www.nina.no

1657

NINA Report

Arctic-breeding seabirds' hotspots in space and time - A methodological framework for year-round modelling of environmental niche and abundance using light-logger data

Per Fauchald, Arnaud Tarroux, Vegard Sandøy Bråthen, Sebastien Descamps, Morten Ekker, Hálfdán Helgi Helgason, Benjamin Merkel, Børge Moe, Jens Åström, Hallvard Strøm

NINA

Norwegian Institute for Nature Research

NINA Publications

NINA Report (NINA Rapport)
This is NINA's ordinary form of reporting completed research, monitoring or review work to clients. In addition, the series will include much of the institute's other reporting, for example from seminars and conferences, results of internal research and review work and literature studies, etc. NINA Report may also be issued in a second language where appropriate.

NINA Special Report (NINA Temahefte)
As the name suggests, special reports deal with special subjects. Special reports are produced as required and the series ranges widely: from systematic identification keys to information on important problem areas in society. NINA special reports are usually given a popular scientific form with more weight on illustrations than a NINA Report.

NINA Factsheet (NINA Fakta)
Factsheets have as their goal to make NINA's research results quickly and easily accessible to the general public. Fact sheets give a short presentation of some of our most important research themes.

Other publishing
In addition to reporting in NINA's own series, the institute's employees publish a large proportion of their scientific results in international journals, popular science books and magazines.

Arctic-breeding seabirds' hotspots in space and time

A methodological framework for year-round modelling of abundance and environmental niche using light-logger data

Per Fauchald, Arnaud Tarroux, Vegard Sandøy Bråthen, Sebastien Descamps, Morten Ekker, Hálfdán Helgi Helgason, Benjamin Merkel, Børge Moe, Jens Åström, Hallvard Strøm

Norwegian Institute for Nature Research

Fauchald, P., Tarroux, A., Bråthen, V. S., Descamps, S., Ekker, M., Helgason, H. H., Merkel, B., Moe, B., Åström, J., Strøm, H. 2019. Arctic-breeding seabirds' hotspots in space and time -a methodological framework for year-round modelling of abundance and environmental niche using light-logger data. NINA Report 1657. Norwegian Institute for Nature Research.

Tromsø, April 2019

ISSN: 1504-3312
ISBN: 978-82-426-3401-6

AVAILABILITY
Open

PUBLICATION TYPE
Digital document (pdf)

QUALITY CONTROLLED BY
Signe Christensen-Dalsgaard

SIGNATURE OF RESPONSIBLE PERSON
Research director Cathrine Henaug (sign.)

COVER PICTURE
Northern fulmar, Northeastern Svalbard © Arnaud Tarroux 2017

KEY WORDS
- Barents Sea, Norwegian Sea, Svalbard, Northeast Atlantic Ocean, Arctic
- Northern fulmar, Black-legged kittiwake, Little auk, Common guillemot, Brünnich's guillemot, Atlantic puffin, Alcids
- Geolocation, Light-loggers
- Habitat modelling, abundance estimation, movement modelling
- Migration, Non-breeding season

CONTACT DETAILS

NINA head office
P.O.Box 5685 Torgarden
NO-7485 Trondheim
Norway
P: +47 73 80 14 00

NINA Oslo
Gaustadalléen 21
NO-0349 Oslo
Norway
P: +47 73 80 14 0

NINA Tromsø
P.O.Box 6606 Langnes
NO-9296 Tromsø
Norway
P: +47 77 75 04 00

NINA Lillehammer
Vormstuguvegen 40
NO-2624 Lillehammer
Norway
P: +47 73 80 14 00

NINA Bergen:
Thormøhlensgate 55
NO-5006 Bergen.
Norway
P: +47 73 80 14 00

www.nina.no

Abstract

Fauchald, P., Tarroux, A., Bråthen, V. S., Descamps, S., Ekker, M., Helgason, H. H., Merkel, B., Moe, B., Åström, J., Strøm, H. 2019. Arctic-breeding seabirds' hotspots in space and time -a methodological framework for year-round modelling of abundance and environmental niche using light-logger data. NINA Report 1657. Norwegian Institute for Nature Research.

By positioning a large number of seabirds throughout the year using miniaturized geolocators (GLS), the SEATRACK program provides a unique dataset on the seasonal distribution of seabirds from colonies in Russia (Barents and White Seas), Norway (incl. Svalbard and Jan Mayen), Iceland, Faroe Islands and the British Isles. Combining this extensive dataset with data on population sizes has for the first time made it possible to develop seasonal estimates of the spatial distribution of Northeast Atlantic seabirds.

In this report, we document the workflow and methods used to develop monthly estimates of the distribution of seabirds from colonies covered by the SEATRACK design. The workflow presented here consists of three steps, starting from pre-processed GLS data. First, because the position data from the loggers represent “presence-only” data, it is vital to remove sampling biases before using the data to make interpretations of the spatial distribution. Therefore, in step 1 we developed a tailored algorithm, IRMA (Informed Random Movement Algorithm), to reduce biases and fill gaps in the dataset due to various factors such as polar day/night, equinox and positions over land. IRMA uses available information and data to triangulate new positions and does ultimately provide a dataset where sampling biases has been reduced to a minimum. In the next step, we combined the position dataset with environmental data to model the habitat of each SEATRACK colony throughout the year. Environmental variables included remote sensing data of oceanography and primary production, and data on bathymetry. We used standard Species Distribution Models (SDM) on presence-only data to model the habitat used by each SEATRACK colony in each month. Finally, in step 3 we combined the predictions from the habitat models with available data on the populations covered by the SEATRACK design to provide predictions on seabird spatial distribution and abundance. A colony database was compiled to address the population sizes, and spatial analyses were conducted to justify a distance-rule for assigning the colonies in the colony database to the nearest SEATRACK colony. Based on the distance rule, we predicted the habitat for each colony covered by the SEATRACK design and weighted the estimates with population size. According to the distance-rule, the SEATRACK design covered from 74% to 96% of the Northeast Atlantic populations, depending on species.

Analyses and predictions were done for six common pelagic seabirds: Northern fulmar (*Fulmarus glacialis*), black-legged kittiwake (*Rissa tridactyla*), common guillemot (*Uria aalge*), Brünnich's guillemot (*Uria lomvia*), little auk (*Alle alle*) and Atlantic puffin (*Fratercula arctica*). The resulting datasets represent monthly estimates of the number of birds from a specific breeding population in each cell of a 0.1° x 0.1° raster grid covering the entire North Atlantic. Monthly outputs were produced for each combination of species and colony, resulting in a dataset of more than 9619 raster maps. The gridded data are provided NetCDF files, one per species, and a short R-script is provided for reading, plotting and aggregating the data. An interactive mapping tool will be made available through the SEATRACK website. Applications for the new tool include marine spatial planning, environmental impact- and risk assessments as well as assessments of seabird responses to environmental and climate change.

Per Fauchald, NINA, per.fauchald@nina.no
Arnaud Tarroux, NINA, arnaud.tarroux@nina.no
Vegard Sandøy Bråthen, NINA, vegard.brathen@nina.no
Sebastien Descamps, Norsk Polarinstitutt, sebastien.descamps@npolar.no
Morten Ekker, Miljødirektoratet, morten.ekker@miljodir.no
Hálfdán Helgi Helgason, Norsk Polarinstitutt, halfdan.helgi.helgason@npolar.no
Benjamin Merkel, Norsk Polarinstitutt, benjamin.merkel@npolar.no
Børge Moe, NINA, borge.moe@nina.no
Jens Åström, NINA, jens.astrom@nina.no
Hallvard Strøm, Norsk Polarinstitutt, hallvard.strom@npolar.no

Contents

Foreword

Conservation of seabirds is high on the agenda for national and international environmental management and organizations. In concert with climate warming, the development of shipping, tourism, fisheries and oil and gas exploitation has the potential of severely impacting seabird populations in the Arctic. To improve marine spatial planning there is a need for detailed and unbiased maps of the distribution of seabirds throughout the year. This has been one of the goals for the SEATRACK program, and this report describes the methods used to provide a new map product that can be utilized by management and industries. The overall results from the SEATRACK program are presented in a joint report (Strøm et al. 2019).

The SEATRACK project has been funded by the Norwegian Ministry of Climate and Environment, the Norwegian Ministry of Foreign Affairs, the Norwegian Oil and Gas Association and seven oil companies: Equinor, Det norske oljeselskap ASA, Eni Norge AS, Total E&P Norge AS, ConocoPhillips Skandinavia AS, Neptune Energy and DEA Norge AS. In addition, the Norwegian Coastal Administration and SEAPOP have provided funding for the development of the map tool.

The map tool has been developed using GLS data collected by the members of the SEATRACK consortium in seabird colonies throughout the Northeast Atlantic.

Per Fauchald, 10 April 2019

1 Introduction

Data on the spatial distribution of vulnerable marine resources are used in a wide range of management applications related to marine conservation, impact assessments and environmental risk assessments (Hays et al. 2019). The SEATRACK program was designed and launched in 2014 to provide data and knowledge on the seasonal distribution of Northeast Atlantic seabird populations (Strøm et al. 2019). To achieve this goal, a large number of seabirds from breeding colonies in Russia (Barents and White Seas), Norway (incl. Svalbard and Jan Mayen), Iceland, Faroe Islands and the British Isles were instrumented with miniaturized geolocators (GLS) loggers (see Fig. 1.1). The loggers use the shifting daylight (i.e. timing of sunrise and sunset events) to record the approximate position of the birds twice a day, and the large-scale approach of the program made it possible to discern the annual migration pattern of the different seabird populations inhabiting the area. The dataset provides new and highly valuable knowledge on the distributional patterns of seabirds (see http://seatrack.seapop.no), however the SEATRACK dataset also offers an unprecedented opportunity to develop seasonal estimates of the distribution of the populations that can be utilized in quantitative environmental assessments and analyses.

In this report, we describe the methods used to develop unbiased monthly maps of the abundance of six Northeast Atlantic populations of pelagic seabirds. The species modelled are: Northern fulmar (*Fulmarus glacialis*), black-legged kittiwake (*Rissa tridactyla*), common guillemot (*Uria aalge*), Brünnich's guillemot (*Uria lomvia*), little auk (*Alle alle*) and Atlantic puffin (*Fratercula arctica*). The methods combine position data of individual birds using GLS loggers, data on the marine environment, and population data from breeding colonies. Our approach includes algorithms to reduce biases and fill data gaps in the position dataset,

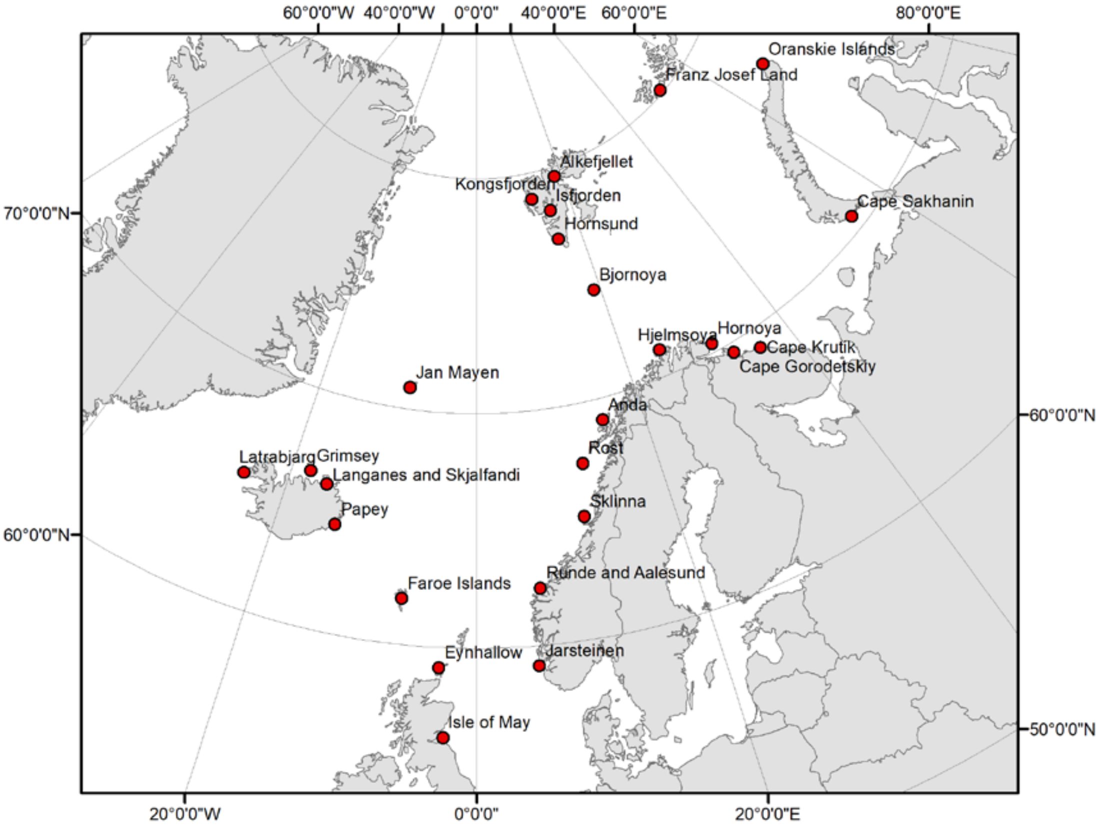


Figure 1.1 – Sample of seabird colonies from the SEATRACK dataset used in population distribution modelling.

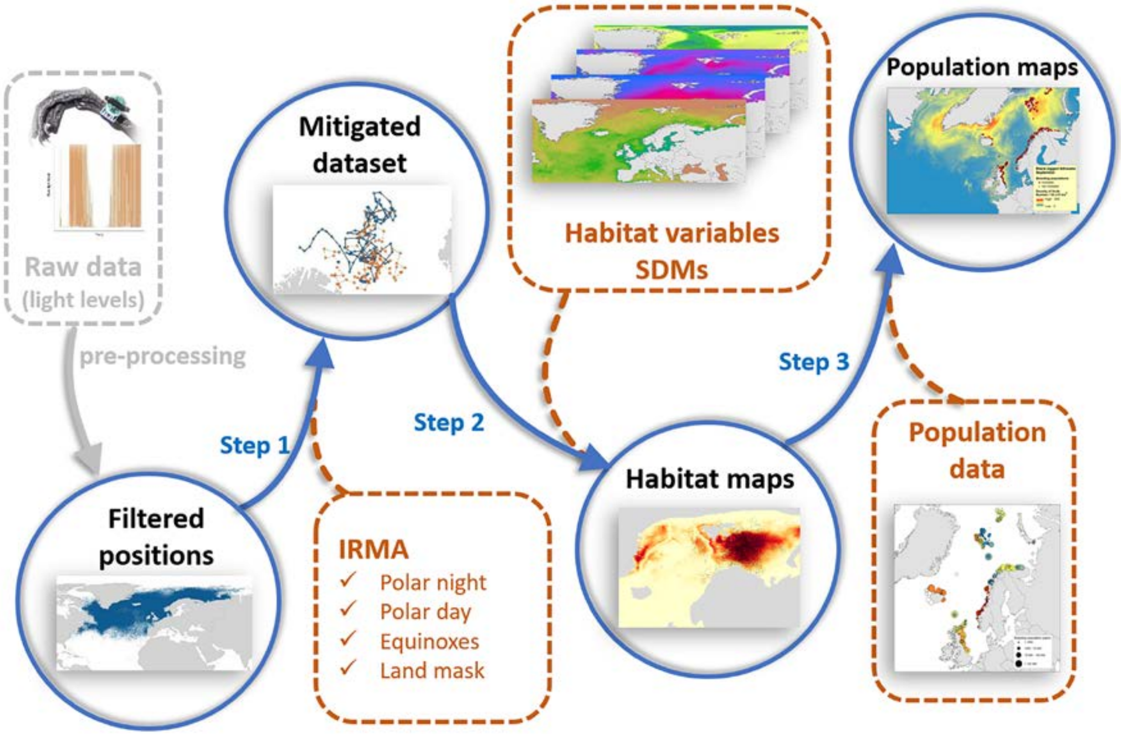
Species Distribution Models (SDM) to predict the habitats of the seabirds, and methods to link habitat maps with population data. The resulting datasets give monthly estimates of the distribution and abundance of breeding populations of pelagic seabirds from the Northeast Atlantic. The datasets can be used in environmental risk and impact assessments, the mapping of important and vulnerable marine areas and in the planning of marine protected areas. Finally, the datasets can be used to assess the responses of seabird populations to environmental and climate change.

The development of miniaturized devices for geographical positioning, has enabled researchers and managers to track migrating animals throughout their life cycle. The immense increase in the quantity of data on the geographical distribution of individuals combined with increased processing capacity and development of user-friendly statistical tools for analysing spatiotemporal datasets, has provided invaluable tools for the conservation of threatened species as well as planning of activities. However, this development has also created new challenges owing to the very nature of the tracking data that are nowadays collected en masse. The size, weight, and life duration of geolocators (light-loggers, or GLS) have made them particularly suited for studies about the distribution and large-scale movements of seabirds (Amélineau et al. 2018, Yurkowski et al. 2019). However, GLS-based positional data are characterized by important biases due to the impossibility to determine reliable locations during certain periods of the year. Such biases need to be addressed before statistical approaches, such as habitat and abundance modelling, can be implemented. Here, we propose a three-step framework that will allow to reduce the bias in the positional dataset and to ultimately model the spatiotemporal variation in seabird abundance (Box 1).

Wildlife positional data represent “presence-only” data. This means that we know the positive presence of a bird (or animal), but we have no information of true absences. The use of presence-only data to predict the spatial distribution of a population is highly sensitive to sampling bias. For example, the assessment of the migration patterns of birds using ringing might be biased by the fact that the chance of recollecting a ring is higher in some areas than in others. Thus, a map derived from a ringing study could reveal the distribution of bird observers (or ring collectors) rather than the actual distribution of birds. Similarly, the lack of registration of positions during equinoxes and polar night/day and the removal of false positions over land represent sampling biases in the GLS dataset that might introduce severe biases in the interpretation of the distribution. Indeed, geolocator data is characterized by large gaps that are not randomly distributed, creating biases and preventing adequate statistical analyses and interpretations. To alleviate those biases, it is possible to model new locations that will replace the missing ones. One such approach consists in using a movement model that determines plausible locations based on a limited set of parameters and user-defined constraints. Methods are available to generate random locations between two known locations along a track and the approach proposed by Technitis et al. (2015) is particularly interesting as it is both efficient and flexible. In Chapter 2, we build on this method to develop an algorithm that replaces missing locations using additional information (Informed Random Movement Algorithm, or IRMA). In order to generate the most realistic locations, it is advantageous to extract additional information from the dataset and use it to parameterize IRMA to constrain its outputs. This includes extracting information on longitude during equinox periods when latitude data are unreliable; determining realistic movement rates for each species; extracting and processing raw activity data (wet/dry sensor) when available; and estimating periods of attendance to the colony based on activity data.

Having removed spatial and temporal biases in the position dataset, the next step of the workflow is to model the habitats of each seabird population (Chapter 3). The presences of seabirds were modelled by Species Distribution Models (SDM). SDM has become a widely used tool for mapping the habitats of wild animals and plants and are used in various management applications (Elith & Leathwick 2009). In short, SDMs are empirical models that relate data of species occurrence to data of relevant environmental predictors (Guisan & Zimmermann 2000). The relationship is estimated by various statistical methods and is expected to reflect the environmental niche utilized by the species. In SDMs of presence-only data, it is necessary to introduce background points to contrast the recorded presences in the analyses (Elith & Leathwick 2009; Barbet-Massin 2012). Several methodological challenges related to the SDM of the SEATRACK dataset has been dealt with and is presented in the present report. The challenges include: Dealing with migratory behaviour and non-stationary distribution; selection of statistical modelling method; definition of representative background points; selection of relevant environmental predictors; and model selection and specifications.

Based on the SDMs, it is possible to predict the spatial distribution of likelihoods of occurrence in the study area. To translate these values into abundance estimates, it is necessary to weight the model predictions with a factor representing population sizes. In other words, the predictions from each SDM should be weighted by the size of the population represented by the model (Chapter 4). To accomplish this task we: Compiled available data of breeding populations to generate a colony dataset for the Northeast Atlantic; analysed how the



Box 1. Workflow to produce SEATRACK abundance maps. Filtered data of positions of individual seabirds are provided by the SEATRACK programme. In Step 1, an algorithm; IRMA (Informed Random Movement Algorithm) is used to remove biases and fill gaps in the dataset due to polar day/night, equinox and positions over land. The algorithm uses available information and data to triangulate new positions. In Step 2, the resulting position data is combined with environmental data to model the habitat of each seabird population throughout the year. Environmental variables include remote sensing data on oceanography and primary production and data on bathymetry. The modelling approach involves standard Species Distribution Models with presence-only data. In Step 3, the predictions from the habitat models are combined with population data to provide predictions on the spatial distribution of each population covered by the SEATRACK design.

overlap in winter habitat changed with distance between breeding colonies and used the results to justify distance-rules for assigning colonies in the colony database to the nearest model colony and; used the SDMs to predict habitat maps for each assigned colony and weighted the predictions with breeding population size.

The results from the analyses are maps of the estimated monthly distribution of the breeding population of each colony covered by the SEATRACK design. The maps can easily be summed to cover different parts of the Northeast Atlantic populations (e.g., national or regional populations, or populations within different ocean areas). Several caveats and limitations regarding the datasets are identified and shortly described in section 4.3.2. One important limitation is that the sample of tagged birds did only consist of breeding individuals. In other words, the maps do not include information on the distribution of non-breeders, including immatures. If the distribution of immatures deviates from that of adults, this might imply an important limitation concerning the present datasets.

The main purpose of the present report is to document the rationale behind the workflow and methods used to map the seasonal distribution of pelagic seabird populations. The presentation of some major results can be found in Strøm et al. (2019) and the abundance maps will be made available on the SEATRACK web page (www.seapop.no/en/seatrack; see section 6.3). Two important tasks remain. First, we will validate and investigate the fit of the SDMs using an independent GLS dataset collected in 2017-2018. Second, we will investigate possible biases in the abundance maps by comparing the estimates with count data collected on seabird-at-sea surveys.

2 Reducing bias in a geolocator-based positional dataset

2.1 Processing of the raw light data

2.1.1 Estimating coordinates from light-level data

Our approach starts from geographic data obtained through a pre-processing phase that involves downloading and decompressing the raw light-level data from the loggers and converting them into estimated geographic locations. The procedure is described in detail in Strøm et al. (2019).

The preliminary processing yielded up to two geolocator-based locations (hereafter GLS locations) per bird and per day (one location at noon and one at midnight), with a total of **1 184 012** unique pre-processed locations and **1 610** individuals in total for the six species considered (see Table 2.1 and Fig. 2.1). All processing and analyses described in the following sections were conducted using R Statistical Software v.3.5.1 (R Development Core Team 2018).

Table 2.1. Summary of the positional dataset for six pelagic species. GLS locations are the filtered locations derived directly from the light-logger data. IRMA locations are the locations that have been determined using the informed random movement algorithm described in this report.

			Number of locations			
Species	Individuals	Colonies	GLS (pre-processed)	GLS (filtered)	IRMA	Tracking period
Little auk	149	5	68 025	34 998	29 780	2010-2018
Atlantic puffin	272	13	181 892	104 452	122 576	2009-2017
Northern fulmar	173	7	130 667	74 289	93 104	2011-2017
Black-legged kittiwake	506	15	414 658	244 988	223 383	2009-2017
Common guillemot	249	11	206 890	109 085	139 356	2011-2017
Brünnich guillemot	261	12	181 880	97 001	128 681	2012-2017
Total	**1 610**	**27**	**1 184 012**	**664 813**	**736 880**	

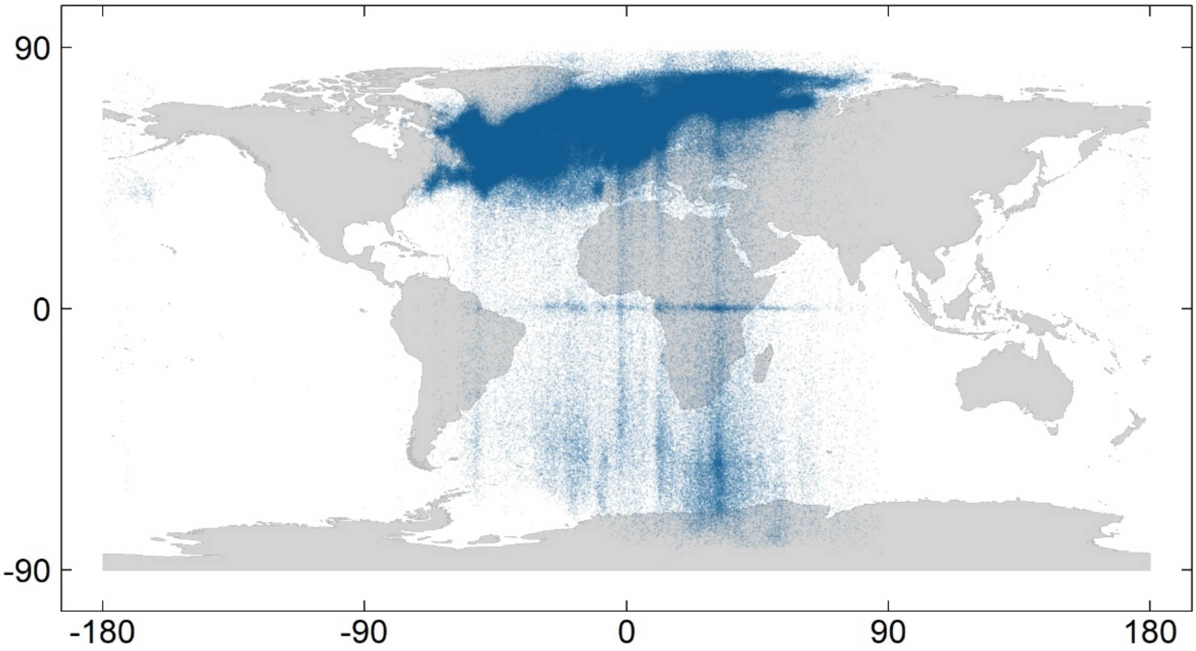


Figure 2.1. World map showing the unfiltered geolocator-based locations obtained from light-logger data on six pelagic species breeding in the North, Norwegian, and Barents seas. Locations with obviously wrong latitudes are clearly visible and correspond to the equinox periods. Map coordinate system is EPSG:4327.

2.1.2 Filtering of the raw positional dataset

Raw (i.e. unfiltered) GLS locations include a lot of erroneous positions that first need to be removed from the dataset. Locations over landmasses were deemed unreliable for all pelagic species and thus systematically eliminated from the dataset. In addition, many locations could not be determined owing to several causes:

- Latitude cannot be estimated adequately throughout the equinox periods, when the day length is virtually the same everywhere on Earth (during ca. one month in March and one month in September each year). However, the longitude derived from light loggers is still reliable during those periods and that information was thus kept for use at later stages (see section 2.2.2.1);
- Neither latitude nor longitude can be determined during polar day/night periods, when the light loggers do not detect large enough variation in light level;
- Occasionally, light loggers cannot properly record sunset/sunrise events because of the outfitted bird’s behaviour (e.g., when a logger is hidden under the plumage for extended periods of time).

Consequently, the filtered dataset is made of individual tracks characterized by frequent gaps (missing locations), ranging from one day to several weeks during the equinoxes or even months during polar day/night periods. A total of **664 813** locations (56%) were retained in the dataset after filtering.

2.2 Mitigating biases in the positional dataset using an Informed Random Movement Algorithm (IRMA)

Gaps are not randomly distributed within the positional dataset, creating biases that prevent adequate statistical analyses. It is necessary to mitigate those biases by modelling new locations that will replace all missing ones, i.e. using a movement model that determines possible locations based on a set of parameters and user-defined constraints. The bias-mitigation process relies primarily on the a(Lisovski et al. 2012)pplication of an algorithm generating locations when those are missing (called Informed Random Movement Algorithm, or IRMA), and which uses the algorithm proposed by Technitis et al. (2015). This algorithm and its application are described in section 2.2.1. It was necessary to extract additional information from the dataset and use it to parameterize IRMA before running it. This included extracting information on longitude during equinox periods (section 2.2.2.1), determining realistic movement rates for each species (section 2.2.2.2), extracting and processing raw activity data (wet/dry sensor) when available (section 2.2.2.3), and estimation periods of attendance to the colony based on activity data (section 2.2.2.4 and 2.2.2.5). The workflow leading to the mitigation of biases can be summarized as follows:

1. Filtering the pre-processed locations and identifying gaps in the positional dataset
2. Extracting longitude information during the equinoxes
3. Determining species-specific movement rates using available GLS-based locations
4. Extracting wet/dry data
5. Determining periods of colony attendance
6. Applying IRMA to replace missing locations based on available additional information

2.2.1 Underlying algorithm

Our approach builds on the random-track generator algorithm recently proposed by Technitis et al. (2015). In short, this algorithm is based on the determination of so-called space-time prisms, which are 3-dimensional volumes defined by the coordinates (x,y) and time (z). The space-time prism delineates all the potential paths that can be followed by an individual moving from point A to point B, given three parameters: the distance from A to B, the time budget available, and the maximum rate of movement (Miller 1991). When projected onto a 2-dimensional plane, the space-time prism becomes the potential point area, (hereafter Ppa; Technitis et al. 2015). Although the 3-dimensional representation of the space-time prism is useful to understand its concept (Neutens et al. 2007), it is naturally more convenient to work with only two dimensions when dealing with discrete time steps, as is the case with tracking studies where locations are obtained at specific time intervals. Computing the Ppa in this context is straightforward (Technitis et al. 2015), given that the three above-mentioned parameters are known. Let us consider a start point (A) and start time (t_{i-1}), and an end point (B) and end time (t_{i+1}). Knowing the maximum rate of movement and the time t_i at which a new location (N_i) is to be created, one can determine the circle defining the maximum range (r_{i-1}) from point A to the new location, and that defining the maximum range (r_{i+1}) from the new location to point B. The Ppa corresponds to the area of overlap between those two circles of maximum range (Fig. 2.2), i.e. the area delimiting all locations that are reachable from both A and B, given the time budget and maximum movement rate. This process can be repeated any number of times, depending on the number of new locations that need to be generated. The new locations are generated in a random order (i.e. not chronological; Fig. 2.3), thus creating a sort of correlated random walk respecting the constraints set by the relative position of A and B, the time budget, and the maximum movement rate. Determining the maximum movement rate was done separately for each species, using the available location data (described in detail in section 2.2.2.2).

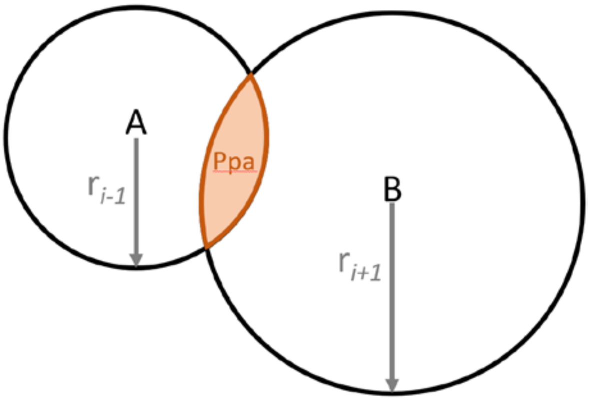


Figure 2.2. Illustration of the calculation of the Potential Point Area (Ppa), i.e. the area where any location at time *i* could be, given the parameters (distance from A to B, maximum speed, and time elapsed from A to B). It corresponds to the overlap between the circles defined by maximum range from point A to the new location (r_{i-1}) and from the new location to point B (r_{i+1}). Adapted from Technitis et al. (2015).

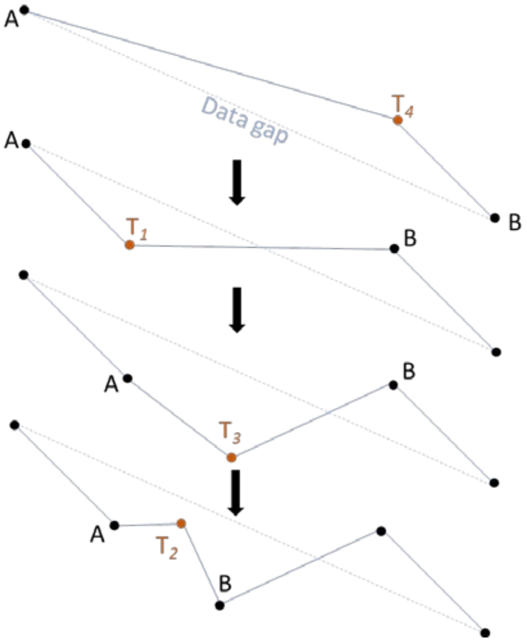


Figure 2.3. Overview of the sequential process for generating new locations within a given data gap. A and B represent the start and end points used to determine the new location at time T_i. The first location is necessarily based on the locations defining the start and end of the gap in the filtered dataset. For each subsequent iteration A (or B) corresponds to the location that is nearest in time before (or after) T_i. In this example, the new locations were created in the following random order: T_4, T_1, T_3, T_2. Adapted from Technitis et al. (2015).

2.2.2 Determining parameters and constraints for IRMA

2.2.2.1 Extracting longitude information during equinoxes

Latitudes estimated from geolocator data are characterized by very large errors during the equinoxes (Lisovski et al. 2012, Merkel et al. 2016) and are thus virtually unusable. The estimation of longitudes, however, is not affected by equinoxes, and therefore such data represent valuable information that can be used to constrain modelling of new locations. For each individual bird, we extracted the available longitude and timestamp information from all raw locations that corresponded to a period of equinox, thus obtaining a time series of longitude. This information is stored separately and retrieved when using IRMA to generate a new location for the corresponding timestamp and bird. More specifically, each new location xy_i is constrained within a given range of longitudes (thereafter longitude buffer; Fig. 2.4) which corresponds to the min/max longitudes recorded at time t_i ± 1.5 day in the time series, to allow for some flexibility. In cases where no solution can be found (i.e the Ppa and the longitude buffer did not overlap), the time window from which the max/min longitude values are extracted is increased by 0.5-day increments, up to a maximum of 10 days. If no solution can be found, the new location is created without any constraint on the longitude.

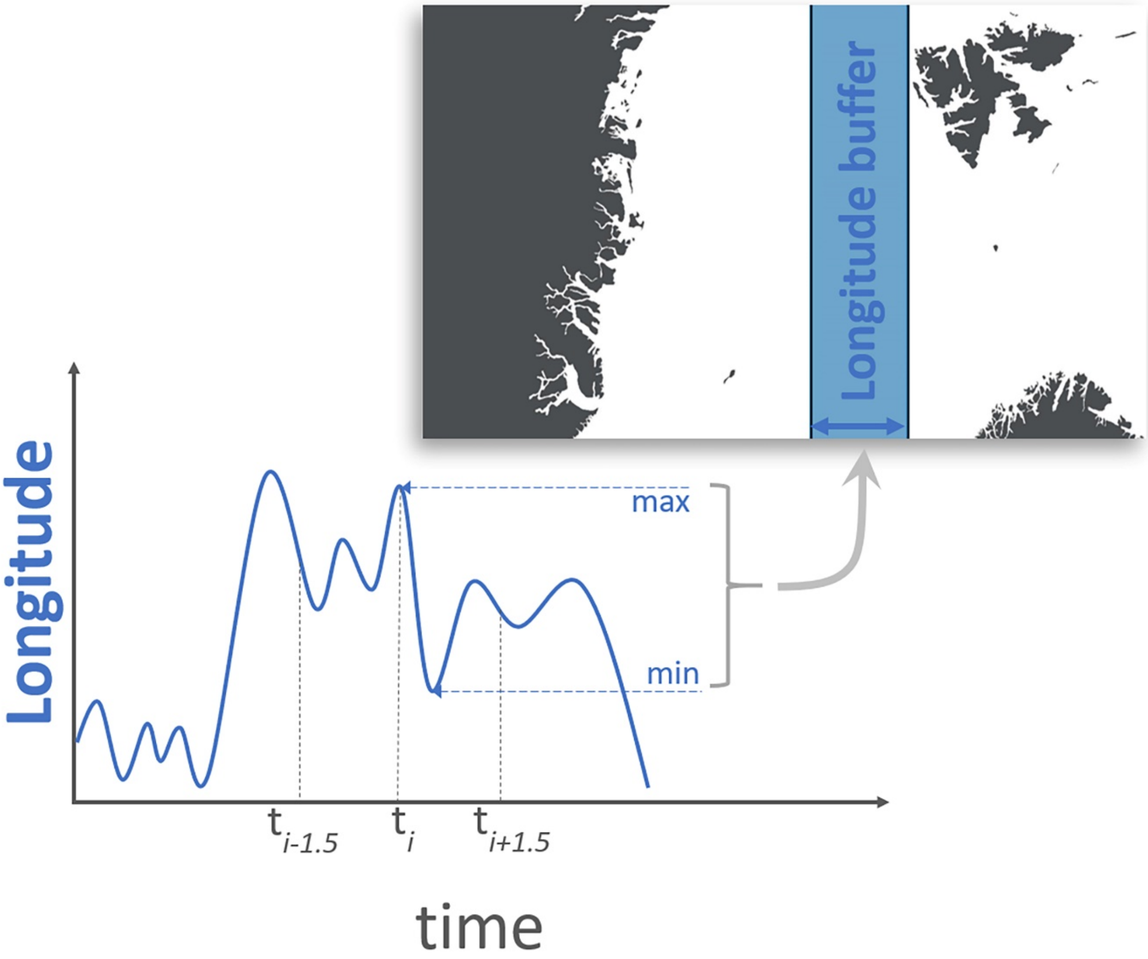


Figure 2.4. Schematic view describing how the longitude buffers are created, starting from the time series of longitudes data that were derived from the light-logger data during equinoxes.

2.2.2.2 Determining species-specific movement rates

As mentioned above, the random movement algorithm on which IRMA is based requires only three parameters (Technitis et al. 2015): the distance between the start point A and the endpoint B, the time budget ($\Delta t_{i,i+1}$) and the maximum rate of movement ($rm_{i,i+1}$) between two successive locations at time *i* and *i+1*. The time budget is always known, as it is the time elapsed between the last and next known locations (i.e. the locations at time *i-1* and *i+1*). The maximum movement rate is not known and is likely to vary among species but also as a function of the time elapsed between two successive locations, with movement rates decreasing as the elapsed time increases. Instead of using a constant parameter value for $rm_{i,i+1}$ we derived models predicting the movement rate as a function of the time elapsed between two successive locations. We first calculated the time elapsed and net displacement among random combinations of pairs of locations. To avoid bias we stratified the dataset by individual and used the same number of locations per individual, i.e. we randomly selected the same number of locations from each individual in the dataset. Individuals with fewer than 180 relocations were not used at this stage in order to have sufficient sample size. All data were then merged into a single dataset before modelling the movement rate as a smoothed function of the time elapsed between two successive locations. We modelled the rate of movement such that $rm_{i,i+1} = bs(\Delta t_{i,i+1})$, where *bs*() is a B-spline polynomial smoother. Because we were interested in movement rates situated in the upper range of possible values, and not average movement rates, we used quantile regression with the 75th percentile as response variable (Fig. 2.5). In addition, the 75th percentile constitutes a rather conservative value for the average movement rate, thus providing IRMA with some flexibility to increase this movement rate in cases where no geographic solutions are found (i.e. when the ranges do not overlap; Appendix 6.1). We used function *rq* from package {quantreg} (Koenker 2018) and function *bs* from package {splines} (R Development Core Team 2018) to run the quantile regression with a smoothing polynomial function. Each model was stored and retrieved later on to provide IRMA with $rm_{i,i+1}$ corresponding to the species considered. This approach is based on the GLS-derived locations and thus integrates the large error associated to this type of positioning. Therefore, the obtained predicted movement rates might differ from the actual movement rates of a given species, although they will reflect the movement rates obtained through GLS positioning and thus yield values that are consistent with the rest of our dataset.

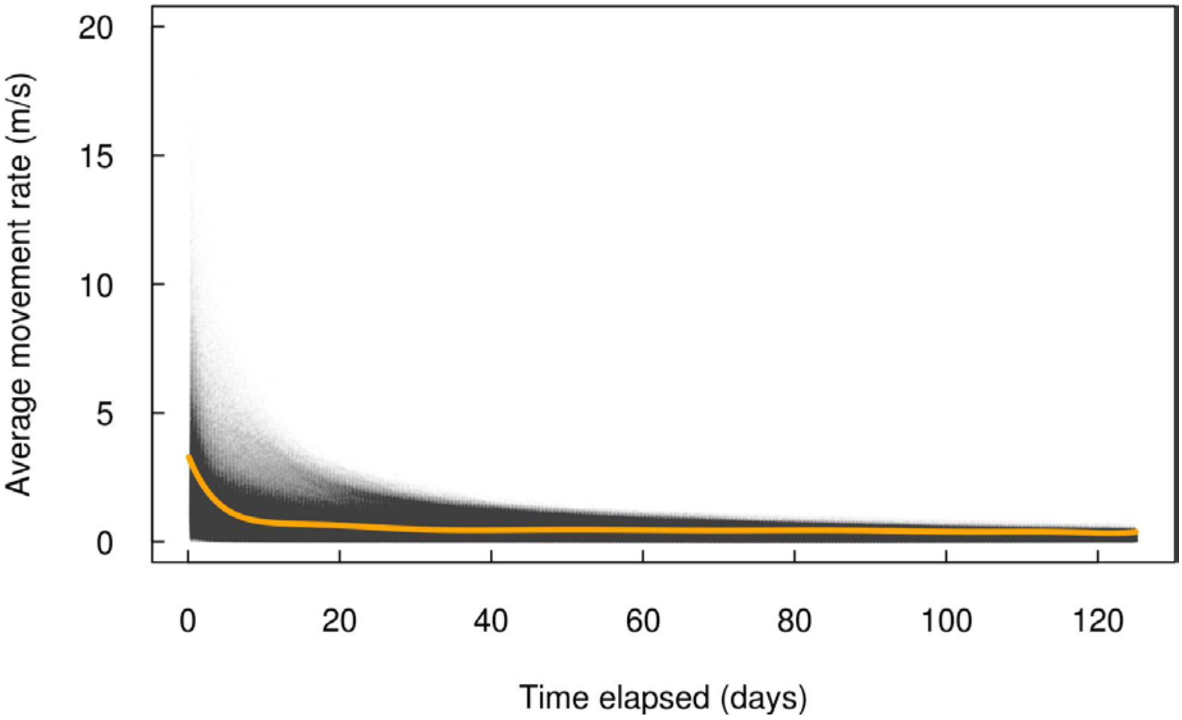


Figure 2.5. Movement rate (orange curve, representing the 75th percentile predicted from a quantile regression model) as a function of time elapsed between two locations, for black-legged kittiwakes. The same modelling approach was used for each species.

2.2.2.3 Extracting wet/dry data

Most of the geolocator models used were equipped with binary conductivity sensors which could be used to record the state (wet, 1 or high conductivity, or dry, 0 or no conductivity) of the logger. Depending on the logger model and programming mode, the wet/dry state was measured at varying intervals (from 3 to 30 sec) and summed over different periods (5 to 240 min), leading to measurements on different scales. All wet/dry data were first standardized such that $x_{st}=x/x_{max}$, where x_{st} is the standardized value and x_{max} is the maximum value that can be measured, i.e.

when the wet/dry sensor indicates “wet” (1) over the entire summing period. Consequently, the standardize wet/dry data can be compared among individuals independently from the logger model. Due to the loggers’ limited storage space, the wet/dry data were sometimes no longer recorded even when light-level data were recorded (i.e. positions were still obtained). In addition, some loggers failed to record any wet/dry data. Therefore, wet/dry data were not available for all individuals and tracking periods. All the available wet/dry data were extracted and stored separately for each individual.

2.2.2.4 Determining breeding dates at colony level

Wet/dry data were used first to determine the start and end dates of the breeding period, thus assuming that periods characterized by longer period of dry data represent the periods when birds are sitting on their nest. Because the wet/dry data are recorded at relatively short intervals, it was necessary to smooth them before further analysis. First, we calculated the daily average for each individual time series of wet/dry data. Then, we calculated a 5-day running mean and running minimum, using functions *runmean* and *runmin* from package {caTools} *(Tuszynski 2018)*. Finally, we applied the Lavielle partitioning algorithm (Barraquand & Benhamou 2008) to identify transitions within each time series (i.e., transitions in running mean and running minimum). We used the function *ts.LaviellePart* from package {adehabitatLT} *(Calenge 2006)*, which partitions time series into *k* sections with similar mean for the response variable. In other words, partitioning was used to distinguish between periods of mostly wet data and periods of mostly dry data. Periods with mostly dry data were assumed to reflect colony attendance. We could then identify the dates of transition between periods with different wet/dry states. For the sake of simplicity, only transition dates between 1 March and 30 September were considered, thereby assuming that the dates identified outside this period did not indicate colony attendance in relation to nesting activity. This can be adjusted for each species. The output from this algorithm is a list, for each colony, species, and individuals, of all breeding dates (start and end) that could be identified. It is important to note here that it was not possible to confirm the breeding status of individual birds in situ in most cases. This means that our results rely on the assumption that all birds returned to the colony and attempted to breed during the breeding season.

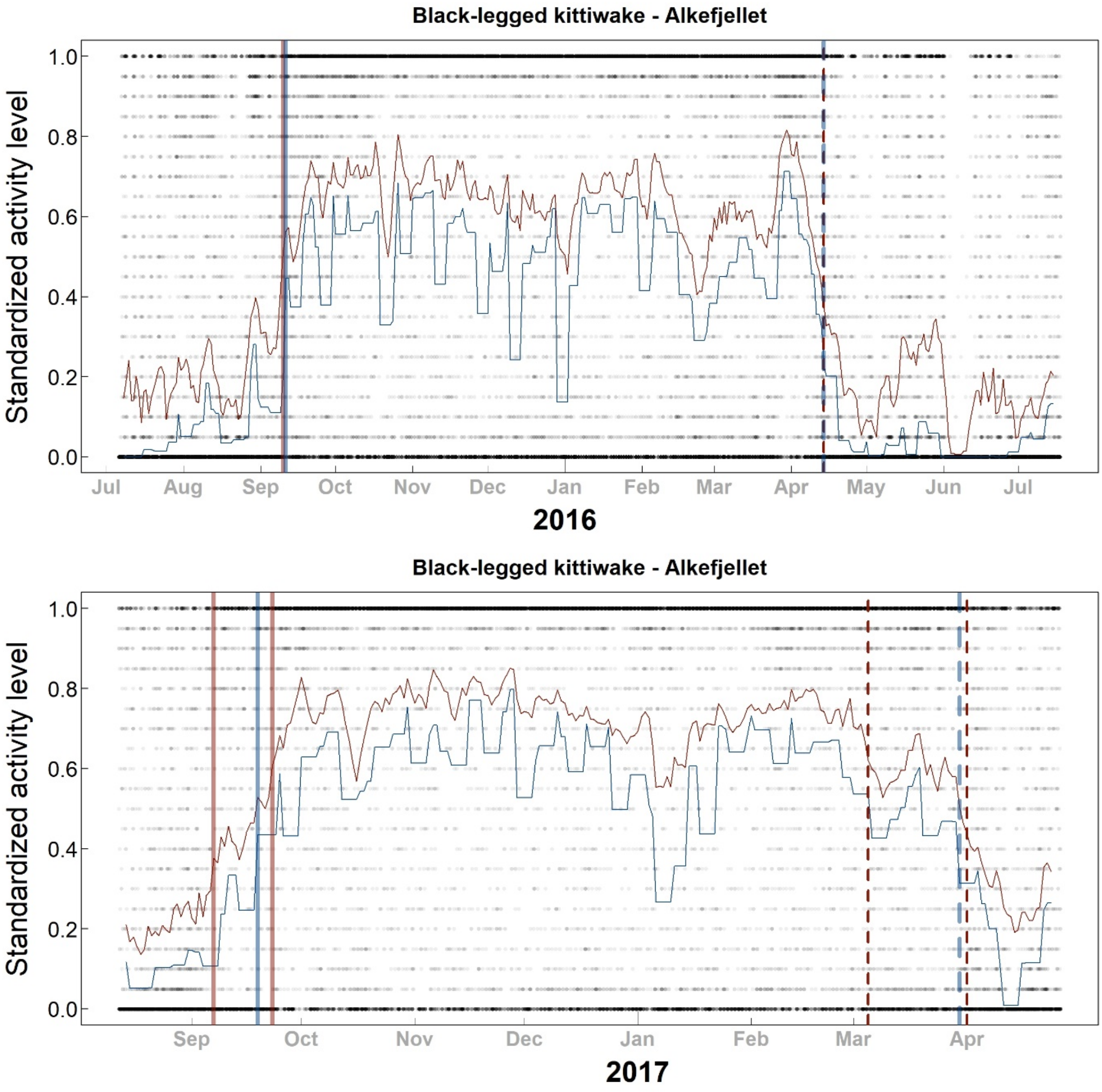


Figure 2.6. Examples of the estimated periods of colony attendance for two individuals of black-legged kittiwakes. The grey dots show the activity level data from the GLS-loggers, after standardizing. The curves represent the 5-day running mean (red) and minimum (blue) of the standardized activity data. The vertical bars indicate the estimated dates of start (dashed lines) and end (continuous lines) of the breeding period, based on the Lavielle partitioning analysis. The upper panel shows an example where the running mean and running minimum yielded the same estimates, while the lower panel showed the discrepancy that can occur when estimates are based on the running mean vs running minimum. In the present approach, all dates were considered for a given colony and year, and the median date among all individuals was used to define the start or end of the breeding period.

2.2.2.5 Determining colony attendance at individual level

During the breeding season, the mobility of birds that are actively breeding is limited, as they behave as central place foragers. Applying IRMA without any kind of constraint would lead to the generation of movements that extend unrealistically far from the breeding colony. Once the breeding period was determined for each colony and year (previous step), a breeding state (potentially breeding/not breeding) could subsequently be attributed to each individual bird, based on the date and its breeding colony. Before running IRMA to create a new location for a given individual at a given time t_i, individual wet/dry data time series were inspected whenever that individual was in a “potentially breeding” state. At this stage, the formatted wet/dry data (see section 2.2.2.3) are used to determine whether that individual was more likely sitting on its nest (mostly dry) or at sea (mostly wet) at a given time. This was done by summarizing the wet/dry data within a 12-hour period around that time (t_i ± 6h). The size of the buffer was chosen to be small enough to be representative of the behaviour around t_i. We used the median as a summary statistic of the wet/dry values over the time period defined by the buffer. A median wet/dry value of <0.4 was considered to represent a “mostly dry” period. This threshold value can be adjusted, but in our case, it means that we work under the assumption that a median wet/dry value of 0.4 or higher indicated that a given bird spent enough time at sea to be considered as not being sitting on its nesting during the 12-h period considered. Every time a period is categorised as mostly dry, the bird’s location is restricted to a buffer area of 50-km radius around the colony at t_i. This buffer area was thus used as the Ppa, within which the new location was then randomly drawn (Fig. 2.7). The value of 50-km represents a compromise between large distance that can be covered by some species, while other species stay only in the vicinity of their colony during the entire breeding season. It could in the future be possible to include species-specific buffer areas in order to achieve higher precision. It is however important to keep in mind that we are dealing with GLS locations with an inherently low precision.

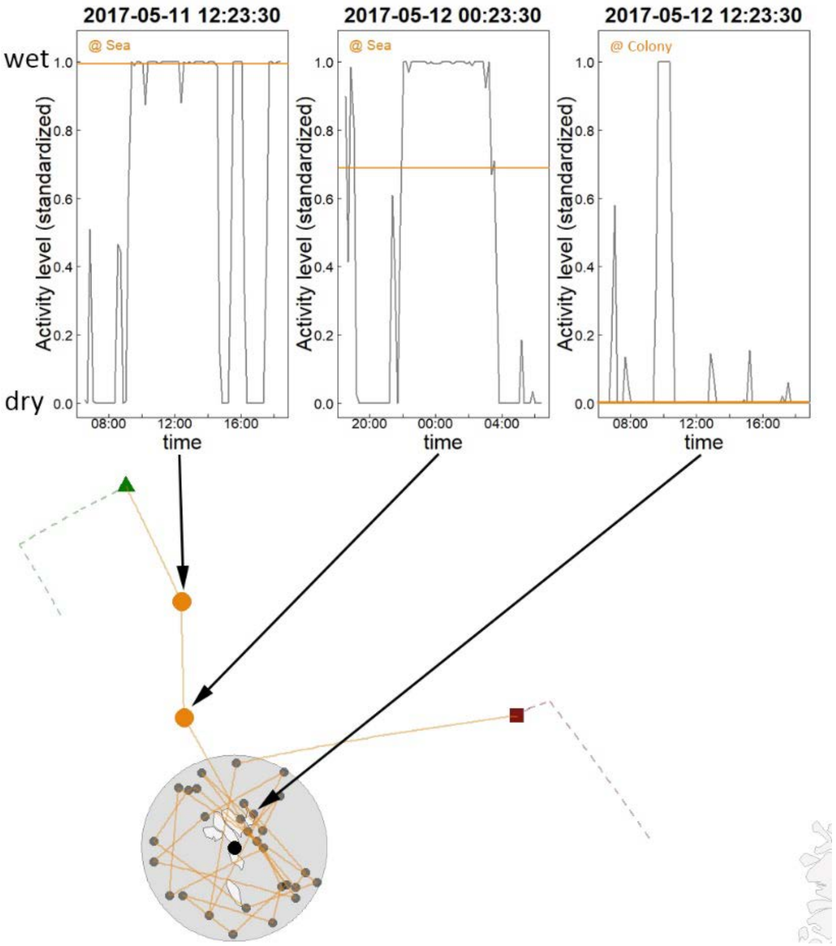


Figure 2.7. Example for a black-legged kittiwake from the Faroe Islands illustrating the approach used to generate new locations during the breeding period. The start and end points of the gap are represented by a green triangle and a red square, respectively. These are GLS-based locations, and the dashed lines show a few steps before and after the gap. The orange points indicate locations that were not constrained within the 50-km buffer around the colony (shaded area) because the corresponding activity data indicated mostly wet (median activity level > 0.4, as shown by the horizontal lines on the two graphs on the left). In contrast, dark grey points indicate locations that were constrained to be close to the colony (median activity level <= 0.4, as shown by the horizontal line on the two graphs on the right). Whether each new location should be constrained or not is based on the standardized activity data (wet/dry) at the corresponding timestamp (± 6 hours).

Conversely, a median wet/dry value of ≥0.4 was considered to represent birds spending a substantial amount of time away from its nest and therefore assumed to be at sea: in this latter situation that individual was allowed to wander farther from its colony by applying IRMA to generate the new location. In other words, in such instances the position of the new location was not constrained in the vicinity of the colony (Fig. 2.7).

2.2.2.6 Excluding areas above landmasses

We used a high-resolution vectorized land mask (NOAA 2018, Wessel & Smith 1996) to exclude all GLS locations that occurred above land and constrain the creation of new locations above ocean areas. In addition, we adjusted the land mask to also exclude small or almost closed seas such as the Mediterranean Sea (closed at the Strait of Gilbraltar) and the Baltic Sea (closed at the level of Gdansk, Poland), where our study species and populations should only rarely be observed. It is safe to assume that the very rare occurrence of a few individuals in those areas during limited periods did not influence our results. It was necessary to restrain the access to these areas to IRMA, to prevent large numbers of individuals from being “trapped” indefinitely in these closed areas once new locations were created there.

2.2.2.7 Excluding areas with high sea-ice cover

Daily sea ice concentration data were retrieved from the NOAA OI SST V2 High-resolution dataset, at a resolution of 0.25°x0.25°, provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. Here, we used the contour lines delimiting areas with > 50% concentration of sea ice to create daily polygons that were used as exclusion areas, similarly to the land mask. We therefore assumed that areas with high concentrations of sea-ice represented unsuitable habitat for the six study species.

2.2.2.8 Excluding areas during polar day/night periods

Locations cannot be derived from light loggers during polar night/day periods, due to the lack of sunset and sunrise events. However, light levels are still being recorded during those periods, and can thus indicate whether a given bird remained continuously north of the Arctic circle, i.e. in the polar night area during winter (continuous low light level) or in the polar day area during summer (continuous high light level). This information can thus help us constrain the creation of new locations to areas within or without the polar night/day area.

2.2.3 Applying IRMA: exceptions and limitations

Once a Ppa has been determined, it will be geographically constrained (i.e. clipped) using additional information extracted either from the light-loggers, e.g. wet/dry data, or from environmental datasets, e.g. sea ice concentration. IRMA integrates all the available extra information to further limit the geographic extent of the Ppa to the smallest possible area (Fig. 2.8 and Appendix 6.1). The following set of rules was used to determine when and how IRMA should be applied:

1. Ideally, all gaps should be filled in, i.e. all missing locations should be replaced in order to reduce the bias in the dataset as much as possible. In practice, however, IRMA was not applied when gaps in the positional dataset were longer than three months. Although it is technically possible to generate movements over any period of time, longer gaps are more likely to cover phenological periods with very different movement behaviours and can thus not provide enough information to IRMA to produce relevant tracks. The tracks for some species (e.g. Atlantic puffins) were characterized by large data gaps spanning both the breeding season and equinox periods, i.e. > 3 months. Tracks in such instances could thus not be reconstructed throughout the entire annual cycle.

2. Because our approach is based on conservative values for the starting set of parameters (e.g. maximum speed, longitude buffer), IRMA cannot always find any geometric solution given that original set of parameters. The entire approach relies on the possibility to calculate a Ppa between two locations A and B (Fig. 2.2). When the maximum ranges r_{i-1} and r_{i+1} (Fig. 2.2) do not overlap, there is naturally no solution for determining a Ppa. This can occur whenever the movement rates used are too low. In other instances, e.g. during an equinox period when longitude data are available, a Ppa can be created that does not overlap the longitude buffer (this can occur owing to GLS-based errors on the longitude data and/or on the known locations A and B). It is therefore necessary to provide some flexibility to IRMA so that it can calculate a Ppa in such instances, given the information at hand. This flexibility is provided by allowing the algorithm to incrementally increase the value of key parameters until it can determine a Ppa from which a new location can be randomly drawn. This is done first by increasing the movement rates until a solution is found or until the speed threshold is met, then by increasing the longitude buffer until a solution is found, or until a maximum period of ten days is used. If that threshold of ten days is reached and still no solution is found, the longitude buffer is ignored. The maximum movement rate has a threshold of 20 m/s, and no location is created if it leads to movement rates beyond that threshold. It is important to remember that because we deal with GLS-based locations having an inherent positional error of several hundreds of km, some locations will generate unrealistic movement rates. However, IRMA will always use the most conservative values for the parameters, starting from the movement rate value estimated from the GLS data, and increasing it to higher values only in cases where no geometric solution is possible (see Appendix 6.1).

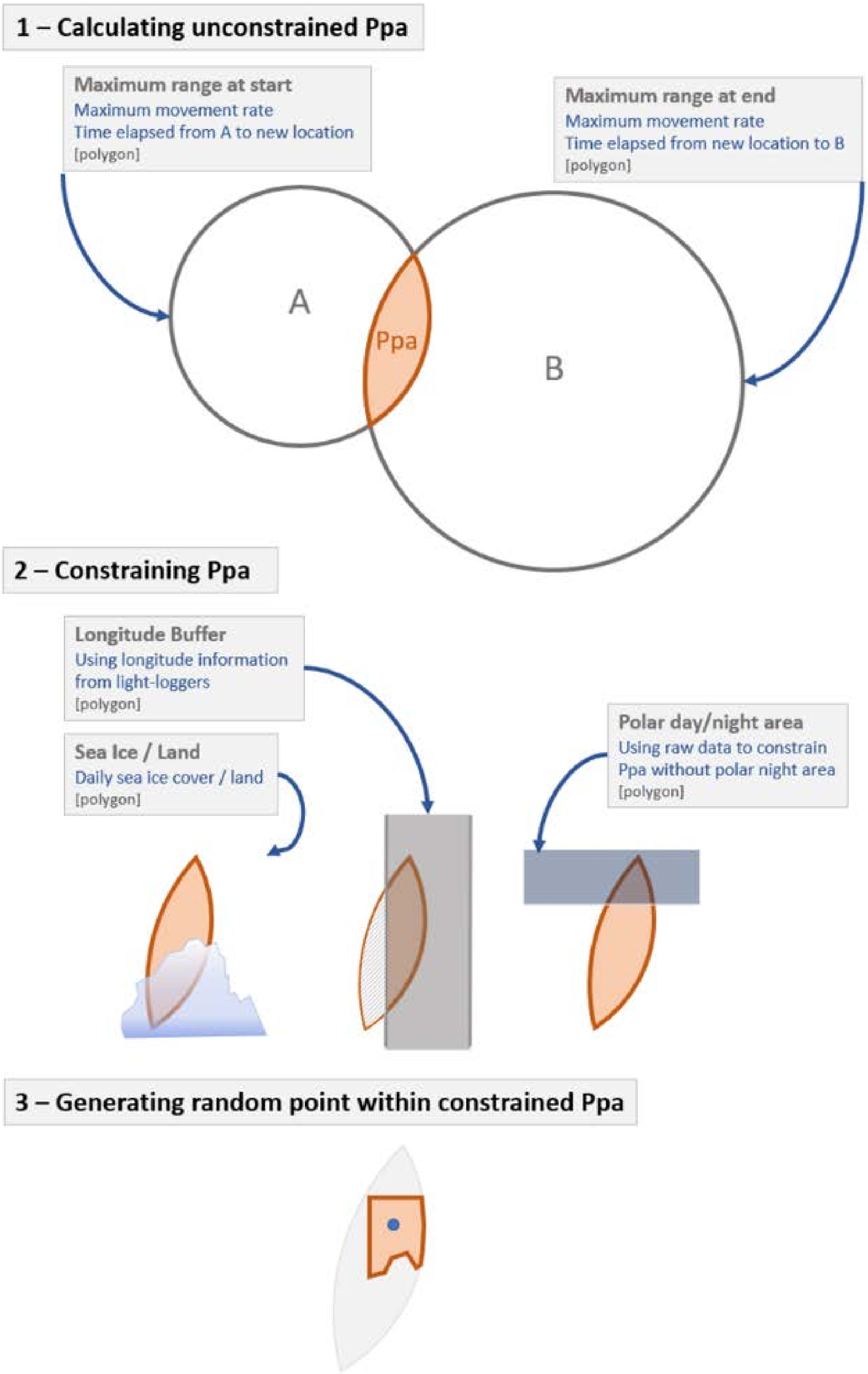


Figure 2.8. Schematic workflow detailing the successive steps undertaken by IRMA when generating a new location. First the Potential Point Area (Ppa) following Technitis et al. (2015) and then various constraints are applied before a new location is randomly drawn within the resulting area. All constraints are based on additional available information used to generate masks (i.e. polygons). Each mask is applied directly onto the Ppa, thus excluding sections where it is assumed a given location could not occur at a given time. See Appendix 6.1 for more details on the logical steps within IRMA.

3 Habitat modelling

3.1 Introduction to Species Distribution Models (SDM)

Species Distribution Models (SDMs) have become a widely used tool for mapping the habitats of wild animals and plants and are used in various management applications such as conservation planning of endangered species, impact assessments of human activities and spatial planning of protected areas (Guisan & Thuiller 2005; Elith & Leathwick 2009). In short, SDMs are empirical models that relate data of species occurrence to data of relevant environmental predictors (Guisan & Zimmermann 2000). The relationship is estimated by various statistical methods and is expected to reflect the environmental niche utilized by the species. If the realized niche is constant across space, the relationship can be used to predict the spatial distribution of the species in areas where the environmental variables are known. Given the environmental conditions, the resulting map represents a quantitative estimate of the distribution of the species and is frequently used to inform decision makers in management processes (Guisan & Thuiller 2005; Elith & Leathwick 2009).

The interpretation and applicability of the spatial predictions derived from SDMs depend on several assumptions as well as on the nature of the input data (Guisan & Thuiller 2005; Elith & Leathwick 2009; Guillera-Arroita et al. 2015). To capture the actual habitat, it is vital that the environmental predictors cover the most important factors that influence the distribution of the species. Failure to identify important predictors will result in poor model fit and low predictive power. The environmental predictors can exert direct or indirect impacts on the distribution of the species. Most often, predictors used in SDMs encompass open-access data of physical or biotic parameters available for extensive areas by for example remote sensing (e.g., temperature, salinity, bathymetry, currents, primary production). In many cases, it is assumed that these predictors exert indirect effects on the species through for example affecting the availability of food resources (for seabirds and GLS see e.g. Torres et al. 2016; Krüger et al. 2017; 2018; Legrand et al. 2016). Biotic interactions such as competition, predation and facilitation are likely to have strong direct impact on species distribution (e.g. Lima 2002), however such data are often difficult to collect, and in cases where they are missing and not indirectly accounted for by other variables, the result could be poor model fit and misleading habitat predictions.

Predictions from SDMs can either involve extrapolations, in which the predictions are done for areas or time periods not covered by data, or they can be interpolations, in which the predictions are done within the area and time period covered by the data. Habitat predictions from extrapolations can only provide reliable estimates when the environmental niche estimated by the model is transferrable from the modelled habitat to the new setting. Because new areas or time periods represent a novel situation that could alter the relationship between the environment and the distribution of the species, the assumption of transferability is not always met (e.g., Torres et al. 2018). However, given that the SDM provides a realistic estimate of the environmental niche utilized by the species, extrapolations can give valuable information of habitat suitability in new areas or time periods. Such analyses have become a highly important tool in the assessments of how climate change might impact the habitats of vulnerable species (e.g., Elith et al. 2010; Krüger et al. 2018).

Related to the issue of interpolation and extrapolation, is the equilibrium assumption in SDMs (Guisan & Thuiller 2005; Elith et al. 2010). To be valid, the predictions from SDMs presuppose that the spatial distribution of the species has reached an equilibrium with the environment. A species has *not* reached an equilibrium in cases in which the abundance is still increasing in favourable areas and decreasing in unfavourable areas. The equilibrium assumption is accordingly not met for migrating or range-shifting species where the species continuously invade and encounter new habitats. Because the individuals have not settled in the habitat, the SDM do not reflect the species' "true" environmental niche. The predictions from the SDM will consequently be biased by space and will at best only be valid for a limited time period. This setting does certainly apply to migrating seabirds that migrate between the breeding colony and seasonal foraging areas at sea. The birds' selection of environmental niches is constrained (or biased) by a geographical area (the colony), and it might therefore be important to control for distance to the colony and season in the SDMs.

Species data in SDMs can either be "presence-only data", "presence-absence data" or "abundance data". In the case of "presence-only data", the data are observations of the occurrence of a species. Tracking data such as GLS data are examples of presence-only data because the data consists of geographical positions of presences only. Survey data where the species is observed or counted in specified areas or along transects, represent another type of data since the data consists of presence-absence or abundance recordings. In SDMs of survey data, the abundance or presence/absence at a given location is related to the environmental predictors. Data from seabirds at sea surveys were for example used in SDMs to map the distribution of different seabird species in Norwegian waters (Fauchald et al. 2011; Fauchald 2011). The predictions from such SDMs can, if the observations are unbiased, represent the true distribution of the abundance of the species.

In contrast to SDMs of survey data, it is necessary to introduce background points to contrast the recorded presences in SDMs of presence-only data (Phillips et al. 2009). Such models are therefore often referred to as presence-background models (Guillera-Arroita et al. 2015). The intention of the background data is to provide a representative sample of the set of environmental conditions available to the species, and then compare this set of available environmental niches to the measured presences (Phillips et al. 2009). Commonly, the background points are a random sample of locations within the distributional range of the species. Because the background points do not contain any information of occurrence or abundance, predictions from presence-background models cannot represent the true abundance or probability of occurrence of a species. It is therefore important to note that the predictions from such models represent relative values and should be interpreted as the relative likelihood of species occurrence (Guillera-Arroita et al. 2015). However, the relative likelihood is proportional to the probability of occurrence (Guillera-Arroita et al. 2015), and for seabird tracking data, where the occurrence of individuals from a specific breeding population is recorded, it is possible to translate the predictions from the SDM into estimates of abundance by weighting the relative likelihood values from the SDM with the size of the population which the tracked individuals represent (see Chapter 3).

When collecting spatial data for SDMs, sampling bias occur when the probability of sampling a location is not the same across the habitat. For abundance and presence-absence data, sampling bias will result in less precise estimates of occurrences in areas that have less probability of being sampled (i.e., areas that are under-sampled). Although the precision of the estimate is negatively affected, sampling bias will in this case not result in a biased estimate. However, for presence-only data, sampling bias has a more critical consequence (e.g. Guillera-Arroita et al. 2015). This is because the probability of sampling a location has a direct positive impact on the probability of recording a presence, and hence on the estimates from the presence-background model. Accordingly, sampling bias result in bias in the estimates from the SDM and consequently wrong (or biased) predictions. For GLS data on seabirds, spatial sampling bias arise when Arctic areas are not sampled during polar night and midnight sun conditions, and when positions over land and fast sea ice are removed. Temporal sampling bias arise during equinox periods when positions are unreliable and removed from the sample. It is essential to remove such biases, and the purpose of IRMA, described in the previous chapters, is to remove as much of this sampling bias as possible.

3.2 SDMs of Northeast Atlantic seabirds

3.2.1 General modelling approach

We applied SDMs on the SEATRACK datasets of six pelagic seabird species breeding in the Northeast Atlantic: Northern fulmar, black-legged kittiwake, common guillemot, Brünnich's guillemot, little auk and Atlantic puffin.

GLS loggers were attached and retrieved from breeding birds in 25 different colonies in the Northeast Atlantic (see Figure 1.1) from 2012 to 2017 (note that two colonies from the original dataset (Table 2.1) were excluded owing to a too small number of birds tracked (< 4 individuals)). In total 1523 annual tracks were recorded (Table 3.1). Note also that not all tracks completed the entire annual cycle and that the same individuals in several cases were recorded in more than one year. Detailed sample sizes (number of individuals, number of GLS positions and number of IRMA positions for each SDM is given in Appendix 6.2. Pelagic seabirds have a seasonal migratory behaviour, migrating between the area around their breeding colony where they stay during spring and summer, and various feeding grounds at sea where they stay during the non-breeding period (see e.g., Frederiksen et al. 2012; Fort et al. 2013; Tranquilla et al. 2013; Fayet et al. 2017; Linnebjerg et al. 2018). Migratory behaviour makes SDMs challenging because when birds enter new habitats, they are less likely to be in spatial equilibrium with the environment (see section 3.1). To account for seasonal changes, we conducted separate SDMs for each month. Moreover, to account for possible differences in migratory pattern among birds from different breeding colonies, we conducted separate SDMs for each colony. Finally, because we wanted the models to reflect the seasonal distribution irrespective of year, we pooled data from all years in the analyses. In other words, SDMs were run for all combinations of month/colony with data from all years pooled.

Table 2.1 - Number of annual tracks for each colony and species. Colonies with fewer than 4 birds are not included.

Species	Colony	Number of tracks
Northern fulmar	Alkefjellet	5
	Bjørnøya	26
	Eynhallow	44
	Faroe Islands	11
	Jan Mayen	33
	Jarsteinen	8
	Langanes and Skjalfandi	46
Black-legged kittiwake	Alkefjellet	20
	Anda	54
	Bjørnøya	40
	Cape Krutik	33
	Cape Sakhanin	5
	Faroe Islands	21
	Franz Josef Land	54
	Hornøya	49
	Isfjorden	29
	Isle of May	35
	Kongsfjorden	34
	Langanes and Skjalfandi	27
	Røst	43
	Runde and Ålesund	25
	Sklinna	36
Common guillemot	Bjørnøya	35
	Cape Gorodetskiy	4
	Faroe Islands	7
	Grimsey	9
	Hjelmsøya	28
	Hornøya	37
	Isle of May	32
	Jan Mayen	27
	Langanes and Skjalfandi	27
	Sklinna	38
Brünnich's guillemot	Alkefjellet	23
	Bjørnøya	29
	Cape Gorodetskiy	16
	Cape Sakhanin	41
	Franz Josef Land/Oranskie Islands	11
	Grimsey	12
	Hornøya	46
	Isfjorden	17
	Jan Mayen	42
	Langanes and Skjalfandi	17
Little auk	Bjørnøya	29
	Franz Josef Land	36
	Hornsund	53
	Isfjorden	15
	Kongsfjorden	16
Atlantic puffin	Anda	15
	Faroe Islands	6
	Grimsey	28
	Hjelmsøya	20
	Hornøya	51
	Isle of May	39
	Papey	19
	Røst	57
	Runde and Ålesund	13

3.2.2 Modelling method

Several modelling techniques are available for fitting SDMs of presence-only data (see e.g., Guisan & Thuiller 2005; Elith & Leathwick 2009). These include regression techniques such as Generalized Linear Models (GLM) and Generalized Additive Models (GAM) and machine learning techniques such as Boosted Regression Tree (BRT) and Maximum Entropy (MaxEnt). Ensemble models that combine the output from several modelling techniques has also been advocated (Scale et al. 2016).

In the present study we decided to use the GAM technique (Hastie & Tibshirani 1990; Wood 2006). GAM is a well-proven and computational efficient regression method that relates a linear response variable to smooth functions of predictor variables. The smooth functions can model non-linear relationships, and GAM is accordingly more flexible than GLM which is based on linear relationships. Compared to the machine learning techniques, GAM offers a more straightforward interpretation of the fitted model. Machine learning techniques will in some cases involve complex relationships and interactions. This could yield higher precisions in the predictions but might on the other hand not yield any intelligible functional relationships. In studies comparing the various methods, GAM has a similar performance as machine learning techniques (see Wisz et al. 2008; Elith et al. 2010; Scales et al. 2016). GAM performs somewhat poorer with small sample sizes (< 20 presences) but is among the best methods for larger samples (>50 presences) (Wisz et al. 2008). The present models were based on from 47 to 6260 presences (median: 1884) (see Appendix 6.2.)

We used logistic regression GAM to fit the binomial presence/background data to the environmental predictors. Models where fitted using the {mgcv} package in R. We used thin plate regression spline as smoothing basis and generalized cross-validation was used to optimize the degree of smoothing. The *predic*t function was used to predict the output from the model on a grid with known environmental variables.

3.2.3 Presence and background data

GLS data are presence-only data, and sampling bias related to factors discussed in chapter 1 will accordingly generate bias in the predictions from the SDMs (see section 3.1). To minimize sampling bias, we applied IRMA (see section 2) to the GLS dataset and used the combined dataset in the SDMs. The aim of IRMA is to provide informed random presences in areas and periods where the sampling process fail to allow recordings and thus introduces biases in the dataset. Accordingly, the proportion of IRMA points in the sample was high during the equinox periods and for species residing in northern areas during polar night and midnight sun. Number of IRMA and GLS presences in each SDM is given in Appendix 6.2.

In SDMs of presence-only data, it is necessary to introduce background points to contrast the recorded presences in the analyses (Elith & Leathwick 2009; Barbet-Massin 2012). It is essential that the background data is a representative set of environmental conditions available to the species (Phillips et al. 2009). In the present analyses, background data points were selected randomly from an area defined by the minimum convex polygon of all recorded positions of the species (including IRMA points) using function *mcp* from package {adehabitatHR}. To the polygon, we added a buffer with a width equal to 10% of the radius of a circle defined by the area of the minimum convex polygon (see Figure 3.1). We assumed that the seabirds did not occur over land and fast sea ice, and we accordingly removed areas covered by land and areas with a sea-ice concentration > 50% using monthly sea-ice data from NSIDC (see section 3.2.4).

The SDMs were conducted separately for different months but data could include observations from more than one year (see section 3.2.1). Because some of the environmental variables varies between years (i.e., ice-cover, sea surface temperature, sea surface height, primary production; see section 3.2.4), we included background data from each year represented by the

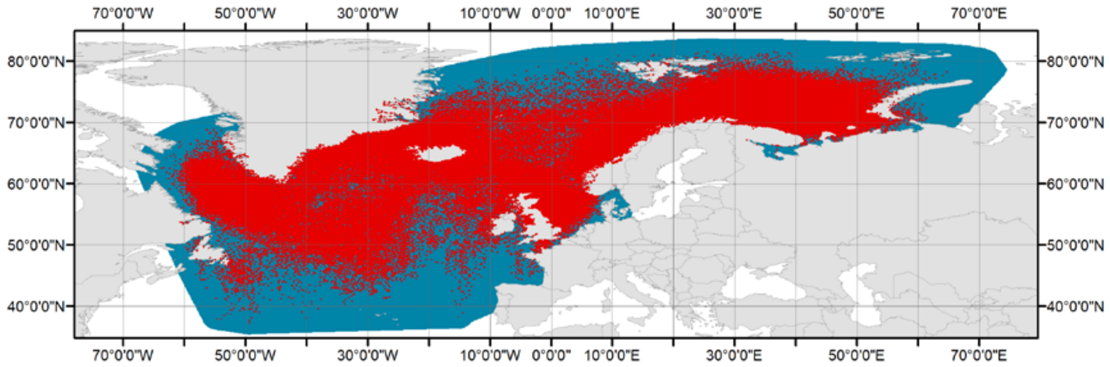


Figure 3.1. All presence recordings of Atlantic puffin (red), and the defined habitat from where background points where drawn (blue). The habitat was defined by the minimum convex polygon of all presence point plus a 10% buffer (see text). In addition, land masses and sea ice concentration >50% (depending on year and month) where removed. Note that ice-covered areas were not removed in this figure.

presence data. The performance of regression techniques such as GAM increases with the number of background points used in the analyses, and Barbet-Massin et al. (2012) recommended the use of a large number (e.g., 10000 points) without weighing for presences and backgrounds. Accordingly, we drew 10000 background points randomly from each habitat/year represented in the presence data. Number of background points in each SDM is given in the Appendix 6.2. Note that some background points might be removed due to missing environmental variables, thus the background sample size might be slightly less than a multiply of 10000.

3.2.4 Model data: Environmental predictors

Ideally the environmental predictors should cover the most important environmental features impacting the spatial distribution of the species (see section 3.1). We used available datasets that cover the study region and that could impact the habitat suitability for seabirds. The predictors encompass several oceanographic features that are important for shaping the productivity and other characteristics of the marine pelagic ecosystem (Table 3.2). During the annual cycle, the birds migrate between the breeding colony and the marine pelagic habitat. The breeding colony is therefore a spatial constraint that is important for shaping the spatial distribution of the birds, and we consequently included distance to the colony (ColD) as a predictor in the analyses. Finally, during the initial analyses, we realized that we in some cases predicted high probability of occurrence north in the Barents Sea without having any observed presences in the area. We concluded that the environmental conditions in this area could be similar to the preferred habitat further south, but that the birds were limited by other factors not accounted for in the analyses. Seabirds are visual predators, and daylength could be one factor that restricts the birds' foraging activity during winter in the far north. We therefore included daylength (Daylen) as a predictor in the analyses, but only in the period from October to February. Environmental data (from January 2016) are plotted in Figures 3.2-3.4.

All predictors were spatially adjusted to fit a geographic spatial raster of the study area from -78°E to 80°E (longitude) and from 35°N to 85°N (latitude) with a spatial resolution of 0.1 x 0.1 degrees. Data with a lower spatial resolution or configuration were either disaggregated using the function *disaggregate* from the package {raster} with a bilinear local interpolation or resampled using the function *resample* with a bilinear local interpolation. We investigated colinearity in the predictor dataset by calculating the correlation matrix between predictors from background data for different months and species (Table 3.3). Some co-linearity was present; SST was more or less correlated with all other variables except Front. Maximum correlation was found between CoastD and Depth (r =0.81; R^2 = 0.66). However, to maximize the predictive value of the models, we decided to keep all variables in the model when feasible (see modelling approach in section 3.2.5).

Table 3.2 - Environmental variables used in the Species Distribution Models (SDM).

SST -Sea Surface Temperature (°C)
Front -Gradient in SST. Local standard deviation in SST in the 0.3°N X 0.7°E area around 0.1°N X 0.1°E focal cells.
Dataset:
Product Id: NOAA_OI_SST_V2
Organization: NOAA/OAR/ESRL PSD, Boulder, Colorado, USA
Url: https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html
Reference: Reynolds et al. (2002)
Resolution: 1°N X 1°E, monthly

Dice -Shortest distance to ice edge (km). Ice edge defined as sea ice concentration between 30% and 70%. Transformation: Log10(x+1).
Icemask -A mask for masking ice covered areas (ice concentration > 50%) from the habitat.
Dataset:
Product Id: NOAA_OI_SST_V2
Org: NOAA/OAR/ESRL PSD, Boulder, Colorado, USA
Url: https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html
Resolution: 1°N X 1°E, monthly

Adt -Absolute dynamic topography (m). Sea surface height above geoid. Monthly average.
Dataset:
Product Id: SEALEVEL_GLO_PHY_L4_REP_OBSERVATIONS_008_047
Org: Copernicus Marine Environment Monitoring Service
url: http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=SEALEVEL_GLO_PHY_L4_REP_OBSERVATIONS_008_047
Resolution: 0.25°N X 0.25°E, daily

Prim -Annual net primary production (mg C/m^2 / day). Annual sum. Transformation: Log10(x/1000+1).
Dataset:
Product Id: Standard VGPM
Org: Oregon State University, Ocean Productivity
url: http://www.science.oregonstate.edu/ocean.productivity/index.php
Reference: Behrenfeld & Falkowski (1997)
Resolution: 1/6°N X 1/6°E, monthly

Depth -Bottom depth (m). Transformation: Log10(x+1).
Edge -Gradient in bottom depth. Smoothed local standard deviation in Depth in the 1.1°N X 2.1°E area around 0.1°N X 0.1°E focal cells.
Dataset:
Product Id: ETOPO2v2 2006
Org: DOC/NOAA/NESDIS/NGDC
url: https://www.ngdc.noaa.gov/mgg/fliers/06mgg01.html
Reference: National Geophysical Data Center (2006)
Resolution: 2/60°N X 2/60°E

CoastD -Shortest distance to coast (km). Transformation: Log10(x+1).

ColD -Shortest distance to breeding colony (km). Transformation: Log10(x+1).

Daylen -Mid-month daylength from nautical dawn to nautical dusk (hours). Nautical dawn and dusk defined by sun 6 degrees below horizon. Calculation: {suncalc} R

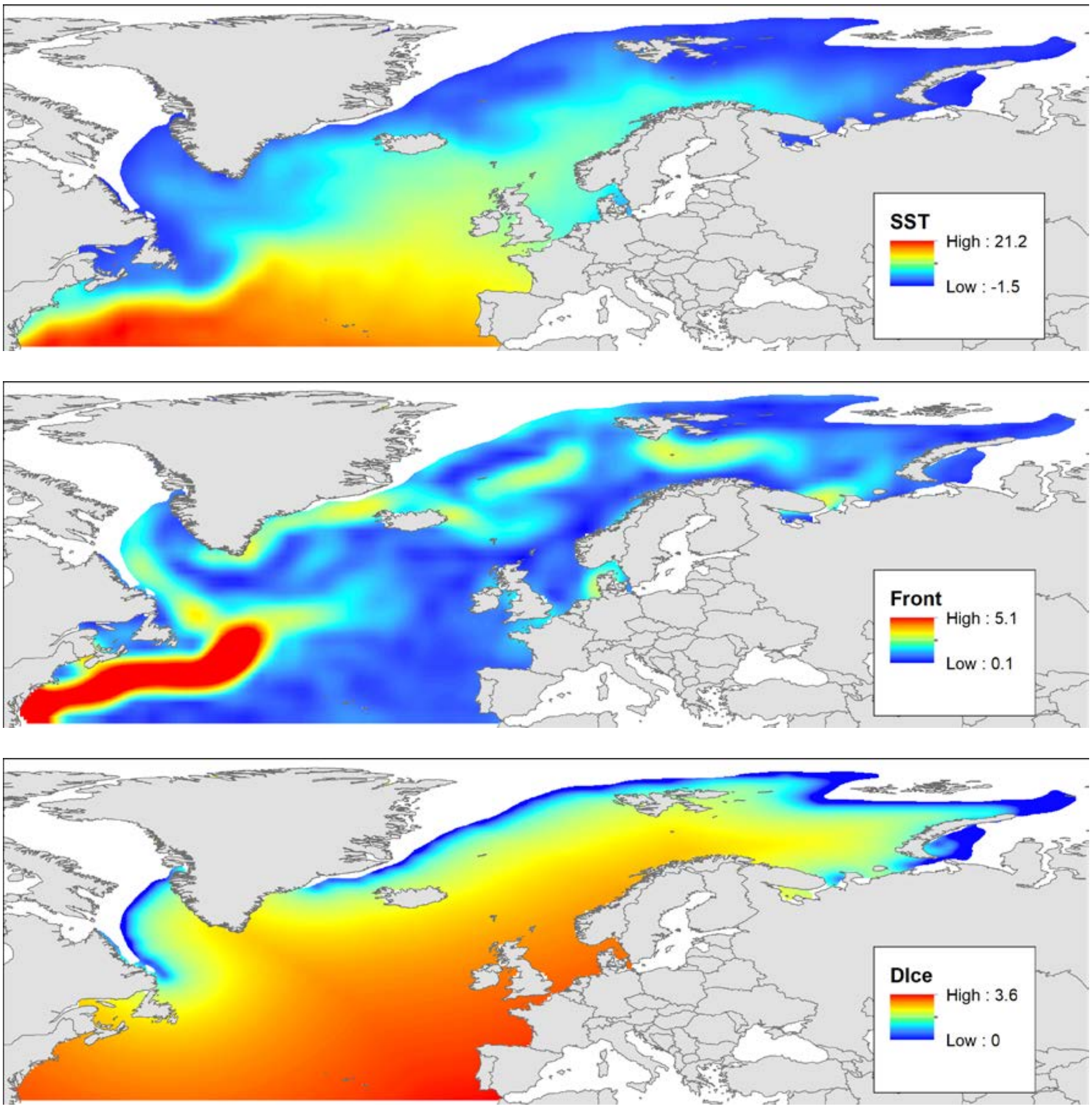


Figure 3.2. Environmental predictors, January 2016. From top: Sea surface temperature (SST), Gradient in SST (front), Distance to sea ice (Dice). Definitions and datasets are given in Table 3.2.

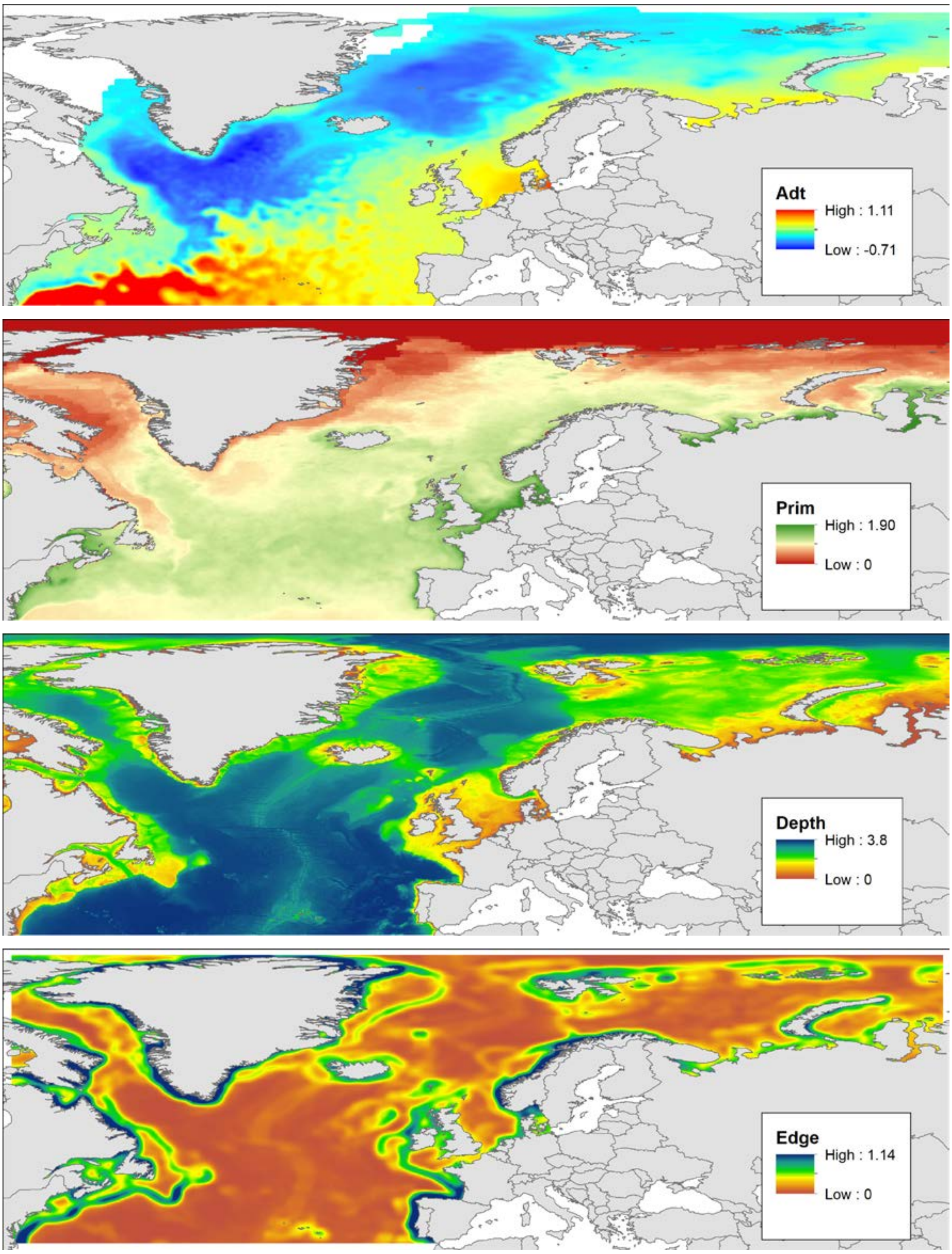


Figure 3.3. Environmental predictors, January 2016. From top: Absolute dynamic topography (sea level above geoid) (Adt), Annual net primary production (Prim), Bottom depth (Depth), gradient in bottom depth (Edge), Definitions and datasets are given in Table 3.2.

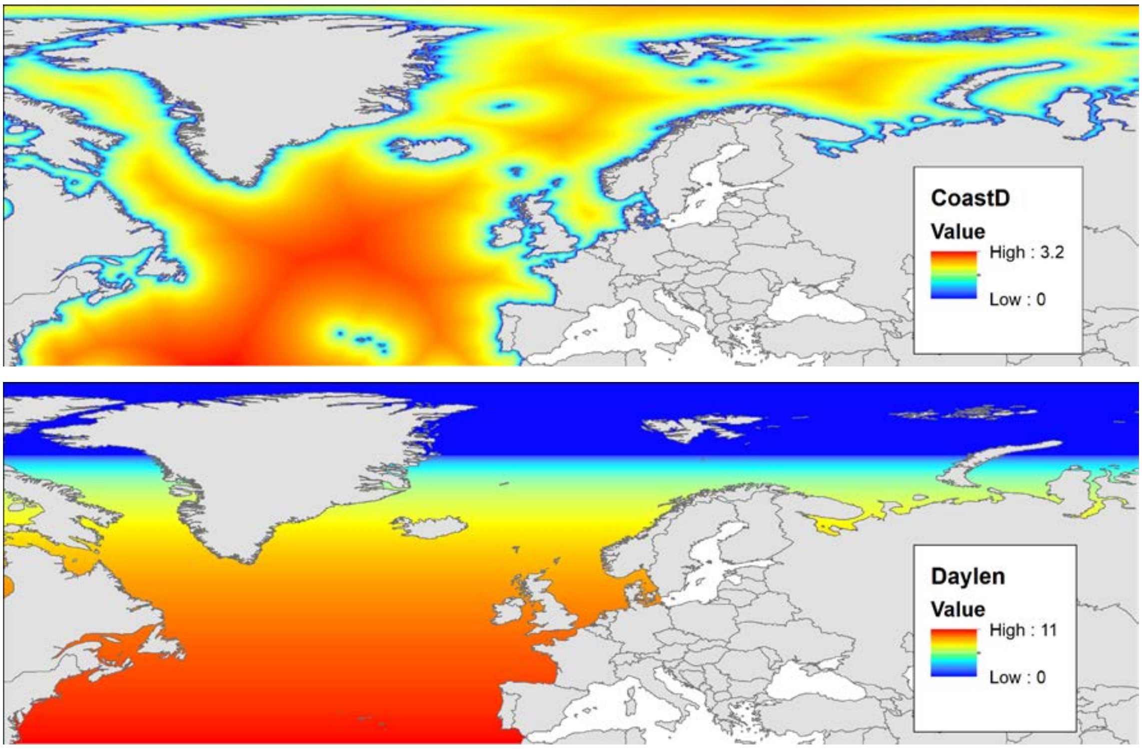


Figure 3.4. Environmental predictors. From top: Distance to coast (CoastD), Daylength (January) (Daylen). Definitions and datasets are given in Table 3.2.

Table 3.3. Pearson's correlation matrices of environmental predictors. A) Correlations calculated on background points from Brünnich's guillemot's habitat in January (n=30000). B) Correlations calculated on background points from black-legged kittiwake's habitat in July (n=40000).

A)	SST	Front	Dice	Adt	Prim	Daylen	Depth	Edge	CoastD
SST	1	0.1	0.73	0.61	0.46	0.59	0.35	-0.25	0.43
Front	0.1	1	0.03	0.01	0.09	0.21	0.17	0.01	0.22
Dice	0.73	0.03	1	0.36	0.58	0.63	0.2	-0.13	0.31
Adt	0.61	0.01	0.36	1	0.32	0.24	-0.28	0.15	-0.09
Prim	0.46	0.09	0.58	0.32	1	0.49	0.05	-0.09	0.16
Daylen	0.59	0.21	0.63	0.24	0.49	1	0.27	-0.04	0.34
Depth	0.35	0.17	0.2	-0.28	0.05	0.27	1	-0.62	0.8
Edge	-0.25	0.01	-0.13	0.15	-0.09	-0.04	-0.62	1	-0.73
CoastD	0.43	0.22	0.31	-0.09	0.16	0.34	0.8	-0.73	1

B)	SST	Front	Dice	Adt	Prim	Depth	Edge	CoastD
SST	1	-0.05	0.79	0.76	0.42	0.42	-0.2	0.4
Front	-0.05	1	-0.06	-0.18	0.07	-0.07	0.15	-0.02
Dice	0.79	-0.06	1	0.45	0.56	0.35	-0.16	0.38
Adt	0.76	-0.18	0.45	1	0.16	0.16	-0.08	0.18
Prim	0.42	0.07	0.56	0.16	1	0.02	-0.03	0.12
Depth	0.42	-0.07	0.35	0.16	0.02	1	-0.64	0.81
Edge	-0.2	0.15	-0.16	-0.08	-0.03	-0.64	1	-0.74
CoastD	0.4	-0.02	0.38	0.18	0.12	0.81	-0.74	1

3.2.5 Model specifications and diagnostics

SDMs of presence-background data for each species, colony and month were fitted using the function *bam* from package {mgcv} in R. The probability of presence/background was modelled using a logit link with a binomial distribution.

Initially a full model including all predictors was fitted (Model 1; Table 3.4). In some cases, the model predicted clearly false presences in the far east and/or the far west of the study area (see example of diagnostic plots in Figure 3.5). This was likely due to the presence of favourable environmental conditions in areas the birds were unable to reach. In these instances, we included longitude as a variable in the full model (Model 2). In cases where the models failed to converge, we tried a simplified model with fewer predictors. In addition, we introduced a limit to the degree of smoothing by limiting the maximum number of knots to four (k=4) for each predictor (Model 3). Finally, sample sizes were often small during the breeding period, and in cases were Model 1 and 3 failed, we used a simple model including only distance to the colony (ColD) with k=3 as a predictor (Model 4).

In total, we conducted 636 successful SDMs for the six focal species (Table 3.5). Models are summarized in the Appendix 6.2. For some colonies and periods, sample size was too small for modelling. This was particularly true during the summer months for the high-arctic colonies of little auk, Brünnich's guillemots, northern fulmar and black-legged kittiwakes. For months and colonies with a small sample size, we adopted the following strategies:

1) Include data 15 days before and 15 days after the month modelled: Applied in 29 cases.
2) Include data from the nearest colony: Data on Brünnich's guillemots from Oranskie Island and Franz Josef Land were pooled.
3) Use the model from the nearest colony or month: This was applied in 25 cases (*No model* in Table 3.5) when modelling was not feasible (during summer only).

For little auks, we were unable to model the period from May to August. Thus, for this species, we have not developed distribution maps for the summer months.

Table 3.4. Model formulations and variables used in the SDMs.

Model 1:	pb~s(sst)+s(front)+s(dice)+s(adt)+s(prim)+s(daylen)+s(depth)+s(edge)+s(coastd)+s(cold), family=binomial(link="logit")
Model 2:	Model 1 + the term s(east)
Model 3:	pb~s(sst,k=4)+s(cold,k=4)+s(coastd,k=4)+s(east,k=4), family=binomial(link="logit")
Model 4:	pb~s(cold,k=3), family=binomial(link="logit")

Variables	Explanation
pb	Presence, Background (1,0); response variable
sst	Sea surface temperature (SST)
front	Gradient in sst (Front)
dice	Distance to sea-ice (Dice)
adt	Absolute dynamic topography (sea level above geoid) (Adt)
prim	Annual primary production (Prim)
daylen	Daylength: Only included for the period October-February (Daylen)
depth	Bottom depth (Depth)
edge	Gradient in depth (Edge)
coastd	Distance to coast (CoastD)
cold	Distance to colony (ColD)
east	Degree longitude

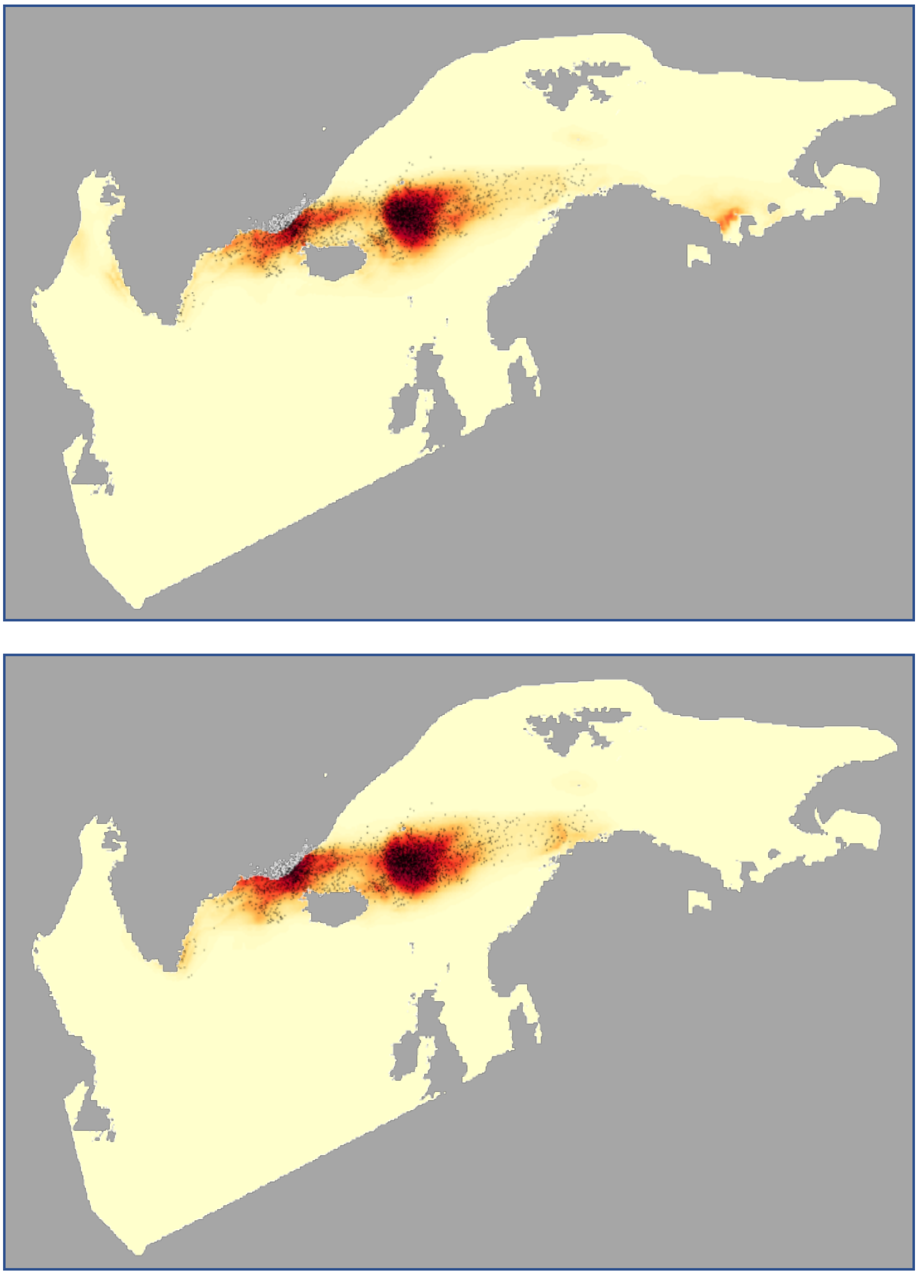


Figure 3.5. Diagnostics of SDMs of Brünnich's guillemots from Bjørnøya in December. Black dots are recorded presences and increasing intensity of red are the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat. Top: Model 1; full model. Note false predictions in Davies Strait (west) and Kola coast (east). Bottom: Model 2; full model plus an east-west variable.

To evaluate the fit of the model and guide the use of the alternative models, we inspected, for each model, a diagnostic plot showing the observed presences and predicted probabilities. In addition, we investigated the adjusted R^2 and the proportion of deviance explained (Appendix 6.2; Figure 3.6). The proportion of deviance explained by the SDMs ranged from 0.18 to 0.98 with a median of 0.63. A larger proportion of deviance was explained when the birds were concentrated around the breeding colony during May to July (Figure 3.6). The models of fulmars and kittiwakes explained slightly less variation during autumn (fulmars) and winter (kittiwakes) compared to the other species. This pattern is probably due to a more widespread pelagic distribution among fulmars and kittiwakes.

For each species, examples of diagnostic plots are shown in Figures 3.7-3.17. Note that the predictions were done on environmental variables for 2017. In general, the predictions fit the observations well, however, in a few cases the model made non-zero predictions in areas where there were no observed presences. This was for example the case for common guillemots from Jan Mayen in December, where the model predicted relatively high probabilities in the North Sea (Figure 3.12). Apparently, the environmental conditions in this area were suitable, but we had no observations of common guillemots from Jan Mayen there. Common guillemots from Isle of May reside in the North Sea during winter, confirming that this area is suitable for common guillemots. This example illustrates the problems associated with using environmental models to predict the spatial distribution of migrating animals or animals that for some reasons are restricted to a specific geographical area: They will not occupy all areas that have favourable environmental conditions. Another example that nicely illustrates this point, is the Atlantic puffins from Røst in September (Figure 3.16). Most recorded occurrences in September were found in the Barents Sea with a few observations in the Greenland/Norwegian Sea. Besides high probabilities in the Barents Sea, the model also predicted relatively high probabilities along the coast of East Greenland. Interestingly, this is also the direction of the migration; the birds migrate from the Barents Sea across the Greenland Sea, along the coast of East Greenland and over winter in the Denmark Strait, the Irminger Sea and in Icelandic waters (Figure 3.17).

Table 3.5. Species distribution models (SDMs). Number of colonies, months and model type for each species. For specification of model type see Table 3.4.

			Number of models				
Species	Colonies	Months	Model 1	Model 2	Model 3	Model 4	No model
Northern fulmar	7	12	78	0	0	2	4
Black-legged kittiwake	15	12	156	1	6	10	7
Common guillemot	10	12	98	0	17	5	0
Brünnich's guillemot	10	12	83	10	8	9	10
Little auk	5	8	40	0	0	0	0
Atlantic puffin	10	12	94	7	7	8	4
Sum			549	18	38	34	25

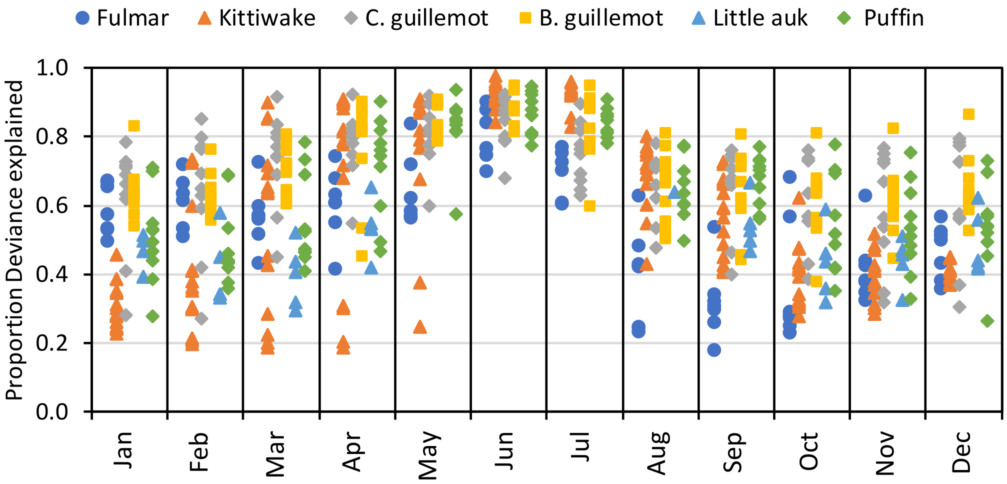


Figure 3.6. Proportion of deviance explained from monthly Species Distribution Models of six seabird species.

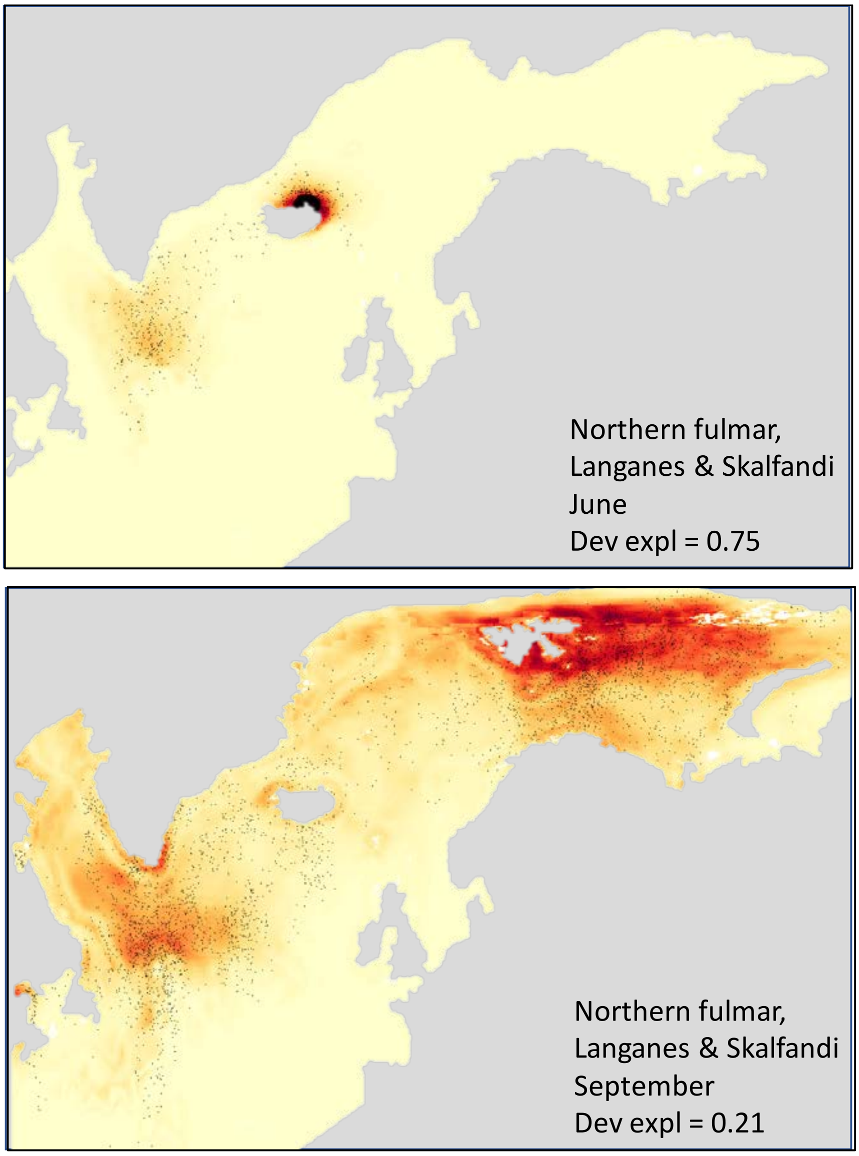


Figure 3.7. Observed presence and predictions from SDMs of Northern fulmar from Langanes and Skalfandi (Iceland) in June (top) and September (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

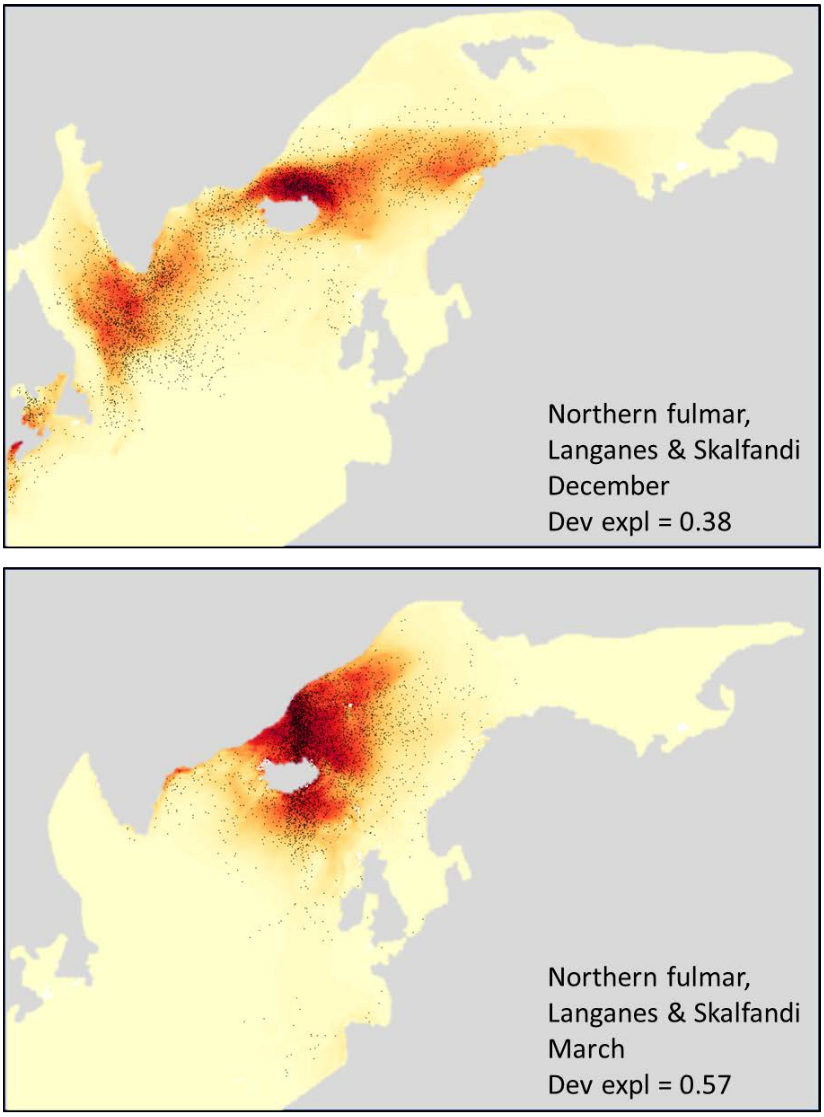
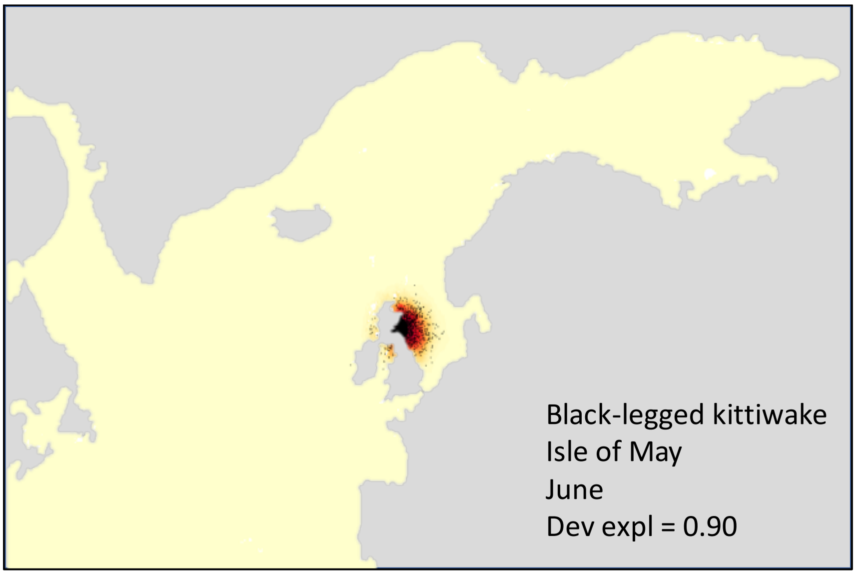


Figure 3.8. Observed presence and predictions from SDMs of Northern fulmar from Langanes and Skalfandi (Iceland) in December (top) and March (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.



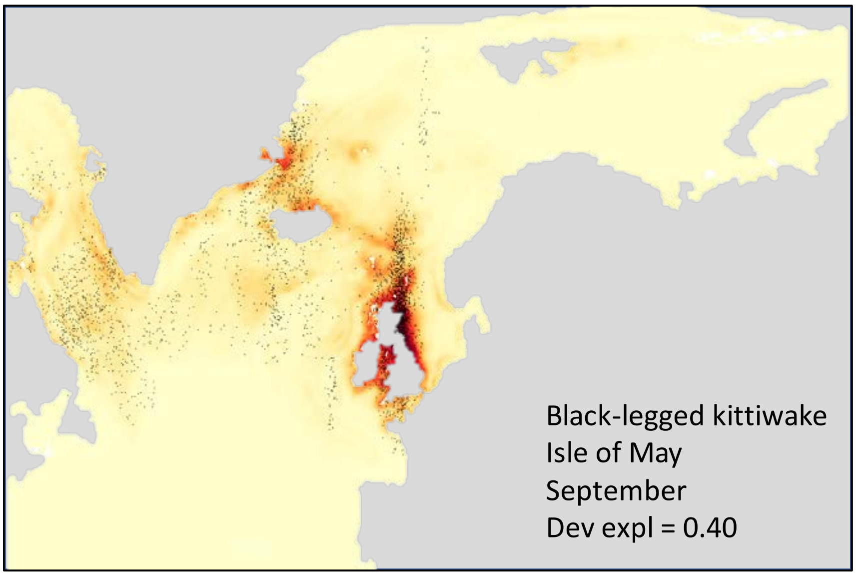


Figure 3.9. Observed presence and predictions from SDMs of black-legged kittiwake from Isle of May (Scotland) in June (top) and September (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

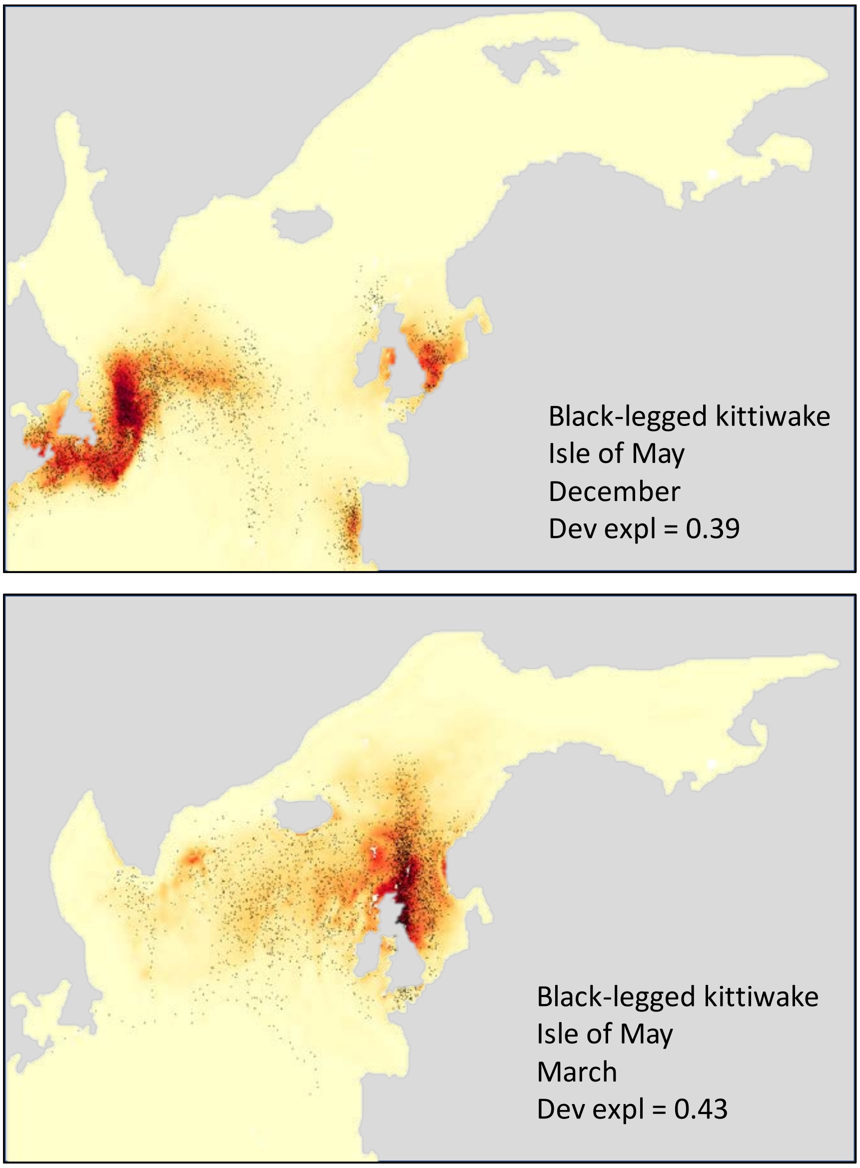


Figure 3.10. Observed presence and predictions from SDMs of black-legged kittiwake from Isle of May (Scotland) in December (top) and March (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

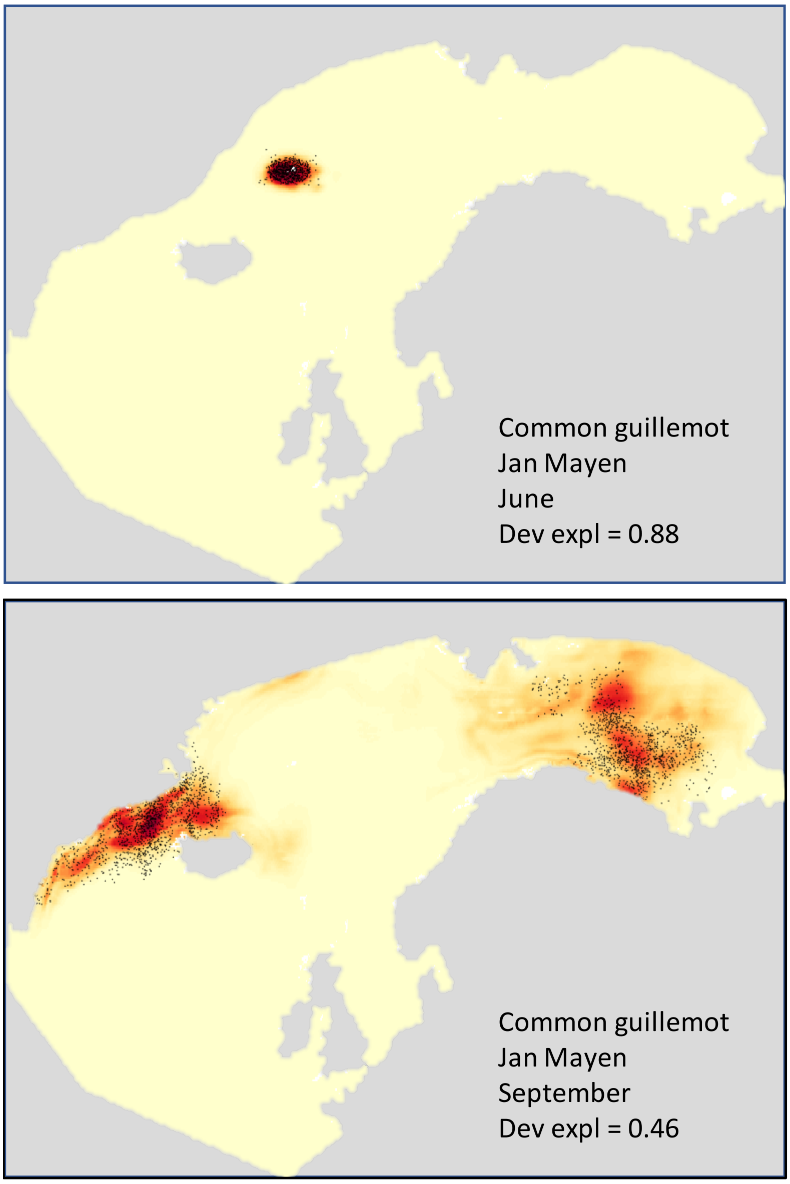


Figure 3.11. Observed presence and predictions from SDMs of common guillemot from Jan Mayen (Norwegian Sea) in June (top) and September (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

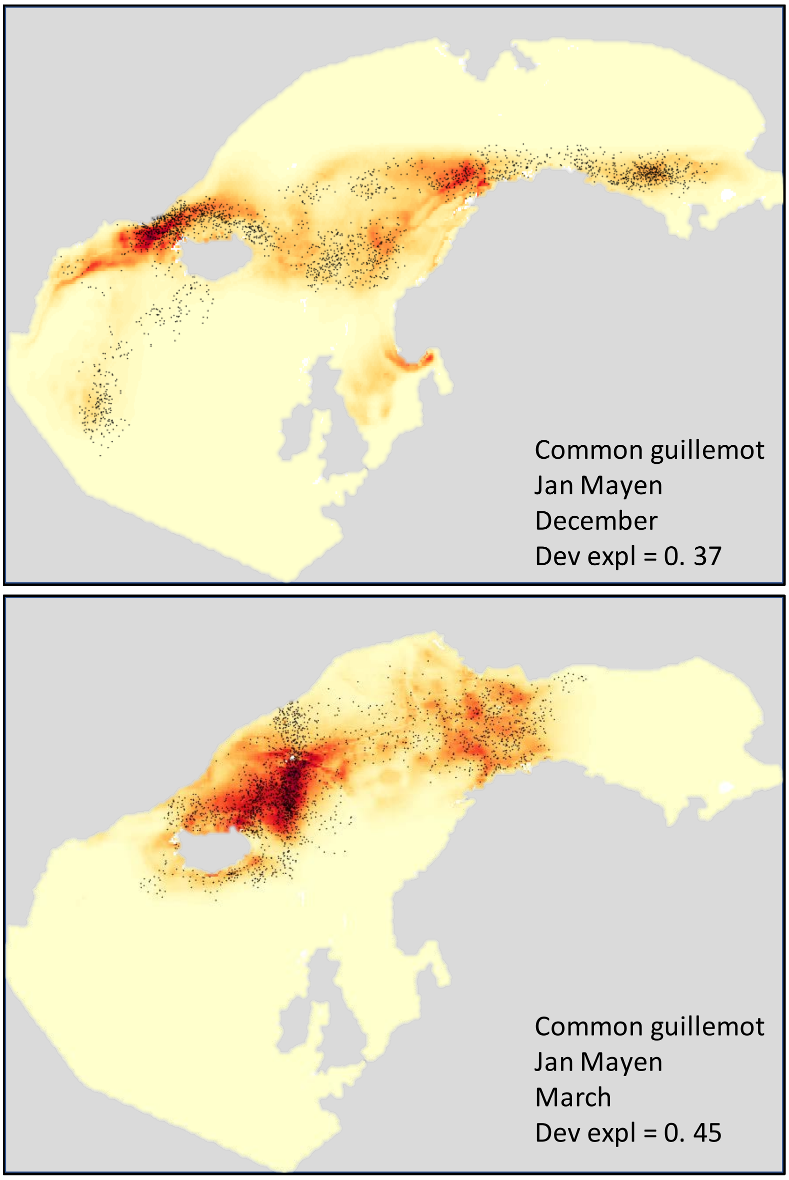


Figure 3.12. Observed presence and predictions from SDMs of common guillemot from Jan Mayen (Norwegian Sea) in December (top) and March (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

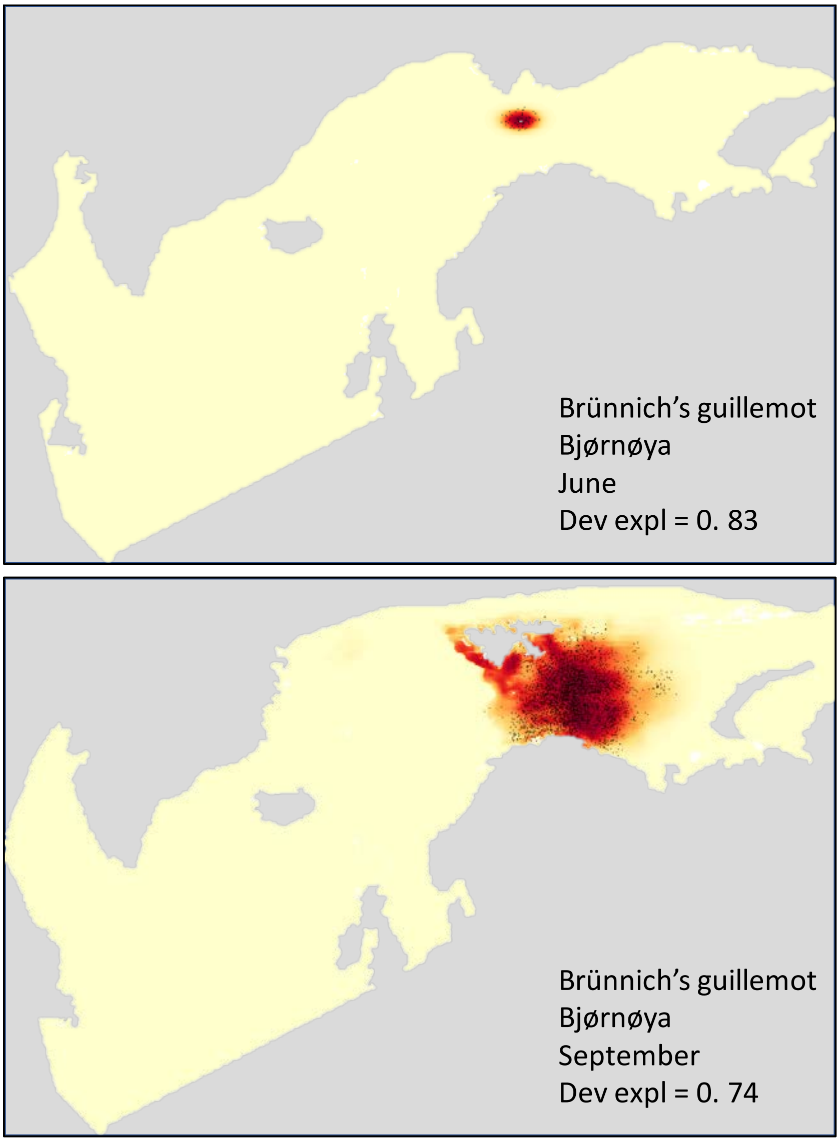


Figure 3.13. Observed presence and predictions from SDMs of Brünnich's guillemot from Bjørnøya (Barents Sea) in June (top) and September (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

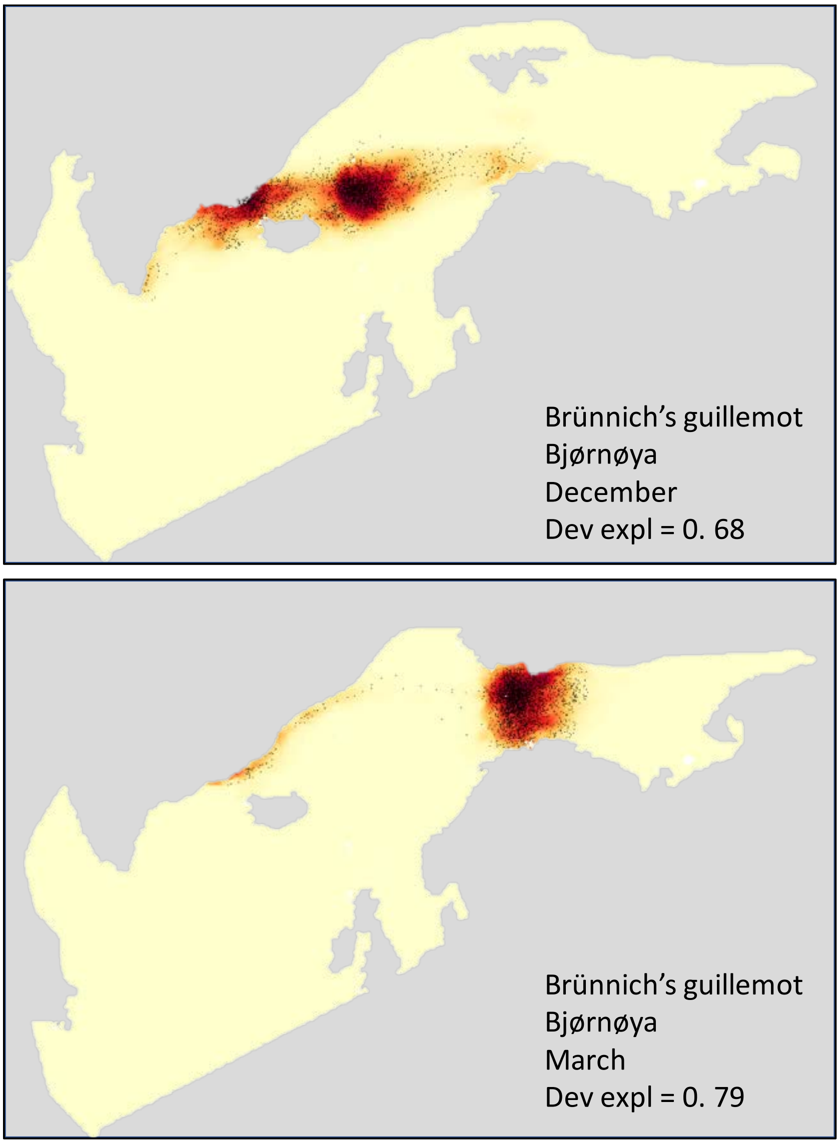


Figure 3.14. Observed presence and predictions from SDMs of Brünnich's guillemot from Bjørnøya (Barents Sea) in December (top) and March (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

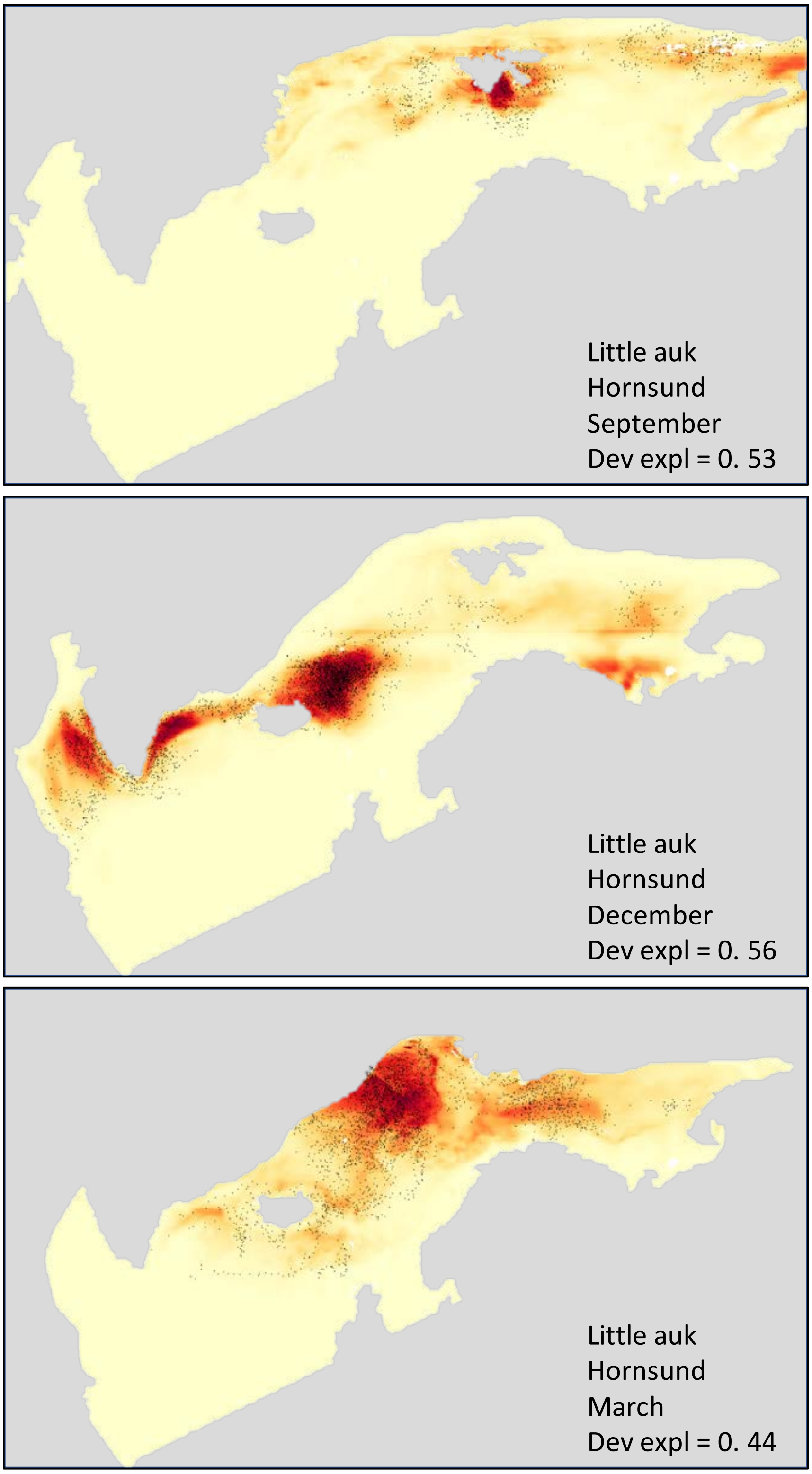


Figure 3.15. Observed presence and predictions from SDMs of little auk from Hornsund (Svalbard) in September (top), December (middle) and March (bottom). Note that models were not conducted for little auks from May to August. Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

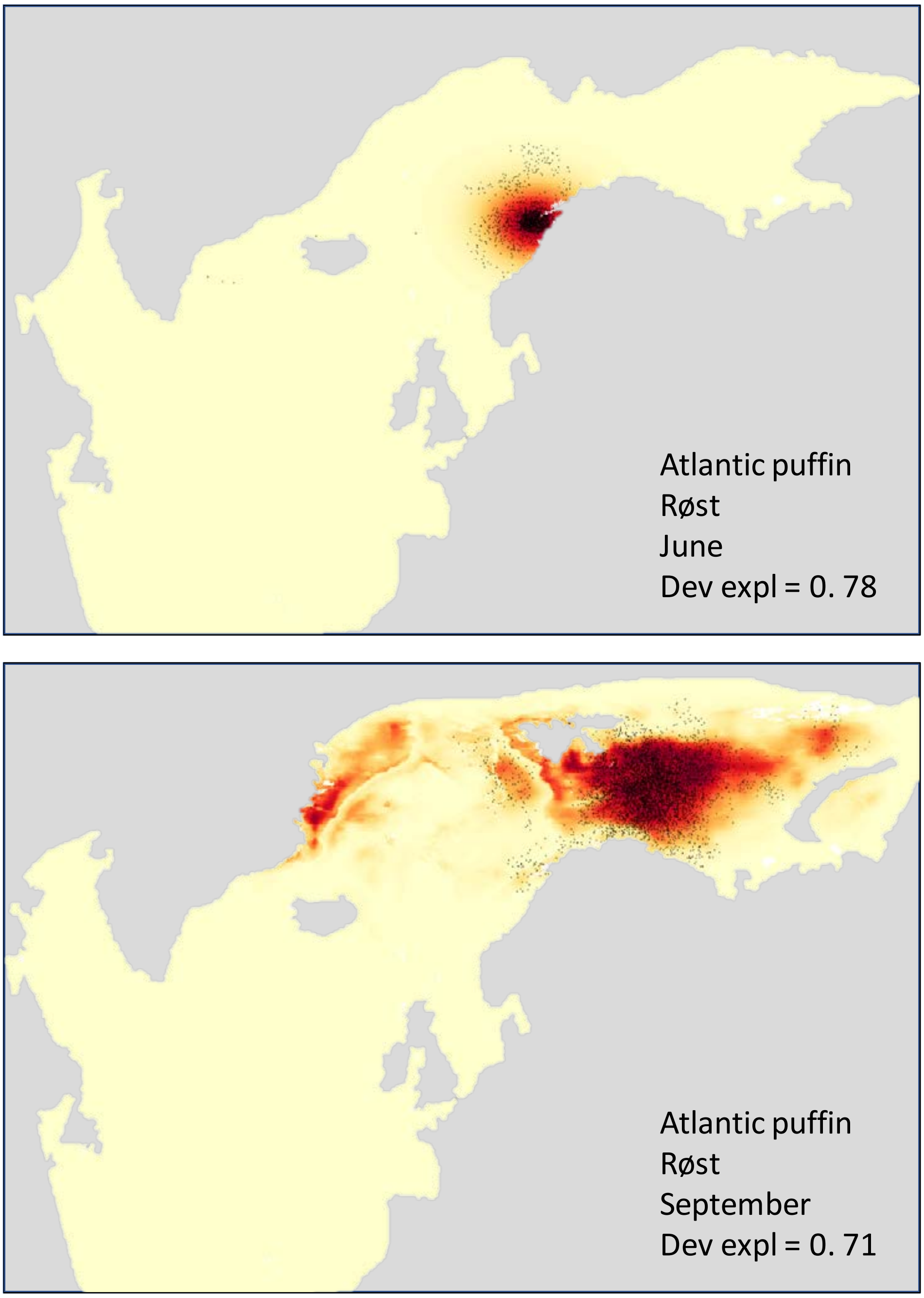
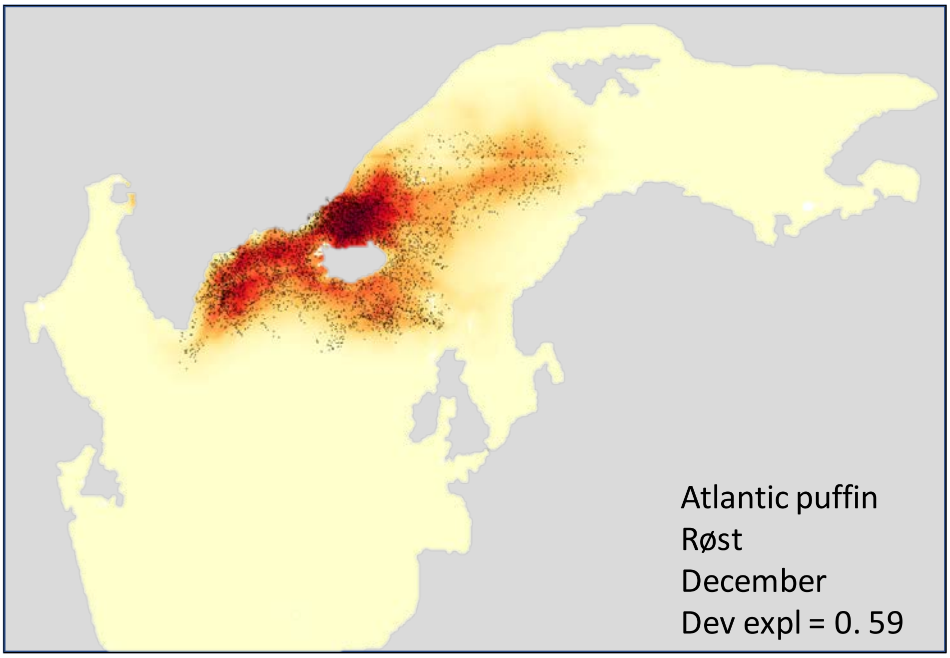


Figure 3.16. Observed presence and predictions from SDMs of Atlantic puffin from Røst (Norway) in June (top) and September (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.



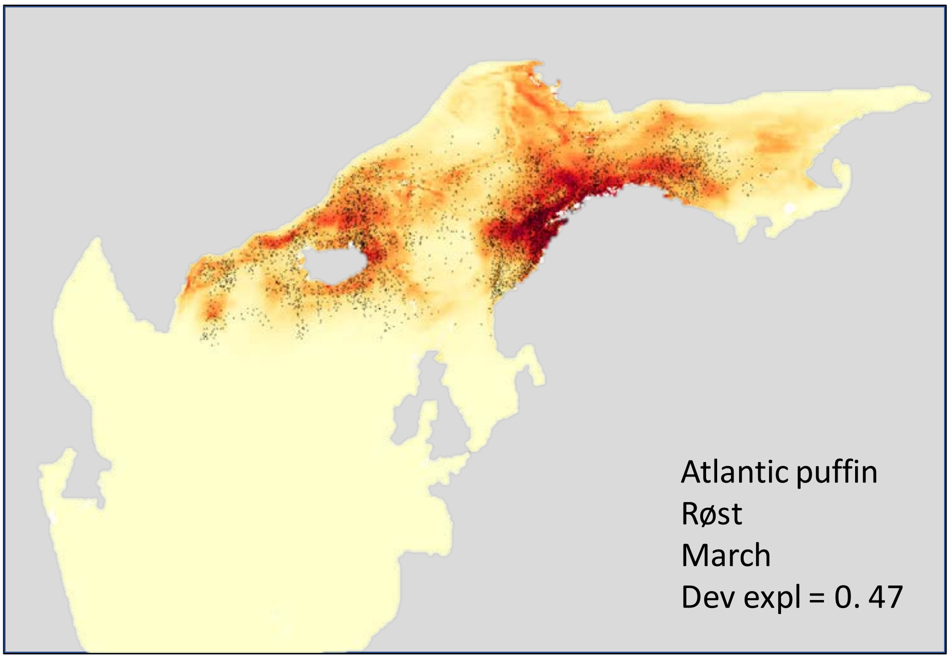


Figure 3.17. Observed presence and predictions from SDMs of Atlantic puffin from Røst (Norway) in December (top) and March (bottom). Black dots are recorded presences and increasing intensity of red is the predicted likelihood of occurrence from the models. Non-grey area is the species' habitat.

4 Abundance maps

Species Distribution Models of presence-background data cannot predict the true probability of occurrence or abundance of a species. The predicted values from the models represent the *relative likelihood of occurrence*, which is proportional to the probability of occurrence (Guillera-Arroita et al. 2015). In other words, the red colouring in Figures 3.7-3.17 are habitat maps showing the relative occurrence of birds from the different breeding populations. In order to obtain realistic and unbiased estimates of abundances, we need to weight the predictions from the models with a factor representing population sizes. Accordingly, the predictions from each SDM must be weighted by the size of the population which the model represent.

Obviously, each model represents the breeding population of the colony where the birds were tagged. In addition, it is likely that neighbouring colonies have similar habitat preferences and migration patterns. Thus, the model could also represent a larger population of colonies close to the model colony. The network of SEATRACK colonies was designed to be sufficiently fine-meshed and representative of a large part of the Northeast Atlantic populations of the different species.

To generate representative abundance maps for each seabird species we

1. Compiled data of breeding populations to generate a colony dataset for the Northeast Atlantic
2. Used the position data from the SEATRACK colonies to investigate how the overlap in winter habitat changed with distance between breeding colonies
3. Based on the analyses in 2, we assigned colonies in the colony database to model-colonies
4. Used the SDMs to predict habitat maps for each assigned colony
5. Weighted each habitat map with the corresponding breeding population size

The resulting maps represent estimates of the monthly density distribution of breeding birds from each colony. By adding the densities from different maps, it is possible to generate maps of the distribution of birds belonging to breeding populations from different nations, regions or ocean areas.

4.1 Colony data

We compiled data from different sources to generate a representative dataset of the breeding populations of the six seabird species in the Northeast Atlantic. Data sources are given in Table 4.1. and a map showing all locations is given in Figure 4.1. Note that we did not include seabird populations breeding in the Baltic Sea, Kategatt, and Southern North Sea.

In several cases, especially for inaccessible places on Franz Josef Land and Novaya Zemlja, the counts are based on rough estimates from anecdotal observations put together from several sources. Estimates from Iceland, Norway, Svalbard and Jan Mayen are all updated and less than ten years old. Estimates from UK and Ireland are from the Seabird 2000 census (Mitchell et al. 2004), and counts were done between 1998 and 2003. Data from Faroe Islands were collected from table 3a in Frederiksen (2010) and were based on counts done between 1981 and 2003. Note also that the definition of a “colony” differs among the regions. For example, Faroe Islands is included as “one” colony, while colonies on Svalbard and Norway have a much finer division.

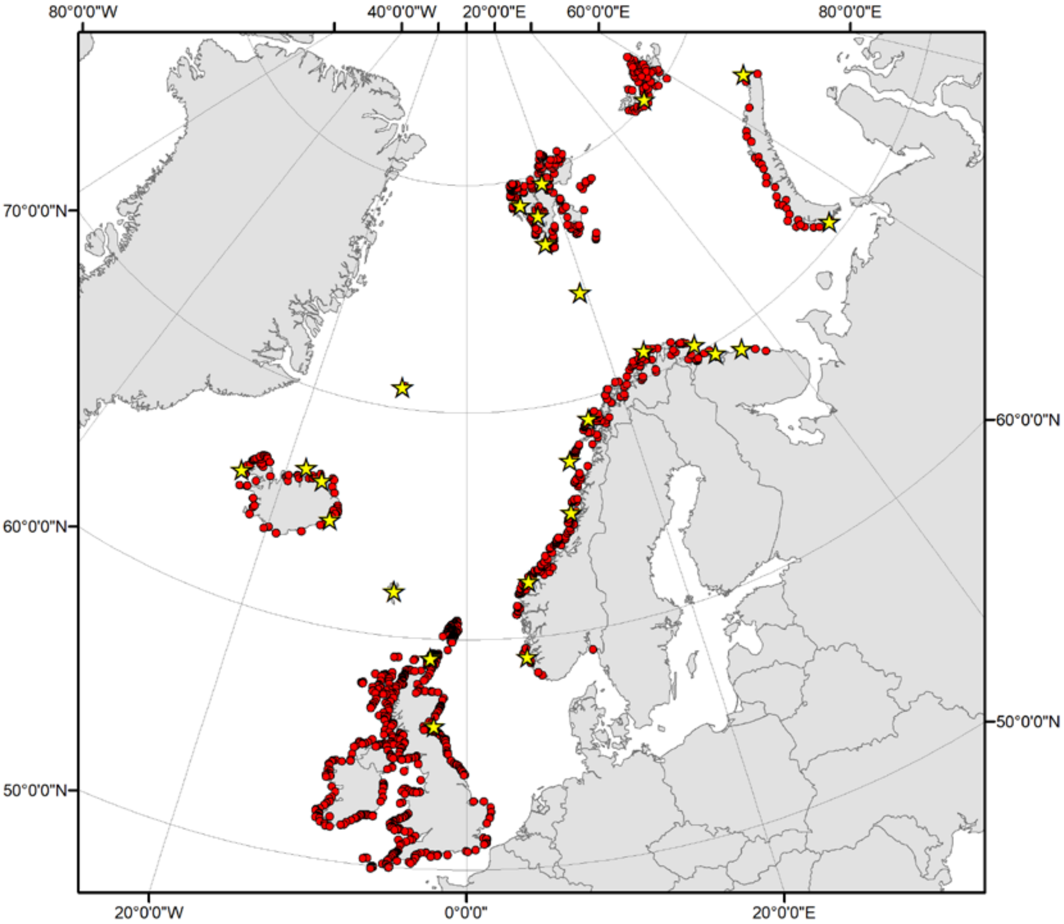


Figure 4.1. Northeast Atlantic seabird colonies (red circles) and SEATRACK model colonies (yellow stars).

Table 4.1. Data sources for the seabird colony data in the Northeast Atlantic

Region	Datasource	Url	Reference
Norway and Svalbard	SEAPOP	http://www.seapop.no/no/	Fauchald et al. 2015
Jan Mayen	SEAPOP	http://www.seapop.no/no/	Gabrielsen & Strøm 2013
Russia	The Seabird Colony Registry of the Barents and White Seas	https://www.barentsportal.com/barentsportal/index.php/en/maps/99-biodiversity/677-seabirds-and-colonies-in-the-barents-sea https://www.barentsportal.com/barentsportal/index.php/en/general-description/109-biotic-components/709-seabirds	Anker-Nilssen et al. 2000 Strøm et al. 2009
Britain and Ireland	Joint Nature Conservation Committee Seabird 2000, Seabird Monitoring Programme Database (SMP)	www.jncc.gov.uk/smp	Mitchell et al. 2004
Iceland	Icelandic Institute of Natural History	https://ni.is/dyr/fuglar/mikilvaeg-fuglasvaedi/sjofuglabyggdir	
Faroe Islands	Table 3a in Frederiksen (2010)	http://norden.diva-portal.org/smash/record.jsf?pid=diva2%3A701212&dswid=-8433	Frederiksen (2010)

4.2 Assigning breeding populations to SDMs

To assign the populations from the colony dataset to the SDMs of SEATRACK colonies, we had to decide the limit for the maximum distance allowed to the nearest SEATRACK colony. Colonies nearer than the limit would be modelled by the SDM from the SEATRACK colony, while colonies with a longer distance would fall outside the populations we predict. To find a maximum distance, we first investigated how the overlap in winter habitat was related to the distance between SEATRACK colonies. We expected the winter-overlap between proximate colonies to be high, indicating that the birds from the two colonies had similar habitat preferences and migratory pattern, while we expected birds from distant colonies to have less winter-overlap due to different habitat preferences and/or migratory pattern. The function between winter-overlap and distance could, in other words give indications on how far away from a modelled colony an SDM would have validity.

To define the winter habitat, we calculated for each colony, the Utility Distribution (UD) based on the kernel of all positions from November until February, combining IRMA and GLS positions. Kernels were calculated using function kernelUD from package {adehabitatHR} with smoothing parameter (h) equal to 100 km. We used the Volume of Intersection (VI) (Fieberg & Kochanny, 2005) between UDs as a measure of overlap in winter habitat between all pairs of colonies. The VI index ranges between zero, for home ranges with no overlap, and one for home ranges with identical UDs. VIs were calculated using function *kerneloverlap* from package {adehabitatHR}.

The matrix of overlap indices (VIs) between colonies was plotted against the corresponding matrix of distances. To calculate the distance between colonies, we decided to use the shortest ocean-distance (i.e., the shortest distance over sea). We used function *shortestPath* from package {gdistance} to calculate the shortest ocean-distance, using the land mask as a transition layer. Because the VI index has a range between 1 and 0, we expected the relationship between winter habitat overlap and distance to follow a (decreasing) logistic function. Accordingly, we fitted the relationship using the nls function with the formula y ~ 1/(1 + exp(-a * (x - b))), where y is the VI index and x is distance. a and b are parameters to be estimated.

Setting a short limit for the maximum allowed distance would exclude a large proportion of the populations included in the sample. To investigate how the distance limit affected the proportion of populations included, we calculated the shortest ocean-distance between all colonies in the colony dataset and the nearest SEATRACK colony and plotted it against the corresponding proportion of the total population. The plots of overlap and population proportion were combined in the same figure for each of the six species (Figure 4.2-4.7). As expected, the overlap in winter habitat decreased with increasing distance between colonies. There were however some differences between species. The negative trend was especially weak for kittiwakes (Figure 4.3). For this species, the overlap was relatively high and did decrease only slightly for increasing distance, suggesting that the distance between colonies had little impact on differences in habitat preference or migratory pattern.

For all species the overlap ranged between 0.4 and 0.8 for distances less than 500 km. Except for kittiwakes, the overlap was between 0.2 and 0.4 for distances between 500 and 1000 km and decreased further to less than 0.2 for distances above 1000 km. For all species, more than 80% of the Northeast Atlantic population were covered with a maximum distance to nearest model colony equal to 500 km.

Based on the analyses above, we decided to use the same limit for maximum distance to nearest model colony for all species. The limit was set to 400 km. Using this limit, the proportion of the Northeast Atlantic population covered was 0.74 for Northern fulmars, 0.83 for black-legged kittiwakes, 0.77 for common guillemots, 0.96 for Brünnich’s guillemots, 0.89 for little auks and 0.94 for Atlantic puffin. The coverage and assignment of model colonies are shown for each species in Figure 4.8-4.13.

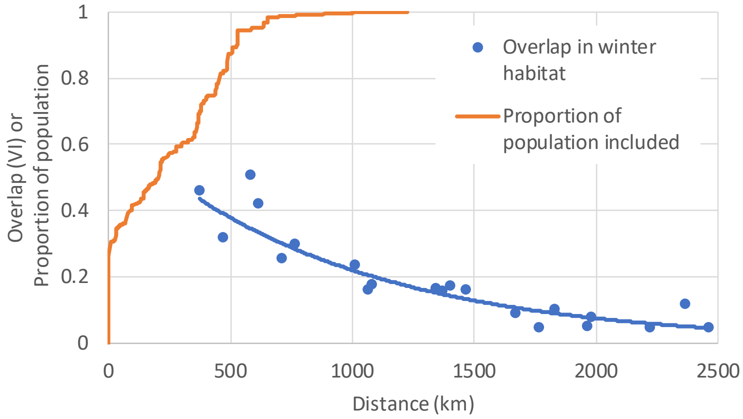


Figure 4.2. Northern fulmar. Overlap in winter habitat as a function of distance between colonies and proportion of colonies included in the analysis for increasing limit of maximum allowed distance to model colony.

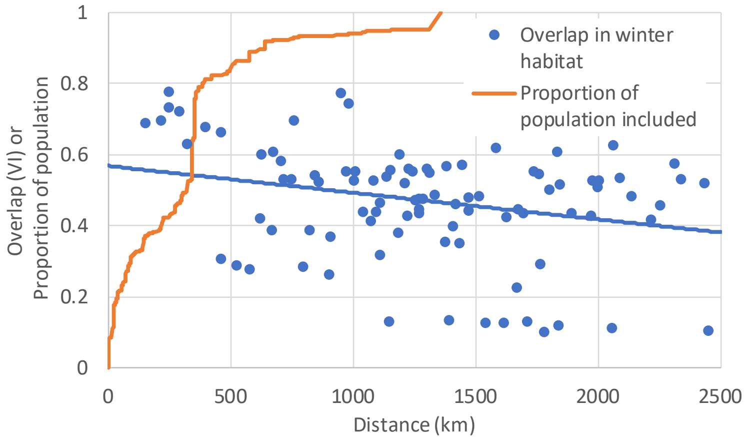


Figure 4.3. Black-legged kittiwake. Overlap in winter habitat as a function of distance between colonies and proportion of colonies included in the analysis for increasing limit of maximum allowed distance to model colony.

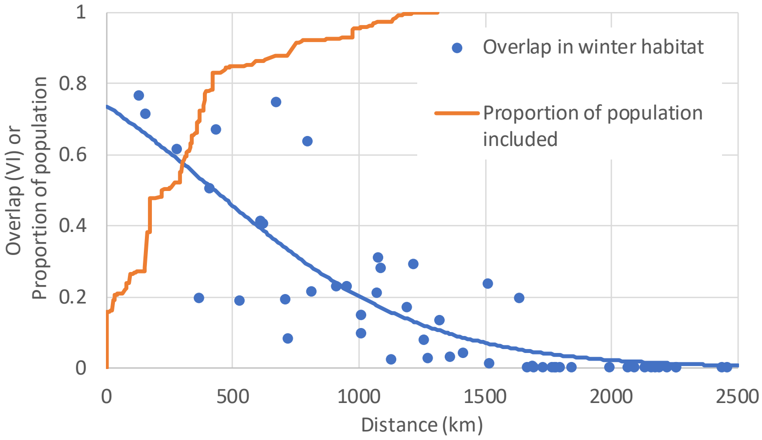


Figure 4.4. Common guillemot. Overlap in winter habitat as a function of distance between colonies and proportion of colonies included in the analysis for increasing limit of maximum allowed distance to model colony.

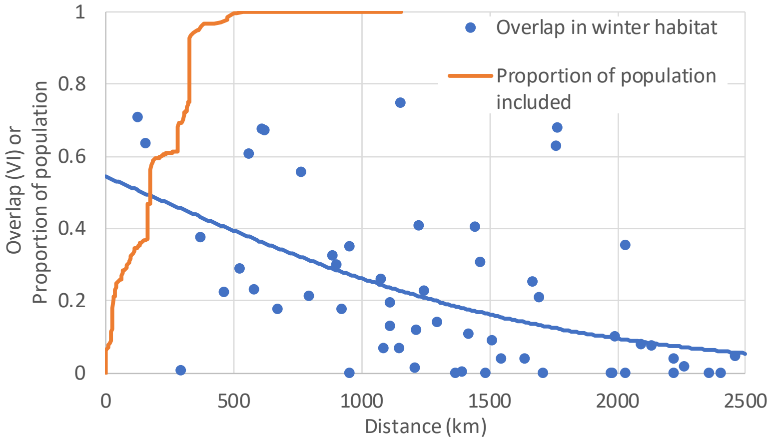


Figure 4.5. Brünnich's guillemot. Overlap in winter habitat as a function of distance between colonies and proportion of colonies included in the analysis for increasing limit of maximum allowed distance to model colony.

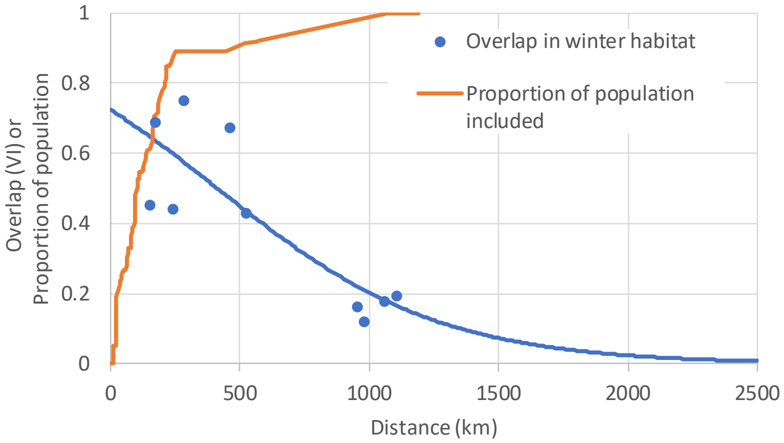


Figure 4.6. Little auk. Overlap in winter habitat as a function of distance between colonies and proportion of colonies included in the analysis for increasing limit of maximum allowed distance to model colony.

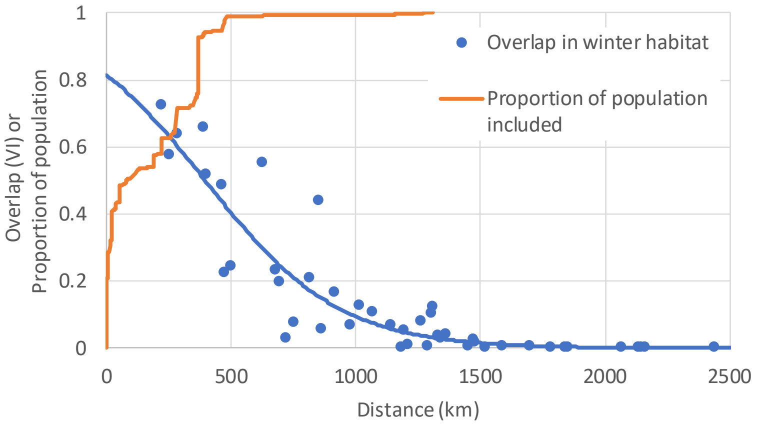


Figure 4.7. Atlantic puffin. Overlap in winter habitat as a function of distance between colonies and proportion of colonies included in the analysis for increasing limit of maximum allowed distance to model colony.

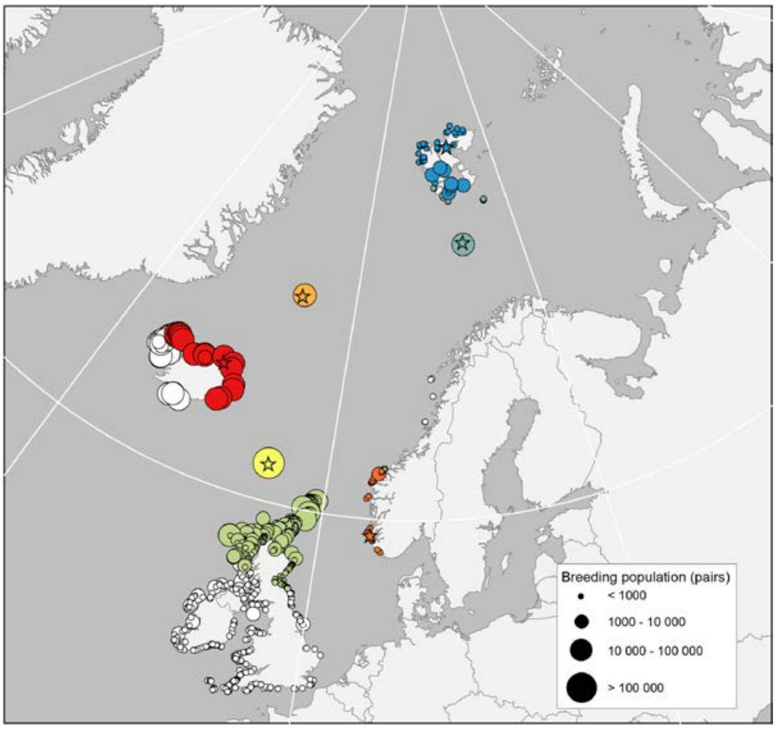


Figure 4.8. Northern fulmar. *Assignment of Northeast Atlantic colonies (coloured circles) to SEATRACK model colonies (coloured stars). Maximum allowed distance to nearest model colony was 400 km. White circles are colonies not covered by the design (distance > 400 km from the nearest model colony). Circle size indicates population size. The design covered 74% of the total population.*

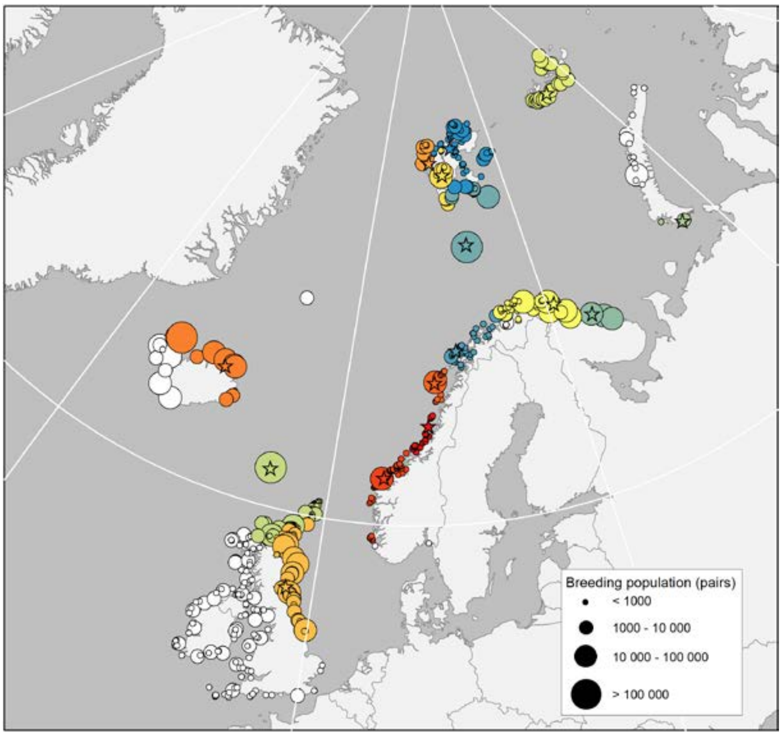


Figure 4.9. Black-legged kittiwake. Assignment of Northeast Atlantic colonies (coloured circles) to SEATRACK model colonies (coloured stars). Maximum allowed distance to nearest model colony was 400 km. White-filled circles are colonies not covered by the design (distance > 400 km from the nearest model colony). Circle size indicates population size. The design covered 83% of the total population.

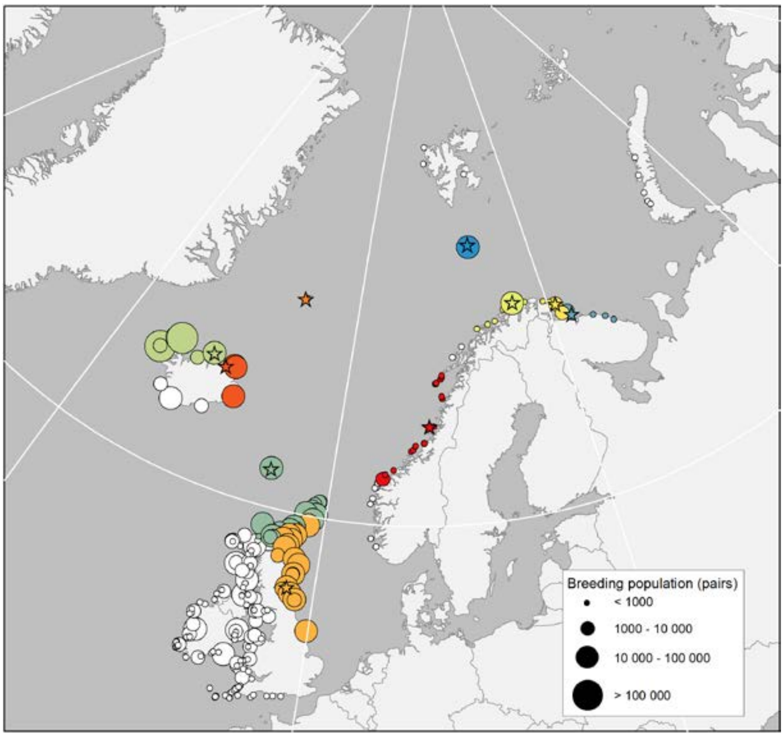


Figure 4.10. Common guillemot. Assignment of Northeast Atlantic colonies (coloured circles) to SEATRACK model colonies (coloured stars). Maximum allowed distance to nearest model colony was 400 km. White-filled circles are colonies not covered by the design (distance > 400 km from the nearest model colony). Circle size indicates population size. The design covered 77% of the total population.

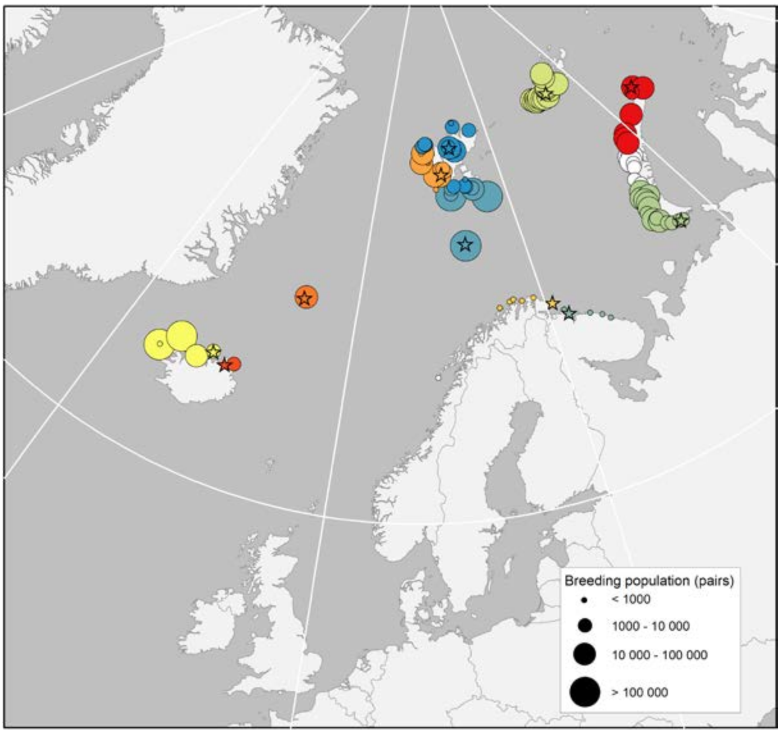


Figure 4.11. Brünnich's guillemot. Assignment of Northeast Atlantic colonies (coloured circles) to SEATRACK model colonies (coloured stars). Maximum allowed distance to nearest model colony was 400 km. White-filled circles are colonies not covered by the design (distance > 400 km from the nearest model colony). Circle size indicates population size. The design covered 96% of the total population.

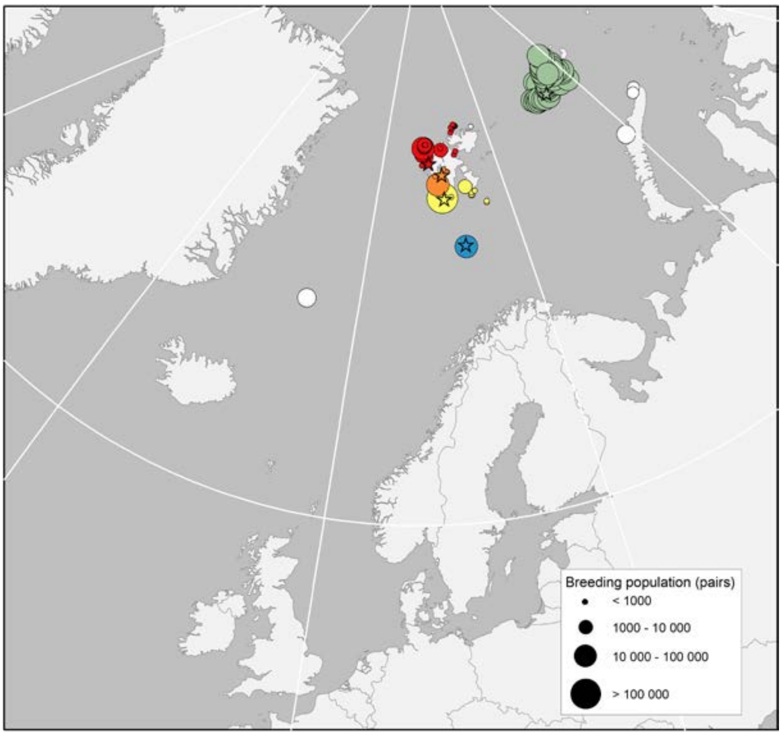


Figure 4.12. Little auk. Assignment of Northeast Atlantic colonies (coloured circles) to SEATRACK model colonies (coloured stars). Maximum allowed distance to nearest model colony was 400 km. White-filled circles are colonies not covered by the design (distance > 400 km from the nearest model colony). Circle size indicates population size. The design covered 89% of the total population.

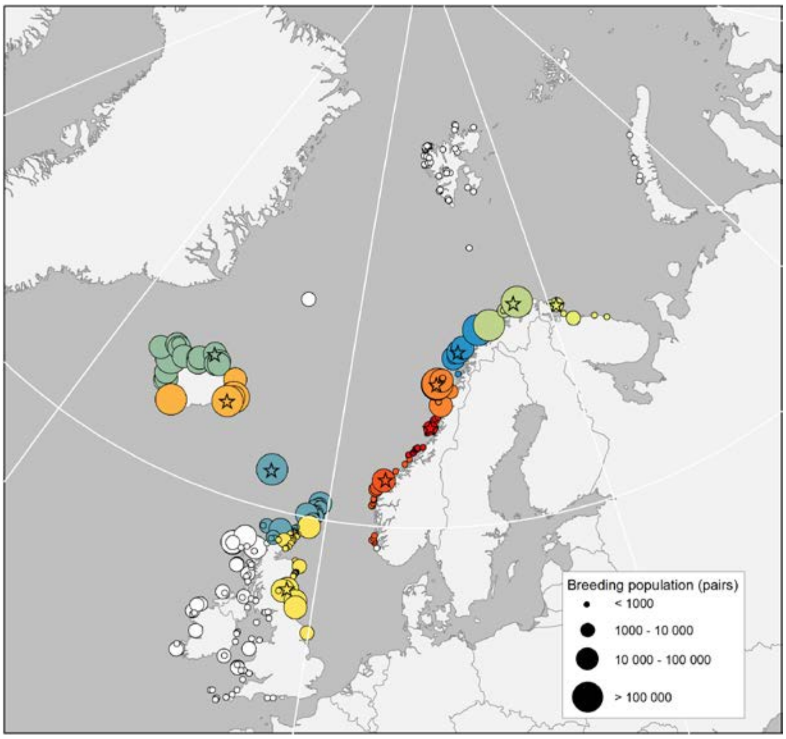


Figure 4.13. Atlantic puffin. Assignment of Northeast Atlantic colonies (coloured circles) to SEATRACK model colonies (coloured stars). Maximum allowed distance to nearest model colony was 400 km. White-filled circles are colonies not covered by the design (distance > 400 km from the nearest model colony). Circle size indicates population size. The design covered 94% of the total population.

4.3 Predictions

In the last step, the Species Distribution Models (SDMs) described in section 3.2 were used to predict the monthly distribution of the six seabird species in the Northeast Atlantic. The predictions were done for all colonies assigned to a model colony (i.e., colonies less than 400 km from a model colony, see Figures 4.8-4.13). This design covered from 74% (Northern fulmar) to 99% (Brünnich's guillemot) of the Northeast Atlantic breeding populations (see 4.2).

Although the models were based on data from 2012-2017, with the most recent positions collected in summer 2017, we decided to use environmental variables from 2017-18 in the predictions. The reason behind this decision was partly to provide predictions based on the most recent environmental conditions, and partly because we wanted to use the GLS data collected in 2017-18 as an independent dataset to validate the SDMs. This work will not be presented in the present report but will be completed as soon as the data collected in 2018 have been compiled and are ready for analyses.

While the environmental predictors were the same for all colonies, one important predictor was colony specific, namely distance to colony. Thus, predictions had to be done separately for each colony, and to save computer time, we did not provide predictions for very small colonies (e.g; colonies with less than 9 pairs).

Predictions were done on a 0.1 x 0.1 degree geographical raster covering the habitat defined for each species (i.e., the area from where background points were drawn, see 3.2.3). Raster cells outside the defined habitat were assigned zero values.

The predictions from the SDMs give the relative likelihood of occurrence, and to scale these values to abundances we used the following equation:

$$\hat{n}_i = N \frac{\hat{p}_i a_i}{\sum_i \hat{p}_i a_i}$$

Where $\hat{n}_i$ is the predicted number of birds from a colony in raster cell *i*, *N* is the number of breeding birds in the colony (i.e, colony size), $\hat{p}_i$ is the relative likelihood of occurrence predicted by the model and a_i is the area of raster cell *i*. Note that the equation controls for the area of the raster cells. This is because the cell-area varies in a geographic grid, with cells having smaller areas in the north.

In total, the resulting raster dataset comprises 9619 distribution maps, one map for each combination of species, month and colony. For most purposes, it will be necessary to aggregate the rasters of colonies belonging to the same region, nation or ocean area. For example, in Figure 4.14, we have aggregated the abundance predictions of Atlantic puffin in September using all colonies covered by the SEATRACK design.

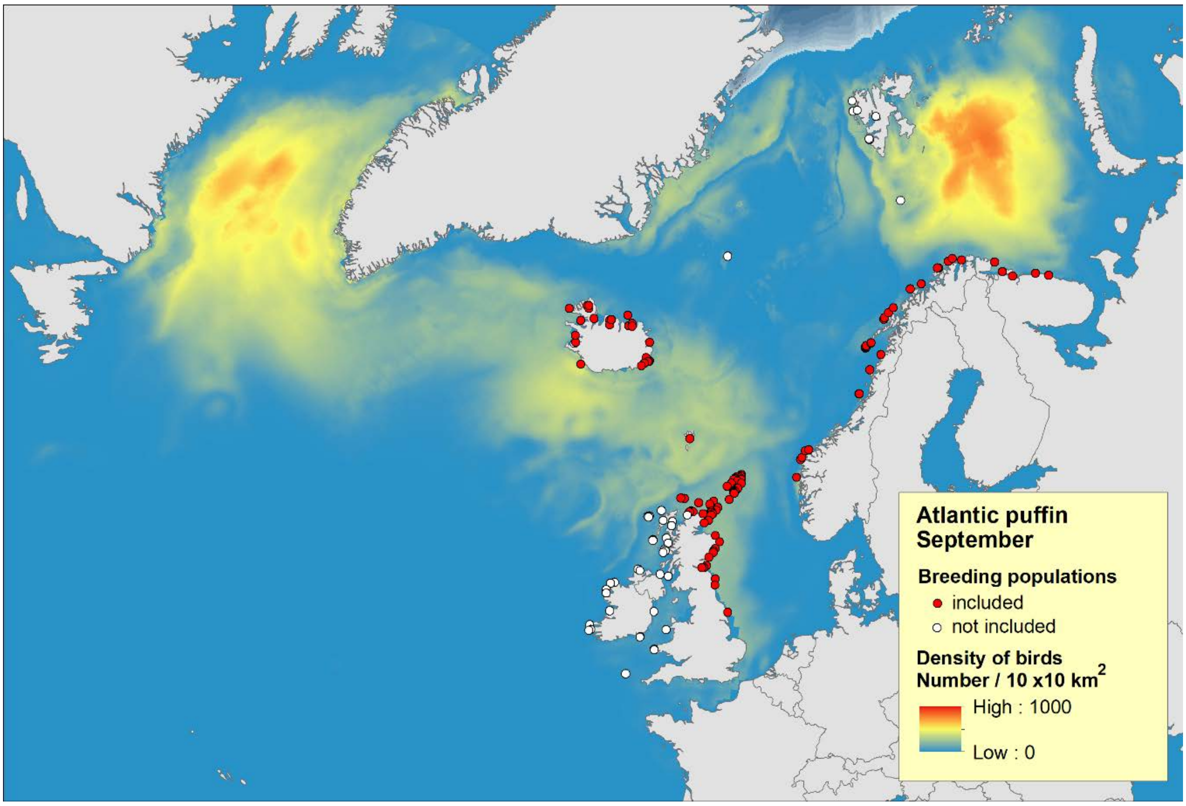


Figure 4.14. The estimated distribution of Northeast Atlantic populations of adult Atlantic puffin (Fratercula arctica) in September. Breeding populations included in the predictions are indicated by red filled circles. The maps cover approximately 94% of the Northeast Atlantic breeding population.

4.3.1 Datasets of modelled seabird abundances

The model predictions were first stored as standard geographical (WGS 84) rasters for each colony covered by the design. Raster extent is -78°E to 80°E, 35°N to 85°N (latitude) with a spatial resolution of 0.1 x 0.1 degrees. Each raster cell is given a value representing the predicted number of breeding birds from a given colony within the raster cell. Note that raster cells outside the defined habitat were assigned zero values. For easier storage and distribution, the rasters were combined and archived as NetCDF[1] files separately for each species. The NetCDF format is self-describing (it includes information about the data content, i.e. metadata), portable (it works across computer platforms), scalable (it is possible to work only on subsets of the whole dataset), and appendable (one can add new data to an existing dataset). The NetCDF file format can be read by most geographic information system software, such as GRASS, Quantum GIS, or ArcGis. It can also easily be read and handled using the free and open source R Statistical Software (R Development Core Team 2018), and here we provide a baseline R-script that will allow reading in the SEATRACK NetCDF output files, extracting specific data, and aggregating seabird modelled abundances among colonies, e.g. to cover a larger geographic region (Appendix 6.3).

---

[1] https://doi.org/10.5065/D6H70CW6

4.3.2 Limitations and caveats

The datasets and analyses are subject to several important limitations and caveats. One important limitation related to the available position dataset is that the sample of tagged birds was limited to breeding individuals. Thus, we do not have information on the distribution of non-breeders, including immature birds. Immatures might constitute a relatively large part of the population, and different migration pattern among life-stages could result in a biased interpretation of the distribution pattern. Another issue that could result in biases, is the assumption that colonies close to each other have similar migration pattern. Although this assumption was supported by the overlap analyses, cases where the assumption do not hold could result in wrong predictions for some colonies.

Improving the coverage of the populations, both with respect to the sample of colonies and with respect to life-stages included in the sample, could reduce biases attributed to these effects. To assess the effect of biases, and thereby assessing the validity of the distribution maps, it is possible to compare the predicted distribution with observations of seabirds at sea. Seabird at sea datasets are available for the Northeast Atlantic (Fauchald 2011), and analyses of how the predictions from the present models fit the data from observations at sea is a task that will be undertaken in a next step. These analyses will give valuable information on possible weaknesses and limitations of the present dataset.

It is important to notice that the error in the positioning of the birds is relatively large (more than 180 km). High observation error makes it difficult to capture fine-scale distribution pattern and furthermore, the precision of the resulting estimated distribution is relatively low (i.e., the predictions will have large confidence intervals). The observation errors can to a certain degree be compensated with large sample sizes. The observations used in the models are however not independent of each other. This is because the observations include positions from tracks of the same individuals. These dependencies will reduce the “effective” sample size and consequently increase the error in the predictions. Models based on a few individuals might accordingly result in erroneous predictions. Because of the dependencies in the datasets and because the dataset consists of both GLS positions and positions from the IRMA model, it is difficult to assess the resulting error from the models using standard statistical procedures. One way to assess the error and predictive power of the models is to compare the SDMs with an independent position dataset. This will be done by comparing the predictions from the SDMs with the dataset collected in 2017-18. These analyses will evaluate the predictive power and performance of each SDM and will give valuable information with respect to optimization of the sampling design.

5 References

Amélineau, F., Fort, J., Mathewson, P.D., Speirs, D.C., Courbin, N., Perret, S., Porter, W.P., Wilson, R.J. & Grémillet, D. 2018. Energyscapes and prey fields shape a North Atlantic seabird wintering hotspot under climate change 5(1): 171883.

Anker-Nilssen T, Bakken V, Strøm H, Golovkin AN, Bianki VV, Tatarinkova IP (2000). The Status of Marine Birds Breeding in the Barents Sea Region. Norsk Polarinstitutt Rapportserie, 113. 213 pp.

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol Evol 3:327–338.

Barraquand F, Benhamou S (2008) Animal movements in heterogeneous landscapes: identifying profitable places and homogeneous movement bouts Ecology 89(12):3336-3348.

Behrenfeld MJ, Falkowski PG (1997) Photosynthetic rates derived from satellite-based chlorophyll concentration. Limnol Oceanogr 42:1–20.

Calenge C (2006) The package ‘adehabitat’ for the R software: a tool for the analysis of space and habitat use by animals. Ecol. Model. 197(3-4):516-519.

Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. Annu Rev Ecol Evol Syst 40:677–697.

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

Fauchald P (2011) Sjøfugl i åpent hav. Utbredelsen av sjøfugl i norske og tilgrensende havområder. NINA Rapport 786.

Fauchald P, Skov H, Skern-Mauritzen M, Hausner VH, Johns D, Tveraa T (2011) Scale-dependent response diversity of seabirds to prey in the North Sea. Ecology 92:228–239.

Fauchald P, Anker-Nilssen T, Barrett R, Bustnes JO, Bårdsen BJ, Christensen-Dalsgaard S, Descamps S, Engen S, Erikstad KE, Hanssen SA, Lorentsen S-H, Moe B, Reiertsen T, Strøm H, Systad GH (2015) The status and trends of seabirds breeding in Norway and Svalbard. NINA report 1151: 84 pp.

Fayet AL, Freeman R, Anker-Nilssen T, Diamond A, Erikstad KE, Fifield D, Fitzsimmons MG, Hansen ES, Harris MP, Jessopp M, Kouwenberg A-L, Kress S, Mowat S, Perrins CM, Petersen A, Petersen IK, Reiertsen TK, Robertson GJ, Shannon P, Sigurðsson IA, Shoji A, Wanless S, Guilford T (2017) Ocean-wide Drivers of Migration Strategies and Their Influence on Population Breeding Performance in a Declining Seabird. Curr Biol 27:3871–3878.e3

Fieberg J, Kochanny CO (2005) Quantifying Home-Range Overlap: The Importance of the Utilization Distribution. J Wildl Manage 69:1346–1359.

Fort J, Moe B, Strøm H, Grémillet D, Welcker J, Schultner J, Jerstad K, Johansen KL, Phillips RA, Mosbech A (2013) Multicolony tracking reveals potential threats to little auks wintering in the North Atlantic from marine pollution and shrinking sea ice cover. Divers Distrib 19:1322–1332.

Frederiksen M (2010) Appendix 1: Seabirds in the North East Atlantic –A review of status, trends and anthropogenic impact. In: Action plan for seabirds in Western-Nordic areas, 587th edn. TemaNord, Norden, p 47–122.

Frederiksen M, Moe B, Daunt F, Phillips RA, Barrett RT, Bogdanova MI, Boulinier T, Chardine JW, Chastel O, Chivers LS, Christensen-Dalsgaard S, Clément-Chastel C, Colhoun K, Freeman R, Gaston AJ, González-Solís J, Goutte A, Grémillet D, Guilford T, Jensen GH, Krasnov Y, Lorentsen SH, Mallory ML, Newell M, Olsen B, Shaw D, Steen H, Strøm H, Systad GH, Thórarinsson TL, Anker-Nilssen T (2012) Multicolony tracking reveals the winter distribution of a pelagic seabird on an ocean basin scale. Divers Distrib 18:530–542.

Gabrielsen GW, Strøm H (2013) Seabird research and monitoring on Jan Mayen. In: Skreslet S (ed) Jan Mayen Island in scientific focus, NATO Science Series, Springer Science & Business Media, p 181–1944.

Goberville E, Beaugrand G, Hautekèete NC, Piquot Y, Luczak C (2015) Uncertainties in the projection of species distributions related to general circulation models. Ecol Evol 5:1100–1116.

Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, Gordon A, Kujala H, Lentini PE, Mccarthy MA, Tingley R, Wintle BA (2015) Is my species distribution model fit for purpose? Matching data and models to applications. Glob Ecol Biogeogr 24:276–292.

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Modell 135:147–186.

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

Hastie TJ, Tibshirani RJ (1990) Generalized Additive Models. Chapman & Hall/CRC.

Hays GC, Bailey H, Bograd SJ, Bowen WD, Campagna C, Carmichael RH, Casale P, Chiaradia A, Costa DP, Cuevas E, Nico de Bruyn PJ, Dias MP, Duarte CM, Dunn DC, Dutton PH, Esteban N, Friedlaender A, Goetz KT, Godley BJ, Halpin PN, Hamann M, Hammerschlag N, Harcourt R, Harrison A-L, Hazen EL, Heupel MR, Hoyt E, Humphries NE, Kot CY, Lea JSE, Marsh H, Maxwell SM, McMahon CR, Notarbartolo di Sciara G, Palacios DM, Phillips RA, Righton D, Schofield G, Seminoff JA, Simpfendorfer CA, Sims DW, Takahashi A, Tetley MJ, Thums M, Trathan PN, Villegas-Amtmann S, Wells RS, Whiting SD, Wildermann NE, Sequeira, AMM (2019) Translating marine animal tracking data into conservation policy and management. Trends Ecol Evol (in press).

Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modelling methods. Ecography (Cop) 29:773–785.

Hijmans RJ (2017) geosphere: Spherical Trigonometry. R package version 1.5-7. https://CRAN.R-project.org/package=geosphere.

Koenker R (2018) quantreg: Quantile Regression. R package version 5.36, available at https://CRAN.R-project.org/package=quantreg.

Kohut J, Bernard K, Fraser W, Oliver MJ, Statscevvich H, Winsor P, Miles T (2014) Studying the impacts of local oceanographic processes on adelie penguin foraging ecology. Marine Technology Society Journal 48(5):25-34

Krüger L, Ramos JA, Xavier JC, Grémillet D, González-Solís J, Kolbeinsson Y, Militão T, Navarro J, Petry M V., Phillips RA, Ramírez I, Reyes-González JM, Ryan PG, Sigurðsson IA, Sebille E Van, Wanless RM, Paiva VH (2017) Identification of candidate pelagic marine protected areas through a seabird seasonal-, multispecific- and extinction risk-based approach. Anim Conserv 20:409–424.

Krüger L, Ramos JA, Xavier JC, Grémillet D, González-Solís J, Petry M V., Phillips RA, Wanless RM, Paiva VH (2018) Projected distributions of Southern Ocean albatrosses, petrels and fisheries as a consequence of climatic change. Ecography 41:195–208.

Legrand B, Benneveau A, Jaeger A, Pinet P, Potin G, Jaquemet S, Corre M Le (2016) Current wintering habitat of an endemic seabird of Réunion Island, Barau’s petrel *Pterodroma baraui*, and predicted changes induced by global warming. Mar Ecol Prog Ser 550:235–248.

Linnebjerg JF, Frederiksen M, Kolbeinsson Y, Snaethórsson AÖ, Thórisson B, Thórarinsson TL (2018) Non-breeding areas of three sympatric auk species breeding in three Icelandic colonies. Polar Biol 41:1951–1961.

Liu C, White M, Newell G (2009) Measuring the accuracy of species distribution models: a review. In: World IMACS/MODSIM Congress. Cairns, Australia, p 4241–4247.

Merkel B, Phillips RA, Descamps S, Yoccoz NG, Moe B, Strøm H (2016) A probabilistic algorithm to process geolocation data. Movement Ecology 4(1): 26.

Miller HJ (1991) Modelling accessibility using space-time prism concepts within geographical information systems. International Journal of Geographical Information Systems 5(3): 287-301.

Mitchell P, Newton S, Ratcliffe NR, Dunn TE (2004) Seabird Populations of Britain and Ireland. T. & AD Poyser, London.

National Geophysical Data Center (2006) 2-minute Gridded Global Relief Data (ETOPO2) v2. National Geophysical Data Center, NOAA. doi:10.7289/V5J1012Q [September, 2018]

Neutens T, Witlox F, Demaeyer P (2007) Individual accessibility and travel possibilities: A literature review on time geography. European Journal of Transport and Infrastructure Research 7(4):335-352.

NOAA (2018) Shoreline / Coastline Resources, National Oceanic & Atmospheric Administration. https://www.ngdc.noaa.gov/mgg/shorelines/. Downloaded 27-11-2018.

Oppel S, Meirinho A, Ramírez I, Gardner B, O’Connell AF, Miller PI, Louzao M (2012) Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biol Conserv 156:94–104.

Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. Ecol Appl 19:181–197.

R Development Core Team (2018) R: a language and environment for statistical computing. - In: R Foundation for Statistical Computing, ed., Vienna, Austria.

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang WQ (2002) An improved in situ and satellite SST analysis for climate. J Clim 15:1609–1625.

Scales KL, Miller PI, Ingram SN, Hazen EL, Bograd SJ, Phillips RA (2016) Identifying predictable foraging habitats for a wide-ranging marine predator using ensemble ecological niche. Divers Distrib 22:212–224.

Strøm et al. (2019) Large-scale tracking of seabirds in the Northeast-Atlantic 2014-2018 -

Final report from SEATRACK phase I. Norsk Polarinst. Rapportserie (Forthcoming).

Strøm H, Gavrilo MV, Krasnov JV, Systad GH (2009). Seabirds. In: Joint Norwegian-Russian Environmental Status 2008 Report on the Barents Sea Ecosystem. Part II – Complete report, pp. 67-73. Ed. by JE Stiansen, O Korneev, O Titov, P Arneberg, A Filin, JR Hansen, Å Høines and S Marasaev. IMR/PINRO Joint Report Series, 3/2009.

Technitis G, Othman W, Safi K, Weibel R (2015) From A to B, randomly: a point-to-point random trajectory generator for animal movement. International Journal of Geographical Information Science 29(6): 912-934.

Torres LG, Sutton PJH, Thompson DR, Delord K, Weimerskirch H, Sagar PM, Sommer E, Dilley BJ, Ryan PG, Phillips RA (2015) Poor Transferability of Species Distribution Models for a Pelagic Predator, the Grey Petrel, Indicates Contrasting Habitat Preferences across Ocean Basins (A Margalida, Ed.). PLoS One 10:e0120014

Tranquilla LAMF, Montevecchi WA, Hedd A, Fifield DA, Burke CM, Smith PA, Regular PM, Robertson GJ, Gaston AJ, Phillips RA (2013) Multiple-colony winter habitat use by murres uria spp. in the northwest atlantic ocean: Implications for marine risk assessment. Mar Ecol Prog Ser 472:287–303.

Tuszynski J (2018) caTools: Tools: moving window statistics, GIF, Base64, ROC AUC, etc.. R package version 1.17.1.1. https://CRAN.R-project.org/package=caTools.

Wessel P, Smith WHF (1996) A global, self-consistent, hierarchical, high-resolution shoreline database 101(B4): 8741-8743.

Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A, Elith J, Dudík M, Ferrier S, Huettmann F, Leathwick JR, Lehmann A, Lohmann L, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JMC, Phillips SJ, Richardson KS, Scachetti-Pereira R, Schapire RE, Soberón J, Williams SE, Zimmermann NE (2008) Effects of sample size on the performance of species distribution models. Divers Distrib 14:763–773

Wood SN (2006) Generalized Additive Models an Introduction with R. Chapman & Hall/CRC, Boca Raton, US.

Yurkowski, D.J., Auger-Méthé, M., Mallory, M.L., Wong, S.N.P., Gilchrist, G., Derocher, A.E., Richardson, E., Lunn, N.J., Hussey, N.E., Marcoux, M., Togunov, R.R., Fisk, A.T., Harwood, L.A., Dietz, R., Rosing-Asvid, A., Born, E.W., Mosbech, A., Fort, J., Grémillet, D., Loseto, L., Richard, P.R., Iacozza, J., Jean-Gagnon, F., Brown, T.M., Westdal, K.H., Orr, J., LeBlanc, B., Hedges, K.J., Treble, M.A., Kessel, S.T., Blanchfield, P.J., Davis, S., Maftei, M., Spencer, N., McFarlane-Tranquilla, L., Montevecchi, W.A., Bartzen, B., Dickson, L., Anderson, C. & Ferguson, S.H. 2019. Abundance and species diversity hotspots of tracked marine predators across the North American Arctic 25(3): 328-345.

6 Appendix

6.1 Summary of the algorithm used to determine a Potential Point Area (Ppa)

The output from this procedure (i.e. the Ppa) will allow generating a new location N_i between two known locations A and B. All distances are calculated as geodesics using function distGeo from package geosphere (Hijmans 2017) and the parameters of the WGS84 ellipsoid.

```
a) IDENTIFY ALL TIME GAPS THAT ARE > 0.5 DAY AND < 120 DAYS

b) LOOP THROUGH EACH GAP SEQUENTIALLY

   Load species-specific Speed Model which provides a value of maximum movement rate given the time elapsed between two consecutive locations

   Load individual-specific Longitude values (previously extracted from the raw original dataset)

   Integrate individual-specific activity data to identify breeding period and colony attendance

   Create timestamps at 12-h intervals to cover the entire period of the data gap

   b.1) GENERATE PPAS AT EACH TIMESTAMP IN A RANDOM SEQUENCE, FOLLOWING TECHNITIS ET AL. (2016), I.E. NOT IN CHRONOLOGICAL ORDER

         b.1.1 - CALCULATE MAXIMUM RANGES

         Calculate maximum ranges at A and B based on two parameters:
         - Maximum movement rates between A and $N_i$ and between $N_i$ and B
         - Time difference between A and $N_i$ and between $N_i$ and B
         This generates two range circles (SpatialPolygons)

         Do maximum ranges overlap?
          | ➔ NO: Increase maximum movement rates by 5% increments if they are below the speed threshold (otherwise no new location is generated)
          | ➔ YES: Calculate Ppa based on maximum ranges

         b.1.2 - EXCLUDE LANDMASSES & AREAS COVERED WITH SEA-ICE

         Crop the Ppa to section that does not overlap landmasses or sea-ice areas (removes the possibility to create any new location above land or sea ice)

         Is Ppa non-NULL after excluding land and sea-ice areas?
          | ➔ NO: Increase maximum movement rates by 5% increments as long as they are below the speed threshold (otherwise no new location is generated)
          | ➔ YES: No problem
```

b.1.3 TAKING POLAR DAY/NIGHT AREA (PDNA) INTO ACCOUNT

The PDNA is defined as the area north of the lowest latitude at which the sun does not rise (polar night) or set (polar day) for a given date. Civil twilight (i.e. a solar angle of 6° below the horizon), was used as threshold for identifying sunrise/sunset events

Are there locations recorded at timestamp T_i in the raw dataset?

- ➔ **YES:** Twilight events thus occurred, and the new location is constrained outside of the PDNA
- ➔ **NO**: IRMA assumes this is due to the absence of twilight events (i.e. the bird was either in complete darkness (polar night) or complete daylight (polar day) 24-h a day. The new location is constrained to be within the PDNA only for short gaps (<10 days) that start and finish inside of the PDNA, otherwise no constraint is applied

  Do Ppa and PDNA overlap?

  - ➔ **NO:** Increase maximum movement rates by 5% increments until Ppa and PDNA overlap and if movement rates < speed threshold. If the speed threshold if reached, increase the size of the longitude buffer, by extracting longitude data over increasingly longer periods (by 1-day increments), until a maximum of 10 days is reached. If the time period threshold is reached, then longitude buffer is set to NULL
  - ➔ **YES:** Crop the Ppa to area intersecting the longitude buffer

b.1.4 - RESTRICT Ppa TO MOST LIKELY LONGITUDES

Extract range of longitudes from actual GLS-based data (+/- 1 day) and create a polygon (SpatialPolygon) delimited by the longitudes (SpatialPolygon) – value is NULL if no data available for the considered period/individual. Following step ignored if no longitude buffer can be determined

Do Ppa and longitude buffer overlap?

- ➔ **NO:** Increase maximum movement rates by 5% increments until Ppa and longitude buffer overlap and if movement rates < speed threshold. If the speed threshold if reached, increase the size of the longitude buffer, by extracting longitude data over increasingly longer periods (by 1-day increments), until a maximum of 10 days is reached. If the time period threshold is reached, then longitude buffer is set to NULL
- ➔ **YES:** Crop the Ppa to area intersecting the longitude buffer

6.2 Model specification, sample size and model diagnostics for each Species Distribution Model (SDM)

Nind is number of individuals included in the sample, *IRMApos* is the number of positions generated from IRMA, *GLSpos* is the number of GLS positions, *Add* indicates whether positions from 15 day before and after *month* was included in the sample (1 = included, 0 = not included). *Model* indicate type of model conducted (1 = full model, 2 = full model plus east variable, 3 = simplified model, 4 = distance to colony only; see 2.2.5 for further specifications). Sample included in the models, after removing positions with missing environmental variables, is given as: *nPres* (number of presence positions) and *nBackval* (number of background positions drawn randomly from the habitat of the species). *RsqAdj* is the adjusted R-square value from the model and *DevExp* is the proportion of deviance explained by the model.

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Little auk	Bjørnøya	1	29	563	2312	0	1	2784	59744	0.238	0.394
Little auk	Bjørnøya	2	29	934	1673	0	1	2567	59808	0.196	0.343
Little auk	Bjørnøya	3	29	2852	0	0	1	2818	59782	0.172	0.320
Little auk	Bjørnøya	4	29	582	950	0	1	1506	59775	0.407	0.547
Little auk	Bjørnøya	5	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Bjørnøya	6	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Bjørnøya	7	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Bjørnøya	8	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Bjørnøya	9	11	698	40	0	1	697	39588	0.219	0.468
Little auk	Bjørnøya	10	29	691	648	0	1	1266	59441	0.141	0.359
Little auk	Bjørnøya	11	29	369	2499	0	1	2744	59633	0.180	0.326
Little auk	Bjørnøya	12	29	742	2234	0	1	2805	59695	0.285	0.417
Little auk	Franz Josef Land	1	36	1748	734	0	1	2454	39845	0.307	0.467
Little auk	Franz Josef Land	2	36	75	1507	0	1	1447	39870	0.420	0.577
Little auk	Franz Josef Land	3	2	124	0	1	1	433	39851	0.233	0.520
Little auk	Franz Josef Land	4	2	26	6	1	1	77	19924	0.294	0.652
Little auk	Franz Josef Land	5	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Franz Josef Land	6	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Franz Josef Land	7	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Franz Josef Land	8	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Franz Josef Land	9	10	496	44	0	1	391	29681	0.421	0.666
Little auk	Franz Josef Land	10	36	531	509	0	1	972	39626	0.361	0.589
Little auk	Franz Josef Land	11	36	1807	591	0	1	2266	39801	0.303	0.471
Little auk	Franz Josef Land	12	36	2268	210	0	1	2440	39796	0.262	0.423
Little auk	Hornsund	1	53	631	2841	0	1	3431	19913	0.459	0.497
Little auk	Hornsund	2	53	1036	2018	0	1	3015	19918	0.409	0.451
Little auk	Hornsund	3	47	3100	0	0	1	3076	19919	0.393	0.438
Little auk	Hornsund	4	47	550	631	0	1	1129	19913	0.414	0.536
Little auk	Hornsund	5	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Hornsund	6	0	0	0	0	NA	NA	NA	NA	NA

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Little auk	Hornsund	7	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Hornsund	8	2	0	2	0	NA	NA	NA	NA	NA
Little auk	Hornsund	9	15	780	59	0	1	802	19820	0.353	0.527
Little auk	Hornsund	10	53	726	772	0	1	1457	19799	0.338	0.460
Little auk	Hornsund	11	53	273	3074	0	1	3322	19856	0.473	0.511
Little auk	Hornsund	12	53	801	2669	0	1	3418	19899	0.541	0.558
Little auk	Isfjorden	1	15	299	791	0	1	1052	29861	0.204	0.394
Little auk	Isfjorden	2	15	310	610	0	1	890	29905	0.169	0.333
Little auk	Isfjorden	3	13	930	0	0	1	917	29883	0.222	0.407
Little auk	Isfjorden	4	13	174	185	0	1	346	29882	0.238	0.530
Little auk	Isfjorden	5	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Isfjorden	6	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Isfjorden	7	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Isfjorden	8	7	16	58	0	NA	NA	NA	NA	NA
Little auk	Isfjorden	9	10	620	87	0	1	640	19769	0.287	0.497
Little auk	Isfjorden	10	15	595	231	0	1	808	29711	0.248	0.437
Little auk	Isfjorden	11	15	290	785	0	1	1055	29825	0.275	0.458
Little auk	Isfjorden	12	15	422	692	0	1	1072	29856	0.228	0.440
Little auk	Kongsfjorden	1	16	107	1319	0	1	649	29869	0.283	0.516
Little auk	Kongsfjorden	2	16	418	862	0	1	622	29908	0.132	0.343
Little auk	Kongsfjorden	3	16	1364	0	0	1	643	29908	0.087	0.296
Little auk	Kongsfjorden	4	16	245	284	0	1	283	29892	0.137	0.419
Little auk	Kongsfjorden	5	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Kongsfjorden	6	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Kongsfjorden	7	0	0	0	0	NA	NA	NA	NA	NA
Little auk	Kongsfjorden	8	3	8	43	1	1	142	29674	0.306	0.639
Little auk	Kongsfjorden	9	10	646	106	0	1	239	29658	0.213	0.547
Little auk	Kongsfjorden	10	16	637	319	0	1	316	29711	0.079	0.318
Little auk	Kongsfjorden	11	16	110	1327	0	1	526	29825	0.167	0.430
Little auk	Kongsfjorden	12	16	186	1276	0	1	502	29836	0.378	0.623
Puffin	Anda	1	15	131	1050	0	1	1117	29744	0.342	0.530
Puffin	Anda	2	15	385	643	0	1	986	29752	0.214	0.433
Puffin	Anda	3	13	992	0	0	1	938	29751	0.356	0.533
Puffin	Anda	4	13	461	493	0	1	866	29719	0.682	0.744
Puffin	Anda	5	8	189	21	0	4	197	29755	0.654	0.818
Puffin	Anda	6	3	180	0	0	4	161	9920	0.854	0.903
Puffin	Anda	7	3	186	0	0	4	179	9900	0.784	0.860
Puffin	Anda	8	15	258	280	0	1	535	29710	0.384	0.636
Puffin	Anda	9	15	970	170	0	1	1140	29750	0.607	0.732
Puffin	Anda	10	15	932	240	0	1	1153	29738	0.345	0.517
Puffin	Anda	11	15	175	960	0	1	1075	29753	0.349	0.533

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Puffin	Anda	12	15	315	863	0	1	1128	29752	0.407	0.574
Puffin	Faroe Islands	1	6	0	558	1	3	1116	29753	0.120	0.279
Puffin	Faroe Islands	2	6	169	343	1	3	1070	19851	0.225	0.377
Puffin	Faroe Islands	3	6	558	0	1	1	1070	19833	0.365	0.523
Puffin	Faroe Islands	4	6	117	423	1	1	1098	19829	0.288	0.468
Puffin	Faroe Islands	5	6	184	374	1	1	1077	19804	0.450	0.575
Puffin	Faroe Islands	6	6	330	120	1	1	901	19826	0.735	0.803
Puffin	Faroe Islands	7	4	277	95	1	4	753	29727	0.751	0.820
Puffin	Faroe Islands	8	6	129	364	1	1	935	29725	0.549	0.668
Puffin	Faroe Islands	9	6	458	82	1	1	1058	19827	0.419	0.561
Puffin	Faroe Islands	10	6	436	122	1	3	1066	19815	0.187	0.354
Puffin	Faroe Islands	11	6	10	530	1	3	1094	19803	0.152	0.327
Puffin	Faroe Islands	12	6	0	558	1	3	1098	29715	0.089	0.266
Puffin	Grimsey	1	27	8	2286	0	1	2294	29740	0.329	0.467
Puffin	Grimsey	2	27	669	1433	0	1	2100	29735	0.295	0.437
Puffin	Grimsey	3	27	2294	0	0	2	2256	29734	0.317	0.469
Puffin	Grimsey	4	27	567	1653	0	1	2136	29739	0.397	0.494
Puffin	Grimsey	5	27	702	543	0	1	1205	29745	0.791	0.845
Puffin	Grimsey	6	8	600	0	0	1	574	19815	0.936	0.946
Puffin	Grimsey	7	20	666	45	0	1	682	29701	0.875	0.910
Puffin	Grimsey	8	28	727	1607	0	1	2215	29705	0.515	0.605
Puffin	Grimsey	9	28	1873	369	0	2	2223	29708	0.591	0.687
Puffin	Grimsey	10	27	1795	499	0	1	2254	29684	0.459	0.571
Puffin	Grimsey	11	27	16	2204	0	1	2220	29732	0.461	0.569
Puffin	Grimsey	12	27	16	2278	0	1	2291	29709	0.431	0.541
Puffin	Hjelmsøya	1	20	456	909	0	1	1332	29738	0.367	0.548
Puffin	Hjelmsøya	2	20	432	798	0	1	1196	29740	0.353	0.535
Puffin	Hjelmsøya	3	19	1302	0	0	1	1197	29748	0.633	0.732
Puffin	Hjelmsøya	4	19	384	499	0	1	838	29730	0.727	0.818
Puffin	Hjelmsøya	5	0	0	0	0	NA	NA	NA	NA	NA
Puffin	Hjelmsøya	6	0	0	0	0	NA	NA	NA	NA	NA
Puffin	Hjelmsøya	7	0	0	0	0	NA	NA	NA	NA	NA
Puffin	Hjelmsøya	8	19	98	247	0	NA	NA	NA	NA	NA
Puffin	Hjelmsøya	9	20	1151	172	0	1	1323	29706	0.613	0.731
Puffin	Hjelmsøya	10	20	1059	305	0	1	1361	29748	0.568	0.702
Puffin	Hjelmsøya	11	20	514	804	0	1	1318	29752	0.492	0.637
Puffin	Hjelmsøya	12	20	970	393	0	1	1298	29753	0.357	0.539
Puffin	Hornøya	1	51	2930	2835	0	1	5708	39660	0.677	0.709
Puffin	Hornøya	2	51	1847	3393	0	1	5126	39698	0.642	0.686
Puffin	Hornøya	3	50	5642	0	0	1	5551	39655	0.765	0.785
Puffin	Hornøya	4	50	2328	2691	0	1	4805	39644	0.900	0.903

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Puffin	Hornøya	5	27	2249	10	0	1	2108	39657	0.927	0.936
Puffin	Hornøya	6	26	2160	0	0	1	2006	29727	0.921	0.932
Puffin	Hornøya	7	26	2232	0	0	1	2104	29736	0.858	0.881
Puffin	Hornøya	8	51	1994	1851	0	1	3767	39672	0.730	0.773
Puffin	Hornøya	9	51	4851	789	0	1	5604	39638	0.749	0.771
Puffin	Hornøya	10	51	4564	1264	0	1	5806	39612	0.762	0.778
Puffin	Hornøya	11	51	1413	4220	0	1	5607	39633	0.730	0.755
Puffin	Hornøya	12	51	4848	921	0	1	5733	39637	0.661	0.695
Puffin	Isle of May	1	38	313	2614	0	3	2726	59478	0.607	0.699
Puffin	Isle of May	2	37	989	1569	0	1	2389	59508	0.565	0.690
Puffin	Isle of May	3	36	2728	0	0	3	2496	59469	0.559	0.691
Puffin	Isle of May	4	36	1354	1191	0	1	2256	59487	0.796	0.845
Puffin	Isle of May	5	33	1770	566	0	1	1985	59439	0.830	0.873
Puffin	Isle of May	6	32	1041	676	0	1	1417	59450	0.794	0.862
Puffin	Isle of May	7	39	1575	1502	0	1	2576	59450	0.828	0.866
Puffin	Isle of May	8	39	971	2320	0	1	2946	49550	0.745	0.772
Puffin	Isle of May	9	39	2746	388	0	1	2831	49552	0.622	0.704
Puffin	Isle of May	10	39	2632	618	0	1	2900	49572	0.605	0.694
Puffin	Isle of May	11	39	449	2564	0	1	2744	49594	0.578	0.684
Puffin	Isle of May	12	38	365	2600	0	1	2714	49584	0.645	0.730
Puffin	Papey	1	19	4	1794	0	1	1798	29749	0.343	0.495
Puffin	Papey	2	19	521	1131	0	2	1652	29753	0.302	0.461
Puffin	Papey	3	19	1798	0	0	1	1785	29715	0.247	0.408
Puffin	Papey	4	19	765	975	0	1	1682	29706	0.537	0.600
Puffin	Papey	5	19	923	340	0	1	1232	29734	0.839	0.877
Puffin	Papey	6	9	600	0	0	1	586	19806	0.899	0.924
Puffin	Papey	7	14	663	77	0	1	720	39614	0.683	0.782
Puffin	Papey	8	19	477	1261	0	1	1664	39642	0.354	0.499
Puffin	Papey	9	19	1457	283	0	2	1721	29707	0.426	0.568
Puffin	Papey	10	19	1414	384	0	1	1796	29735	0.393	0.518
Puffin	Papey	11	19	38	1702	0	1	1733	29717	0.294	0.461
Puffin	Papey	12	19	14	1784	0	2	1797	29727	0.363	0.522
Puffin	Røst	1	57	257	4836	0	1	4870	39638	0.440	0.527
Puffin	Røst	2	57	1705	2937	0	1	4441	39701	0.332	0.439
Puffin	Røst	3	57	5022	0	0	1	4695	39630	0.380	0.473
Puffin	Røst	4	57	1808	3075	0	1	4752	39639	0.706	0.712
Puffin	Røst	5	53	1416	776	0	1	2151	39653	0.792	0.835
Puffin	Røst	6	19	1317	3	0	4	1306	29703	0.707	0.776
Puffin	Røst	7	25	1356	21	0	1	1367	29723	0.730	0.798
Puffin	Røst	8	56	1416	1012	0	1	2413	29734	0.628	0.700
Puffin	Røst	9	57	4091	816	0	1	4871	29727	0.698	0.712

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Puffin	Røst	10	57	4098	983	0	1	4893	39646	0.408	0.488
Puffin	Røst	11	57	617	4310	0	1	4810	39663	0.431	0.513
Puffin	Røst	12	57	746	4336	0	2	4944	39674	0.491	0.568
Puffin	Runde and Ålesund	1	13	39	830	0	1	853	29750	0.164	0.386
Puffin	Runde and Ålesund	2	13	247	495	0	1	722	29741	0.146	0.360
Puffin	Runde and Ålesund	3	12	806	0	0	1	750	29756	0.239	0.450
Puffin	Runde and Ålesund	4	12	450	339	0	1	731	29720	0.677	0.781
Puffin	Runde and Ålesund	5	12	404	181	0	1	550	29716	0.765	0.852
Puffin	Runde and Ålesund	6	2	63	1	0	4	57	19837	0.581	0.813
Puffin	Runde and Ålesund	7	9	167	105	0	4	242	19816	0.698	0.811
Puffin	Runde and Ålesund	8	13	519	207	0	1	718	29713	0.369	0.574
Puffin	Runde and Ålesund	9	13	739	101	0	1	818	29744	0.492	0.651
Puffin	Runde and Ålesund	10	13	706	159	0	1	821	29756	0.245	0.421
Puffin	Runde and Ålesund	11	13	36	801	0	1	823	29722	0.260	0.485
Puffin	Runde and Ålesund	12	13	23	845	0	2	850	29755	0.264	0.493
Puffin	Sklinna	1	19	152	1211	0	1	1323	29731	0.266	0.439
Puffin	Sklinna	2	19	476	773	0	1	1211	29753	0.226	0.421
Puffin	Sklinna	3	19	1364	0	0	1	1323	29736	0.292	0.464
Puffin	Sklinna	4	19	670	602	0	1	1186	29720	0.704	0.762
Puffin	Sklinna	5	17	357	238	0	1	579	29746	0.709	0.816
Puffin	Sklinna	6	3	180	0	0	4	176	19833	0.791	0.879
Puffin	Sklinna	7	6	218	18	0	1	228	29700	0.739	0.849
Puffin	Sklinna	8	19	567	392	0	1	945	29714	0.429	0.602
Puffin	Sklinna	9	19	1148	172	0	1	1317	29736	0.514	0.607
Puffin	Sklinna	10	19	1089	273	0	1	1288	29760	0.242	0.417
Puffin	Sklinna	11	19	121	1195	0	1	1288	29745	0.216	0.394
Puffin	Sklinna	12	19	102	1263	0	1	1332	29714	0.279	0.454
Fulmar	Alkefjellet	1	5	239	134	0	1	369	19660	0.265	0.534
Fulmar	Alkefjellet	2	5	197	143	0	1	316	19673	0.267	0.536
Fulmar	Alkefjellet	3	5	372	0	0	1	319	19586	0.350	0.597
Fulmar	Alkefjellet	4	5	134	61	0	1	144	19634	0.212	0.550
Fulmar	Alkefjellet	5	0	0	0	0	NA	NA	NA	NA	NA
Fulmar	Alkefjellet	6	0	0	0	0	NA	NA	NA	NA	NA
Fulmar	Alkefjellet	7	0	0	0	0	NA	NA	NA	NA	NA
Fulmar	Alkefjellet	8	2	0	3	0	NA	NA	NA	NA	NA
Fulmar	Alkefjellet	9	3	208	20	0	4	226	19593	0.080	0.323
Fulmar	Alkefjellet	10	5	221	55	0	1	265	19585	0.296	0.567
Fulmar	Alkefjellet	11	5	245	115	0	1	353	19632	0.173	0.439
Fulmar	Alkefjellet	12	5	318	55	0	1	370	19643	0.301	0.525
Fulmar	Bjørnøya	1	26	2402	887	0	1	3227	29493	0.612	0.670
Fulmar	Bjørnøya	2	26	1192	1796	0	1	2927	29478	0.670	0.721

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Fulmar	Bjørnøya	3	26	3224	0	0	1	3159	29478	0.679	0.726
Fulmar	Bjørnøya	4	26	977	1027	0	1	1990	29434	0.679	0.745
Fulmar	Bjørnøya	5	3	186	0	0	1	186	9801	0.752	0.839
Fulmar	Bjørnøya	6	3	180	0	0	1	180	9812	0.814	0.880
Fulmar	Bjørnøya	7	3	186	0	0	1	183	9806	0.662	0.772
Fulmar	Bjørnøya	8	23	254	352	0	1	597	29415	0.441	0.630
Fulmar	Bjørnøya	9	26	2594	356	0	4	2892	29364	0.463	0.538
Fulmar	Bjørnøya	10	26	2498	646	0	1	3052	29396	0.631	0.684
Fulmar	Bjørnøya	11	26	1817	1363	0	1	3151	29440	0.560	0.628
Fulmar	Bjørnøya	12	26	3220	66	0	1	3226	29420	0.397	0.515
Fulmar	Eynhallow	1	44	701	3324	0	1	3919	29480	0.656	0.674
Fulmar	Eynhallow	2	44	1597	2097	0	1	3565	29449	0.586	0.615
Fulmar	Eynhallow	3	44	4030	0	0	1	3822	29449	0.451	0.518
Fulmar	Eynhallow	4	44	1363	2549	0	1	3791	29419	0.577	0.610
Fulmar	Eynhallow	5	44	1598	2159	0	1	3683	29448	0.598	0.623
Fulmar	Eynhallow	6	29	1799	1028	0	1	2777	39238	0.653	0.699
Fulmar	Eynhallow	7	44	1234	2085	0	1	3228	39196	0.674	0.703
Fulmar	Eynhallow	8	44	890	3274	0	1	4046	29326	0.476	0.483
Fulmar	Eynhallow	9	44	3408	571	0	1	3863	29428	0.277	0.307
Fulmar	Eynhallow	10	44	3268	820	0	1	3943	29387	0.216	0.280
Fulmar	Eynhallow	11	44	422	3535	0	1	3865	29450	0.393	0.426
Fulmar	Eynhallow	12	44	650	3380	0	1	3927	29434	0.545	0.567
Fulmar	Faroe Islands	1	11	45	700	0	1	739	19645	0.323	0.496
Fulmar	Faroe Islands	2	11	248	435	0	1	676	19631	0.321	0.511
Fulmar	Faroe Islands	3	11	744	0	0	1	727	19611	0.230	0.435
Fulmar	Faroe Islands	4	11	227	494	0	1	708	19626	0.254	0.418
Fulmar	Faroe Islands	5	11	351	301	0	1	650	19637	0.453	0.571
Fulmar	Faroe Islands	6	8	330	91	0	1	419	19609	0.680	0.767
Fulmar	Faroe Islands	7	10	185	304	0	1	484	29406	0.405	0.608
Fulmar	Faroe Islands	8	11	125	612	0	1	733	19585	0.297	0.429
Fulmar	Faroe Islands	9	11	627	94	0	1	712	19591	0.064	0.180
Fulmar	Faroe Islands	10	11	597	150	0	1	726	19604	0.089	0.230
Fulmar	Faroe Islands	11	11	109	610	0	1	706	19627	0.154	0.325
Fulmar	Faroe Islands	12	11	118	625	0	1	733	19670	0.189	0.359
Fulmar	Jan Mayen	1	33	1093	2519	0	1	3585	29514	0.602	0.655
Fulmar	Jan Mayen	2	33	1240	2043	0	1	3247	29487	0.553	0.634
Fulmar	Jan Mayen	3	33	3596	0	0	1	3580	29437	0.470	0.562
Fulmar	Jan Mayen	4	33	988	2325	0	1	3286	29457	0.590	0.633
Fulmar	Jan Mayen	5	22	936	357	0	1	1287	29422	0.556	0.584
Fulmar	Jan Mayen	6	12	924	11	0	1	926	29423	0.879	0.903
Fulmar	Jan Mayen	7	14	904	95	0	1	983	29405	0.711	0.727

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Fulmar	Jan Mayen	8	21	616	1052	0	1	1649	29373	0.140	0.249
Fulmar	Jan Mayen	9	29	2599	358	0	1	2903	29371	0.247	0.297
Fulmar	Jan Mayen	10	33	2429	807	0	1	3128	29383	0.208	0.290
Fulmar	Jan Mayen	11	33	568	3038	0	1	3579	29411	0.298	0.384
Fulmar	Jan Mayen	12	33	2236	1468	0	1	3668	29442	0.441	0.500
Fulmar	Jarsteinen	1	8	339	532	0	1	832	39288	0.436	0.576
Fulmar	Jarsteinen	2	8	417	378	0	1	722	39272	0.414	0.637
Fulmar	Jarsteinen	3	8	868	0	0	1	807	39271	0.336	0.573
Fulmar	Jarsteinen	4	8	426	415	0	1	773	39259	0.518	0.681
Fulmar	Jarsteinen	5	8	495	377	0	1	776	39325	0.585	0.721
Fulmar	Jarsteinen	6	8	707	167	0	1	706	39286	0.766	0.843
Fulmar	Jarsteinen	7	8	497	327	0	1	535	39198	0.618	0.753
Fulmar	Jarsteinen	8	8	348	518	0	1	523	29365	0.267	0.423
Fulmar	Jarsteinen	9	8	734	106	0	1	517	29384	0.128	0.342
Fulmar	Jarsteinen	10	8	731	139	0	1	524	29400	0.090	0.275
Fulmar	Jarsteinen	11	8	302	541	0	1	523	29409	0.193	0.379
Fulmar	Jarsteinen	12	8	351	520	0	1	539	29467	0.255	0.432
Fulmar	Langanes and Skjalfandi	1	46	841	3435	0	1	4169	29514	0.509	0.530
Fulmar	Langanes and Skjalfandi	2	46	1814	2095	0	1	3750	29496	0.643	0.665
Fulmar	Langanes and Skjalfandi	3	46	4278	0	0	1	4153	29440	0.538	0.572
Fulmar	Langanes and Skjalfandi	4	46	2187	1950	0	1	3902	29487	0.601	0.609
Fulmar	Langanes and Skjalfandi	5	45	1824	1642	0	1	3279	29438	0.513	0.565
Fulmar	Langanes and Skjalfandi	6	32	1357	267	0	1	1525	29400	0.729	0.749
Fulmar	Langanes and Skjalfandi	7	32	1458	452	0	1	1807	29385	0.569	0.605
Fulmar	Langanes and Skjalfandi	8	43	1364	1953	0	1	3234	29382	0.154	0.233
Fulmar	Langanes and Skjalfandi	9	45	3406	490	0	1	3756	29363	0.208	0.260
Fulmar	Langanes and Skjalfandi	10	46	3243	833	0	1	3851	29363	0.188	0.252
Fulmar	Langanes and Skjalfandi	11	46	677	3463	0	1	4088	29449	0.287	0.349
Fulmar	Langanes and Skjalfandi	12	46	981	3299	0	1	4206	29425	0.317	0.383
Kittiwake	Alkefjellet	1	20	10	1602	0	1	1584	19530	0.248	0.345
Kittiwake	Alkefjellet	2	20	524	948	0	1	1459	19533	0.188	0.305
Kittiwake	Alkefjellet	3	20	1612	0	0	1	1603	19518	0.178	0.284
Kittiwake	Alkefjellet	4	20	299	593	0	1	855	19497	0.066	0.186
Kittiwake	Alkefjellet	5	1	0	11	1	3	154	19534	0.034	0.249

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Kittiwake	Alkefjellet	6	0	0	0	0	NA	NA	NA	NA	NA
Kittiwake	Alkefjellet	7	0	0	0	0	NA	NA	NA	NA	NA
Kittiwake	Alkefjellet	8	3	2	3	1	3	231	19436	0.509	0.720
Kittiwake	Alkefjellet	9	9	526	23	0	1	524	19417	0.483	0.663
Kittiwake	Alkefjellet	10	20	528	459	0	1	967	19416	0.360	0.478
Kittiwake	Alkefjellet	11	20	195	1361	0	1	1546	19481	0.232	0.311
Kittiwake	Alkefjellet	12	20	8	1604	0	1	1604	19479	0.305	0.418
Kittiwake	Anda	1	52	1575	4065	0	1	5391	38999	0.373	0.385
Kittiwake	Anda	2	52	2661	2340	0	1	4770	39016	0.757	0.735
Kittiwake	Anda	3	51	5394	0	0	1	5117	38997	0.863	0.857
Kittiwake	Anda	4	51	2107	3096	0	1	4793	39020	0.906	0.904
Kittiwake	Anda	5	48	2351	331	0	1	2394	39022	0.903	0.905
Kittiwake	Anda	6	26	2220	0	0	1	1991	29194	0.927	0.924
Kittiwake	Anda	7	26	2289	5	0	1	2060	29163	0.958	0.957
Kittiwake	Anda	8	53	2858	1805	0	1	4373	38811	0.756	0.772
Kittiwake	Anda	9	53	4744	627	0	1	5202	38781	0.529	0.566
Kittiwake	Anda	10	54	3828	1791	0	1	5364	38876	0.251	0.311
Kittiwake	Anda	11	54	406	5104	0	1	5421	38979	0.350	0.380
Kittiwake	Anda	12	54	484	5157	0	1	5564	39056	0.361	0.410
Kittiwake	Bjørnøya	1	40	994	3163	0	1	3911	39044	0.247	0.311
Kittiwake	Bjørnøya	2	40	1726	1905	0	1	3474	39066	0.252	0.304
Kittiwake	Bjørnøya	3	39	3906	0	0	1	3804	38973	0.613	0.642
Kittiwake	Bjørnøya	4	39	990	1388	0	1	2342	39000	0.724	0.778
Kittiwake	Bjørnøya	5	3	124	1	0	4	125	19500	0.636	0.791
Kittiwake	Bjørnøya	6	2	120	0	0	4	120	19485	0.687	0.842
Kittiwake	Bjørnøya	7	2	124	0	0	4	124	19476	0.724	0.856
Kittiwake	Bjørnøya	8	28	202	295	0	1	491	38828	0.510	0.721
Kittiwake	Bjørnøya	9	39	3309	492	0	1	3747	38822	0.676	0.727
Kittiwake	Bjørnøya	10	40	3082	1031	0	1	3916	38811	0.373	0.417
Kittiwake	Bjørnøya	11	40	859	3211	0	1	3888	38968	0.239	0.312
Kittiwake	Bjørnøya	12	40	911	3303	0	1	3965	39009	0.323	0.396
Kittiwake	Cape Krutik	1	31	595	3452	0	1	3946	29326	0.191	0.246
Kittiwake	Cape Krutik	2	31	1517	2131	0	1	3538	29345	0.278	0.297
Kittiwake	Cape Krutik	3	31	3906	0	0	1	3763	29227	0.692	0.695
Kittiwake	Cape Krutik	4	31	2092	1566	0	1	3444	29227	0.880	0.884
Kittiwake	Cape Krutik	5	29	1247	116	0	1	1262	29235	0.843	0.870
Kittiwake	Cape Krutik	6	17	1080	0	0	1	1002	19521	0.893	0.909
Kittiwake	Cape Krutik	7	17	1116	0	0	1	1032	19422	0.910	0.927
Kittiwake	Cape Krutik	8	29	1362	793	0	1	2092	29102	0.688	0.760
Kittiwake	Cape Krutik	9	33	3162	431	0	1	3515	29147	0.626	0.669
Kittiwake	Cape Krutik	10	33	2609	1321	0	1	3768	29096	0.272	0.327

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Kittiwake	Cape Krutik	11	33	243	3787	0	1	4004	29246	0.328	0.368
Kittiwake	Cape Krutik	12	31	363	3670	0	1	3976	29237	0.308	0.369
Kittiwake	Cape Sakhanin	1	5	42	269	0	1	296	19482	0.251	0.456
Kittiwake	Cape Sakhanin	2	5	115	167	0	1	264	19516	0.165	0.411
Kittiwake	Cape Sakhanin	3	5	310	0	0	1	286	19519	0.185	0.452
Kittiwake	Cape Sakhanin	4	5	75	212	0	1	266	19513	0.473	0.681
Kittiwake	Cape Sakhanin	5	1	1	1	0	NA	NA	NA	NA	NA
Kittiwake	Cape Sakhanin	6	0	0	0	0	NA	NA	NA	NA	NA
Kittiwake	Cape Sakhanin	7	0	0	0	0	NA	NA	NA	NA	NA
Kittiwake	Cape Sakhanin	8	5	105	90	0	1	186	19380	0.578	0.749
Kittiwake	Cape Sakhanin	9	5	272	28	0	1	295	19411	0.339	0.591
Kittiwake	Cape Sakhanin	10	5	224	86	0	1	291	19428	0.384	0.621
Kittiwake	Cape Sakhanin	11	5	98	202	0	1	283	19476	0.317	0.518
Kittiwake	Cape Sakhanin	12	5	136	174	0	1	300	19512	0.198	0.448
Kittiwake	Faroe Islands	1	21	43	2685	0	1	2716	29238	0.144	0.247
Kittiwake	Faroe Islands	2	21	961	1535	0	1	2474	29265	0.237	0.353
Kittiwake	Faroe Islands	3	21	2728	0	0	1	2684	29310	0.603	0.653
Kittiwake	Faroe Islands	4	21	358	2283	0	1	2622	29242	0.659	0.716
Kittiwake	Faroe Islands	5	21	664	2039	0	1	2676	29253	0.726	0.767
Kittiwake	Faroe Islands	6	21	1478	414	0	1	1869	29182	0.856	0.879
Kittiwake	Faroe Islands	7	21	432	1744	0	1	2157	29195	0.807	0.827
Kittiwake	Faroe Islands	8	21	729	1999	0	1	2564	29111	0.330	0.430
Kittiwake	Faroe Islands	9	21	2199	441	0	1	2403	29104	0.301	0.414
Kittiwake	Faroe Islands	10	21	1848	881	0	2	2616	29150	0.383	0.475
Kittiwake	Faroe Islands	11	21	74	2566	0	1	2634	29238	0.417	0.518
Kittiwake	Faroe Islands	12	21	17	2711	0	1	2723	29257	0.334	0.448
Kittiwake	Franz Josef Land	1	54	175	4911	0	1	4868	39037	0.173	0.228
Kittiwake	Franz Josef Land	2	54	1697	2946	0	1	4494	38995	0.153	0.209
Kittiwake	Franz Josef Land	3	54	5084	0	0	1	4957	39004	0.136	0.186
Kittiwake	Franz Josef Land	4	54	1104	2099	0	1	3048	38987	0.109	0.205
Kittiwake	Franz Josef Land	5	7	68	53	1	1	778	38972	0.067	0.249
Kittiwake	Franz Josef Land	6	1	60	0	1	4	47	9747	0.921	0.940
Kittiwake	Franz Josef Land	7	1	62	0	1	4	117	9744	0.871	0.919
Kittiwake	Franz Josef Land	8	1	62	0	1	3	668	38858	0.705	0.800
Kittiwake	Franz Josef Land	9	26	1410	83	0	1	1424	38794	0.606	0.721
Kittiwake	Franz Josef Land	10	54	1427	1582	0	1	2934	38849	0.332	0.412
Kittiwake	Franz Josef Land	11	54	641	4277	0	1	4887	38976	0.245	0.300
Kittiwake	Franz Josef Land	12	54	112	4972	0	1	5055	39066	0.305	0.370
Kittiwake	Hornøya	1	49	379	3671	0	1	3954	39050	0.179	0.262
Kittiwake	Hornøya	2	49	1635	2062	0	1	3581	39028	0.315	0.364
Kittiwake	Hornøya	3	49	4030	0	0	1	3760	39021	0.698	0.718

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Kittiwake	Hornøya	4	49	1853	1855	0	1	3445	39023	0.905	0.908
Kittiwake	Hornøya	5	25	245	63	0	1	286	39042	0.658	0.791
Kittiwake	Hornøya	6	3	180	0	0	4	170	9715	0.874	0.912
Kittiwake	Hornøya	7	3	186	0	0	4	170	9719	0.927	0.940
Kittiwake	Hornøya	8	28	301	219	0	1	505	29130	0.320	0.550
Kittiwake	Hornøya	9	42	2775	361	0	1	3072	38852	0.621	0.675
Kittiwake	Hornøya	10	49	2392	1238	0	1	3510	38844	0.236	0.318
Kittiwake	Hornøya	11	49	210	3689	0	1	3875	38927	0.350	0.428
Kittiwake	Hornøya	12	49	146	3883	0	1	4003	39029	0.273	0.371
Kittiwake	Isfjorden	1	29	73	3213	0	1	3168	29316	0.168	0.242
Kittiwake	Isfjorden	2	29	1090	1901	0	1	2899	29276	0.156	0.216
Kittiwake	Isfjorden	3	29	3224	0	0	1	3085	29257	0.123	0.201
Kittiwake	Isfjorden	4	29	1082	1206	0	1	2087	29293	0.200	0.303
Kittiwake	Isfjorden	5	12	284	61	0	1	273	29231	0.599	0.677
Kittiwake	Isfjorden	6	4	240	0	0	4	177	9759	0.973	0.977
Kittiwake	Isfjorden	7	4	248	0	0	4	191	9736	0.952	0.958
Kittiwake	Isfjorden	8	19	283	92	0	1	313	29065	0.546	0.719
Kittiwake	Isfjorden	9	29	2713	353	0	1	2931	29099	0.574	0.637
Kittiwake	Isfjorden	10	29	2282	998	0	1	3138	29144	0.350	0.424
Kittiwake	Isfjorden	11	29	350	2824	0	1	3095	29205	0.218	0.284
Kittiwake	Isfjorden	12	29	36	3250	0	1	3192	29225	0.292	0.369
Kittiwake	Isle of May	1	35	225	3183	0	1	3365	29298	0.208	0.284
Kittiwake	Isle of May	2	35	1272	1839	0	1	3055	29311	0.130	0.216
Kittiwake	Isle of May	3	35	3410	0	0	1	3266	29292	0.364	0.428
Kittiwake	Isle of May	4	35	1093	2220	0	1	3106	29239	0.803	0.820
Kittiwake	Isle of May	5	35	1740	1624	0	1	2970	29254	0.901	0.908
Kittiwake	Isle of May	6	34	1536	794	0	1	2046	38911	0.881	0.903
Kittiwake	Isle of May	7	35	1775	1542	0	1	2877	29170	0.920	0.924
Kittiwake	Isle of May	8	35	1594	1819	0	1	3082	29128	0.697	0.689
Kittiwake	Isle of May	9	35	2867	436	0	1	3018	29123	0.345	0.406
Kittiwake	Isle of May	10	35	2359	1063	0	1	3204	29117	0.204	0.279
Kittiwake	Isle of May	11	35	275	3022	0	1	3257	29200	0.280	0.346
Kittiwake	Isle of May	12	35	236	3174	0	1	3363	29260	0.315	0.390
Kittiwake	Kongsfjorden	1	34	37	3125	0	1	3114	48800	0.176	0.280
Kittiwake	Kongsfjorden	2	34	1007	1865	0	1	2852	48780	0.101	0.197
Kittiwake	Kongsfjorden	3	33	3100	0	0	1	3053	48772	0.126	0.226
Kittiwake	Kongsfjorden	4	33	1166	938	0	1	2032	48780	0.209	0.307
Kittiwake	Kongsfjorden	5	7	17	29	1	3	680	48737	0.177	0.377
Kittiwake	Kongsfjorden	6	0	0	0	0	NA	NA	NA	NA	NA
Kittiwake	Kongsfjorden	7	0	0	0	0	NA	NA	NA	NA	NA
Kittiwake	Kongsfjorden	8	8	14	15	1	3	792	48475	0.357	0.603

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Kittiwake	Kongsfjorden	9	24	1611	164	0	1	1695	48506	0.449	0.597
Kittiwake	Kongsfjorden	10	34	1501	906	0	1	2312	48594	0.190	0.325
Kittiwake	Kongsfjorden	11	34	229	2831	0	1	3050	48664	0.268	0.369
Kittiwake	Kongsfjorden	12	34	23	3139	0	1	3153	48732	0.288	0.410
Kittiwake	Langanes and Skjalfandi	1	26	72	2284	0	1	2336	29344	0.220	0.352
Kittiwake	Langanes and Skjalfandi	2	26	841	1319	0	1	2083	29340	0.244	0.380
Kittiwake	Langanes and Skjalfandi	3	26	2356	0	0	1	2227	29256	0.579	0.637
Kittiwake	Langanes and Skjalfandi	4	26	547	1733	0	1	2210	29240	0.746	0.789
Kittiwake	Langanes and Skjalfandi	5	26	695	774	0	1	1397	29267	0.765	0.819
Kittiwake	Langanes and Skjalfandi	6	12	720	0	0	1	672	19456	0.945	0.954
Kittiwake	Langanes and Skjalfandi	7	21	772	86	0	1	787	29121	0.902	0.930
Kittiwake	Langanes and Skjalfandi	8	27	742	1542	0	1	2187	29128	0.657	0.711
Kittiwake	Langanes and Skjalfandi	9	27	1962	378	0	1	2241	29121	0.402	0.487
Kittiwake	Langanes and Skjalfandi	10	27	1641	780	0	1	2351	29170	0.322	0.434
Kittiwake	Langanes and Skjalfandi	11	27	102	2238	0	1	2326	29216	0.305	0.418
Kittiwake	Langanes and Skjalfandi	12	27	36	2355	0	1	2388	29269	0.326	0.451
Kittiwake	Røst	1	43	1160	5078	0	1	6035	39048	0.386	0.387
Kittiwake	Røst	2	43	2559	2890	0	1	5326	39028	0.758	0.724
Kittiwake	Røst	3	42	5642	0	0	1	5532	39005	0.859	0.850
Kittiwake	Røst	4	42	1642	3740	0	1	5294	38966	0.902	0.897
Kittiwake	Røst	5	42	2582	848	0	1	3390	38958	0.880	0.882
Kittiwake	Røst	6	30	2700	0	0	4	2675	29215	0.919	0.907
Kittiwake	Røst	7	33	2853	72	0	1	2894	38870	0.943	0.942
Kittiwake	Røst	8	42	2634	2294	0	1	4843	38777	0.647	0.664
Kittiwake	Røst	9	43	5132	865	0	1	5809	38830	0.414	0.450
Kittiwake	Røst	10	43	4169	2173	0	1	6078	38852	0.248	0.304
Kittiwake	Røst	11	43	272	5936	0	1	6122	38971	0.448	0.470
Kittiwake	Røst	12	43	438	5941	0	1	6260	39021	0.337	0.377
Kittiwake	Runde and Ålesund	1	25	189	1801	0	1	1948	19499	0.184	0.278
Kittiwake	Runde and Ålesund	2	25	835	924	0	1	1675	19519	0.366	0.413
Kittiwake	Runde and Ålesund	3	24	1860	0	0	3	1736	19529	0.582	0.637
Kittiwake	Runde and Ålesund	4	24	433	1379	0	1	1734	19529	0.791	0.815
Kittiwake	Runde and Ålesund	5	24	642	1122	0	1	1649	19481	0.868	0.879
Kittiwake	Runde and Ålesund	6	17	1100	47	0	1	994	29197	0.935	0.948

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
Kittiwake	Runde and Ålesund	7	25	1116	776	0	1	1710	29144	0.905	0.918
Kittiwake	Runde and Ålesund	8	25	815	1173	0	1	1858	29109	0.654	0.692
Kittiwake	Runde and Ålesund	9	25	1646	274	0	1	1801	19423	0.352	0.425
Kittiwake	Runde and Ålesund	10	25	1339	642	0	1	1911	19449	0.255	0.341
Kittiwake	Runde and Ålesund	11	25	126	1791	0	1	1899	19460	0.408	0.483
Kittiwake	Runde and Ålesund	12	25	140	1845	0	1	1956	19463	0.355	0.451
Kittiwake	Sklinna	1	36	627	3458	0	1	3989	29287	0.260	0.301
Kittiwake	Sklinna	2	36	1930	1758	0	1	3532	29271	0.619	0.599
Kittiwake	Sklinna	3	36	3968	0	0	1	3723	29220	0.909	0.899
Kittiwake	Sklinna	4	36	1823	2021	0	1	3648	29229	0.904	0.904
Kittiwake	Sklinna	5	36	1614	1603	0	1	3039	29316	0.897	0.903
Kittiwake	Sklinna	6	20	1679	1	0	1	1569	19453	0.960	0.961
Kittiwake	Sklinna	7	34	2067	677	0	1	2553	29184	0.931	0.932
Kittiwake	Sklinna	8	35	1822	2148	0	1	3827	29078	0.754	0.758
Kittiwake	Sklinna	9	36	3431	514	0	1	3796	29104	0.479	0.525
Kittiwake	Sklinna	10	36	2943	1147	0	1	3910	29127	0.361	0.394
Kittiwake	Sklinna	11	36	369	3576	0	1	3888	29175	0.354	0.410
Kittiwake	Sklinna	12	36	205	3887	0	1	4066	29265	0.347	0.408
C. guillemot	Bjørnøya	1	35	1734	2668	0	1	4285	29887	0.697	0.727
C. guillemot	Bjørnøya	2	35	1435	2558	0	1	3935	29893	0.745	0.772
C. guillemot	Bjørnøya	3	35	4278	0	0	1	4273	29874	0.796	0.811
C. guillemot	Bjørnøya	4	35	1059	1501	0	1	2552	29863	0.733	0.780
C. guillemot	Bjørnøya	5	8	558	0	0	1	558	19919	0.668	0.778
C. guillemot	Bjørnøya	6	8	540	0	0	1	540	19890	0.691	0.791
C. guillemot	Bjørnøya	7	8	558	0	0	1	558	19893	0.738	0.820
C. guillemot	Bjørnøya	8	35	397	1765	0	1	2149	29846	0.622	0.709
C. guillemot	Bjørnøya	9	35	3560	760	0	3	4278	29831	0.709	0.737
C. guillemot	Bjørnøya	10	35	3590	874	0	1	4412	29847	0.740	0.761
C. guillemot	Bjørnøya	11	35	1018	3302	0	1	4196	29871	0.743	0.766
C. guillemot	Bjørnøya	12	35	2998	1425	0	1	4260	29857	0.759	0.774
C. guillemot	Cape Gorodetskiy	1	4	86	224	0	1	299	29881	0.474	0.710
C. guillemot	Cape Gorodetskiy	2	4	132	150	0	1	266	29896	0.331	0.630
C. guillemot	Cape Gorodetskiy	3	4	310	0	0	1	306	29872	0.524	0.742
C. guillemot	Cape Gorodetskiy	4	4	180	107	0	1	266	29852	0.684	0.823
C. guillemot	Cape Gorodetskiy	5	2	63	1	0	4	62	19893	0.444	0.749
C. guillemot	Cape Gorodetskiy	6	1	60	0	0	4	56	9962	0.591	0.800
C. guillemot	Cape Gorodetskiy	7	1	62	0	0	4	61	9950	0.318	0.645
C. guillemot	Cape Gorodetskiy	8	4	49	145	0	1	194	29855	0.272	0.622
C. guillemot	Cape Gorodetskiy	9	4	266	34	0	1	283	29841	0.382	0.670
C. guillemot	Cape Gorodetskiy	10	4	245	65	0	3	308	29834	0.232	0.569
C. guillemot	Cape Gorodetskiy	11	4	89	211	0	1	290	29865	0.484	0.723

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
C. guillemot	Cape Gorodetskiy	12	4	161	149	0	1	295	29845	0.591	0.783
C. guillemot	Faroe Islands	1	7	23	471	0	3	493	19925	0.113	0.282
C. guillemot	Faroe Islands	2	7	165	294	0	3	458	19919	0.079	0.273
C. guillemot	Faroe Islands	3	7	496	0	0	3	496	19922	0.322	0.564
C. guillemot	Faroe Islands	4	7	140	340	0	3	471	19887	0.315	0.550
C. guillemot	Faroe Islands	5	7	139	298	0	3	436	19910	0.367	0.600
C. guillemot	Faroe Islands	6	3	72	32	0	4	103	9942	0.443	0.678
C. guillemot	Faroe Islands	7	7	88	166	0	4	254	19900	0.385	0.628
C. guillemot	Faroe Islands	8	7	39	459	0	3	497	19900	0.318	0.536
C. guillemot	Faroe Islands	9	7	395	85	0	3	476	19905	0.147	0.400
C. guillemot	Faroe Islands	10	7	390	106	0	3	491	19898	0.175	0.388
C. guillemot	Faroe Islands	11	7	12	468	0	3	479	19889	0.113	0.317
C. guillemot	Faroe Islands	12	7	16	481	0	3	497	19894	0.111	0.305
C. guillemot	Grimsey	1	9	93	588	0	1	673	19912	0.476	0.619
C. guillemot	Grimsey	2	9	279	355	0	1	614	19912	0.496	0.649
C. guillemot	Grimsey	3	9	682	0	0	1	660	19920	0.613	0.740
C. guillemot	Grimsey	4	9	299	361	0	1	605	19903	0.638	0.747
C. guillemot	Grimsey	5	9	186	166	0	1	326	19880	0.685	0.811
C. guillemot	Grimsey	6	2	120	0	0	1	118	9949	0.838	0.903
C. guillemot	Grimsey	7	8	171	29	0	1	182	19893	0.420	0.673
C. guillemot	Grimsey	8	9	198	482	0	1	623	19895	0.583	0.710
C. guillemot	Grimsey	9	9	569	91	0	1	619	19897	0.564	0.705
C. guillemot	Grimsey	10	9	548	135	0	1	656	19909	0.377	0.569
C. guillemot	Grimsey	11	9	58	603	0	1	647	19912	0.372	0.563
C. guillemot	Grimsey	12	9	61	620	0	1	671	19901	0.391	0.575
C. guillemot	Hjelmsøya	1	28	1070	1392	0	1	2332	49781	0.602	0.717
C. guillemot	Hjelmsøya	2	26	907	1220	0	1	1997	49835	0.718	0.799
C. guillemot	Hjelmsøya	3	26	2294	0	0	1	2178	49788	0.781	0.836
C. guillemot	Hjelmsøya	4	26	691	1094	0	1	1697	49765	0.717	0.805
C. guillemot	Hjelmsøya	5	8	679	3	0	1	658	29850	0.640	0.774
C. guillemot	Hjelmsøya	6	8	660	0	0	1	599	29858	0.766	0.850
C. guillemot	Hjelmsøya	7	8	682	0	0	1	597	29827	0.748	0.840
C. guillemot	Hjelmsøya	8	28	543	1099	0	1	1571	39772	0.569	0.709
C. guillemot	Hjelmsøya	9	28	2034	374	0	1	2332	39816	0.685	0.761
C. guillemot	Hjelmsøya	10	28	2002	478	0	1	2382	39778	0.650	0.739
C. guillemot	Hjelmsøya	11	28	662	1738	0	1	2217	39794	0.642	0.733
C. guillemot	Hjelmsøya	12	28	1998	482	0	1	2272	39807	0.629	0.726
C. guillemot	Hornøya	1	37	1281	3307	0	1	4371	29879	0.770	0.783
C. guillemot	Hornøya	2	37	1786	2399	0	1	4040	29891	0.845	0.851
C. guillemot	Hornøya	3	37	4526	0	0	1	4327	29884	0.924	0.916
C. guillemot	Hornøya	4	37	1694	2355	0	1	3833	29842	0.925	0.922

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
C. guillemot	Hornøya	5	26	2232	0	0	1	2103	19888	0.917	0.919
C. guillemot	Hornøya	6	26	2160	0	0	1	2020	19891	0.925	0.922
C. guillemot	Hornøya	7	26	2232	0	0	1	2078	19894	0.888	0.896
C. guillemot	Hornøya	8	37	1289	2276	0	1	3541	29838	0.753	0.781
C. guillemot	Hornøya	9	37	3705	795	0	1	4467	29821	0.722	0.747
C. guillemot	Hornøya	10	37	3703	947	0	1	4578	29827	0.702	0.731
C. guillemot	Hornøya	11	37	779	3664	0	1	4265	29859	0.735	0.750
C. guillemot	Hornøya	12	37	2031	2557	0	1	4372	29834	0.786	0.796
C. guillemot	Isle of May	1	30	1243	2332	0	1	3276	29870	0.630	0.689
C. guillemot	Isle of May	2	30	1894	1392	0	1	2972	29901	0.628	0.693
C. guillemot	Isle of May	3	30	3596	0	0	1	3207	29877	0.739	0.772
C. guillemot	Isle of May	4	30	1871	1615	0	1	3160	29858	0.802	0.817
C. guillemot	Isle of May	5	30	2291	1310	0	1	3033	29851	0.892	0.897
C. guillemot	Isle of May	6	32	2003	1215	0	1	2777	39790	0.848	0.867
C. guillemot	Isle of May	7	32	1314	2452	0	1	3528	29834	0.760	0.762
C. guillemot	Isle of May	8	32	944	2783	0	1	3553	29852	0.696	0.723
C. guillemot	Isle of May	9	31	3011	482	0	1	3314	29842	0.624	0.681
C. guillemot	Isle of May	10	30	3129	491	0	1	3280	29847	0.580	0.635
C. guillemot	Isle of May	11	30	1306	2169	0	1	3077	29824	0.612	0.670
C. guillemot	Isle of May	12	30	1239	2334	0	1	3275	29844	0.682	0.726
C. guillemot	Jan Mayen	1	27	332	2147	0	1	2415	39818	0.271	0.409
C. guillemot	Jan Mayen	2	27	847	1411	0	1	2195	39846	0.296	0.419
C. guillemot	Jan Mayen	3	27	2480	0	0	1	2461	39838	0.309	0.449
C. guillemot	Jan Mayen	4	27	674	1537	0	1	2196	39787	0.657	0.716
C. guillemot	Jan Mayen	5	11	691	16	0	1	703	19906	0.805	0.855
C. guillemot	Jan Mayen	6	8	660	0	0	1	650	19897	0.840	0.882
C. guillemot	Jan Mayen	7	8	680	2	0	1	677	19882	0.590	0.694
C. guillemot	Jan Mayen	8	27	418	1453	0	1	1853	39791	0.321	0.477
C. guillemot	Jan Mayen	9	27	2003	397	0	1	2365	39752	0.313	0.465
C. guillemot	Jan Mayen	10	27	1959	522	0	1	2415	39776	0.290	0.429
C. guillemot	Jan Mayen	11	27	225	2174	0	1	2350	39794	0.209	0.346
C. guillemot	Jan Mayen	12	27	646	1832	0	1	2414	39787	0.233	0.370
C. guillemot	Langanes and Skjalfandi	1	27	346	2008	0	1	2285	29851	0.607	0.663
C. guillemot	Langanes and Skjalfandi	2	27	968	1180	0	1	2076	29898	0.719	0.766
C. guillemot	Langanes and Skjalfandi	3	27	2356	0	0	1	2322	29872	0.758	0.797
C. guillemot	Langanes and Skjalfandi	4	27	942	1338	0	1	2185	29861	0.758	0.799
C. guillemot	Langanes and Skjalfandi	5	27	789	497	0	1	1177	29851	0.745	0.810

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
C. guillemot	Langanes and Skjalfandi	6	7	540	0	0	1	473	19891	0.884	0.917
C. guillemot	Langanes and Skjalfandi	7	8	557	12	0	1	518	29836	0.648	0.779
C. guillemot	Langanes and Skjalfandi	8	27	297	1852	0	1	2121	29828	0.648	0.715
C. guillemot	Langanes and Skjalfandi	9	27	1922	358	0	1	2127	29838	0.581	0.674
C. guillemot	Langanes and Skjalfandi	10	27	1887	469	0	1	2260	29815	0.416	0.555
C. guillemot	Langanes and Skjalfandi	11	27	238	2046	0	1	2219	29850	0.367	0.495
C. guillemot	Langanes and Skjalfandi	12	27	218	2136	0	1	2294	29851	0.473	0.562
C. guillemot	Sklinna	1	37	994	2536	0	1	3337	39824	0.559	0.631
C. guillemot	Sklinna	2	37	1394	1796	0	1	3047	39846	0.479	0.591
C. guillemot	Sklinna	3	37	3472	0	0	1	3052	39826	0.617	0.689
C. guillemot	Sklinna	4	37	1319	2002	0	1	3120	39826	0.807	0.836
C. guillemot	Sklinna	5	37	1146	987	0	1	2057	39806	0.778	0.827
C. guillemot	Sklinna	6	14	1080	0	0	1	1040	29856	0.699	0.789
C. guillemot	Sklinna	7	26	1179	104	0	1	1218	29852	0.655	0.749
C. guillemot	Sklinna	8	38	1278	1570	0	3	2694	39806	0.546	0.658
C. guillemot	Sklinna	9	38	3000	480	0	3	3407	39789	0.570	0.658
C. guillemot	Sklinna	10	38	2908	689	0	3	3511	39782	0.457	0.568
C. guillemot	Sklinna	11	38	620	2827	0	3	3312	39791	0.433	0.538
C. guillemot	Sklinna	12	37	1077	2457	0	3	3364	39798	0.476	0.564
B. guillemot	Alkefjellet	1	23	1705	218	0	1	1911	19942	0.533	0.624
B. guillemot	Alkefjellet	2	23	119	1088	0	1	1199	19931	0.594	0.693
B. guillemot	Alkefjellet	3	0	0	0	1	1	262	19932	0.393	0.647
B. guillemot	Alkefjellet	4	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Alkefjellet	5	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Alkefjellet	6	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Alkefjellet	7	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Alkefjellet	8	5	3	6	1	3	384	19791	0.379	0.611
B. guillemot	Alkefjellet	9	13	723	72	0	1	793	19843	0.587	0.712
B. guillemot	Alkefjellet	10	23	753	349	0	1	1075	19834	0.528	0.662
B. guillemot	Alkefjellet	11	23	1558	302	0	1	1830	19864	0.565	0.656
B. guillemot	Alkefjellet	12	23	1922	0	0	1	1904	19916	0.504	0.606
B. guillemot	Bjørnøya	1	29	487	2507	0	1	2950	29900	0.467	0.541
B. guillemot	Bjørnøya	2	29	913	1750	0	1	2547	29896	0.487	0.558
B. guillemot	Bjørnøya	3	28	2728	0	0	2	2700	29907	0.748	0.786
B. guillemot	Bjørnøya	4	28	581	900	0	1	1465	29876	0.755	0.814
B. guillemot	Bjørnøya	5	1	62	0	0	4	62	9947	0.673	0.830
B. guillemot	Bjørnøya	6	1	60	0	0	4	60	9942	0.664	0.832

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
B. guillemot	Bjørnøya	7	1	62	0	0	4	62	9939	0.310	0.598
B. guillemot	Bjørnøya	8	29	141	842	0	1	972	29726	0.550	0.706
B. guillemot	Bjørnøya	9	29	2578	493	0	1	3043	29759	0.684	0.738
B. guillemot	Bjørnøya	10	29	2522	628	0	1	3109	29719	0.581	0.635
B. guillemot	Bjørnøya	11	29	594	2325	0	1	2767	29833	0.417	0.527
B. guillemot	Bjørnøya	12	29	498	2474	0	2	2822	29858	0.610	0.678
B. guillemot	Cape Gorodetskiy	1	16	436	1275	0	1	1697	29890	0.568	0.676
B. guillemot	Cape Gorodetskiy	2	16	553	958	0	1	1483	29912	0.408	0.567
B. guillemot	Cape Gorodetskiy	3	16	1612	0	0	1	1588	29914	0.597	0.700
B. guillemot	Cape Gorodetskiy	4	16	881	544	0	1	1360	29878	0.841	0.878
B. guillemot	Cape Gorodetskiy	5	9	565	11	0	1	551	29878	0.677	0.786
B. guillemot	Cape Gorodetskiy	6	8	540	0	0	1	510	19862	0.736	0.816
B. guillemot	Cape Gorodetskiy	7	8	558	0	0	1	525	19838	0.735	0.826
B. guillemot	Cape Gorodetskiy	8	16	461	657	0	1	1098	29727	0.524	0.689
B. guillemot	Cape Gorodetskiy	9	16	1447	228	0	1	1661	29736	0.533	0.669
B. guillemot	Cape Gorodetskiy	10	16	1353	382	0	1	1728	29754	0.533	0.653
B. guillemot	Cape Gorodetskiy	11	16	167	1510	0	1	1672	29842	0.577	0.673
B. guillemot	Cape Gorodetskiy	12	16	697	1038	0	1	1709	29843	0.638	0.730
B. guillemot	Cape Sakhanin	1	41	627	2845	0	1	3199	19949	0.829	0.831
B. guillemot	Cape Sakhanin	2	41	1050	2103	0	1	3127	19932	0.747	0.765
B. guillemot	Cape Sakhanin	3	40	3286	0	0	1	3259	19931	0.698	0.724
B. guillemot	Cape Sakhanin	4	40	1161	1737	0	1	2544	19916	0.704	0.738
B. guillemot	Cape Sakhanin	5	14	755	9	0	4	718	19929	0.874	0.904
B. guillemot	Cape Sakhanin	6	12	720	0	0	4	678	9937	0.950	0.949
B. guillemot	Cape Sakhanin	7	12	744	0	0	4	700	9917	0.949	0.948
B. guillemot	Cape Sakhanin	8	41	949	1301	0	1	2158	19812	0.792	0.813
B. guillemot	Cape Sakhanin	9	41	2923	437	0	1	3271	19832	0.806	0.809
B. guillemot	Cape Sakhanin	10	41	2777	695	0	1	3380	19839	0.806	0.810
B. guillemot	Cape Sakhanin	11	41	592	2768	0	1	3220	19859	0.820	0.826
B. guillemot	Cape Sakhanin	12	41	1799	1673	0	1	3300	19903	0.869	0.866
B. guillemot	Franz J/Oranskie	1	10	445	177	0	1	620	19933	0.437	0.633
B. guillemot	Franz J/Oranskie	2	10	175	368	0	1	527	19944	0.437	0.646
B. guillemot	Franz J/Oranskie	3	9	558	0	0	3	518	19940	0.397	0.607
B. guillemot	Franz J/Oranskie	4	9	97	67	1	3	363	19914	0.245	0.535
B. guillemot	Franz J/Oranskie	5	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Franz J/Oranskie	6	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Franz J/Oranskie	7	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Franz J/Oranskie	8	2	4	14	1	3	168	9916	0.228	0.525
B. guillemot	Franz J/Oranskie	9	6	296	34	0	1	329	9920	0.422	0.624
B. guillemot	Franz J/Oranskie	10	11	313	128	0	1	410	29764	0.394	0.650
B. guillemot	Franz J/Oranskie	11	11	506	104	0	1	588	29816	0.467	0.663

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
B. guillemot	Franz J/Oranskie	12	10	620	0	0	1	614	19910	0.435	0.634
B. guillemot	Grimsey	1	12	310	868	0	1	1095	19924	0.519	0.627
B. guillemot	Grimsey	2	12	548	534	0	1	1015	19935	0.546	0.628
B. guillemot	Grimsey	3	12	1178	0	0	1	1133	19930	0.717	0.759
B. guillemot	Grimsey	4	12	489	651	0	1	1046	19936	0.797	0.834
B. guillemot	Grimsey	5	12	449	327	0	1	723	19926	0.862	0.891
B. guillemot	Grimsey	6	7	420	0	0	4	406	9936	0.934	0.936
B. guillemot	Grimsey	7	10	422	23	0	1	428	19842	0.749	0.789
B. guillemot	Grimsey	8	12	245	835	0	1	1056	19836	0.429	0.554
B. guillemot	Grimsey	9	12	958	182	0	3	1061	19841	0.295	0.456
B. guillemot	Grimsey	10	12	968	210	0	1	1093	19834	0.399	0.551
B. guillemot	Grimsey	11	12	101	1039	0	2	1103	19884	0.448	0.589
B. guillemot	Grimsey	12	12	103	1075	0	1	1153	19917	0.408	0.530
B. guillemot	Hornøya	1	46	1843	3491	0	1	5238	39834	0.583	0.611
B. guillemot	Hornøya	2	46	1899	2974	0	1	4725	39871	0.626	0.654
B. guillemot	Hornøya	3	46	5332	0	0	1	5128	39854	0.818	0.806
B. guillemot	Hornøya	4	46	2486	2329	0	1	4500	39854	0.909	0.904
B. guillemot	Hornøya	5	23	2476	4	0	1	2301	29885	0.896	0.909
B. guillemot	Hornøya	6	23	2400	0	0	4	2249	29832	0.857	0.876
B. guillemot	Hornøya	7	23	2480	0	0	1	2341	29708	0.902	0.911
B. guillemot	Hornøya	8	45	1703	2199	0	1	3870	49498	0.652	0.726
B. guillemot	Hornøya	9	45	4329	853	0	1	5158	39622	0.693	0.732
B. guillemot	Hornøya	10	46	4191	1152	0	1	5306	39637	0.655	0.680
B. guillemot	Hornøya	11	46	982	4235	0	1	5196	39771	0.613	0.636
B. guillemot	Hornøya	12	46	3459	1875	0	1	5243	39828	0.564	0.592
B. guillemot	Isfjorden	1	17	187	991	0	1	1069	39850	0.367	0.576
B. guillemot	Isfjorden	2	17	382	589	0	1	791	39888	0.380	0.604
B. guillemot	Isfjorden	3	12	806	0	0	1	746	29899	0.482	0.616
B. guillemot	Isfjorden	4	12	149	33	1	3	478	29883	0.179	0.455
B. guillemot	Isfjorden	5	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Isfjorden	6	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Isfjorden	7	0	0	0	0	NA	NA	NA	NA	NA
B. guillemot	Isfjorden	8	14	25	193	1	3	799	39660	0.237	0.505
B. guillemot	Isfjorden	9	18	981	202	0	3	1165	39634	0.212	0.443
B. guillemot	Isfjorden	10	18	1033	212	0	1	1120	39657	0.204	0.379
B. guillemot	Isfjorden	11	18	346	841	0	1	1130	39756	0.251	0.448
B. guillemot	Isfjorden	12	17	124	1054	0	2	1100	39806	0.374	0.589
B. guillemot	Jan Mayen	1	42	1070	4200	0	1	4815	29893	0.626	0.646
B. guillemot	Jan Mayen	2	42	2010	2712	0	2	4133	29903	0.542	0.593
B. guillemot	Jan Mayen	3	42	5084	0	0	1	4941	29898	0.688	0.697
B. guillemot	Jan Mayen	4	42	1421	3158	0	1	4535	29877	0.869	0.866

Species	Colony name	Month	Nind	IRMApos	GLSpos	Add	Model	nPres	nBackval	RsqAdj	DevExpl
B. guillemot	Jan Mayen	5	28	2422	7	0	4	2410	29885	0.783	0.808
B. guillemot	Jan Mayen	6	27	2340	0	0	1	2318	19861	0.885	0.888
B. guillemot	Jan Mayen	7	27	2418	0	0	1	2391	19825	0.875	0.883
B. guillemot	Jan Mayen	8	42	1414	2503	0	2	3888	29733	0.750	0.774
B. guillemot	Jan Mayen	9	42	4353	877	0	2	4967	29762	0.728	0.738
B. guillemot	Jan Mayen	10	42	4318	1084	0	1	4848	29735	0.496	0.536
B. guillemot	Jan Mayen	11	42	601	4619	0	2	5048	29832	0.594	0.616
B. guillemot	Jan Mayen	12	42	602	4720	0	2	5047	29859	0.627	0.647
B. guillemot	Langanes and Skjalfandi	1	17	338	1175	0	1	1411	29920	0.523	0.642
B. guillemot	Langanes and Skjalfandi	2	17	633	721	0	1	1226	29897	0.476	0.626
B. guillemot	Langanes and Skjalfandi	3	17	1488	0	0	1	1452	29907	0.638	0.719
B. guillemot	Langanes and Skjalfandi	4	17	545	895	0	1	1396	29891	0.770	0.824
B. guillemot	Langanes and Skjalfandi	5	17	523	313	0	1	773	29876	0.765	0.836
B. guillemot	Langanes and Skjalfandi	6	6	420	0	0	1	373	19866	0.916	0.937
B. guillemot	Langanes and Skjalfandi	7	7	436	16	0	1	431	29761	0.638	0.766
B. guillemot	Langanes and Skjalfandi	8	17	401	747	0	2	1125	29759	0.528	0.667
B. guillemot	Langanes and Skjalfandi	9	17	1286	215	0	1	1426	29734	0.501	0.593
B. guillemot	Langanes and Skjalfandi	10	17	1243	307	0	1	1479	29741	0.458	0.563
B. guillemot	Langanes and Skjalfandi	11	17	144	1356	0	1	1440	29838	0.418	0.567
B. guillemot	Langanes and Skjalfandi	12	17	156	1394	0	1	1474	29871	0.505	0.630

6.3 R-script - Procedures for reading SEATRACK output files

SEATRACK - Models of pelagic seabird abundance

Reading model output data from netCDF files using R Statistical Software

For more information about the project, please visit SEATRACK's website

This page provides example R code to:

- connect to a netCDF file
- examine structure and read metadata information of the netCDF file
- extract and map monthly abundance data for a specific colony
- extract, aggregate, and map monthly abundance data for all colonies from a specific ocean area and country

1) Preparing workspace and loading data

```
# Loading necessary libraries
  library(ncdf4); library(raster); library(maps); library(knitr)

# Selecting one of the six species
# c("Alle_alle", "Fratercula_arctica", "Fulmarus_glacialis",
# "Rissa_tridactyla", "Uria_aalge", "Uria_lomvia")
  sp <- "Uria_aalge"

# Opening connection to netCDF file
  nc <- nc_open(paste("outputs/SEATRACK_Abundance_Model_", sp, ".nc", sep = ""))
```

2) Extracting metadata from the NetCDF file

```
# Printing general information about the netCDF file, variables and dimensions
  print(nc)
```

```
## File outputs/SEATRACK_Abundance_Model_Uria_aalge.nc (NC_FORMAT_NETCDF4):
##
##      11 variables (excluding dimension variables):
##         float PredictedAbundanceMean[lon,lat,month,colonyCode] (Chunking:[264,84,1,21]) (Compression: level 9)
##             units: birds/pixel
##             _FillValue: NaN
##             long_name: Mean of the predicted abundance
##         char colonyName[nchar,colonyCode]   (Contiguous storage)
##             long_name: Name of the colony
##         float colonyLatitude[colonyCode]   (Contiguous storage)
##             units: dd.mmmmm
##             long_name: Latitude of the colony
##         float colonyLongitude[colonyCode]   (Contiguous storage)
##             units: ddd.mmmmm
##             long_name: Longitude of the colony
##         char colonyOceanArea[nchar,colonyCode]   (Contiguous storage)
##             long_name: Ocean area of the colony
##         char colonyRegion[nchar,colonyCode]   (Contiguous storage)
##             long_name: Region of the colony
##         char colonyCountry[nchar,colonyCode]   (Contiguous storage)
##             long_name: Country of the colony
##         int colonyNpairs[colonyCode]   (Contiguous storage)
##             units: number of breeding pairs
##             long_name: Number of breeding pairs
##         char SmcolCode[nchar,colonyCode]   (Contiguous storage)
##             long_name: Code of the corresponding SEATRACK model colony
##         char SmcolName[nchar,colonyCode]   (Contiguous storage)
##             long_name: Name of the corresponding SEATRACK model colony
##         float SmcolDistance[colonyCode]   (Contiguous storage)
##             units: km
##             long_name: Distance to the corresponding SEATRACK model colony
```

```
##
##      5 dimensions:
##         lon  Size:1580
##             units: degrees east
##             long_name: Longitude
##         lat  Size:500
##             units: degrees north
##             long_name: Latitude
##         month  Size:12
##             long_name: Month of the year (integer)
##         colonyCode  Size:126
##             long_name: Colony identification code (integer)
##         nchar  Size:50
##
##     7 global attributes:
##         Dataset source: SEATRACK - distribution and abundance model outputs - Common guillemot (Uria aalge)
##         Dataset version: v1.0
##         Dataset last update: 2019-04-07
##         Coordinate system (proj-string): '+proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0'
##         Created with: R version 3.4.4 (2018-03-15), package ncdf4 version 1.16 (2017-04-01)
##         Authors: Per Fauchald & Arnaud Tarroux, Norwegian Institute for Nature Research
##         Please cite:
##
##  Fauchald et al. (2019) Arctic-breeding seabird's hotspots in space and time - a methodological
##  framework for year-round modelling of abundance and environmental niche using light-logger data.
##  NINA report 1657, ISBN 978-82-426-3401-6. Norwegian Institute for Nature Research.
##
##
##  For information about the SEATRACK project, please visit: http://www.seapop.no/en/seatrack/
##  __________________________________________________________________________________________________________________
_______
```

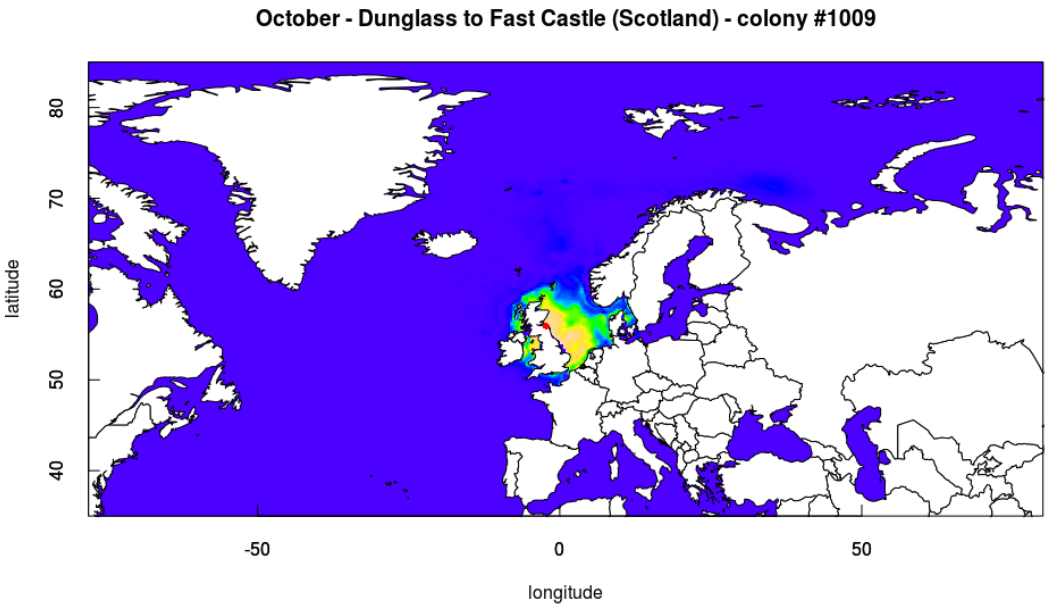
3) Creating a table summarizing the metadata

```r
# Combining the metadata into a table
  colonies <- cbind.data.frame(code        = ncvar_get(nc, "colonyCode"),
                               name        = ncvar_get(nc, "colonyName"),
                               lat         = ncvar_get(nc, "colonyLatitude"),
                               lon         = ncvar_get(nc, "colonyLongitude"),
                               oceanArea   = ncvar_get(nc, "colonyOceanArea"),
                               region      = ncvar_get(nc, "colonyRegion"),
                               country     = ncvar_get(nc, "colonyCountry"),
                               nPairs      = ncvar_get(nc, "colonyNpairs"),
                               modelColony = ncvar_get(nc, "SmcolName"))

# Examining the first rows of the table
  kable(head(colonies, 10))
```

code	name	lat	lon	oceanArea	region	country	nPairs	modelColony
43	Runde	62.40	5.63	Norwegian_Sea	More_og_Romsdal	Norway	9000	Sklinna
119	Kolvaeret	64.33	10.32	Norwegian_Sea	Sor_Trondelag	Norway	11	Sklinna
136	Maaoya	65.22	10.96	Norwegian_Sea	Nord_Trondelag	Norway	750	Sklinna
157	Buvaer-Gjelfruvaer	67.42	11.90	Norwegian_Sea	Nordland	Norway	30	Sklinna
158	Hernyken	67.43	11.88	Norwegian_Sea	Nordland	Norway	50	Sklinna
160	Trenyken og holmer omkring	67.44	11.89	Norwegian_Sea	Nordland	Norway	300	Sklinna
162	Ellefsnyken	67.45	11.91	Norwegian_Sea	Nordland	Norway	200	Sklinna
164	Storfjellet	67.46	11.94	Norwegian_Sea	Nordland	Norway	500	Sklinna
167	Vedoy	67.48	12.02	Norwegian_Sea	Nordland	Norway	670	Sklinna
170	Knappen-Buneset-Maastadfjord	67.64	12.59	Norwegian_Sea	Nordland	Norway	125	Sklinna

4) Producing maps for specific colony and month

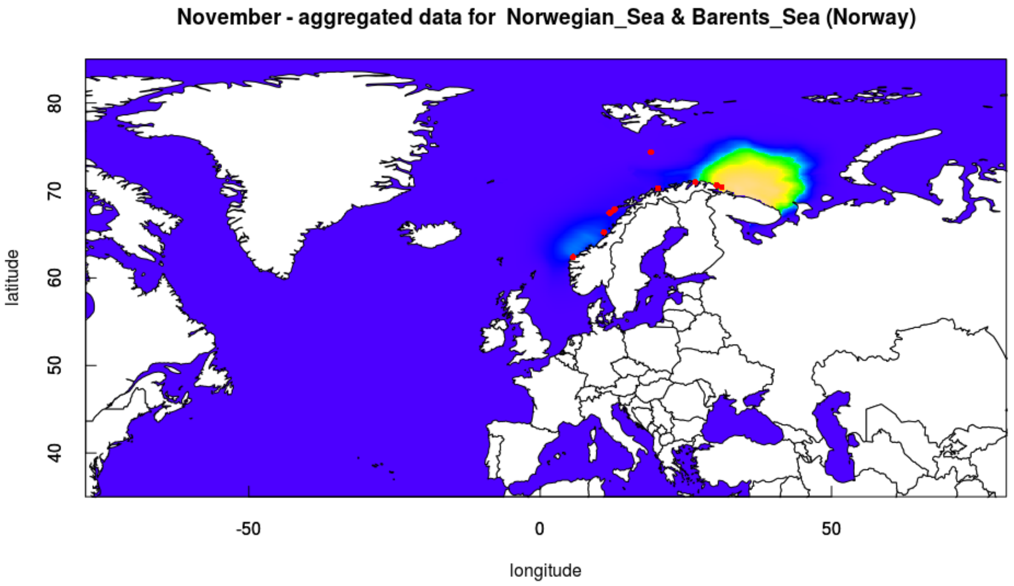
```
# Selecting a specific colony (based on its name)
  selname <- "Dunglass to Fast Castle"
  codecol <- ncvar_get(nc, "colonyCode")[ncvar_get(nc, "colonyName") == selname]
  icol    <- which(ncvar_get(nc, "colonyCode") == codecol)

# Selecting month (e.g. October)
  iMth <- 10

# Reading corresponding slice from the netCDF dataset
  rast <- raster(nc$filename, varname="PredictedAbundanceMean", level = iMth, band = icol)

# Producing a quick map
  image(rast, tcl = .4, xlab = "longitude", ylab = "latitude", col = topo.colors(10000),
     main=paste(month.name[as.numeric(ncvar_get(nc,"month")[iMth])], " - ", ncvar_get(nc,"c
olonyName")[icol],
     " (", ncvar_get(nc, "colonyRegion")[icol], ") - colony #", ncvar_get(nc, "colonyCode")
[icol], sep = ""))
  map("world", add = T, col = "white", fill = TRUE)
# Adding colony location
  points(colonies$lon[colonies$code == codecol], colonies$lat[colonies$code == codecol],
         pch = 16, col = "red", cex = .8)
  axis(1, tcl = .5); axis(2, tcl = .5); box()
```



5) Aggregating data

```
# Selecting grouping parameter value (e.g. all Norwegian colonies from Barents Sea and
Norwegian Sea)
  selarea <- c("Norwegian_Sea", "Barents_Sea")
  selctry <- "Norway"

  icols    <- which(ncvar_get(nc, "colonyOceanArea") == selarea & ncvar_get(nc,
"colonyCountry") == selctry)
  codecols <- unique(ncvar_get(nc, "colonyCode")[icols])

# selecting month (e.g. november)
```

```
  iMnth <- 11

# Aggregating (here cumulative sum) the data for all selected colonies and month
  sum.fun <- function(ii, jj) raster(nc$filename, varname="PredictedAbundanceMean", band = ii,
level = jj)
  rast.gr <- sum(stack(mapply(FUN = sum.fun, ii = icols, jj = iMth)))

# Producing a quick map
  image(rast.gr, tcl = .4, xlab = "longitude", ylab = "latitude", col = topo.colors(10000),
     main=paste(month.name[as.numeric(ncvar_get(nc,"month")[iMnth])], "- aggregated data for ",
       paste(paste(unique(selarea), collapse = " & "), " (",
       paste(unique(selctry), collapse = " & "), ")", sep = "")))
  map("world", add = T, col = "white", fill = TRUE)
# Adding colony locations
  points(colonies$lon[colonies$code %in% codecols], colonies$lat[colonies$code %in% codecols],
         pch = 16, col = "red", cex = .8)
  axis(1, tcl = .5); axis(2, tcl = .5); box()
```



```
# closing connection to netCDF file
  nc_close(nc)
```

The Norwegian Institute for Nature Research, NINA,
is as an independent foundation focusing on environmental research, emphasizing the interaction between human society, natural resources and biodiversity.

NINA was established in 1988. The headquarters are located in Trondheim, with branches in Tromsø, Lillehammer, Bergen and Oslo. In addition, NINA owns and runs the aquatic research station for wild fish at Ims in Rogaland and the arctic fox breeding center at Oppdal.

NINA's activities include research, environmental impact assessments, environmental monitoring, counselling and evaluation. NINA's scientists come from a wide range of disciplinary backgrounds that include biologists, geographers, geneticists, social scientists, sociologists and more. We have a broad-based expertise on the genetic, population, species, ecosystem and landscape level, in terrestrial, freshwater and coastal marine ecosystems.

ISSN: 1504-3312
ISBN: 978-82-426-3401-6

Norwegian Institute for Nature Research
NINA head office
Postal address: P.O. Box 5685 Torgarden,
NO-7485 Trondheim, NORWAY
Visiting address: Høgskoleringen 9, 7034 Trondheim
Phone: +47 73 80 14 00
E-mail: firmapost@nina.no
Organization Number: 9500 37 687
http://www.nina.no

NINA

Cooperation and expertise for a sustainable future