis, resembling the strict movements of rooks on a chess board.

In the western NA Mitchell and Chapman [8] suggested that whales belonging to the purported Nova Scotia stock migrate along the North American shelf-break from and to unidentified wintering grounds at lower latitudes. For the eastern NA, the most widely cited theory is based on Ingebrigtsen [9], who suggested an offshore movement between wintering grounds placed somewhere off northwest Africa and feeding grounds around Iceland, Scotland, western and Norwegian Sea. No migration theory for the purported Labrador stock exists.

The existing data on the NA sei whale are outdated and probably biased for being based chiefly on whaling records [4]. Moreover, in the decades that followed the moratorium on commercial whaling the stock structure of the species may have been affected by population fluctuations, competition with other whale species, distribution shifts of prey and environmental changes. To appropriately manage this species in the NA it is imperative to obtain up-to-date information on all aspects of its biology and ecology, particularly regarding the stock structure.

In the present work the movements and behaviour of sei whales were studied using satellite telemetry technology, in order to characterize their migratory movements and identify foraging areas. The results presented here are valuable for the creation of plausible hypothesis about the population dynamics of this species in the NA as required by the RMP and other population management strategies.

MATERIAL AND METHODS

Ethics Statement

Fieldwork and whale tagging were developed under research permits by the administrative authorities of the Autonomous Region of the Azores. All procedures followed the guidelines of the American Society of Mammalogists [20].

Data collection

Satellite-monitored platform transmitter terminals (PTT) housed in surgical grade stainless steel (model SPOT5-implantable, Wildlife Computers, Redmond, Washington, USA) were implanted in 14 sei whales off Faial and Pico islands, in the Azores archipelago, Portugal (38ºN 28ºW; Figure 5-1 A). Whales were tagged from May to June in 2008 and 2009, except for one whale that was tagged in late September 2009 (Table 5-1). Tagging procedures were similar to those described in Olsen et al. [15] and Silva et al. [18]. Tags were deployed from a 12-m fiberglass boat using a compressed air gun (model ARTS/RN, Restech, Norway), similar to that described in Heide-Jørgensen et al. [19], and set at 10-13 bar pressure. All tags were cleaned with 70% ethanol to remove contaminants from manufacture and handling and then sterilized under UV light and stored in sterilized sealed bags until use. Prior to implantation the tip was covered with gentamicin sulphate cream to act both as antibiotic and lubricant. Tags were attached to the back of the individuals, anterior to the dorsal fin, with a four-bladed tip and held in place with 4 sets of barbs and 6 backward-facing petals.

In 2008 the tags were programmed to transmit every hour of the day, every other day, in an effort to prolong battery life. In 2009 no duty cycle was applied to the tags, since the experience gained in the previous year demonstrated that battery duration was not critical for deployment longevity. All tags were programmed to transmit a maximum of 500 messages per day.

All procedures followed the guidelines of the American Society of Mammalogists [20]. Fieldwork and tagging were developed under research permits by the administrative authorities of the Autonomous Region of the Azores.

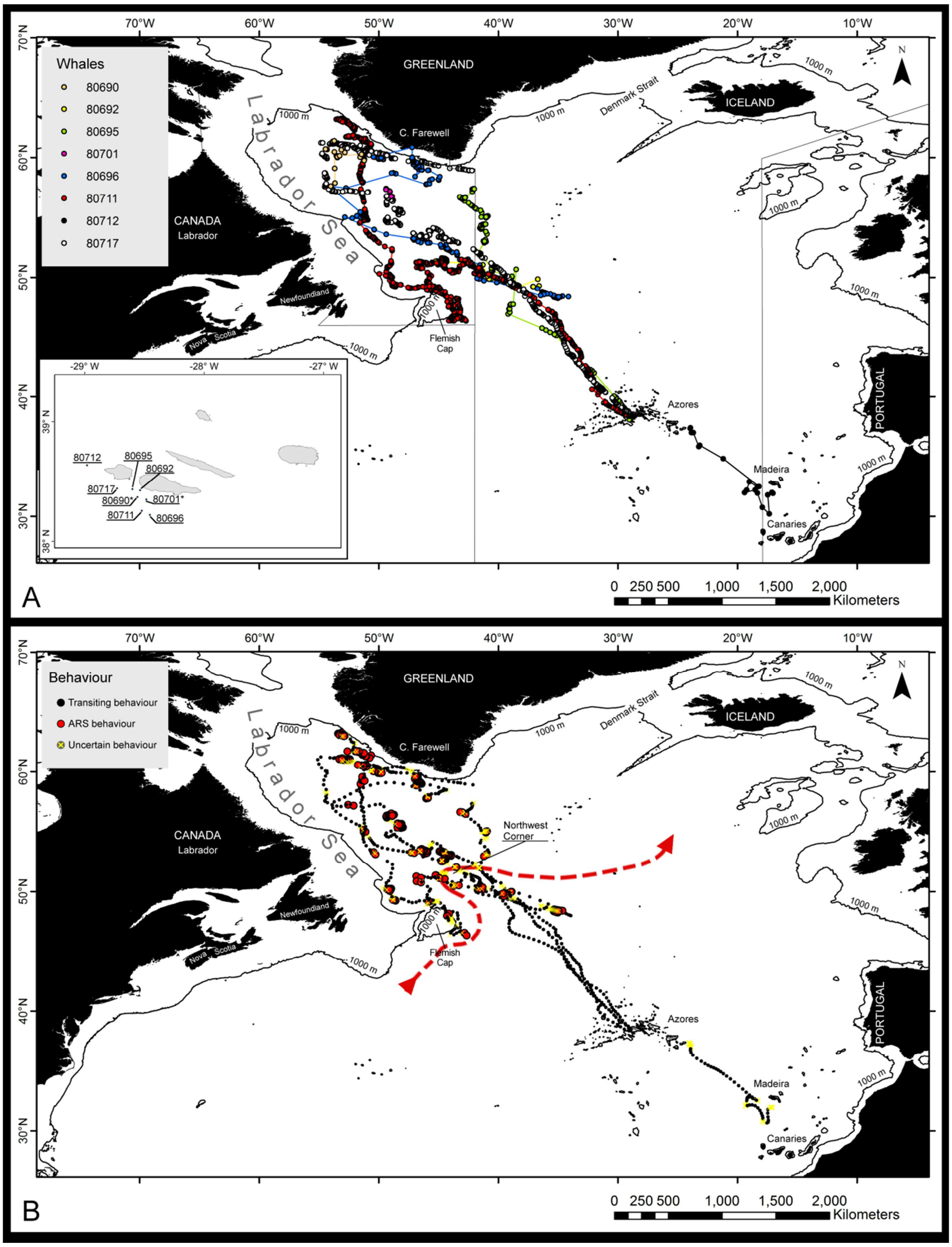

Figure 5-1: A – Sei whale, *Balaenoptera borealis*, tracks derived from raw ARGOS locations. Tagging locations are shown in the inset. International Whaling Commission stock boundaries for the species in the North Atlantic are shown as narrow lines. B – Hierarchical switching state-space model derived locations of sei whales, *Balaenoptera borealis*, showing inferred behavioural modes. The red dashed line is a schematic representation of the main branch of the North Atlantic Current, showing the quasi-stationary large meander called *Northwest Corner* and referred in the text

Table 5-1: Sei whale, *Balaenoptera borealis*, satellite tag deployment results.

ARGOS #	Tagging date	Tagging position		First transmission	First Position received		Last transmission	Last position received		Nº of valid positions	Delay from tagging to first transmission	Tag lifetime (days)
		Lat	Long		Lat	Long		Lat	Long			
80696	26/05/2008	38.2203	-28.4538	06/06/2008	49.0692	-36.7285	24/07/2008	59.9880	-49.7940	226	11	62
80690*	02/06/2008	38.3482	-28.4837	21/08/2008	38.3482	-28.4837	28/09/2008	54.8888	-51.5532	25	80	119
80692	06/06/2008	38.4250	-28.5358	14/06/2008	49.2300	-37.1016	22/06/2008	51.3497	-45.2351	43	8	17
80695	07/06/ 2008	38.4343	-28.6001	07/06/2008	38.0697	-30.3996	28/06/2008	57.3839	-42.0129	172	0	22
80711	02/05/2009	38.3689	-28.5578	02/05/2009	38.3886	-28.6070	29/06/2009	62.9760	-53.0349	935	0	59
80701*	15/06/2009	38.2544	-28.5218	15/07/2009	38.2544	-28.5219	17/07/009	57.3219	-49.3946	8	30	33
80717	17/06/2009	38.4407	-28.7285	17/06/2009	37.6330	-24.6526	19/08/2009	58.9193	-42.2075	753	0	64
80712	21/09/2009	38.6359	-28.9803	09/10/2009	37.3972	-23.9318	25/10/2009	31.9122	-17.0369	20	18	35

* whales not included in fitting the hSSSM

Switching state-space model

In order to analyse the movements and behaviour of tagged whales a model-based stochastic approach was adopted, by fitting a Bayesian switching state-space model (SSSM; [17,21]) to Argos-derived locations processed with the Kalman filter [22]. Locations processed with the Kalman filter were preferred over those processed with the classical nonlinear Least Square algorithm [22] since Silva et al. [23] showed that state-space models fit to locations processed with the Kalman filter perform better.

The SSSM allows for simultaneous correction of location errors and estimation of a behavioural state. In order to accomplish that, the SSSM couples two models. The first is a process model (a first difference correlated random walk) that predicts the future state of an individual given its current state, based on the previous state of the data. The process model is complemented by an observation model of the time series data that relates the unobserved location states predicted by the process model to the observed positional data.

In heterogeneous landscapes it is assumed that predators will change their movement path in response to prey density, increasing their turning rates to stay within the prey patch, a behaviour that is often referred as area-restricted search (ARS; [24,25]). The SSSM allows indexing the movement parameters by behavioural states through the inclusion of a process model for each behavioural state [21]. By assuming that the correlated random walk describing two distinct behavioural states (transiting or ARS) differ only in values of mean turning angle and move persistence (autocorrelation in speed and direction), the SSSM can model switches between those states by calculating the probability for these changes through a Markov chain model (see Jonsen et al. [17] for further background details).

In this study the SSSM was fitted as a single hierarchical model to the combined data of several individuals [17]. By using a hierarchical switching state-space model (hSSSM) information from all individuals can be pooled to more efficiently estimate parameters at both the individual and population levels [26]. In this hierarchical framework, the movement process for each individual whale k is modelled by

$$\mathbf{d}_{t,k} \sim \mathcal{N}_2(\gamma_{bt,k}\mathbf{T}(\theta_{bt,k})\mathrm{d}_{t-1,k},\Sigma)$$

where d_{t-1} and d_t are the differences between the unobserved locations at x_{t-1}, x_{t-2} and x_t, x_{t-1}, respectively, $\mathbf{T}(\theta)$ is a behaviour specific transition matrix describing the mean turning angle (θ) necessary to move from d_{t-1} to d_t, γ the move persistence parameter, $\mathcal{N}_2$ is a bivariate Gaussian random variable describing the stochasticity of the movement process with covariance matrix Σ; the subscript *bt* denotes the behavioural state. At each displacement *t*, the behavioural state *b* takes a value corresponding to the set of parameters θ and γ providing the best model fit.

The same priors on movement parameters as in Jonsen et al. [17] were used, under the assumption that transit behaviour is characterized by turning angles close to 0º and a higher autocorrelation in speed and direction than during ARS behaviour. This assumption is rooted in behavioural studies (e.g.: [27,28]) and have been broadly utilized in similar studies (e.g.: [18,29,30]).

The hSSSM was fitted using a time step of 4 hours, a value that comprises 90% of the transmission intervals recorded in the raw data (Figure 5-2). Models were fit using R (R Development Core Team 2008) code provided in Jonsen et al. [17]. The code implements the hSSSM using Markov Chain Monte Carlo (MCMC) methods via the software Just Another Gibbs Sampler (JAGS). For each hSSSM two MCMC chains were run for 50.000 iterations, dropping the first 45.000 samples as a burn-in and retaining every 5th sample from the remaining 5.000 assumed post-convergence samples to reduce sample autocorrelation. Thus, model parameters and estimates of whales' locations and behaviours were calculated using a total of 2.000 MCMC samples. Model convergence and sample autocorrelation were assessed by visually inspecting trace and autocorrelation plots and using the Gelman and Rubin diagnostic available in R package 'boa'.

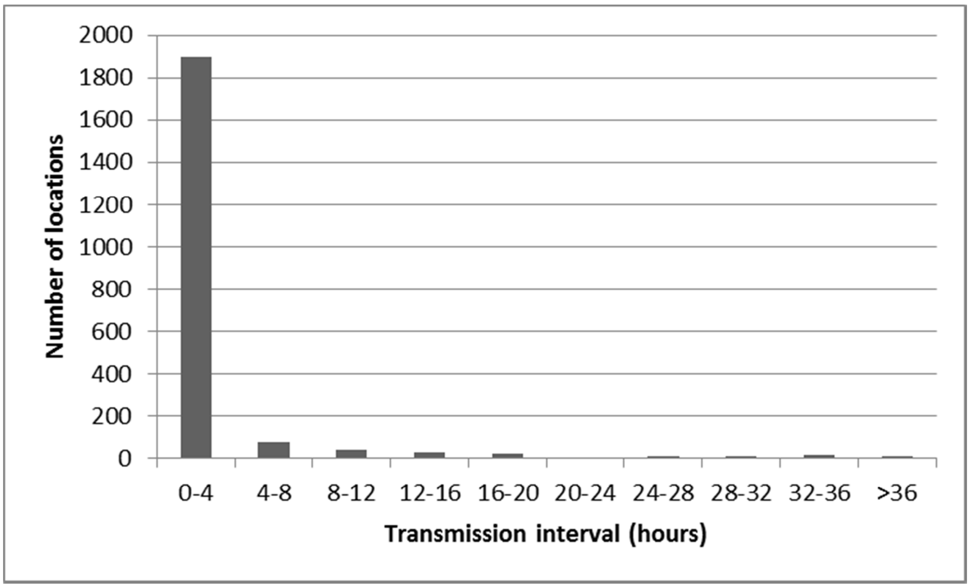

Figure 5-2: Transmission intervals of Argos derived surface locations from the sei whales, *Balaenoptera borealis*, used in fitting the hierarchical switching state-space model.

Whale track analysis

Whale behaviour at each modelled location was inferred from the output of the hSSSM. Although only two modes are modelled by the hSSSM, the means of the MCMC are presented as continuous values between 1 and 2. The same cut-off values as Jonsen et al. [31] were used and it was assumed that a mean behavioural mode $b<1.25$ represented transiting behaviour, and a value $b>1.75$ represented ARS. Any locations with mean behavioural mode $1.25 \leq b \leq 1.75$ were considered uncertain. These cut-off values are considered conservative and were used to ensure good distinction between the two behaviours [31].

In the lack of real time confirmation, it is impossible to be sure that all ARS classifications translate into active foraging. Nevertheless, it is well established that ARS generally corresponds to foraging in marine predators [28,32-34]. Another possibility is that ARS indicates resting or breeding behaviours [29]. However, since all but one animal were tagged during the known feeding season, it is unlikely that breeding behaviour was captured in the data. Thus in the present work ARS was considered to be predominantly indicative of foraging for all animals tagged during their northbound migration. Thus for the sake of readability, ARS will be hereafter referred to as ‘foraging’ for all northbound whales.

Based on data about travel direction and speed collected during fieldwork all individuals were assumed to be on migration when they were tagged. The migratory phase was assumed to have ceased when three or more consecutive locations within a track with a behavioural mode $b \geq 1.25$ were encountered. Foraging bouts were defined as three or more consecutive locations with behavioural mode $b > 1.75$ [29].

Movement parameters and statistics were calculated in R software using packages ‘geosphere’ and ‘circular’. Means are presented with ±1 standard deviation (SD).

RESULTS

Satellite tags were deployed on eight sei whales in 2008 and on six in 2009 but data were received from only four whales in each year. Deployment data for working tags are summarized in Table 5-1.

Argos position estimates are assigned a level of accuracy within seven classes: 3, 2, 1, 0, A, B, and Z. Locations are tagged as class Z when the algorithm fails to converge [22] and thus were considered invalid. Locations in all other classes were considered valid and used in the analysis. A mean of 273±364 locations were received from each tag. Mean tag lifetime (from implantation to last transmission received) was 51.4±32.9 days and valid locations were received for a mean period of 31.0±23.2 days. Reception of the first transmission from the working tags had a mean delay of 18.4±27.0 days in five of the tagged whales, rendering incomplete tracks for those individuals. All individuals but one (80712) were tagged between May and June, during the migration to higher latitudes. Hereafter we will refer to those as northbound individuals. Individual 80712 was tagged in late September, presumably during the migration to wintering grounds. All northbound individuals moved to the Labrador Sea and the signal of the whale tagged in September was lost in the region between Madeira and Canary archipelagos (Figure 5-1 A).

The data from two individuals were not included when fitting the hSSSM: individual 80701 due to the small number of valid locations (n=8) and individual 80690 because the data included several large temporal gaps, resulting in a mean transmission interval of 37.2±48.6 hours with a maximum interval of 182 hours.

In general there was a good distinction between two behavioural modes, as indicated by the parameter estimates that aggregated in two non-overlapping groups (Figure 5-3).

Transiting comprised most of the 1292 locations inferred by the hSSSM (43% of locations), with foraging accounting for 29% of those locations and 28% of locations remaining as uncertain. No ARS behaviour was detected for individual 80712, although some behavioural estimates were classified as uncertain (Figure 5-1 B). For the northbound individuals, foraging was only detected north of 48º N and west of 34º W. In total, 27 foraging bouts were identified, lasting from eight to 132 hours (mean: 48.3±39 hours). Eventually all individuals crossed the North Atlantic Current into the Labrador Sea basin where 63% of the foraging bouts were detected (Figure 5-1 B). At least five of the individuals crossed the North Atlantic Current between 51°N and 53°N in the vicinity of a quasi-stationary large meander of that current known as *Northwest Corner*, positioned at ~51ºN, 44ºW just outside the 4000 m isobath (Figure 5-1 B; [35-37]).

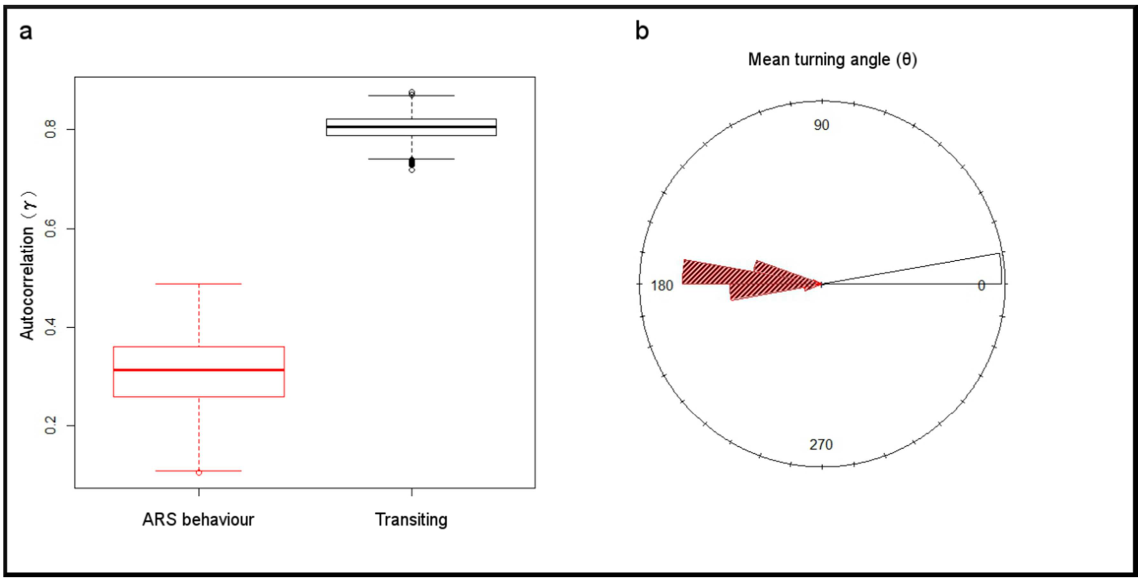

Figure 5-3: Distribution of movement parameters of sei whales, Balaenoptera borealis, in transiting (black) and area-restricted search (ARS; red) for (a) combined autocorrelation in speed and direction (γ), and (b) mean turning angle (θ), based on the means from the Markov Chain Monte Carlo (MCMC) samples.

Comparison of movement parameters on- and off-migration was conducted for only the individuals with complete tracks (80695, 80711, 80712, and 80717). Movement directionality was calculated separately for whales tagged in late spring and early summer (northbound) and the whale tagged in late summer (southbound). Transit speeds were similar between the migration and off-migration phases (Figure 5-4). In contrast and as expected, the movement directionality was considerably higher during migration (Table 5-2; Figure 5-5).

Table 5-2: Sei whale, *Balaenoptera borealis*, movement parameters during migration and off-migration. Please refer to text about the northbound and southbound annotations. Means are presented with ±1 standard deviation.

	Migratory phase	Off-migration phase
Mean transiting speed (km/h)	7.4±0.4	6.2±0.8
Direction of travel during transiting (northbound)	328.8±21.3º	345.5±80.0º
Direction of travel during transiting (southbound)	134.1±14.8º	145.4±102.2º

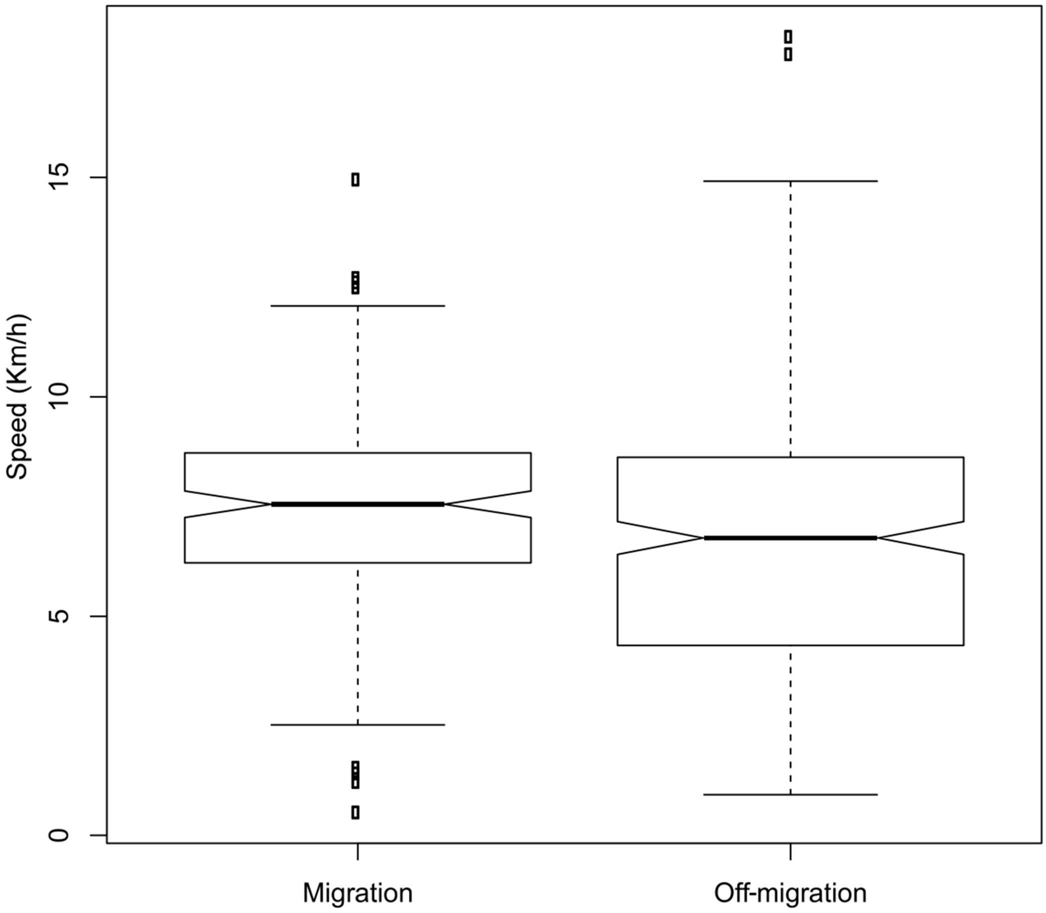

Figure 5-4: Notched boxplots of transit speeds of sei whales, *Balaenoptera borealis,* during migration and off-migration. The notch defines the 95% confidence interval on the median.

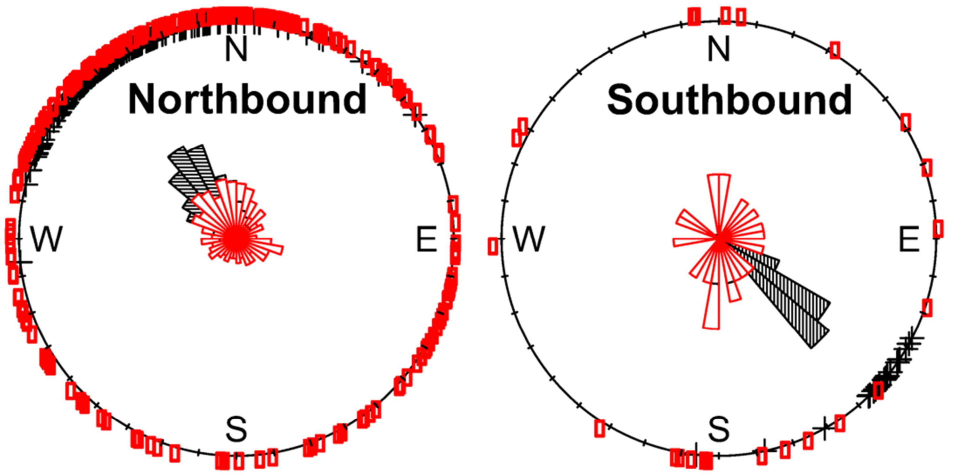

Figure 5-5: Distribution of direction of travel of sei whales, *Balaenoptera borealis*, during migration and off-migration. Crosses and hatched bins correspond, respectively, to data values and frequencies of travel direction during migration. Squares and hollow bins refer to off-migration travel direction. Area of bins is proportional to the group frequency. Refer to text for explanation of Northbound and Southbound annotations.

DISCUSSION AND CONCLUSIONS

Behavioural phases

The results clearly show two distinct phases in the tracking data of northbound individuals with uninterrupted tracks. Before the occurrence of state estimates with values $b \geq 1.25$ in northbound individuals, whales kept a nearly constant travel orientation, consistent within and among individuals. This type of persistent, straightened-out movement is characteristic of migration [38]. Once state estimates with values $b \geq 1.25$ occurred, overall movement directionality decreased and travel direction among individuals ceased to be similar. Even when on transit, whales showed frequent changes of travel direction and transiting behaviour was often interrupted by foraging bouts (Figure 5-1 B), indicating that animals were no longer migrating [38].

These results indicate a well-defined migratory corridor between the Azores and the Labrador Sea. Whales tagged in completely different occasions followed very similar tracks ending their migration in a relatively restricted area. Those results are further backed by Olsen et al. [15], who report on a sei whale tagged in the Azores in April 2005 that also moved to the Labrador Sea.

Interpretation of the data from individual 80712 is more problematic. The directional movement between the Azores and Madeira is also consistent with migratory behaviour. Between Madeira and the Canaries the animal changed the travel direction several times and although no ARS was detected some modelled locations were classified as uncertain. Given that these data are from a single individual, trying to interpret this apparent change in behaviour would be highly speculative and we will refrain from it.

Locations with behavioural state values $b > 1.75$, interpreted as foraging, were only detected north of 48ºN in the vicinity of the well-known oceanographic feature called *Northwest Corner*. The *Northwest Corner* is characterized by the interaction of the warm waters from the North Atlantic Current with colder waters from the Labrador Sea [35-37]. It has been hypothesized that whales can use ocean current patterns and water mass characteristics as directional cues or “guideposts” [39]. The sudden change in the flow characteristics of the current at the vicinity of the *Northwest Corner* in combination with the drastic temperature and salinity variations may serve as cues that trigger the initiation of foraging behaviour.

The role of the Labrador Sea as feeding ground and relation to other known feeding grounds

It is believed that the reason for the large migratory movements of baleen whales from calving/breeding to feeding grounds is related to their requirement of large quantities of high-energy prey found only in highly productive cold waters [40]. In the North Atlantic sei whales feed primarily on the copepod *Calanus finmarchicus* and it is expected that they will favour areas with high concentrations of that prey as preferred feeding grounds [3].

In the past the Labrador Sea was among the known feeding grounds used by sei whales in the North Atlantic [41]. However, since the early 1970’s information on sei whale presence and ecology in that region is virtually non-existent. Sei whales disappeared from or became rare in other known feeding grounds in the northeastern North Atlantic where they were once common [3] but little information was available for the Labrador Sea in that respect.

The data from this study data not only position sei whales inside the Labrador Sea but also show that they spend considerable time foraging while there. Most of the 27 ARS bouts detected occurred when whales were well within the Labrador Sea and lasted from two up to five and half days. At least four individuals eventually moved to the waters just south of the Greenland shelf, where *C. finmarchicus*, dominates the mesozooplankton biomass in spring and early summer [42,43]. In light of that evidence it is reasonable to assume that the tagged whales moved to the Labrador Sea because that region still plays an important role in their feeding ecology.

As summer progresses, or in response to changing conditions, it is conceivable that some individuals move to other feeding grounds, leaving the Labrador Sea. High densities of sei whales have been reported in areas southeast of Cape Farewell, Greenland [44,45]. Based on data from summer sighting surveys undertaken southeast and east of Greenland, Sigurjónsson et al. [44] suggested a movement of whales from areas located southwest of Greenland to the Denmark Strait.

Few of the satellite tags transmitted into late summer, resulting in an incomplete representation of the whales’ movements during the feeding season. Nevertheless, two whales (80695, 80717) were apparently heading east of Cape Farewell when their signals were lost respectively on June 28 and August 19, 2008. A link between sei whales in the Labrador Sea and in the Iceland-Denmark Strait cannot be indisputably established based solely on the data of the present study. However, that hypothesis is compelling based on the available data.

One of the arguments used to suggest the existence of two sei whale stocks off the United States and Canadian east coasts was the simultaneous occurrence of this species in the Labrador Sea and off New England and Nova Scotia during spring and early summer [12]. A similar pattern emerged from our data. Tagged sei whales arrived to the Labrador Sea as early as mid-May and some animals stayed in that area at least until mid-September. Sei whales are known to use the waters of the Gulf of Maine as a feeding ground between spring and early summer [46]. Thus the Gulf of Maine and the Labrador Sea seem to comprise two discrete feeding grounds, utilized simultaneously by sei whales. Nevertheless, it is not possible to speculate if whales frequenting those feeding grounds belong to distinct biological stocks or sub-stocks (*sensu* Jackson and Pampoulie [10]) based on existing data.

Relation to purported wintering grounds

It is not possible to know where the whales that were tagged in the Azores originated. However, it seems unlikely that whales originating from hypothesized wintering grounds located somewhere off the American coast would head to the Azores and then veer to the Labrador Sea, which would add thousands of kilometres to the journey. Unless the Azores constituted an important feeding station during migration, this detour would result in important additional energy expenditure to migrating individuals for no evident advantage. In the present study no foraging behaviour was detected for sei whales in the vicinity of the Azores. Additionally, photo-identification and behavioural data suggest that sei whales have short residency times near the islands and seldom engage in foraging activities (author’s unpublished data). In view of that, the most plausible hypothesis is that sei whales migrating through the Azores in spring/summer originate somewhere south or southeast of the archipelago.

The whale that was tagged in late September (80712) moved southeast and signal was lost between Madeira and Canaries archipelagos. Ingebrigtsen [14] was the first to allude to a sei whale wintering ground off northwestern Africa. Between 1976 and 1978 the unregulated catcher boat *Sierra* captured more than 100 sei whales in a small area just north of Cape Verde archipelago during winter and spring months [47], supporting the notion of a northwestern Africa wintering ground. More recently, the presence of sei whales in the same area during early spring was confirmed, offering further evidence for a wintering ground in the region [3]. It is likely that the whale tagged in late September was on its southbound migration towards wintering grounds placed somewhere in that region. However, further work is needed to confirm that result.

It is tempting to draw a line and conclude a link between the presumed wintering area off northwestern Africa and the Labrador Sea feeding ground. Although this seems to be a likely possibility, such assumption may be flawed. In a scenario of multiple biological stocks, due to its position in the central North Atlantic it is conceivable that the Azores is frequented by whales from one stock during the spring migration and from another during late summer and autumn.

Concluding remarks

There are still great gaps in our understanding of the distribution, movements and population structure of the sei whale in the North Atlantic [3,11]. The results herein shed some light on the movements and behaviour of sei whales that migrate through the Azores. From these results, it is possible to dismiss at least one theory about stock structure; the data indisputably shows that sei whales in the Labrador Sea are not isolated as proposed by Schmidly [13].

Understanding the implications of overlap and mixing of biological stocks in feeding grounds is a major consideration for the management of whaling activities [5,48]. The data from this study indicate, although not indisputably, that sei whales using the Labrador Sea feeding ground may also use other known feeding grounds to the east. Additionally, prior to going to the Labrador Sea the whale that was tracked in the study by Olsen et al. [15] visited the Charlie Gibbs Fracture Zone, at a longitude of ~30ºW, where large aggregations of feeding sei whales have been reported [45]. The combined results reported by Olsen et al. [15] and in this study give some support for the idea of a continuum of the feeding grounds located in the Labrador Sea with other feeding grounds located southeast and east of Greenland, up to the Denmark Strait, as suggested by Sigurjónsson et al. [44]. Whether this results in mixing of different stocks remains to be investigated.

The data also support the idea of a discrete feeding ground located off the Gulf of Maine and Nova Scotia. Investigating if there is any degree of genetic differentiation between animals utilizing that feeding ground and other known feeding grounds such as the Labrador Sea and the Irminger Sea (southwest of Iceland) should be a priority.

The telemetry data show a clear migratory corridor for sei whales between the Azores and the Labrador Sea. This result is relevant not only for management within the IWC. Injury and mortality resulting from ship strikes represent a major conservation problem for many cetacean species [49]. Information on the geographical and temporal distribution of migration corridors is essential to reduce ship strikes [50].

This work shows that satellite telemetry can be a powerful tool to study data-poor species such as the sei whale and contribute for the creation of plausible hypotheses about population structuring. In that aspect, it is important to continue and expand studies on the movement and migration of the sei whale in the North Atlantic to better understand the distribution and isolation of breeding and feeding grounds as well as to clearly identify migratory routes and aggregation areas. But despite the great potentialities of satellite telemetry, this technique is limited if not combined with other. In the case of the North Atlantic sei whales investigating the genetic structuring within the ocean basin is essential.

REFERENCES

1. Horwood J (1986) The lengths and distribution of Antarctic sei whales. Scientific Reports of the Whales Research Institute 37: 47-60.

2. Reilly SB, Bannister JL, Best PB, Brown M, Brownell Jr. RL, et al. (2008) Balaenoptera borealis. In: IUCN 2013, editor. IUCN Red List of Threatened Species Version 20132. www.iucnredlist.org. Downloaded on 01 April 2014.: IUCN.

3. Prieto R, Janiger D, Silva MA, Waring GT, Gonçalves JM (2012) The forgotten whale: a bibliometric analysis and literature review of the North Atlantic sei whale *Balaenoptera borealis*. Mammal Review 42: 235-272.

4. Gambell R (1993) International management of whales and whaling: an historical review of the regulation of commercial and aboriginal subsistence whaling. Arctic 46: 97-107.

5. Punt AE, Donovan GP (2007) Developing management procedures that are robust to uncertainty: lessons from the International Whaling Commission. ICES Journal of Marine Science 64: 603-612.

6. Cooke J, Leaper R, Wade P, Lavigne D, Taylor BL (2012) Management rules for marine mammal populations: A response to Lonergan. Marine Policy 36: 389-392.

7. Smith ADM, Sainsbury KJ, Stevens RA (1999) Implementing effective fisheries-management systems – management strategy evaluation and the Australian partnership approach. ICES Journal of Marine Science: Journal du Conseil 56: 967-979.

8. Milner-Gulland EJ, Arroyo B, Bellard C, Blanchard J, Bunnefeld N, et al. (2010) New directions in management strategy evaluation through cross-fertilization between fisheries science and terrestrial conservation. Biology Letters 6: 719-722.

9. Anonymous (2014) Report of the Scientific Committee. Journal of Cetacean Research and Management 15 (Supplement): 1-75.

10. Jackson JA, Pampoulie C (2012) Stock definition: Terminologies revisited. Scientific Committee of the International Whaling commission (ms). IWC doc SC/64/SD3, 4 pp. Available from the Secretariat of the International Whaling Commission, The Red House, 135 Station Rd., Cambridge, UK.

11. Donovan GP (1991) A review of IWC stock boundaries. Report of the International Whaling Commission special issue 13: 39-68.

12. Mitchell ED, Chapman DG (1977) Preliminary assessment of stocks of northwest Atlantic sei whales (*Balaenoptera borealis*). Reports of the International Whaling Commission Special issue 1: 117-120.

13. Schmidly DJ (1981) Marine mammals of the Southeastern United States coast and the Gulf of Mexico (FWS/OBS-80/41). Washington, D.C., USA: U.S. Fish and Wildlife Service, Office of Biological Services. 163 p.

14. Ingebrigtsen A (1929) Whales captured in the North Atlantic and other seas. Rapports et Proces-Verbaux des Reunions Conseil International pour l’Exploration de la Mer 56: 1-26.

15. Olsen E, Budgell P, Head E, Kleivane L, Nottestad L, et al. (2009) First satellite-tracked long-distance movement of a sei whale (*Balaenoptera borealis)* in the North Atlantic. Aquatic Mammals 35: 313-318.

16. Block BA, Jonsen ID, Jorgensen SJ, Winship AJ, Shaffer SA, et al. (2011) Tracking apex marine predator movements in a dynamic ocean. Nature 475: 86-90.

17. Jonsen ID, Basson M, Bestley S, Bravington MV, Patterson TA, et al. (2013) State-space models for bio-loggers: A methodological road map. Deep Sea Research Part II: Topical Studies in Oceanography 88–89: 34-46.

18. Silva MA, Prieto R, Jonsen I, Baumgartner MF, Santos RS (2013) North Atlantic blue and fin whales suspend their spring migration to forage in middle latitudes: building up energy reserves for the journey? PLoS ONE 8: e76507.

19. Heide-JØrgensen MP, Kleivane L, Øien N, Laidre KL, Jensen MV (2001) A new technique for deploying satellite transmitters on baleen whales: Tracking a blue whale (*Balaenoptera musculus*) in the North Atlantic. Marine Mammal Science 17: 949-953.

20. Gannon WL, Sikes RS, The Animal Care and Use Committee of the American Society of Mammalogists (2007) Guidelines of the American Society of Mammalogists for the use of wild mammals in research. Journal of Mammalogy 88: 809-823.

21. Jonsen ID, Flemming JM, Myers RA (2005) Robust state-space modeling of animal movement data. Ecology 86: 2874-2880.

22. Lopez R, Malardé JP, Royer F, Gaspar P (2014) Improving Argos doppler location using multiple-model Kalman filtering. IEEE Transactions on Geoscience and Remote Sensing, 52: 4744-4755.

23. Silva MA, Jonsen I, Russell DJF, Prieto R, Thompson D, et al. (2014) Assessing performance of Bayesian State-Space Models fit to argos satellite telemetry locations processed with Kalman filtering. PLoS ONE 9: e92277.

24. Tinbergen N, Impekoven M, Franck D (1967) An experiment on spacing-out as a defence against predation. Behaviour 28: 307-320.

25. Kareiva P, Odell G (1987) Swarms of predators exhibit "preytaxis" if individual predators use arearestricted search. The American Naturalist 130: 233-270.

26. Gelman A, Carlin JB, Stern HS, Rubin DB (2003) Bayesian data analysis. Boca Ratón, USA: CRC Press. 668 p.

27. Watkins WA, Schevill WE (1979) Aerial observation of feeding behavior in four baleen whales: *Eubalaena glacialis*, *Balaenoptera borealis*, *Megaptera novaeangliae*, *Balaenoptera physalus*. Journal of Mammalogy 60: 155-163.

28. Mayo CA, Marx MK (1990) Surface foraging behaviour of the North Atlantic right whale, *Eubalaena glacialis*, and associated zooplankton characteristics. Canadian Journal of Zoology 68: 2214-2220.

29. Bailey H, Mate BR, Palacios DM, Irvine L, Bograd SJ, et al. (2009) Behavioural estimation of blue whale movements in the Northeast Pacific from state-space model analysis of satellite tracks. Endangered Species Research 10: 93-106.

30. Kennedy AS, Zerbini AN, Rone BK, Clapham PJ (2014) Individual variation in movements of satellite-tracked humpback whales *Megaptera novaeangliae* in the eastern Aleutian Islands and Bering Sea. Endangered Species Research 23: 187-195.

31. Jonsen ID, Myers RA, James MC (2007) Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. Marine Ecology Progress Series 337: 255-264.

32. Fauchald P (2009) Spatial interaction between seabirds and prey: review and synthesis. Marine Ecology Progress Series 391: 139-151.

33. Hamer KC, Humphreys EM, Magalhães MC, Garthe S, Hennicke J, et al. (2009) Fine-scale foraging behaviour of a medium-ranging marine predator. Journal of Animal Ecology 78: 880-889.

34. Papastamatiou YP, DeSalles PA, McCauley DJ (2012) Area-restricted searching by manta rays and their response to spatial scale in lagoon habitats. Marine Ecology Progress Series 456: 233-244.

35. Worthington LV (1976) On the North Atlantic circulation. Baltimore and London: The Johns Hopkins University Press. 110 p.

36. Lazier J (1994) Observations in the Northwest Corner of the North Atlantic Current. Journal of Physical Oceanography 24: 1449-1463.

37. Belkin IM, Levitus S (1996) Temporal variability of the Subarctic Front near the Charlie-Gibbs Fracture Zone. Journal of Geophysical Research 101: 28317-28324.

38. Dingle H (1996) Migration: the biology of life on the move. New York, USA: Oxford University Press. 474 p.

39. Kenney RD, Mayo CA, Winn HE (2001) Migration and foraging strategies at varying spatial scales in western North Atlantic right whales: a review of hypotheses. Journal of Cetacean Research and Management 2: 251-260.

40. Stern SJ (2009) Migration and movement patterns. In: Perrin WF, Würsig B, Thewissen JGM, editors. Encyclopedia of Marine Mammals 2nd ed. London, UK: Academic Press. pp. 726-730.

41. Mitchell ED, Kozicki VM (1974) The sei whale (*Balaenoptera borealis*) in the northwest Atlantic Ocean (ms). IWC doc SC/SP74/32, 48 pp. + 48 figs. Available from the Secretariat of the International Whaling Commission, The Red House, 135 Station Rd., Cambridge, UK.

42. Head EJH, Harris LR, Yashayaev I (2003) Distributions of *Calanus* spp. and other mesozooplankton in the Labrador Sea in relation to hydrography in spring and summer (1995–2000). Progress in Oceanography 59: 1-30.

43. Frajka-Williams E, Rhines PB (2010) Physical controls and interannual variability of the Labrador Sea spring phytoplankton bloom in distinct regions. Deep-Sea Research Part I 57: 541-552.

44. Sigurjónsson J, Gunnlaugsson T, Ensor P, Newcomer MW, Vikingsson GA (1991) North Atlantic sightings survey 1989 (NASS-89): Shipboard surveys in Icelandic and adjacent waters July-August 1989. (*Balenoptera borealis*, *Balaenoptera physalus*, *Balaenoptera acutorostrata*). Report of the International Whaling Commission 41: 559-572.

45. Waring GT, Nottestad L, Olsen E, Skov H, Vikingsson GA (2008) Distribution and density estimates of cetaceans along the mid-Atlantic Ridge during summer 2004. Journal of Cetacean Research and Management 10: 137-146.

46. Baumgartner MF, Lysiak NSJ, Schuman C, Urban-Rich J, Wenzel FW (2011) Diel vertical migration behavior of *Calanus finmarchicus* and its influence on right and sei whale occurrence. Marine Ecology Progress Series 423: 167-184.

47. Best PB (1992) Catches of fin whales in the North Atlantic by the M.V. Sierra (and associated vessels). Report of the International Whaling Commission 42: 697-700.

48. Hoelzel AR (1998) Genetic structure of cetacean populations in sympatry, parapatry, and mixed assemblages: implications for conservation policy. The Journal of Heredity 89: 451-458.

49. Moore MJ (2014) How we all kill whales. ICES Journal of Marine Science: Journal du Conseil doi: 10.1093/icesjms/fsu008.

50. Mullen KA, Peterson ML, Todd SK (2013) Has designating and protecting critical habitat had an impact on endangered North Atlantic right whale ship strike mortality? Marine Policy 42: 293-304.

SIXTH CHAPTER

Discussion & Conclusion

DISCUSSION

This final section reviews the outputs of the research chapters, summarizing the significant findings and ideas that were presented throughout this dissertation. It is often the case with long and complex documents, such as academic theses, that a conducting line is veiled by seemingly disparate sections and it is the author's responsibility to restore some sense of order to the reader. This dissertation focuses on the North Atlantic (NA) sei whale, and obviously that is the coincident element in all chapters. But if there is another word that stands out in this dissertation, it probably is *management* appearing 59 times in the text of the preceding chapters. That is hardly a coincidence. For long, the main driver of research on the great whales (the sei whale included) has been the need to recover and manage depleted whale populations, a need rooted in decades, in some cases centuries, of bad resource management [1].

The Second Chapter clearly shows how management questions have driven the research on the sei whale. The results of the bibliometry for the publications on the NA sei whale show that most of the literature produced during the analysis period is related to whaling issues. It is also paradigmatic that the peak in publications about the species occurred in 1977, as a result of a special meeting of the International Whaling Commission (IWC) on sei and Bryde's (*Balaenoptera edeni*) whales to discuss the management of whaling for those two species [2]. As shown by the bibliometric analysis, once the moratorium on whaling was put in place and management of the species was no longer considered as pressing as before, the scientific investment on the sei whale became stagnant.

However, in the following decades to the whaling moratorium the conditions in the marine environment degraded, human activities in the oceans intensified and the climate changed considerably, posing new threats to cetaceans [3-5]. It has been argued that these new threats comprise a new type of whaling [5]. Still in the Second Chapter another potential threat is advanced: limited or non-existent research. The title of that chapter encloses that concept by suggesting that the lack of research can bring a poorly studied species, such as the sei whale, into oblivion.

The sei whale is probably as likely to be affected by the threats mentioned above as any other whale species. Nevertheless, since significantly fewer resources were devoted to studying the sei whale in comparison to other whale species, there is no way of knowing how the species is coping with those threats. Even if there are clear signs that the species is being negatively affected by some, or all, of those threats, we are unable to notice since we are not looking for these signs.

With the progressive and increasing degradation of the marine ecosystems research priorities and fund allocation for marine (and cetacean) research became chiefly conservation-oriented [6,7]. In today's world, the terms *ecology* and *management* come hand in hand, more so in the case of cetaceans, and it is tacitly expected that ecological studies can in some way contribute to the management of the subject species or ecosystems.

However, restricting resource allocation only to the most endangered species is counterproductive and does not minimise the total number of extinctions in the long term [8]. Wilson et al. [8] advocate that we should instead focus on preventive conservation, allocating resources to protect the greatest number of non-endangered species. This is why exercises such as that done in the Second Chapter are essential. The first step in preventive conservation is to understand if we have enough information to properly manage the resource. If not, we must understand the reasons that are impairing us of gaining that information and act to solve them.

Unfortunately, despite the shift of ecological studies focus to conservation issues, much of the policy and management decisions are still taken without the support of scientific literature [9,10]. Much of the time, relevant information exist in the scientific literature but policy/decision makers find it difficult to locate or simply intractable [10].

Bridging the gap between conservation policy and science is one of the greatest challenges in natural resource management. That goal can only be achieved if a good communication between policy/decision makers and scientist is established, guaranteeing that management policies are supported by rigorous, relevant science [11,12]. To that goal it is essential that information is presented in a simple but accurate way, and that problems are accompanied by realistic suggestions on lines of action to solve them. Following that spirit, in the end of the Second Chapter the main research needs necessary for a proper management of the sei whale are listed and solutions for obtaining the data are offered. Some of these solutions point to the use of recent techniques to derive and analyse those data in a cost-effective manner. Consequently, the Second Chapter becomes not only an in-depth review of the current knowledge on the NA sei whale, but it also offers clear guidance for future actions in the study and management of the species, becoming a valuable document for decision and policy makers.

One of the most cost-effective ways of gaining new knowledge about a resource is to mine for existing data and use innovative techniques to extract new, valuable information from that data. That is what was attempted and accomplished in the Third Chapter.

Species distribution modelling (SDM) can be a powerful management tool [13]. In fact, recently the focus of cetacean habitat modelling has changed from model development and refinement to using models as a heuristic tool to support ecological insight and marine spatial planning [14]. However, obtaining presence-absence (PA) cetacean data for modelling is difficult and extremely expensive. Additionally, obtaining those data can be unrealistically time-consuming for some management issues [12]. In a recent study [15], it has been shown that the accumulated area covered by dedicated line-transect surveys for cetaceans undertaken over a period of 30 years (1975-2005) covered less than 25% of the world's ocean surface. In the Atlantic, coverage is concentrated in the coastal waters, clearly misrepresenting the potential habitat of pelagic species, such as the sei whale [15].

Nevertheless other sources of data for modelling exist. With the growing usage of the marine habitat by scientific and non-scientific users, the number of reliable cetacean sightings is increasing and those become available through diverse sources (e.g. http://seamap.env.duke.edu; http://mms.data.npolar.no; http://www.marine-life.org.uk/; http://data.marinemammals.gov.au). Much of those data are presence-only (PO) and although not being as desirable as PA data for density and habitat modelling, still carry valuable ecological information [16]. The advantages and drawbacks of using PO approaches to SDM have been discussed in the Third Chapter. Here it suffices to say that as long as the data and the PO-SDM limitations are understood, these can offer cost-effective, ready to use approaches for management decision and policy making.

In the Third Chapter that concept is exemplified by the fitting of a credible SDM to sei whale records in the waters off New England and Nova Scotia. When discussing the results of that model, great emphasis is put on its limitations and possible improvements.

This brings us back to the importance of communication between policy/decision makers and scientists. For some management applications simple approaches may be desirable, since they are cost-effective and often faster to implement [12]. Nevertheless, it is of great importance that the limitations of those approaches are clearly understood and judged against the goals that are being sought. For instance, in setting the boundaries of a management area, managers would probably be only interested in the state model (suitable versus unsuitable habitat) developed in the Third Chapter. Nevertheless, it has been shown that the threshold used to define suitability may have created some amount of commission errors. Managers must be aware of that issue and the options that exist to tackle the problem. Thus the options to deal with the problem are also clearly stated in that chapter: 1) refining the model, both spatially and temporally, preferably with the inclusion of another ecologically meaningful variable (prey); 2) collecting more data to improve model evaluation; 3) changing the suitability threshold to down-weight marginal habitats and 4) (not mentioned but implicit) to combine any or all of the prior. All options have advantages and drawbacks. For example, the first option would come at the expense of time and transferability and the second would come at the expense of time and added financial costs. Without further knowledge, the third option could create the opposite problem (omission errors). We can think of scenarios in which the time or funds involved in improving model construction and checking are not available, or in which commission errors would be preferable to omission errors and vice-versa.

In essence, the point being made in the Third Chapter and reinforced here was brilliantly summarized in only nine words by the late and influential statistician George Box: “Essentially all models are wrong, but some are useful” [17]. That is a fundamental concept in evidence-based conservation policy. Still in Box’s words, “the practical question is how wrong do they [the models] have to be to not be useful”. Here is where the importance of understanding the full implications of model limitation becomes essential. As long as there is a good communication between scientists and managers and model limitations are understood by everyone, decisions will be made in an informed way and will represent a great improvement over otherwise subjective resolutions [9,10,12].

Understanding the limitations and improving models is accomplished by continuously challenging our assumptions. In the Fourth Chapter the applicability of well-established Bayesian state-space models (SSMs) to fit satellite telemetry data derived from a new positioning algorithm is challenged. At the root of that work was a very simple question: can we have the same confidence in the results of fitting SSMs to the data of the new positioning algorithm (Kalman filter; KF) as we had with the old algorithm (Least squares; LSq)? Or, using Box's terminology, will our models become more wrong? This question was fundamental if we are to take full advantage of the improvements reported for the new algorithm: more position fixes and improved position accuracy.

Satellite telemetry is a powerful way of investigating the ecology of whales (e.g. [18]) but due to the time they spend submerged, satellite uplinks are restricted which reflects in the number of positions calculated by the positioning algorithms. Sei whales present further challenges for this technique. The study presented in the Fifth Chapter was the first at a worldwide level to systematically use satellite telemetry to study sei whales. On average, sei whales produce less daily transmissions than blue and fin whales (author's unpublished data). That is attributable to the smaller size of the sei whale in comparison with the blue and fin whales and to species specific swimming behaviours. Sei whales are shallow swimmers that show only a small part of their backs when surfacing and dive without a pronounced arching, tending to sink instead of wheel-over as other rorqual whale species usually do [19]. That behaviour combined with a less than optimal positioning of the tag at the top of the back probably decrease the opportunities for transmissions from the tags, an effect that will be more pronounced while the whales are on migration.

Using a positioning algorithm that produces more fixes is highly desirable when fitting the SSMs to sei whale data. Nevertheless, since the changes in the autocorrelation structure of the data produced by the new algorithm were unknown it would be inappropriate to fit those models without first testing their performance in relation to that of models fit using the LSq algorithm. The results show that it not only is appropriate, as in some instances there are improvements in doing so. Nevertheless, it is also shown that the increased number of positions granted by the KF algorithm comes at the cost of a great number of positions with low spatial precision (1-message positions). Again, the limitations of the data and the model are an important issue that must be understood for proper use of the results. If the desired application of the model is at a reasonably large scale in relation to the low spatial precision of the 1-message positions, the model is still a useful one. But when a higher level of accuracy and precision are desirable, the number of fixes available for modelling may have to be sacrificed in favour of a better model.

The results of the investigation in the Fourth Chapter paved the way for the work presented in the Fifth Chapter. By knowing that SSMs could be safely fitted to data derived from the KF processing algorithm, it was possible to take the most of the sei whale satellite telemetry.

The Second Chapter shows that we are ignorant of the stock structure of the North Atlantic sei whale, the location and extension of their current breeding or feeding areas (apart from some known feeding grounds) as well as their migratory behaviour. Satellite telemetry is one of the new research techniques that can provide a cost-effective mean for obtaining data to partially or entirely answer some of those questions. Knowing *where* and *when* are fundamental questions in ecological and population dynamics studies and since cetaceans often move to remote regions, telemetry studies are invaluable in obtaining data on those aspects [20,21]. Furthermore, movement metrics can often be used to infer behaviour even if only positional data is being collected [22]. The value of satellite telemetry in answering fundamental ecological and management questions with a relatively low level of investment is well illustrated in the Fifth Chapter.

Two seasons of work using satellite telemetry to gain insight into the movements and behaviour of sei whales were enough to answer some questions that remained unanswered for decades. More importantly, these data served to disproof prior beliefs about the migratory habits of the species in the NA, showing that their migrations are not restricted to latitudinal displacements. The implication of that discovery for management of the population is tremendous. It is shown that even with the partial information obtained, plausible scenarios of stock structure can be drawn from the telemetry data. Those scenarios can be improved by further investment on this technique in combination with molecular population techniques.

Another way in which satellite telemetry can be useful is in characterizing habitat preferences and even in creating SDMs. That was not attempted in the present dissertation because the available data is clearly insufficient for that. Nevertheless, with sufficient data, modelling spatial patterns and producing reliable SDMs from satellite telemetry is entirely possible (e.g. [23]). As discussed above, obtaining accessory or alternative data to high cost PA data is sometimes crucial for management purposes and satellite telemetry can be one of these data sources.

However, satellite telemetry can also be invaluable in making dedicated sighting surveys viable. For instance, it was previously mentioned that sighting survey effort covering the pelagic waters of the NA is nearly non-existent [15]. As a result detections of sei whales by many sighting surveys probably represent only a fraction of the population and of their habitat [24]. Realistically, even if investment in surveying was increased it would be unfeasible to cover all the potential habitat of the sei whale which, in practical terms, corresponds virtually to the entire NA. But there is no need to cover that vast area to obtain a representative sample of the population. That is where distribution models based on satellite telemetry may become instrumental, by yielding data that can be used to plan and conscribe survey effort only to representative areas.

CONCLUSION

Perhaps the most advisable and the expected in this section would be that I argue for the relevance of this dissertation in advancing of the knowledge on the sei whale ecology and its contribution for the proper management of that species. I believe, however, that it is much more significant that others do so. Instead, in lieu of a conclusion, I prefer to express how the work conducting to this dissertation contributed to the realization of the dimension of my ignorance by transcribing Thoreau's lumberjack song [25]:

"*Men say they know many things; But lo! they have taken wings – The arts and sciences, And a thousand appliances; The wind that blows Is all that anybody knows.*"

REFERENCES

1. Schevill WE (1974) The whale problem. A status report. Cambridge, MA, USA: Harvard University Press. 419 p.

2. International Whaling Commission (1977) Report of the special meeting of the scientific committee on sei and Bryde’s whales. Cambridge, UK: International Whaling Commission. 62 p.

3. Elizabeth Alter S, Simmonds MP, Brandon JR (2010) Forecasting the consequences of climate-driven shifts in human behavior on cetaceans. Marine Policy 34: 943-954.

4. Hofman RJ (1995) The changing focus of marine mammal conservation. Trends in Ecology & Evolution 10: 462-465.

5. Moore MJ (2014) How we all kill whales. ICES Journal of Marine Science: Journal du Conseil doi: 10.1093/icesjms/fsu008.

6. Rose NA, Janiger D, Parsons ECM, Stachowitsch M (2011) Shifting baselines in scientific publications: A case study using cetacean research. Marine Policy 35: 477-482.

7. Stachowitsch M (2003) Research on intact marine ecosystems: a lost era. Marine Pollution Bulletin 46: 801-805.

8. Wilson HB, Joseph LN, Moore AL, Possingham HP (2011) When should we save the most endangered species? Ecology Letters 14: 886-890.

9. Sutherland WJ, Pullin AS, Dolman PM, Knight TM (2004) The need for evidence-based conservation. Trends in Ecology & Evolution 19: 305-308.

10. Pullin AS, Knight TM, Stone DA, Charman K (2004) Do conservation managers use scientific evidence to support their decision-making? Biological Conservation 119: 245-252.

11. Sutherland WJ, Fleishman E, Mascia MB, Pretty J, Rudd MA (2011) Methods for collaboratively identifying research priorities and emerging issues in science and policy. Methods in Ecology and Evolution 2: 238-247.

12. Sutherland WJ, Freckleton RP (2012) Making predictive ecology more relevant to policy makers and practitioners. Philosophical Transactions of the Royal Society B: Biological Sciences 367: 322-330.

13. Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, et al. (2013) Predicting species distributions for conservation decisions. Ecology Letters 16: 1424-1435.

14. Gregr EJ, Baumgartner MF, Laidre KL, Palacios DM (2013) Marine mammal habitat models come of age: the emergence of ecological and management relevance. Endangered Species Research 22: 205-212.

15. Kaschner K, Quick NJ, Jewell R, Williams R, Harris CM (2012) Global coverage of cetacean line-transect surveys: status quo, data gaps and future challenges. PLoS ONE 7: e44075.

16. Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231-259.

17. Box GEP, Draper NR (1986) Empirical model-building and response surface: John Wiley \& Sons, Inc. 669 p.

18. Silva MA, Prieto R, Jonsen I, Baumgartner MF, Santos RS (2013) North Atlantic blue and fin whales suspend their spring migration to forage in middle latitudes: building up energy reserves for the journey? PLoS ONE 8: e76507.

19. Horwood J, editor (1987) The sei whale: Population biology, ecology & management. London: Croom Helm. 375 p.

20. McConnell BJ, Fedak M, Hooker S, Patterson TA (2010) Telemetry. In: Boyd I, Don Bowen W, Iverson SJ, editors. Marine mammal ecology and conservation A handbook of techniques. Oxford, UK: Oxford University Press. pp. 222-242.

21. Krebs CJ (1994) Ecology: the experimental analysis of distribution and abundance. New York, USA: Harper Collins College Publishers. 801 p.

22. Jonsen ID, Myers RA, James MC (2007) Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. Marine Ecology Progress Series 337: 255-264.

23. Edrén SMC, Wisz MS, Teilmann J, Dietz R, Söderkvist J (2010) Modelling spatial patterns in harbour porpoise satellite telemetry data using maximum entropy. Ecography 33: 698-708.

24. Cattanach KL, Sigurjónsson J, Buckland ST, Gunnlaugsson T (1993) Sei whale abundance in the North Atlantic, estimated from NASS-87 and NASS-89 data. Report of the International Whaling Commission 43: 315-321.

25. Thoreau HD (2009) Walden or Life in the woods; Manis J, editor. State College, PA, USA: Pennsylvania State University. 260 p.

APPENDICES

APPENDIX I: SUPPLEMENTARY FIGURE

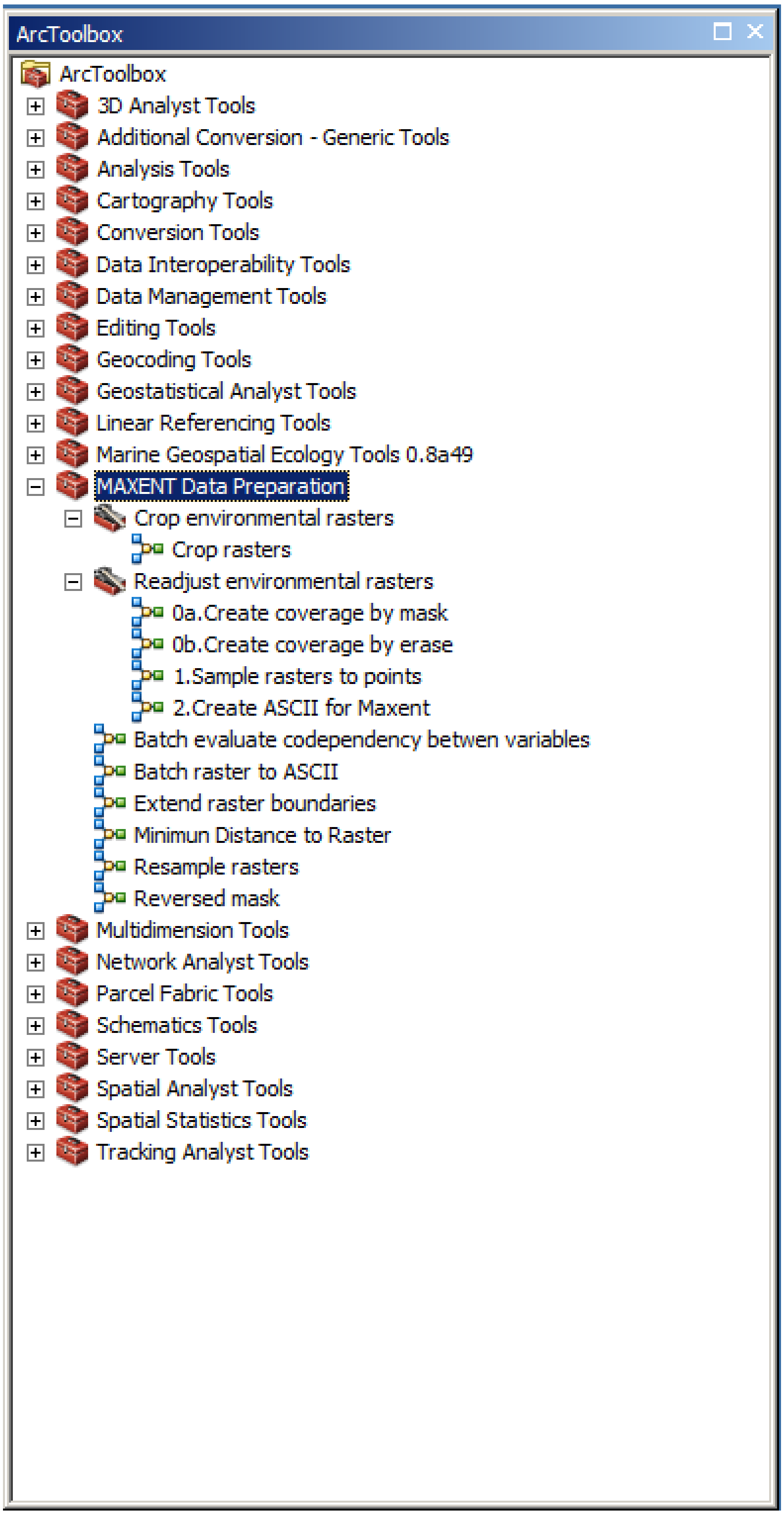

Figure A-I.1: Set of custom tools (expanded tree view) developed for the software ARcGIS 10.1 (ESRI©) to aid in data preparation for habitat modelling.

APPENDIX II: SUPPLEMENTARY TABLES

Table A-II.1: AICc scores of 76 competing models for choice of best regularization multiplier in MAXENT starting model. The models are ranked in ascending order by AICc score. The best model, carried over to the next step in the modelling process, is that with lowest AICc value.

Regularization multiplier	AICc score
0.6	19057.82561
0.4	19072.68445
1	19111.10388
0.8	19130.85241
1.2	19191.05751
1.4	19192.0377
1.8	19196.54977
1.6	19218.29417
2	19230.17725
2.2	19233.15815
2.6	19255.9552
2.4	19256.26834
2.8	19265.01326
3.2	19265.81407
3.6	19278.54801
3	19282.8583
3.4	19288.23847
3.8	19309.00544
5	19326.98729
4.6	19328.15169
4.4	19330.39441
4	19333.32204
4.2	19333.55801
5.2	19336.83566
5.6	19344.67011
4.8	19346.58652
5.4	19347.49091
5.8	19349.84496
6.2	19354.20989
6	19363.64311
6.4	19368.09245
6.6	19374.25135
7	19379.58925
7.4	19386.70518
7.2	19392.51429
6.8	19392.9634
7.6	19405.82958

Regularization multiplier	AICc score
7.8	19410.38031
8	19426.81892
8.2	19428.83932
8.4	19442.05887
8.6	19445.99686
8.8	19448.67653
9	19455.78026
9.2	19465.5606
9.4	19470.29946
9.6	19485.56611
9.8	19488.49868
10	19501.83676
10.2	19508.03215
10.4	19518.64471
10.6	19537.64248
11	19540.97192
11.6	19544.4545
10.8	19545.29964
11.4	19545.44073
11.2	19550.31205
11.8	19554.61577
12	19556.91287
12.2	19558.81706
12.4	19558.97128
12.8	19569.88212
13.2	19571.92427
12.6	19574.00095
13.6	19578.94203
13.8	19579.18359
13	19583.82746
13.4	19585.79368
14.2	19592.81109
14.4	19594.46094
14	19594.46547
14.6	19605.74507
14.8	19607.56282
15	19612.28759
0.2	19935.96905
0	24102.46104

Table A-II.2: AICc scores of 76 competing models for choice of best regularization multiplier in MAXENT final model. The models are ranked in ascending order by AICc score. The best model, used in the final model, is that with lowest AICc value.

Regularization multiplier	AICc score
0.4	19236.33416
0.6	19266.84255
0.8	19319.63652
1.0	19347.56471
1.2	19359.57164
0.2	19364.53634
1.4	19396.18215
1.6	19447.8216
1.8	19448.35144
2.0	19467.97905
2.2	19499.36973
2.4	19518.03649
2.6	19530.1969
2.8	19555.75593
3.2	19573.81929
3.0	19577.77114
3.6	19592.71477
3.4	19600.89412
4.0	19609.1645
3.8	19621.8182
4.2	19623.09025
5.0	19624.55954
4.6	19634.28248
4.8	19638.01
4.4	19640.00728
5.4	19656.09675
4.6	19664.6823
5.2	19670.69725
5.8	19684.07708
6.0	19684.73175
6.4	19702.49415
6.2	19702.64747
6.6	19704.54692
6.8	19712.88077
7.0	19722.6006
7.2	19732.55665
7.6	19752.15805

Regularization multiplier	AICc score
8.0	19754.3049
7.8	19755.43017
7.4	19761.65418
8.6	19772.96468
8.2	19774.99206
8.4	19780.1381
8.8	19782.86121
9.0	19796.7877
9.2	19806.57427
9.4	19807.80474
9.6	19819.57087
9.8	19826.05603
10.0	19844.76206
10.2	19854.54989
10.4	19862.47139
10.6	19874.49989
10.8	19875.45438
11.0	19887.48228
11.2	19890.33507
111.6	19893.8308
11.4	19900.34296
11.8	19904.00416
12.0	19905.0456
12.4	19907.51876
12.2	19907.80067
13.2	19912.84707
13.0	19915.29113
12.6	19916.63655
13.6	19920.53827
13.8	19922.26088
12.8	19923.00844
14.4	19927.24273
14.0	19927.80189
14.2	19928.50148
14.6	19941.66856
13.4	19948.80196
14.8	19973.55955
15.0	19974.48818
0.0	19995.9475

Table A-II.3: GPS, Least-Squares (LS) and Kalman filtered (KF) data obtained for each seal trip.

Seal	Trip	GPS data			LS-Argos data			KF-Argos data		
		Trip duration (h)	N*	Time step (h)**	N*	Time step (h)**	LC 0-B***	N*	Time step (h)**	LC 0-B***
1545	11	54	28	2.0	42	1.3	76.2	66	0.8	80.3
	21	70	25	2.9	46	1.5	65.2	66	1.1	74.2
	23	69	21	3.5	34	2.0	41.2	43	1.6	48.8
	27	87	50	1.8	54	1.6	51.9	66	1.3	53.0
	28	49	32	1.6	37	1.3	43.2	45	1.1	55.6
28503	11	56	36	1.6	44	1.3	81.8	68	0.8	88.2
	18	42	31	1.4	31	1.4	77.4	46	0.9	82.6
	19	59	43	1.4	41	1.5	85.4	63	0.9	88.9
	23	30	31	1.0	33	0.9	87.9	47	0.6	87.2
	42	46	23	2.1	33	1.3	72.7	55	0.8	80.0
43844	4	96	33	3.0	30	3.3	66.7	54	1.8	79.6
	8	47	30	1.6	32	1.5	50.0	43	1.1	55.8
	14	39	35	1.1	31	1.3	61.3	45	0.9	71.1
	16	70	63	1.1	61	1.2	49.2	77	0.9	57.1
	22	81	44	1.9	44	1.9	65.9	55	1.5	72.7
43871	7	97	77	1.3	77	1.3	87.0	124	0.8	91.9
	8	63	57	1.1	43	1.5	86.0	76	0.8	90.8
	13	82	76	1.1	58	1.4	82.8	102	0.8	85.3
	19	67	63	1.1	48	1.4	83.3	79	0.9	88.6
120346	24	50	41	1.2	47	1.1	74.5	65	0.8	80.0
	25	38	43	0.9	48	0.8	81.3	68	0.6	86.8
	26	29	34	0.9	37	0.8	75.7	51	0.5	82.4
	30	45	37	1.2	44	1.0	75.0	61	0.7	82.0
	32	26	21	1.3	33	0.8	60.6	37	0.7	62.2
120349	3	91	18	5.4	39	2.4	87.2	62	1.5	95.2
	4	100	20	5.2	35	2.9	88.6	69	1.5	95.7
	5	83	20	4.4	41	1.9	90.2	97	0.8	95.9
	6	49	20	2.6	34	1.5	88.2	79	0.6	94.9
120350	3	108	35	3.2	51	2.2	72.5	83	1.3	83.1
	4	73	29	2.6	39	1.9	79.5	78	1.0	87.2
	5	66	58	1.2	72	0.9	81.9	113	0.6	90.3
Total		**1957**	**1174**	**1.7**	**1339**	**1.5**	**73.8**	**2083**	**0.9**	**81.8**

*N: Number of locations. **Time step: Average time between locations per trip. ***Proportion of locations of LC 0, A and B.

Table A-II.4: Argos Least-Squares (LS) and Kalman filtered (KF) data obtained for each fin whale used to fit the switching state-space models.

Whale	LS-Argos data			KF-Argos data		
	Track duration (days)	N*	Mean nº positions/day	Track duration (days)	N*	Mean nº positions/day
80702	18	213	11.8	18	253	14.1
80704	23	258	11.2	34	333	12.3
80707	19	215	11.3	19	254	13.4
80713	3	17	5.7	3	30	10.0
80716	18	29	1.6	18	108	6.8
89969	52	1604	30.9	55	2125	39.4
Total	**133**	**2336**	**12.1**	**147**	**3103**	**16**

*Number of locations

Table A-II.5: Parameter estimates from the best fitting linear mixed-effects model for errors in modelled locations.

		Fixed effects			Random effect	
Variable	Factor	Estimate	SE	*P*	Variance	SD
Trip					0.09	0.303
Intercept		-1.780	0.393	<0.001		
Algorithm	LS	0.439	0.078	<0.001		
Proportion LC 0-B		2.865	0.483	<0.001		