

1893

NINA Report

## An automated procedure (v2.0) to obtain positions from light-level geolocators in large-scale tracking of seabirds

A method description for the SEATRACK project

Vegard Sandøy Bråthen, Børge Moe, Françoise Amélineau, Morten Ekker, Per Fauchald, Hálf dán Helgi Helgason, Malin Kjellstadli Johansen, Benjamin Merkel, Arnaud Tarroux, Jens Åström, Hallvard Strøm



# **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

## **NINA Special Report (NINA Temahefte)**

Special reports are produced as required and the series ranges widely: from systematic identification keys to information on important problem areas in society. Usually given a popular scientific form with weight on illustrations.

## **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 research results in international scientific journals and in popular academic books and journals.

# An automated procedure (v2.0) to obtain positions from light-level geolocators in large-scale tracking of seabirds

A method description for the SEATRACK project

Vegard Sandøy Bråthen, Børge Moe, Françoise Amélineau, Morten Ekker, Per Fauchald, Hálfván Helgi Helgason, Malin Kjellstadli Johansen, Benjamin Merkel, Arnaud Tarroux, Jens Åström, Hallvard Strøm

Bråthen, V.S., Moe, B., Amélineau, F., Ekker, M., Fauchald, P., Helgason, H.H., Johansen, M.K., Merkel, B., Tarroux, A., Åström, J. & Strøm, H. 2021. An automated procedure (v2.0) to obtain positions from light-level geolocators in large-scale tracking of seabirds.

A method description for the SEATRACK project. NINA Report 1893. Norwegian Institute for Nature Research.

Trondheim, March 2021

ISSN: 1504-3312

ISBN: 978-82-426-4666-8

COPYRIGHT

© Norwegian Institute for Nature Research

The publication may be freely cited where the source is acknowledged

AVAILABILITY

Open

PUBLICATION TYPE

Digital document (pdf)

QUALITY CONTROLLED BY

Signe Christensen-Dalsgaard

SIGNATURE OF RESPONSIBLE PERSON

Research director Svein-Håkon Lorentsen (sign.)

COVER PICTURE

Atlantic puffin at Anda © Vegard Sandøy Bråthen

KEY WORDS

- Barents Sea, Norwegian Sea, North Sea, Svalbard, Northeast Atlantic Ocean, Arctic

- Seabirds, Alcids, Atlantic puffin, Black-legged kittiwake, Brünnich's guillemot, Common eider, Common guillemot, European shag, Glaucous gull, Herring gull, Northern fulmar, Lesser-black-backed gull, Little Auk

- Light-level geolocators, Global location sensing, GLS, Migration, Non-breeding season

CONTRIBUTING INSTITUTIONS



CONTACT DETAILS

**NINA head office**

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

**NINA Oslo**

Sognsveien 68  
0855 Oslo  
Norway  
P: +47 73 80 14 00

**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:**

Thormøhlens gate 55  
NO-5006 Bergen.  
Norway  
P: +47 73 80 14 00

[www.nina.no](http://www.nina.no)

## Abstract

Bråthen, V.S., Moe, B., Amélineau, F., Ekker, M., Fauchald, P., Helgason, H.H., Johansen, M. K., Merkel, B., Tarroux, A., Åström, J. & Strøm, H. 2021. An automated procedure (v2.0) to obtain positions from light-level geolocators in large-scale tracking of seabirds. A method description for the SEATRACK project. NINA Report 1893. Norwegian Institute for Nature Research.

Partners in the SEATRACK project have since 2014 deployed >14 000 light-level geolocators on 11 species of seabirds to study their non-breeding distribution in the North Atlantic. Geocator tags are ideal for large-scale tracking of seabirds being cheap and small. The tags contains a clock and a light sensor that register light levels at fixed intervals. These data are stored in the internal memory and are obtained when the tag is recovered from the bird. From these data two positions per day are calculated by estimating latitude from the length of day and night, and longitude from time of mid-day and mid-night. However, positions cannot be obtained from recorded light-data during the polar night or midnight sun. Further, the latitudinal accuracy is unreliable close to spring and autumn equinox when the length of day and night is similar across the earth.

Using a threshold method, we first identify twilight events, which is the time when light-levels cross a predefined threshold that separate day from night. However, the light-level recordings are affected by environmental factors and the behaviour of the bird that may shade the geocator or expose it to artificial light. As such, the accuracy is low compared to GPS or Argos tracking devices. A common approach has therefore been to improve the identification of these twilight events by manually inspecting the light-level data. This process is, however, time-consuming and prone to not being fully consistent and reproducible among different persons applying it.

In this report, we describe an automated procedure (v2.0) for obtaining the basic positional dataset in SEATRACK from raw light-level data. The procedure automatically filters and edits the twilight events used for calculating positions. It further removes unrealistic positions using filters on equinox periods, speed, distribution, angle, distance, variation in timing of twilights and mid-night sun periods, and produces double smoothed positions. Calibration of sun elevation angles, crucial for producing the final positions, is performed on each track and is the only part involving subjective assessment, but we show that it can be performed consistently and with a high repeatability.

SEATRACK processes light data from >1000 geolocators after each field season, and the database has become one of the largest seabird tracking databases in the world. The automated procedure (v2.0) is a very cost-efficient method for such large-scale tracking and is consistent and reproducible. We have recently updated the entire database using this procedure, replacing all previous data based on the manual procedure and the first version of the automated procedure (v1.0).

This report describes the methods used to obtain positions from geolocators in the SEATRACK project. As the described procedure replace our previous manual method, we show comparisons of the two procedures. The report also provides examples of how to read and visualize the positional data and can serve as the reference for the methods and as a tool for using the data.

Vegard Sandøy Bråthen, Norwegian Institute for Nature Research, P.O. Box 5685 Torgarden, NO-7485 Trondheim, [vegard.brathen@nina.no](mailto:vegard.brathen@nina.no)

Børge Moe, Norwegian Institute for Nature Research, P.O. Box 5685 Torgarden, NO-7485 Trondheim, [borge.moe@nina.no](mailto:borge.moe@nina.no)

Françoise Amélineau, Norwegian Polar Institute, FRAM - High North Research Centre for Climate and the Environment, NO-9296 Tromsø, [francoise.amelineau@npolar.no](mailto:francoise.amelineau@npolar.no)

Morten Ekker, Norwegian Environment Agency, Postboks 5672 Torgarden, 7485 Trondheim, [morten.ekker@miljodir.no](mailto:morten.ekker@miljodir.no)

Per Fauchald, Norwegian Institute for Nature Research, FRAM - High North Research Centre for Climate and the Environment, NO-9296 Tromsø, [per.fauchald@nina.no](mailto:per.fauchald@nina.no)

Hálfdrán Helgi Helgason, East Iceland Nature Research Centre, Mýrargötu 10  
740 Neskaupstað, [halfdan@na.is](mailto:halfdan@na.is)

Malin Kjellstadli Johansen, Norwegian Polar Institute, FRAM - High North Research Centre for Climate and the Environment, NO-9296 Tromsø, [malin.johansen@npolar.no](mailto:malin.johansen@npolar.no)

Benjamin Merkel, Akvaplan-niva, FRAM - High North Research Centre for Climate and the Environment, NO-9296 Tromsø, [bme@akvaplan.niva.no](mailto:bme@akvaplan.niva.no)

Arnaud Tarroux, Norwegian Institute for Nature Research, FRAM - High North Research Centre for Climate and the Environment, NO-9296 Tromsø, [arnaud.tarroux@nina.no](mailto:arnaud.tarroux@nina.no)

Jens Åström, Norwegian Institute for Nature Research, P.O. Box 5685 Torgarden, NO-7485 Trondheim, [jens.astrom@nina.no](mailto:jens.astrom@nina.no)

Hallvard Strøm, Norwegian Polar Institute, FRAM - High North Research Centre for Climate and the Environment, NO-9296 Tromsø, [hallvard.strom@npolar.no](mailto:hallvard.strom@npolar.no)

## Sammendrag

Partnere i SEATRACK prosjektet har siden 2014 instrumentert 11 sjøfuglarter med over 14 000 lysloggere for å studere deres utbredelse utenfor hekketiden i Nord-Atlanteren. Lysloggere er godt egnet til slik stor-skala sporing av sjøfugl fordi de er små og er relativt rimelige i innkjøp. Loggerne er utstyrt med en lyssensor og en klokke og lagrer lysnivå i bestemte tidsintervaller som lastes ned fra loggeren når den hentes inn igjen fra fuglen. Disse dataene muliggjør beregninger av to posisjoner per dag ved å estimere breddegrader fra lengden på dag og natt, og lengdegrader fra dagens og nattens midtpunkt. Unntaket er perioder hvor lysloggeren måler gjennomgående høye lysnivåer på grunn av midnattssol, eller gjennomgående lave lysnivåer på grunn av polarnatt. Breddegradene vil også være unøyaktige jo nærmere man kommer høst- og vårjevndøgn, hvor dag og natt er omtrent like lang verden over.

Tidspunkt for hver soloppgang og solnedgang baseres på terskelverdier som skiller det lysnivået vi regner som dag fra det lysnivået vi regner som natt. En utfordring med denne metoden er at lysnivåene blir påvirket av ulike miljøfaktorer, og at adferden til sjøfuglene kan dekke til sensoren eller eksponere den for kunstig lys. Nøyaktigheten til de endelige posisjonene kan derfor reduseres noe i forhold til andre sporingsmetoder, for eksempel GPS-loggere eller satellittbaserte sendere (Argos). For å delvis kompensere for disse faktorene har en vanlig fremgangsmåte vært å inspisere og eventuelt justere eller fjerne feilaktig identifiserte soloppganger og solnedganger manuelt. Denne prosessen er tidkrevende, og er sårbar for ikke å være konsistent og reproduserbar mellom ulike personer som anvender den.

I denne rapporten beskriver vi en automatisert prosedyre (v2.0) for prosessering av posisjonsdata fra lysloggere utviklet av SEATRACK. Prosedyren filtrerer og justerer soloppganger og solnedganger automatisk før disse brukes for å beregne posisjoner. Videre fjerner den urealistiske posisjoner ved bruk av filtre på vår- og høstjevndøgn-perioder, hastighet, utbredelse, vinkel og avstand, variasjon i soloppganger og solnedganger, perioder med midnattssol, og til slutt glattes (snittes) posisjonene i to ulike steg. Kalibrering av solens vinkel til horisonten ved soloppgang og solnedgang er et avgjørende steg for å kalkulere de endelige posisjonene, og en unik solvinkel blir gitt for hvert enkelt spor mellom to hekkesesonger. Dette er det eneste steget i prosedyren som involverer subjektive vurderinger, men våre tester viser at dette steget i høy grad kan utføres på en konsistent og repeterbar måte.

SEATRACK prosesserer årlig lysdata fra over 1000 lysloggere og databasen har med årene vokst til å bli en av verdens største for systematisk sporinger av sjøfugler. Den automatiserte prosedyren (v2.0) er en svært kostnadseffektiv metode for en slik stor-skala sporing, og er konsistent og reproduserbar. Denne prosedyren er nå anvendt på hele SEATRACK-datasettet, og erstatter det som tidligere var basert på en manuelle prosedyre, samt den første versjon av den automatisert prosedyre (v1.0).

Hovedformålet med denne rapporten er å beskrive metodene som er brukt for å oppnå posisjoner fra lysloggere i SEATRACK. Siden prosedyren erstatter tidligere prosedyrer, viser vi resultater av sammenligninger og diskuterer forskjeller mellom den tidligere manuelle og den nåværende automatiserte prosedyren. Rapporten gir også eksempler for hvordan datasettet kan leses og posisjonene visualiseres, og kan videre brukes som en referanse til metodene og i forståelsen av dataene.

# Contents

<b>Abstract</b> .....	<b>3</b>
<b>Sammendrag</b> .....	<b>5</b>
<b>Contents</b> .....	<b>6</b>
<b>1 Introduction</b> .....	<b>7</b>
<b>2 Estimating coordinates from light-level data</b> .....	<b>11</b>
2.1 Setup.....	11
2.2 Download data.....	14
2.3 Split data into years tracked.....	14
2.4 Identify twilight events.....	15
2.4.1 Define model specific thresholds.....	15
2.4.2 Twilight estimation.....	16
2.4.3 Advance sunsets.....	16
2.4.4 Filter twilights.....	17
2.4.5 Edit twilights.....	19
2.5 Estimate geographic coordinates.....	20
2.6 Filter unrealistic positions.....	20
2.6.1 Determining equinox periods for latitudinal filtering.....	20
2.6.2 Speed filter.....	20
2.6.3 Distribution filter.....	22
2.6.4 Angle and distance filter.....	23
2.6.5 Loess filter.....	23
2.6.6 Midnight sun filter.....	25
2.7 Double smoothing of positions.....	25
2.8 Sun angle calibration.....	26
2.8.1 Latitude versus time plots.....	26
2.8.2 Maps for calibration.....	26
2.8.3 Consistency of the calibration assessments.....	29
<b>3 Filter effects on the number of twilights and positions</b> .....	<b>30</b>
<b>4 Comparing SEATRACK procedures; manual versus automatic</b> .....	<b>33</b>
<b>5 Applications and limitations</b> .....	<b>37</b>
<b>6 References</b> .....	<b>38</b>
<b>7 Appendix</b> .....	<b>41</b>



## Preface

This report describes an automated procedure v2.0 for how SEATRACK's positional dataset is obtained from light-level data. The automated procedure is cost-efficient and meet the needs of SEATRACK for processing light data from large numbers of geolocators. It is consistent and reproducible and allows previous and future tracking data to be processed in the same manner.

In 2020, this automated procedure (v2.0) replaced our first version of an automated procedure (v1.0), introduced in 2018, and our manual procedure introduced in 2014. Hence, any older SEATRACK data have also been reprocessed with our latest version from 2020 and onwards. The two previous procedures are described in Strøm et al. (2021). Many of the functions from the 2018 version of the automatic procedure (v1.0) have been kept, but their settings and sequence have changed. See Strøm et al. (2021) for logger model specifications and overviews of sample sizes used in different species and colonies.

To test and quality check the procedure we established a reference group with two external members, Morten Frederiksen (Aarhus University) and Maria Bogdanova (UK Centre for Ecology & Hydrology), in addition to three members of the SEATRACK project-group (Françoise Amélineau, Børge Moe and Vegard Sandøy Bråthen). The reference group tested the script for the procedure and provided feedback. We are indebted to the reference group, and especially Morten Frederiksen and Maria Bogdanova for their input.

On behalf of the authors

Vegard Sandøy Bråthen  
March 2021

# 1 Introduction

Marine biodiversity is under pressure, and seabirds are among the most threatened of all bird groups (Dias et al. 2019). Data on spatial and temporal distribution are therefore important for the conservation and management of seabird populations (Lascelles et al. 2016). The SEATRACK project was started in 2014 to improve knowledge about the non-breeding distribution of seabirds in the Northeast Atlantic (Strøm et al. 2021). In the first phase of SEATRACK (2014-2018), 11 seabird species breeding in colonies around the Barents Sea, Norwegian Sea and North Sea, in Norway (including Svalbard and Jan Mayen), Russia, Iceland, the Faroe Islands and the UK were tracked. In the second phase (2019-2022), SEATRACK was extended to include Greenland, Canada and western parts of the British Isles (**Figure 1.1**).

SEATRACK has been using light-level geolocators (hereafter 'geolocators') to track the seabirds. These small and relative low-cost tags are equipped with a light sensor and a clock, and are useful for large-scale and multi-colony studies (e.g. Frederiksen et al. 2012, 2016, van Bemmelten et al. 2017, Merkel et al. 2020). Geolocators store light-data in regular intervals in their internal memory together with a time stamp, and the data can be downloaded when the loggers are retrieved from the bird.



**Figure 1.1.** Map of locations where birds are instrumented with geolocators in the SEATRACK project.

Using the threshold method, positions are obtained from light-level data by first identifying the time of twilights (sunset and sunrise) using a pre-defined light threshold to separate night from day, using a threshold method. From those twilights, latitudes can be calculated from the daylength, and a longitude from time of midnight or noon, resulting in a maximum of two calculated positions per date (Wilson et al. 1992). The calculation of latitudes also requires that the light-level threshold should be matched with calibrations of the sun's angle to the horizon ('sun angle calibration'). Since the light recordings are affected by environmental factors and the behaviour of the bird which may shade the geolocator or expose it to artificial light (Fox 2010, Lisovski et al. 2012), the accuracy of positions from light-level data is considered low compared to positions recorded with use of GPS-loggers or Argos transmitters (Edwards et al. 2016, Lisovski et al. 2012, Philips et al. 2014, Fudickar et al. 2012), and the use of the data to study movements < 200 km is not recommended (Lisovski et al. 2019). The error is mainly due to error in latitude, which can be particularly inflated closer to spring and autumn equinox and to the Equator (Lisovski et al. 2012). Furthermore, positions are unattainable during periods of constant daylight or darkness at high latitudes.

In SEATRACK, we have tracked six pelagic seabird species (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*) and five coastal species (common eider *Somateria mollissima*, European shag *Phalacrocorax aristotelis*, herring gull *Larus argentatus*, lesser black-backed gull *Larus fuscus* and glaucous gull *Larus hyperboreus*). When seabirds are offshore, the open sea provides no obstructions and weather/cloudiness is the main source of shading. Leg position (e.g. tucking the leg into the feathers) may however shade the geolocator, and light from fishing vessels or offshore installations may expose the geolocator to artificial light. When the birds are inshore or on land in the colony, the geolocators may often be shaded. Cliffs, burrows, vegetation and the bird itself shade the logger during the breeding season, especially when the bird is sitting, incubating eggs or brooding chicks. The birds can also be exposed to artificial light from vessels, harbours, installations and other anthropogenic light sources along the coast.

Due to the many environmental and behavioural factors affecting the light recordings, the identification of twilight is usually aided by manual inspection of the light-level data with different software for the threshold method (Fox 2010, Frederiksen et al. 2012, Fox 2018, Lisovski et al. 2019). In the first phase of SEATRACK, we applied such a manual procedure and carefully inspected light data when identifying the twilight periods before calculating positions. This process is, however, time-consuming and prone to not being fully consistent and reproducible among different persons. We therefore replaced this manual procedure with an automated procedure in 2018. Both the manual and the first version of our automated procedure are described in Strøm et al. (2021). The manual procedure was used for all data obtained from 2013 to 2017, and for European shag, herring gull and common eider up to 2018. The first version of the automated procedure was used on data from the remaining species in 2018. In 2020 we developed the second version of the automated procedure (v2.0), which has now been applied to all the species and all the years of the SEATRACK project.

The procedure automatically filters and edits the light data to identify the twilight events used to calculate positions. It further removes unrealistic positions using filters on equinox periods, speed, distribution, angle and distance, variation in timing of twilights and midnight sun periods, and produces double smoothed positions. The calibration of the sun elevation angle is performed on each track before the final positions are produced. This calibration is the only part involving subjective assessment, but we show that it can be performed consistently and with a high repeatability.

The purpose of the automated procedure (v2.0) is to be a cost- and time-effective and consistent procedure for SEATRACK. It enables us to process data from many geolocators per hour, in a

consistent way, and avoid bias related to use of different logger models from different production years. It represents a substantial advancement from the manual procedure, and a refinement of the first automated procedure (v1.0), without being particularly computation intensive. The resulting positional dataset is stored in the SEATRACK database and forms the basis of SEATRACK products, such as kernel maps in the SEATRACK web application (<http://seatrack.seapop.no/map/>), a dataset with reduced biases and filled data gaps ('IRMA', Fauchald et al. 2019) and abundance maps (Fauchald et al. 2019).

The main aim of this report is to describe and document the automated procedure (v2.0) used to identify twilights and obtain positions from the geolocators in SEATRACK. The report also provides examples of how to read and visualize the positional data in R (R Core Team 2019). This report can therefore serve as a reference for the methods and as a tool for using the data.

## 2 Estimating coordinates from light-level data

The two main methods to identify diurnal patterns in light-level geolocation from which positions can be calculated are the template fit method (also known as curve fitting) and the threshold method (Lisovski et al 2019). The template fit method produces a coordinate from the time of a twilight and the rate of change in the recorded light. The threshold method produces a coordinate from the times of two consecutive twilight events, i.e. when the recorded light level crosses a predefined threshold during dusk and dawn. The coordinate is then estimated from the length of the day or the night (latitude) and the time of noon or midnight (longitude, Lisovski et al. 2019). Template fit is recommended for relatively clean light-level data with little inference from shading, that record the full range of light and at short time intervals (Lisovski et al. 2019). In SEATRACK, there is substantial variation in the degree of shading in the light data, the geolocators rarely record the full range of light (using a clipped light range) and the recording intervals vary. Since there is a need for a standard approach in such a large project, we use the threshold method that is simpler and can deal with a variety of light ranges, logging intervals and degrees of shading in the geocator data.

The calculation of latitudes from twilights depends on a proper match between the sun's angle to the horizon and the time when the light crosses the chosen threshold that separates day from night. Sun angles are obtained by calibration for which several approaches and methods exist (see Lisovski et al. 2019), and usually involves some subjective assessment and manual input (e.g. defining appropriate calibration periods). In chapter 2.8, we describe how we assign the sun angle for a track by manually studying latitude versus time plots generated for a range of sun angles. The calibration increases the time to analyse data from each geocator, but we show that trained personnel can perform the calibration consistently and with a high repeatability. Our calibration method is adapted from the approach by Hanssen et al. (2016) and van Bemmelten et al. (2019). It applies a set of criteria that allow us to calibrate tracks from all the different species and populations in SEATRACK, including colonies at high latitudes where midnight sun (constant daylight) prevent calibration at deployment and recapture. Furthermore, by assigning individual sun angles to each track, we avoid systematic bias in latitude related to potential differences in light sensors among geocator models or year of production. This also allows us to analyse and include geocator data collected over multiple years, also before SEATRACK.

In the following chapters we describe the entire setup of our procedure, which automatically identifies twilight events from raw light data and applies a set of filters on twilight events (remove or edit) and positions (speed, distribution limits, angle filter)

### 2.1 Setup

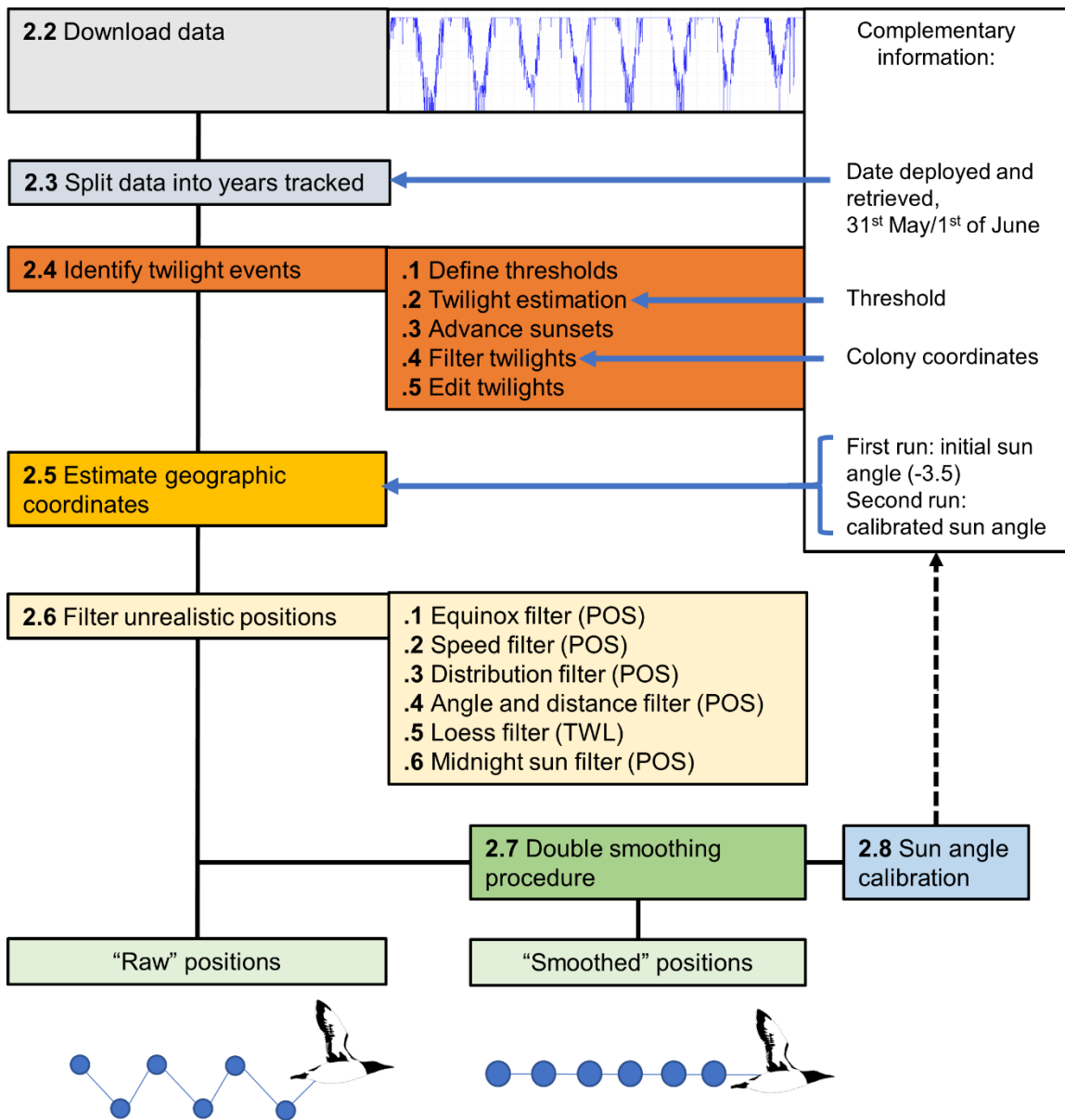
The SEATRACK procedure for estimating positions from light-level recordings consists of an R-script that loops through one year tracked at a time (**Figure 2.1**). One year tracked refers to the track of an individual from one breeding season to the next, representing a full non-breeding season. The procedure automatically identifies twilight events from raw light data and applies a set of filters on twilight events (remove or edit) and positions (speed, distribution limits, angle filter). The procedure relies on the SEATRACK database which stores all the logger data. When the R-script has estimated the positions, they are imported to the database along with relevant parameters.

Logger- and species-specific settings are specified in the script. A file with complementary information provides information about the individual (bird id), logger (logger id and model), deployment dates, colony (name, location) and the final calibrated sun elevation angle. The complementary information consists of one row for each logger and year tracked (appendix, **Table 7.3**).

The R script consists of a two-round loop that calculates positions from raw light-level data using the following steps:

- Read one row from the supporting information at a time and load the corresponding raw light data for one logger year tracked (**Figure 2.1**, chapter 2.2).
- Limit light-level data to one year tracked (**Figure 2.1**, chapter 2.3).
- Determine the times of twilights based of predefined thresholds of light (**Figure 2.1**, chapter 2.4).
- Calculate positions from twilights with an initial sun elevation angle (**Figure 2.1**, chapter 2.5).
- Filter twilights and positions (**Figure 2.1**, chapter 2.6).
- Perform a double smoothing of positions (**Figure 2.1**, chapter 2.7).
- Calibrate sun elevation angles (**Figure 2.1**, chapter 2.8).

The purpose of the first round of the loop is to calibrate the sun elevation angle. In the second round of the loop the geographic positions are estimated with the calibrated sun elevation angle. Hence, the filters act on positions estimated with the calibrate sun elevation angle. The positional data (both smoothed and unsmoothed positions) are stored in the SEATRACK database along with various parameters, metadata and filter codes (e.g. for identifying equinox periods) for later use in SEATRACK and by partners.



TWL = filter positions indirectly from filtering of twilight data, POS = filter positions directly

**Figure 2.1.** Schematic overview of the geolocator data analysis from the downloading the data to the final positions in the SEATRACK dataset. Numbers refer to the chapters where each step is described in this report.

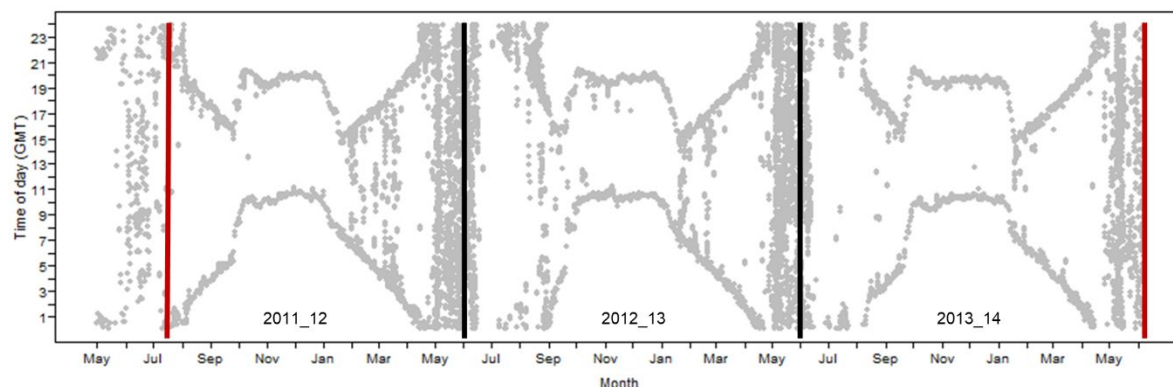
## 2.2 Download data

In SEATRACK, we mainly have used geolocators produced by Lotek, which until 2019 was known as Biotrack, and Migrate Technology. Data from previous projects was usually recorded with geolocators produced by British Antarctic Survey (BAS) with technology that were taken over by Biotrack in 2012.

The raw light data are first downloaded from the geolocators (**Figure 2.1**, step 2.2) using the interface boxes and software provided by Biotrack (BASTrak, BAS, Cambridge, UK, Biotrack 2013) and Migrate Technology (IntegeolF, Fox 2015), respectively. All raw data files are stored in the SEATRACK database, and the files are given unique names after a standardized pattern based on logger id, year retrieved and logger model.

## 2.3 Split data into years tracked

This step splits light-level data into separate years tracked (**Figure 2.1**, step 2.3) to cover a full non-breeding season for the species tracked in SEATRACK. Since many of our colonies experience constant daylight (midnight sun) around the summer solstice, when we cannot estimate twilights and consequently obtain positions, the summer solstice could be a natural end-point of a year tracked. However, we use the 1 June to separate tracks because of the early start of the breeding season in the southern colonies (April). If there is only one year of tracking data, we use all the light-level data from date deployed to date retrieved. If the bird carried the logger over multiple years, the split between the first and second year tracked is 31 May/1 June (**Figure 2.2**). Consequently, over two consecutive years, the data would be split in two years tracked; from deployment (calendar year 1) to 31 May (calendar year 2) and 1 June (calendar year 2) to retrieval (calendar year 3).



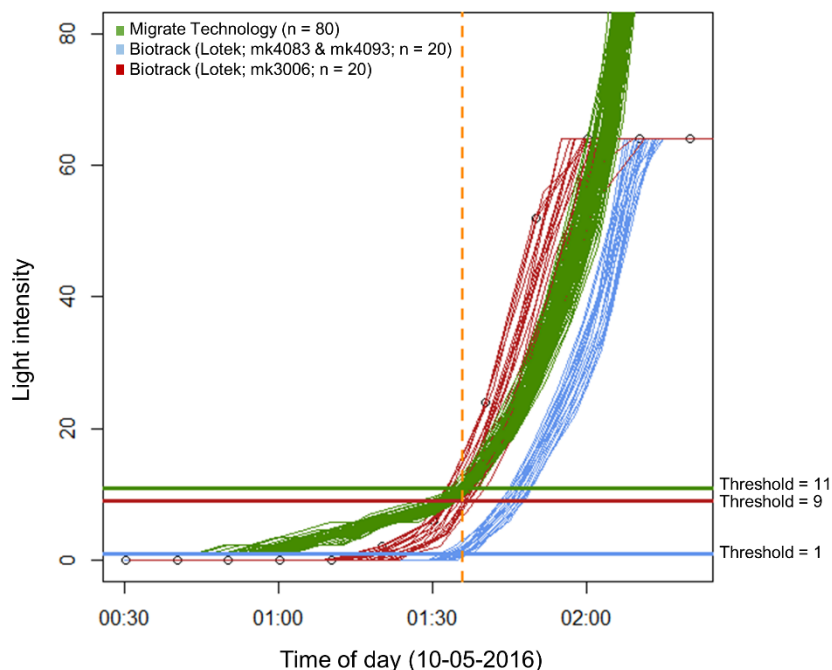
**Figure 2.2.** The figure shows data downloaded from a geolocator deployed on a black-legged kittiwake over several years. Grey points show each time the recorded light-levels crossed our threshold for separating day from night. Vertical lines show an example of how the raw light-level data is split into years tracked. As loggers may have been started months before they are deployed on a bird or may be stored for a long time after retrieval before the data is downloaded, the first step was to limit the data from date deployed to date retrieved (red vertical lines). The second step was to split the data in years tracked, the first year tracked (2011\_12) from date deployed to 31 May 2012, the second year tracked (2012-13) from 1 June 2012 to 31 May 2013, the third year tracked (2013-14) from 1 June 2013 to date retrieved.



## 2.4 Identify twilight events

### 2.4.1 Define model specific thresholds

Geolocators produced by BAS and Biotrack (Biotrack was granted the rights to produce BAS loggers in 2011 and was renamed to Lotek in 2019) and Migrate Technology record light in arbitrary units and lux, respectively, and different models vary in light detection range. To select model-specific thresholds which correspond to approximately the same light intensity (**Table 2.1**), we perform roof-top calibrations lasting seven days when a new logger model is included into SEATRACK (**Figure 2.3**). The threshold is selected using three criteria. First, a threshold should be above the general variation in light during night-time for that model (Lisovski et al. 2019). Second, as the main focus in SEATRACK is to track seabirds throughout the non-breeding season, we select low thresholds of light detection in order to acquire more positions through the darkest periods of the year. Third, timing of twilights should be comparable among models (**Figure 2.3**). Exceptions are made when including data from models that have not been available for calibration. For the LAT2000 model produced by Lotek, we selected a high threshold that mostly accommodated the first criterion. For old BAS models, we used the thresholds provided in **Table 2.1** based on the history of model development and production. We acknowledge that the sensitivity of the light sensors may have changed over the years of production from BAS to Biotrack. However, our procedure for calibrating the sun elevation angle copes with this by assigning sun elevation angles to each individual logger and year tracked. If the sensitivity of the light sensor differs (and all other factors are equal), different sun elevation angles will be assigned. We consequently avoid bias in estimated latitudes arising from altered sensitivity of light sensors in old logger models.



**Figure 2.3.** Light intensity recorded by different models of geolocators in a roof-top calibration. The Y-axis show the light intensity in lux, for models produced by Migrate Technology (max light recorded: inf, or if cut: 1200 lux), or in arbitrary units for models produced by Biotrack (max recorded light is cut at 64 (arbitrary)). Comparisons were made over a period of seven days. While the models provided by Migrate Technology all record a similar lux value at the same time of day (green, f100/c65/c250/c330), the larger geolocators from Biotrack (red, mk3006) recorded similar light values at a different time of day compared to their smaller models (blue, mk4083/mk4093). The orange vertical dotted line illustrates how the respective thresholds (horizontal lines) compare in time.

With the light thresholds listed in **Table 2.1**, the geolocators detect twilight when the sun is approximately  $-3^\circ$  to  $-4^\circ$  below the horizon. Consequently, the southern limit of apparent constant daylight and darkness is roughly  $63^\circ\text{N}$  at the summer solstice and  $70^\circ\text{N}$  at the winter solstice, respectively. The LAT2000 model is a notable exception with a mean sun angle of  $-7.5^\circ$ , which will pull the limit for apparent constant daylight southwards, and opposite for the apparent constant darkness. Note, however, that this model has been rarely used in SEATRACK.

**Table 2.1.** Producers, logger models and their model-specific thresholds used in SEATRACK for identifying timing of twilights.

Producer	Model	Threshold
<b>Migrate Technology</b>	w65, c65, c65_super, f100, c250, c330	11
<b>Biotrack (Lotek*)</b>	mk4093, mk4083	1
	mk3006	9
<b>BAS</b>	mk18, mk19	1
	mk15, mk3, mk4, mk5, mk7, mk13, mk14	9
<b>Lotek</b>	LAT2000	150

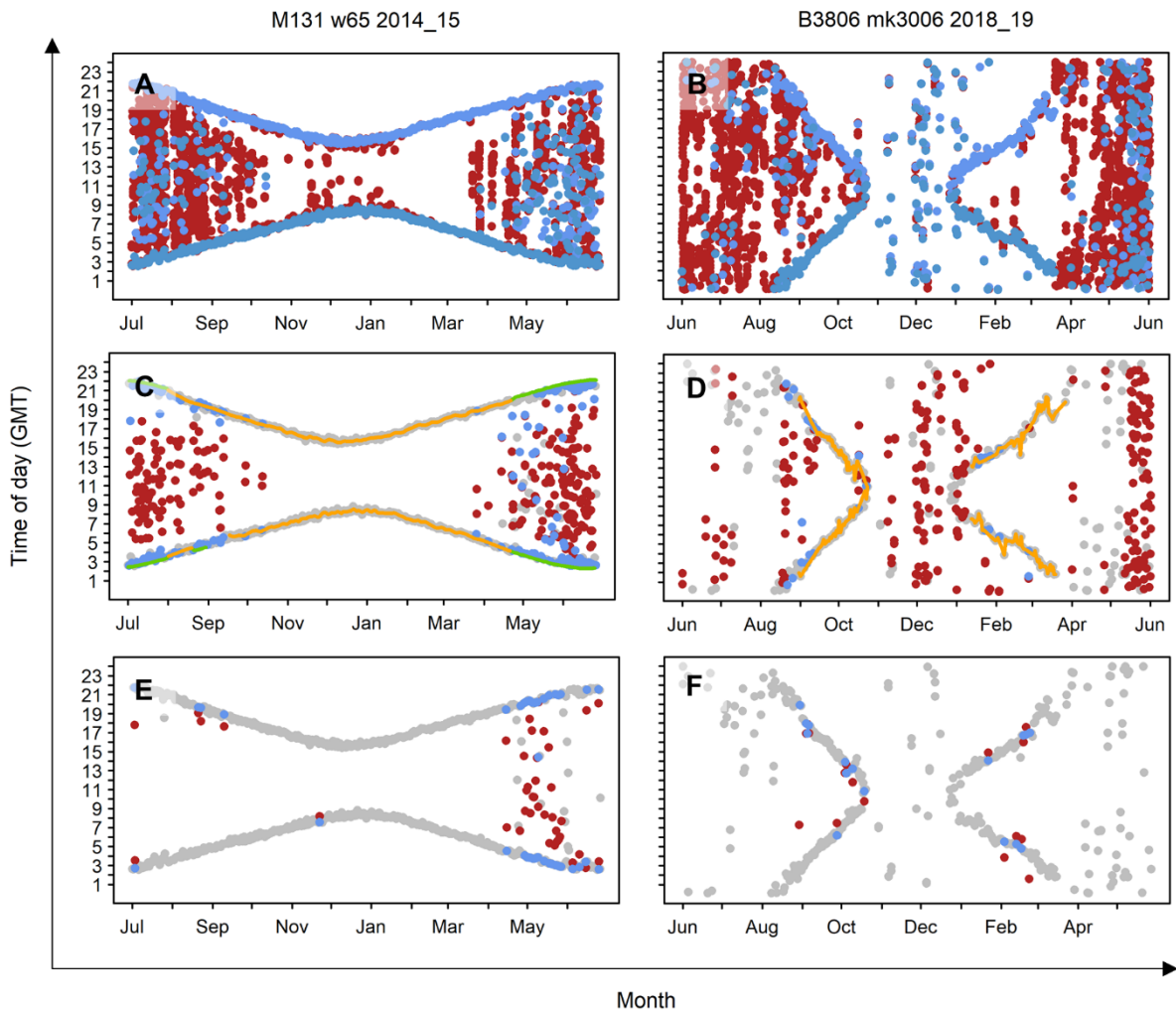
\**Biotrack was renamed Lotek in 2019.*

## 2.4.2 Twilight estimation

To identify twilights using raw light-level data we use the `twilightCalc` function from the `GeolLight` package (Lisovski & Hahn 2012; R Core Team 2019). This function assigns a twilight when the light intensity crosses the light threshold specified for each model. Under optimal conditions this would include one just before an actual sunrise and one just after an actual sunset, for each day. However, shading of the light sensor, from behavioural or environmental factors, or conversely artificial light from anthropogenic sources can cause the light-level to cross the threshold several times a day. To limit the influence of such shading and lighting, `twilightCalc` does not allow nights shorter than one hour. Consequently, it ignores short events of shading or exposure to artificial light and is often capable of identifying the most likely time for a twilight where there are several candidates within a short period, as illustrated in panel A and B in **Figure 2.4**. However, `twilightCalc` does not prevent the inclusion of unrealistic twilights when the light-level cross the threshold at intervals of several hours. Therefore, many unrealistic twilights can be generated at this stage of the process, leading to wrong definitions of days and nights in our dataset (example A and B, **Figure 2.4**). The next filter described in chapter 2.4.4 removes such twilights (example C and D, **Figure 2.4**).

## 2.4.3 Advance sunsets

Geolocators measure light-levels approximately every minute, but store only the maximum sampled light value at the end of a recording interval of five or ten minutes. Consequently, when the light-level is decreasing close to a sunset, the maximum recorded light-level occurs at the very beginning of a recording interval, but is stored at the end. Hence, the sunset needs to be advanced and `twilightCalc` advances sunsets by subtracting the recording interval. We reduce the length of the subtracted interval by an additional minute, because the first light measurement takes place after one minute within the recording interval, following the procedures previously described by Fox (2015) and Biotrack (2013).



**Figure 2.4.** Two examples of twilight estimation for an Atlantic puffin residing at approximately 50-60°N where there are no events of midnight sun or polar night (panels A, C and E) and a glaucous gull residing far north (70-80°N) and experiencing midnight sun and polar night for large parts of the year (panels B, D, F). Panels A and B show the result of the `twilightCalc` function, where blue points are the estimated timing of twilights while red points are the time of day where the recorded light values crossed our defined threshold but were not assigned as twilights by the function. Panels C and D show the results after subsequent filtering of twilights (where too many twilights occurred per day), where sunrise/sunset events corresponding to the red dots were removed from the dataset, while blue points are twilights that were retained. The orange curves are the predicted twilights and the green lines are the colony twilights used to aid the selection of candidate twilights. Panels E and F show the results of the “edit twilights” filter where red times are the original times which later were changed to the average timing of their adjacent sunsets or sunrises, the time they were changed to are shown in blue.

#### 2.4.4 Filter twilights

To remove erroneous and unrealistic twilights generated by the `twilightCalc` function we developed a filter that identifies days with too many twilights and retains the two most likely twilight events sunset and sunrise, respectively, while discarding the rest.

The first step marks days that have an unrealistic number of twilights present. Normally, there are approximately 24 hours between, for example, one sunset and the next, but to account for small deviations (e.g. due to weather or seasonal shifts in daylength) and larger deviations (e.g.

migration bouts/movements causing shorter days or nights, like flying east- or westward), the filter allows days with two sunset or two sunrise events as long as each occurs at a minimum of interval of 22 hours.

In the second step, the filter predicts sunset or sunrise events based on twilights from adjacent days that remain after we have excluded twilights (1) from days that were marked in the first step as having an unrealistic number of twilight events in step one, (2) that were marked as unrealistic after running a `LoessFilter` function (`GeoLight` package, k-value = 3, Lisovski & Hahn 2012) and (3) that were marked as unrealistic as they surpass a standard deviation of 60 min. For the latter, standard deviations were calculated from a moving time window of five consecutive sunsets or sunrises, using the `roll_sd` function from the `roll` package (Foster 2020). We then use the twilights that remain when running a local polynomial-regression fitting (`loess` function from the `stats` package, R Core Team 2019) with span set to 0.05 to predict sunset or sunrise events, and the `fillMissing` function from the `baytrend` package (Murphy et al. 2020) to fill gaps in the predicted times series.

In the third and last step, the filter compares each candidate twilight event within a marked day to a predicted time of sunrise or sunset for that day. It further selects the most likely sunset and sunrise events, and discards the rest (**Figure 2.4**, C & D). However, if predicted sunset- or sunrise events were not produced for that day in the second step, all candidate twilights are deleted from that date.

To avoid selecting twilights generated from shading of the logger under periods of apparent mid-night sun, the filter will only use the `fillMissing` function to fill short gaps in predicated twilights spanning up to five days, between 1 May to 31 July. In cases where a diurnal pattern should be present, but predications are not achievable due to large variations in the twilight times (e.g. due to nest attendance), the filter may aid the selection by using twilights at the breeding colony from the beginning of April to the end of August. Colony twilights are obtained with the `getSunlightTimes` function in the `suncalc` package (Thieurmel & Elmarhraoui 2019), with twilights corresponding to when the sun is  $-6^\circ$  below the horizon (civil dawn and dusk). Using the colony twilight as a substitute should have little influence if the bird at some point leaves the colony, as the number of erroneous and unrealistic twilights usually drops considerably whenever the bird is away from the colony in summer. Consequently, the potential for making predictions improves, and the colony twilight becomes irrelevant. An example where the colony twilight enhances our ability to select candidate twilights is shown for an Atlantic puffin breeding on Isle of May ( $56.2^\circ\text{N}$ ) in example C in **Figure 2.4**. The use of colony twilights has less to no effect in northern colonies with many days of midnight sun. Instead, the filter will correctly remove all twilights from marked dates, as for the glaucous gull breeding in Kongsfjorden ( $78.9^\circ\text{N}$ ) in example D, **Figure 2.4**.

The filter fills gaps in predicted values up to 20 days from the 1 August until the 30 April, which allows more twilights to be retained and more subsequent calculation of positions but avoids selecting twilight events generated from sources of artificial light in longer periods of apparent polar night. Since individuals can exhibit a lot of variation in twilight events in early autumn, probably because of the autumn moult where some species may rest with their leg and therefore also logger tucked into their plumage while resting (e.g. Elliott & Gaston 2014), and the `fillMissing` function is unable to expand the beginning or the end of a dataset, we allow candidate twilights to be compared to the predicated time from the nearest following *date* within a time span of 10 days in a period from beginning of August to the end of October. This can be seen when selecting twilight events in early September (**Figure 2.4 C**) and in late August (**Figure 2.4 D**), where candidate twilights have been retained although there are no predicated twilights available for that day (no predicted twilights available (orange line) or colony twilight (green line)).

## 2.4.5 Edit twilights

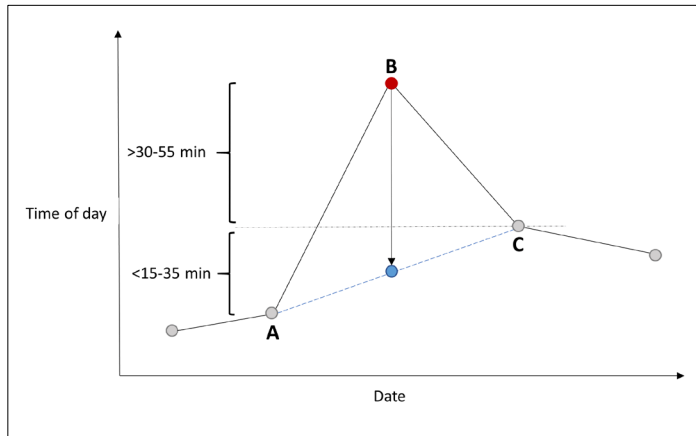
As a final step in assigning twilights, we have made a function that aims to edit some of the twilights that were generated from light-level data influenced by events of shading or artificial light, which in turn can retain more positions in the final dataset. The function identifies and changes the timing of twilights that fulfil two conditions: first, a twilight must be clearly different from two adjacent twilights of the same type (e.g. sunsets), and second, the two adjacent twilight events must occur at a similar time of day. If these conditions are fulfilled, the timing of the identified twilight is changed to the average of its adjacent twilights (**Figure 2.4**, E & F; **Figure 2.5**). The function will have no effect when the variation between twilights is high for several consecutive days, e.g., due to migration or persistent influence of shadows and/or artificial light. We owe credit for this approach to the `twilightEdit` function from the `TwGeos` package (Lisovski et al. 2019), which is based on some of the same principles. However, since the latter failed to run with our data on several occasions, we made our own and slightly different function from scratch.

**Table 2.2.** Overview of daily changes in minutes for sunsets or sunrises (numeric) by month and latitude. These values are used to determine if timing of a twilight differs from twilights on adjacent dates.

Latitude	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
> 70°N	10	10	7.5	10	10	10	10	10	7.5	10	10	10
> 60°N	6	6	5	6	6	6	6	6	5	6	6	6
> 47.5°N	2.5	3	3.5	3.5	3	2.5	2.5	3	3.5	3.5	3	2.5
> 32.5°N	1.5	2	2.5	2.5	2	1.5	1.5	2	2.5	2.5	2	1.5
> 12.5°N	0.5	0.5	1.5	1.5	1	0.5	0.5	1	1.5	1.5	1	0.5
< 12.5°N	0	0.5	1	1	0.5	0	0	0.5	1	1	0.5	0

Adjacent twilights (of the same type) are considered similar if they occur less than 15 to 35 minutes apart in time of day, based on the location of the bird, the time of year and the daily variance in the geolocator data. Here, we consider the daily change in timing of sunsets or sunrises at the monthly median latitude where the bird was located (**Table 2.2**). The threshold is set to twice the daily change to cover three consecutive dates, plus 15 minutes to consider variance that comes with shifting environmental conditions, such as cloud cover. From a black-legged kittiwake, breeding on Sklinna, Norway, this daily variance was estimated to be approximately 15 minutes.

A twilight is considered different from its adjacent twilights when the difference overreaches the daily variance of 15 minutes, plus the daily change of light (range 0 – 10 minutes, **Table 2.2**), and a chosen species-specific threshold. The species-specific threshold should reflect a maximum plausible travel distance within 24 hours. We chose a 15 minute threshold for all species (allowing approximately 750 km straight travel within 24 hours at a 55° latitude), except the northern fulmar, which we set to 30 minutes (allowing approximately 2000 km straight travel within 24 hours at a 55° latitude). Hence, a twilight is considered unrealistic if it is more than 45 to 55 minutes different from its adjacent twilights for northern fulmars, or 30 to 40 minutes for other species in SEATRACK (**Figure 2.5**).



**Figure 2.5.** Illustration for how the ‘Edit twilights’ filter works. Depending on species, month and monthly median latitude, time of twilight B can be replaced by the average time (blue point) of adjacent twilights that are of the same type (A and C). See panel E and F in **Figure 2.4** for examples on how it affects actual twilight data.

## 2.5 Estimate geographic coordinates

Geographic coordinates were calculated with the `coord` function in the `GeoLight` package (Lisovski & Hahn 2012). In a first run of our loop, this function calculates preliminary positions with an initial sun elevation angle of  $-3.5^\circ$ . Preliminary positions are then filtered, smoothed and finally used in an individual sun angle calibration for each year tracked (chapter 2.8, **Figure 2.1**). In the second run of the loop, all steps of the loop are repeated, except the sun angle calibration, and new coordinates from the present function are calculated with individually calibrated sun angles.

## 2.6 Filter unrealistic positions

### 2.6.1 Determining equinox periods for latitudinal filtering

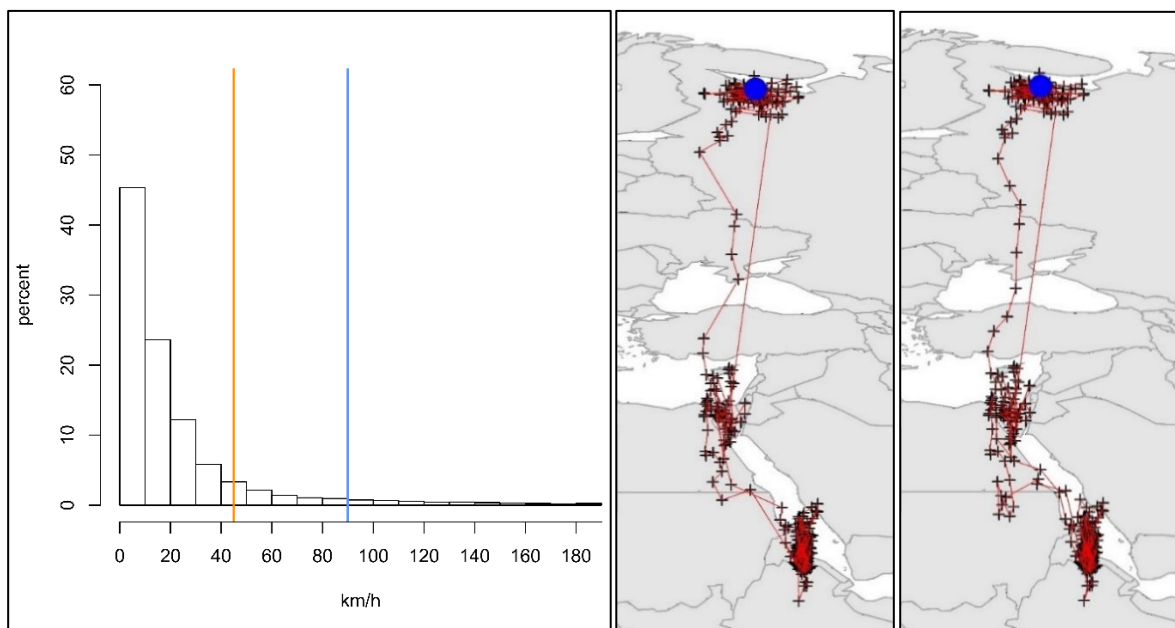
Close to an equinox, the daylength is similar across the world and even a small imprecision in the estimated daylength can have larger consequences for the estimated latitude. Latitudes are therefore unreliable in the weeks around the spring (19-21 March) and autumn (21-24 September) equinoxes. Since we mostly detect twilights 3-4 degrees below the horizon when using our pre-defined thresholds, the apparent time of the equinox is shifted towards winter solstice in the northern hemisphere (Hill & Braun 2001). In agreement with the periods defined in Frederiksen et al. (2012) for black-legged kittiwakes (**Figure 2.1**, step 2.6.1), we defined the equinox periods as a three week period on each side of the apparent equinox: 8 September – 20 October and 20 February – 3 April for all species. In these periods we recommend not to use the latitudes in maps and analyses. To avoid an unnecessary over-filtering of longitudes due to the erroneous latitudinal data, the effect of functions acting on positional data are restricted during the equinox periods.

### 2.6.2 Speed filter

Unrealistic positions that is too far apart are filtered by applying maximum flight speeds over 12 hours for each species using the `distanceFilter` function in the `GeoLight` package (**Figure 2.1**, step 2.6.2). This ‘speed filter’ is effective when working on complete time series, and is put first of the following filters because it becomes less effective as the number of time gaps increases. It is not applied during the equinox periods.

The maximum speed for each species was determined from a 3-step process. First, we looked at published values of flight speeds. Then, we assessed whether the species was likely to sustain high flight speed for 12 hours, which is the expected time interval between positions from geolocator data. Finally, we plotted positions on maps to qualitatively assess the balance between gain in noise and loss of positions when allowing progressively higher speeds. The latter step was necessary when using positions obtained from geolocators, because long distance movements in any particular direction will create a zig-zagging pattern that increases with the bird's actual speed due to apparent prolonged or shortened days/nights (Fox 2010). This zig-zagging will later be reduced when positions are double smoothed.

There is a wide range of speeds reported in the literature of the study species. According to Pennycuick (1997), flight speeds between 30 and 70 km/h are the most common for the focal species. Maximum flight speed for alcids can be between 63 and 79 km/h (Pennycuick 1997, Elliott et al. 2014, Jakubas et al. 2016). Northern fulmars can also fly faster than 60 km/h, but their average travel speed over 55 hours is only about half (Edwards et al. 2013). At one extreme, the overall migration speed of lesser-black backed gulls can be as low as 2 km/h along the western flyway (44 km/d, Klaassen et al. 2012). Movement rates averaged over many hours and days are slower than maximum speed due to inclusion of other activities such as resting, foraging and detours (Klaassen et al. 2012). Furthermore, during sustained flight, birds are limited by aerobic capacity and must fly slower than maximum momentary speed that is mainly determined by muscle power (Pennycuick et al. 1994, Pennycuick 1997).



**Figure 2.6.** Histogram of speed of lesser black-backed gulls. For this species, we used 45 km/h (orange line) as the limit in the first version (v 1.0) of the automated procedure, but it was increased to 90 km/h (blue line) in the current version (v2.0). This speed threshold includes more positions from migration, while not greatly decreasing the quality of the dataset. The maps show an example of a lesser black-backed gull tracked from the Solovetsky archipelago, where positions have been filtered with a speed limit of 45 km/h (left) or 90 km/h (right). The speed filter acts on unsmoothed positions (Figure 2.1), but the maps show smoothed positions (the equinoxes have been excluded). See appendix (Figure 7.1) for histograms of other species.

We assumed black-legged kittiwakes and northern fulmars to be the ones most likely to sustain high movement rates over a 12 h interval. Based on the literature (Pennycuick 1997, Edwards et al. 2013), we first chose 60 km/h as maximum speed. After qualitatively assessing plotted positions, we increased the maximum speed to 90 km/h, which corresponds to a movement rate

of 1080 km over 12 h (see also **Figure 7.1** in the supplementary material). The final speed was chosen based on the filter effects on migration tracks that lasted over several days. The increased speed may reflect an underreported effect of utilizing tailwinds in long-distance migration for these species, but certainly reflects the mentioned zig-zagging effect. We also adjusted the maximum speeds for the lesser black-backed gull, from 45km/h to 90 km/h to account for rapid migration of gulls along the eastern flyway (Klaassen et al. 2012, Bustnes et al. 2013). An example showing the effect of increasing the speed limit is shown in **Figure 2.6**.

Based on morphology and empirical evidence, the other species are less likely to sustain high flight speeds over long periods (e.g. Pennycuik et al. 1994, Fayet 2016), and we chose a 45 km/h maximum speed (540 km over 12 h) for common guillemot, Brünnich's guillemot, Atlantic puffin, little auk, glaucous gull, herring gull and the common eider. For the European shag, we chose a very restrictive speed threshold of 30 km/h because of the level of noise in the light-level data, and the short migratory movements. For the last eight species, we did not find it necessary to further adjust the speed in accordance with long-distance migration, but we acknowledge that higher movement rates can sometimes occur, especially during parts of a directional and rapid migration. However, we found that few positions were added compared to the noise added when speed exceeded those limits.

### 2.6.3 Distribution filter

The distribution filter (**Figure 2.1**, step 2.6.3) removes positions outside the species' expected geographic distribution in the North Atlantic (**Table 2.3**). The distribution limits were mainly defined from distributions previous known from SEATRACK phase I (2014-2018). This type of filter could potentially limit our ability to discover any unexpected distributions. However, in our sun angle calibration, described in chapter 2.8, we study maps for each track where we detect potential tracks that cross these boundaries. The present filter was only applied to longitudes during the equinox periods.

**Table 2.3.** *Expected distribution limits for the different species in the North Atlantic. Since some black-legged kittiwakes breeding at Novaya Zemlya migrate to the Pacific Ocean, we do not define a longitudinal limit of distribution for this breeding population.*

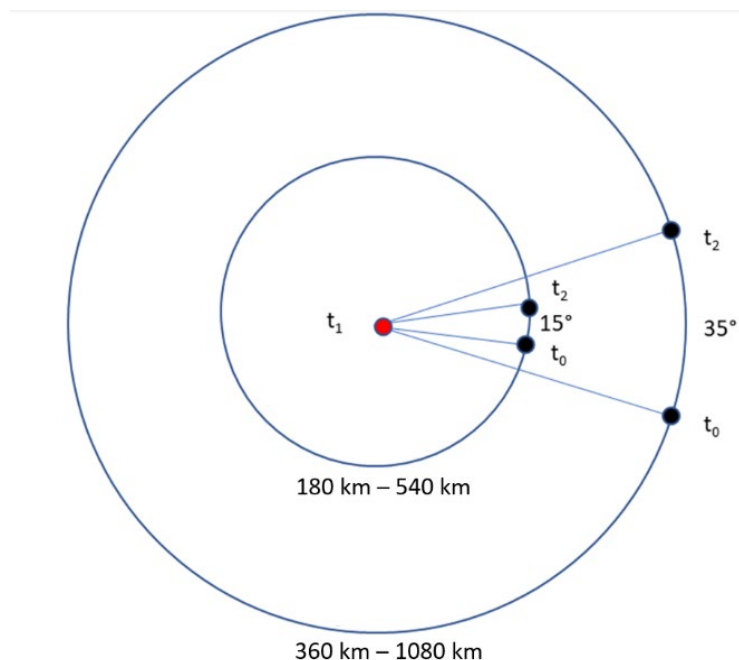
Species	Distribution limits			
	East	West	South	North
Atlantic puffin	-95°	80°	30°	85°
Black-legged kittiwakes	-95°	90°	25°	88°
Black-legged kittiwakes breeding at Novaya Zemlya	-180°	180°	20°	90°
Brünnich's guillemot	-95°	90°	35°	88°
Common eider	-30°	50°	52°	82°
Common guillemot	-95°	70°	30°	85°
European shag	-30°	50°	50°	72°
Glaucous gull	-50°	80°	50°	88°
Herring gull	-35°	55°	45°	75°
Lesser-black backed gull	-30°	80°	-20°	72°
Little auk	-80°	90°	30°	88°
Northern fulmar	-95°	90°	35°	88°



## 2.6.4 Angle and distance filter

Assuming that it is unlikely that an individual would travel a long distance for then to return right away, unrealistic positions that form ‘spikes’ in an individual’s track (Freitas et al. 2007) can be identified and removed by the `sdafilter` from the `argosfilter` package (Freitas 2012, **Figure 2.1**, step 2.6.4). These spikes occur when a position has a long distance to the adjacent positions and the turning angle is low (**Figure 2.7**).

The maximum length of a spike was determined empirically and based on the movement rates applied to the speed filter for each species (**Figure 2.1**, step 2.6.3). Unlike the speed filter, the `argosfilter` was made independent of time and considers the distance only between adjacent positions in the present procedure. The filter will act if any of the following two sets of conditions are fulfilled; it will filter any position where the distance to adjacent positions is above the maximum distance, divided by two, and the turning angle is small, below  $15^\circ$ , or, it will filter any position where the distance to adjacent positions is above the maximum distance and the angle is smaller than  $35^\circ$  (**Figure 2.7**). The function is not applied within the equinox periods and does not work for the first and last position in the dataset.



**Figure 2.7.** Illustration of how the distance + angle filter will filter out the position at  $t_1$  in a path from  $t_0$  to  $t_1$  and  $t_2$ . In the first example, the positions at  $t_0$  and  $t_2$  are a distance specified by species ranging from 180 to 540 km from  $t_1$  and the angle is  $15^\circ$ . In the second example, the positions at  $t_0$  and  $t_2$  are twice the distance specified in the first example (360 to 1080 km) from  $t_1$  and the angle is  $35^\circ$ .

## 2.6.5 Loess filter

As a complementary approach to positional filtering, we apply a `loessFilter` from the `GeoLight` package (Lisovski & Hahn 2012) for removing positions indirectly by using local polynomial regression to find patterns in twilight times and to identify outliers (**Figure 2.1**, step 2.6.5). The filter’s effect is tuned by the variance in the dataset and was put next to last, once many of the most unrealistic twilights have been removed. This is the only filter in our procedure that is completely unaffected by the equinox periods.

Loess filtering can be very effective for filtering positions obtained from species that move slowly and display a predictable pattern movement. It is not, however, well designed for recognizing sudden onset of migration bouts and can sometimes delete twilights occurring at the start or at the end of rapid movements. To account for this, we used species-specific  $k$ -values, which represent how many interquartile ranges it takes before a twilight event is identified as an outlier (**Table 2.4**). The  $k$ -values were based on inspecting plots for >50 individuals per species from over >4 colonies and qualitatively assessing the balance between increase in noise and loss of positions when allowing progressively lower values of  $k$ .

For black-legged kittiwakes, we selected a high  $k$ -value of 6, as this species often undertakes large-scale movements, while their recorded light-level data generally have a clear diurnal pattern which require less filtering. Northern fulmars also undertake large-scale movements, but their light-level data are also often influenced by anthropogenic light from e.g. the ships they follow (Dupuis et al. 2021) and shadow events when attending breeding grounds already after their autumn moult (Grissot et al. 2020). Hence, we chose a  $k$ -value of 3 as a compromise, which provides a clearer identification of where each individual went, but can sacrifice twilights occurring under actual movements. For alcids, the  $k$ -value is also set to 3, as they are less likely to sustain high flight-speeds over a 12-hour period due to the high cost of flapping flight (Hedenström, 1993, Watanabe, 2016), and movements are therefore slow and predictable.

**Table 2.4.** Species and their assigned  $k$ -value that is used under loess filtering.

Species	$k$ -value
Atlantic puffin	3
Black-legged kittiwake	6
Brünnich's guillemot	3
Common eider	2
Common eider breeding in Kongsfjorden	5
Common guillemot	3
European shag	2
Glaucous gull	2
Herring gull	2
Lesser black-backed gull	3
Little auk	3
Northern fulmar	3

Common eiders and European shags are coastal species that can experience shading from the coastal terrain throughout the year. Therefore, we chose a low  $k$ -value of 2. An exception was made for common eiders breeding in Kongsfjorden, Svalbard. This is the only population of common eiders in SEATRACK that migrates over long-distances, flying from Svalbard to Iceland or Norway over a few days (Hanssen et al. 2016). We set a  $k$ -value of 5 for this particular population to avoid losing information about this migration.

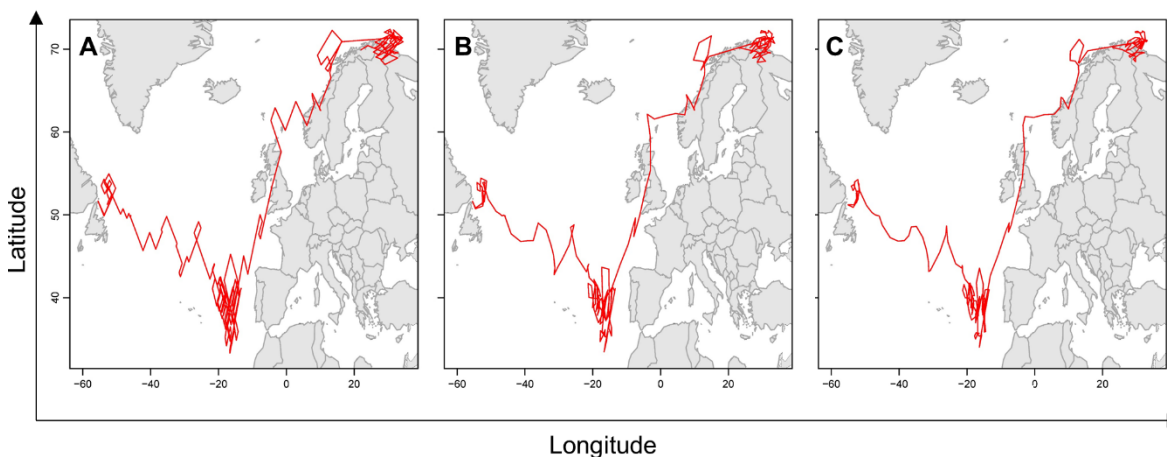
For glaucous gulls and herring gulls, we set a low  $k$ -value of 2. These two species are capable of undertaking large scale movements, but their logger data is also influenced by anthropogenic light or shadows as they can be attracted to human activities in coastal and terrestrial areas through the non-breeding season. For the lesser black-backed gull, we chose a higher  $k$ -value of 3.

## 2.6.6 Midnight sun filter

Shading events may appear frequently during midnight sun conditions, e.g. from breeding activities in the colony. The recorded light-level will cross the threshold and create twilight events and subsequently unrealistic positions. Ideally, such twilight events and positions should have been removed at this point. Yet, filters often fail to remove all of these because the ambiguous nature of the positions often fail to fulfil the necessary conditions needed to be excluded from the dataset. Therefore, we apply a last function that defines a start and an end for the year tracked, which in turn helps remove positions under (or very close to) the midnight sun periods. The function defines the start of a year tracked as the first of the first three consecutive positions within 25 hours after deployment/1 June. The end of a track was defined as the last position of the last three consecutive positions within 25 hours and before retrieval/1 June.

## 2.7 Double smoothing of positions

Raw positions can exhibit a typical noon–midnight zigzag pattern in latitude due to east–west movements, and to lesser extent in longitude due to north–south movements. In order to reduce the influence of inaccurate positions and compensate for movements, we used a double smoothing procedure as described in Hanssen et al. (2016), adapted from Fox (2010, 2015) and Phillips et al. (2004). The first step involves averaging noon–midnight and midnight–noon paired latitudes and longitudes, respectively. This is equivalent to the ‘two point moving average’ procedure in the software Intiproc (Migrate Technology Ltd, Cambridge, Fox 2015). The second step was to calculate a two-point moving average over these positions using spherical trigonometry (Frederiksen et al. 2012). The **Figure 2.8** show a tracked black-legged kittiwake from Hornøya with smoothed and unsmoothed positions.



**Figure 2.8.** The double smoothing applied on positions from a black-legged kittiwake from November to April. A) Raw positions, B) average noon–midnight and midnight–noon paired latitudes and longitudes and, finally, C) two point moving average using spherical trigonometry. Positions during the spring equinox period has been excluded.

## 2.8 Sun angle calibration

During the first run, preliminary positions are used to produce two kinds of plots that can be used in a sun angle calibration. The first kind is a series of latitude versus time plots (**Figure 2.9**) and the second is a series of maps with plotted positions (**Figure 2.10**). Both consist of a series of 11 plots where latitudes have been calculated based on sun angles ranging from  $-1.5^{\circ}$  to  $-6.5^{\circ}$ , with intervals of  $0.5^{\circ}$ . The person running the procedure then selects a sun angle based on the criteria described below.

### 2.8.1 Latitude versus time plots

Plots of latitude versus time for different sun elevation angles are used for selecting an appropriate sun elevation angle to use in calculating positions from twilight events (see examples in **Figure 2.9** and in the appendix, **Figure 7.2 – 7.5**). Such plots show a disruptive pattern before and after the equinoxes, illustrating amplification of latitudinal error with inappropriate sun elevation angles (Lisovski and Hahn 2012). The more inappropriate the sun elevation angle is, the more amplified is the error, and latitude curves take on exponential shapes. Illustratively, these curves are disrupted at the equinoxes, and this pattern is reversed when going from one end of inappropriate sun elevation angles to the other (see plots in appendix **Figure 7.3** for maximum and minimum sun angle as examples). At the appropriate sun elevation angle, there is no clear disruptive pattern, and the amplification of the latitudinal error is at a minimum.

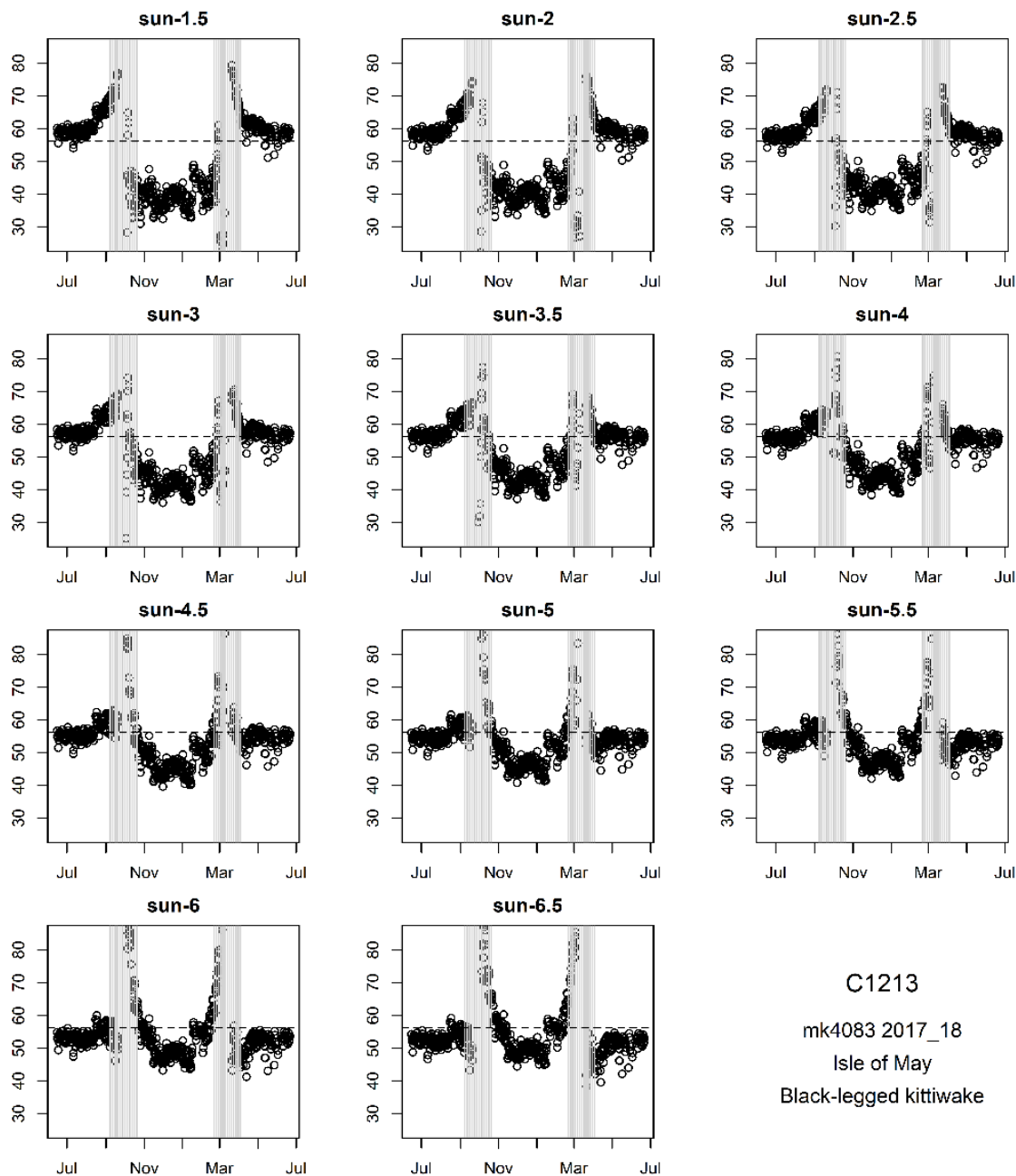
Our calibration process is adapted from the approach by Hanssen et al. (2016) and van Bemelen et al. (2019). Using the latitude versus time plots, we select the sun elevation angle based on three criteria that 1) minimize the amplification of latitudinal error close to the equinoxes, 2) result in matching latitudes at both sides of the equinox and 3) result in positions that fit the latitude of the colony in the beginning and the end of the track. We cannot, however, use all three criteria for all tracks. It is less useful to use criteria 2 when the bird has moved north/south during the equinox period. Nevertheless, logical and realistic latitudes should be expected at each side of the equinox also during latitudinal movements. Criteria 3 is not applicable when the track does not start or end at the colony, which is often the case at high latitudes since birds which may depart or arrive at the colony when constant daylight prevents geolocation. In addition, we plot positions on maps, using different sun elevation angles, to check that positions fitted the shape and position of the oceans/continents. Latitude versus time plots are plotted sequentially at intervals of  $0.5^{\circ}$  and the most appropriate sun elevation angle is chosen. If two adjacent plots/sun elevation angles are considered to be equally good, the average is assumed to be most appropriate.

In some cases, the year tracked is incomplete and not appropriate to calibrate with the above principles. This may happen when a geolocator stopped recording in the summer or autumn. If the geolocator has been recording for  $>1$  year, we use the previous year tracked to calibrate the incomplete track. Otherwise, we use the average sun elevation angle of the same logger model used on conspecifics.

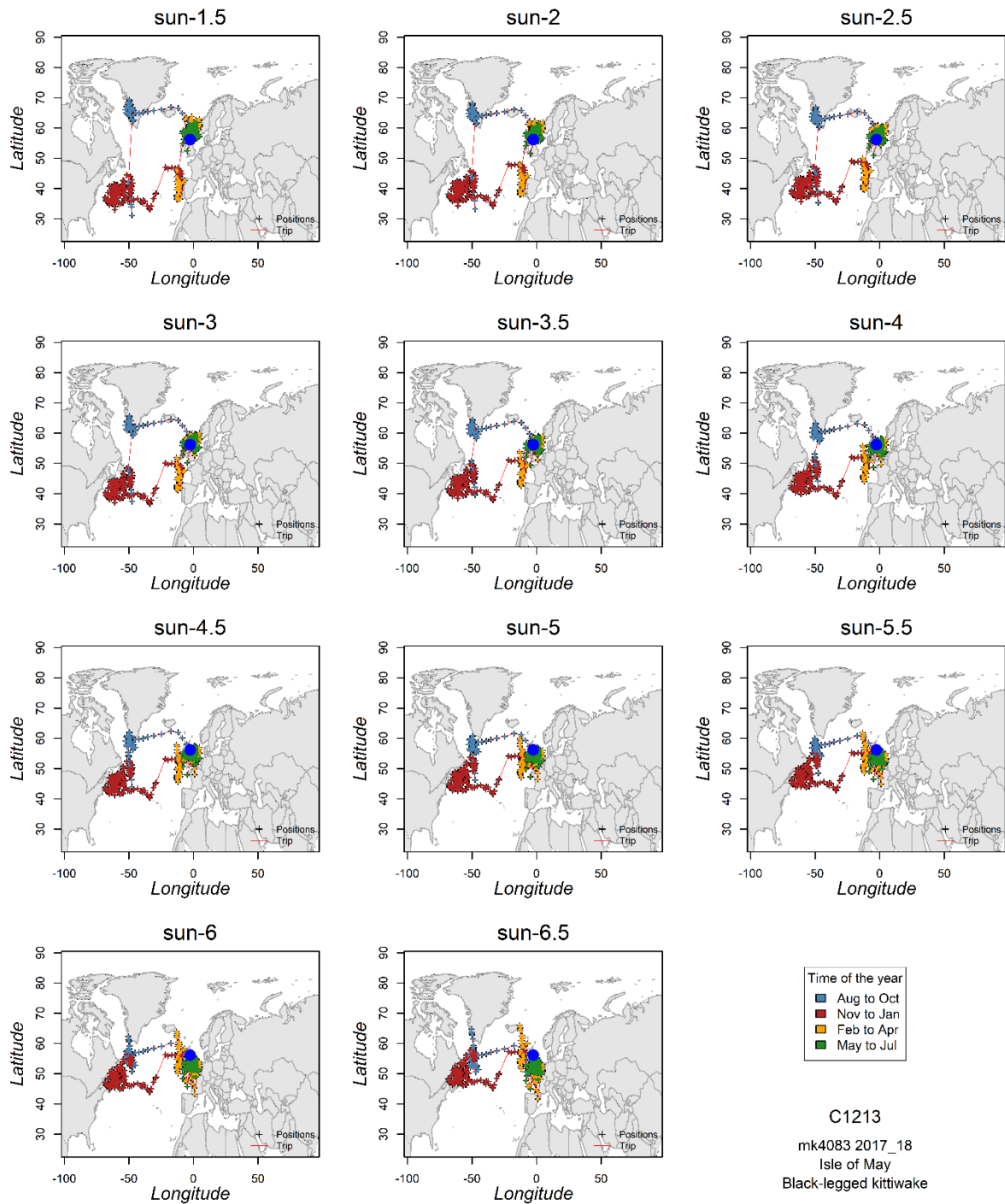
### 2.8.2 Maps for calibration

In addition to latitude versus time plots, we plot positions on a series of maps using the same range and resolution of sun elevation angles as for the latitude versus time plots. The main purpose of these maps is to check that positions fit the shape and position of the oceans/continents, as well as the position of the colony. When a bird is inshore, positions should distribute equally over ocean and land. The maps are therefore used to confirm that the selected sun elevation angle from the previous step is appropriate. They are also used to aid the selection of sun elevation angle if the assessment from the first step is less clear or if the maps indicate that the first

step needed reassessment. Below we provide examples of such maps which correspond to the examples from above (**Figure 2.10**, and in the appendix, **Figure 7.3 & 7.5**).



**Figure 2.9.** Latitude versus time plotted for different sun elevation angles for a black-legged kittiwake tracked from Isle of May. Time spans one non-breeding season from June 2017 to June 2018. The horizontal dotted line shows the latitude of the colony, and the vertical grey lines indicate the periods around autumn and spring equinoxes. In this example, we selected  $-3.5^\circ$  as the appropriate sun elevation angle. We mainly used criteria 1) and 3). It was less useful to use criteria 2) since the bird moved south/north during the equinox periods. However,  $-3.5^\circ$  provided realistic latitudes at both sides of the equinoxes.

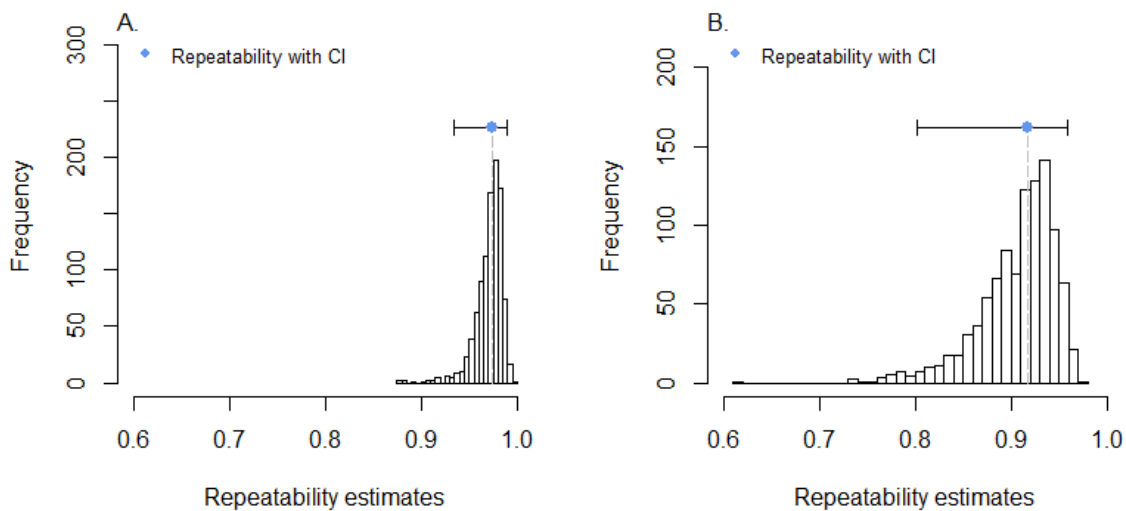


**Figure 2.10.** Smoothed and filtered positions calculated with different sun elevation angles for a black-legged kittiwake tracked from Isle of May in 2017/18 (same individual as in **Figure 2.9**). These maps supported the selection of  $-3.5^\circ$  as sun elevation angle. The location of the colony is marked with a filled blue symbol, and map limits are defined by the species-specific distribution range from the Distribution filter. Positions are coloured by month. Positions from the equinox periods has been excluded from the map.

### 2.8.3 Consistency of the calibration assessments

To test the repeatability of our methods for calibration and the selection of sun elevation angles, we randomly selected latitude versus time plots of four northern fulmars, four black-legged kittiwakes, four common guillemots and four common eiders from one colony each. These 16 sets of plots were triplicated to obtain 48 (3 x 16) sets. They were given unique plot identities and information about species and logger model was removed. These 48 sets were presented to VSB (Researcher 1) and BM (Researcher 2) who assigned sun elevation angles to each set.

VSB assigned an average sun elevation angle of  $-3.65^\circ$  (SE = 0.13), while BM differed by only  $-0.04^\circ$  (SE = 0.041, linear mixed-effects model,  $t_{79} = -1.02$ ,  $p > 0.3$ ). Using the `rptR` package (Stoffel et al. 2017) we calculated repeatability of the assessments with bird identity entered as random factor (Stoffel et al. 2017). Both researchers had high and statistically significant repeatability in their assessments, with 0.97 (95% CI = [0.935, 0.989]) for VSB and 0.92 (95% CI = [0.802, 0.959]) for BM (**Figure 2.11**). This demonstrates consistent assessments. VSB showed the highest consistency, and he also performed most of the calibration assessments in SEATRACK.



**Figure 2.11.** Repeatability with 95% confidence interval from bootstrap repeatability estimates for assessments of sun elevation angles by researcher 1 (A.) and researcher 2 (B.)

### 3 Filter effects on the number of twilights and positions

Our automated procedure described in chapter 2 first removes redundant twilights ('Filter twilights'). Second, it shifts unrealistic twilights to a more likely time of day ('Edit twilights'). Finally, it aims to remove any remaining unrealistic positions by the filters described in chapter 2.6.2 – 2.6.6. When applied to light-level data extracted from our SEATRACK database, 'Filter twilights' removed a median number of 60 twilights per year tracked, 'Edit twilights' moved a median number of 9 twilights per year, while the filters described in chapters 2.6.2 – 2.6.6 removed a median number of 55 positions per year. A species-specific overview of the effects of these filters is provided in **Table 3.1** and examples of how each filter influence the shape of a track, together with the smoothing of the final positions are shown in **Figure 3.1**.

**Table 3.1.** The effect of the various filters shown as median and 50% percentiles (0.25%, 0.75%) for the numbers of twilights or positions removed for each species. A species' winter habitat is characterized in the 'winter habitat' column. Species are ranked by their median number of removed positions.

species*	winter habitat**	# tracks	twilights		positions					Proportion (%)		
			filtered	edited	speed	distribution	angle & distance	loess	midnight sun		final	removed
ES	C	450	55 (26, 129)	21 (15, 28)	94 (71, 118)	10 (5, 18)	24 (17, 31)	15 (9, 24)	6 (0, 7)	516 (476, 554)	153 (121, 179)	30
GG	P- C	275	171 (108, 257)	13 (10, 17)	55 (37, 71)	8 (5, 13)	8 (4, 11)	18 (13, 26)	7 (6, 9)	310 (263, 387)	100 (76, 121)	32
LBBG	P- C- T	24	86 (56, 123)	7 (4, 8)	25 (17, 29)	4 (3, 8)	19 (12, 24)	10 (9, 14)	6 (6, 7)	531 (514, 646)	93 (76, 103)	18
AP	P	904	145 (87, 215)	10 (5, 17)	50 (34, 66)	7 (3, 10)	10 (6, 16)	8 (4, 14)	6 (6, 7)	484 (408, 541)	83 (62, 107)	17
HG	P- C- T	90	80 (41, 151)	20 (9, 29)	34 (26, 54)	4 (1, 10)	12 (8, 18)	14 (7, 20)	6 (0, 7)	552 (459, 643)	77 (64, 101)	14
CE	C	535	188 (117, 262)	11 (7, 16)	26 (17, 39)	7 (4, 12)	10 (6, 14)	21 (7, 33)	7 (6, 8)	501 (443, 546)	77 (54, 101)	15
NF	P- C	814	108 (56, 175)	10 (6, 14)	31 (20, 47)	4 (2, 8)	11 (7, 14)	12 (7, 19)	6 (0, 7)	527 (395, 597)	66 (50, 87)	13
CG	P	1192	30 (4, 98)	8 (4, 13)	26 (10, 46)	2 (0, 5)	7 (4, 12)	7 (3, 11)	6 (6, 7)	497 (450, 549)	52 (30, 80)	11
BG	P	892	27 (8, 64)	7 (4, 11)	19 (11, 33)	3 (1, 6)	7 (3, 12)	6 (2, 10)	6 (6, 7)	460 (377, 496)	45 (30, 66)	10
LA	P	314	74 (30, 111)	7 (4, 11)	22 (12, 35)	5 (3, 8)	4 (2, 6)	5 (2, 8)	6 (6, 7)	417 (317, 429)	45 (30, 62)	11
BLK	P	2207	30 (11, 61)	7 (4, 10)	15 (9, 23)	3 (1, 6)	7 (4, 10)	2 (0, 4)	7 (6, 8)	492 (450, 565)	36 (26, 47)	7
<b>Total</b>		<b>7896</b>	<b>60</b> (19, 137)	<b>9</b> (5, 14)	<b>25</b> (13, 46)	<b>4</b> (2, 8)	<b>8</b> (5, 13)	<b>6</b> (2, 13)	<b>6</b> (6, 8)	<b>483</b> (429, 544)	<b>55</b> (35, 86)	<b>14</b>

\*GG = glaucous gull, NF = northern fulmar, CE = common eider, ES = European shag, HG = herring gull, LBBG = lesser black-backed gull, AP = Atlantic puffin, CG = common guillemot, BG = Brünnich's guillemot, BLK = black-legged kittiwake, LA = little auk

\*\*C = coastal, P = pelagic, T = terrestrial

\*\*\*numbers do not take into account effects of 'filtered twilights' or 'edited twilights'

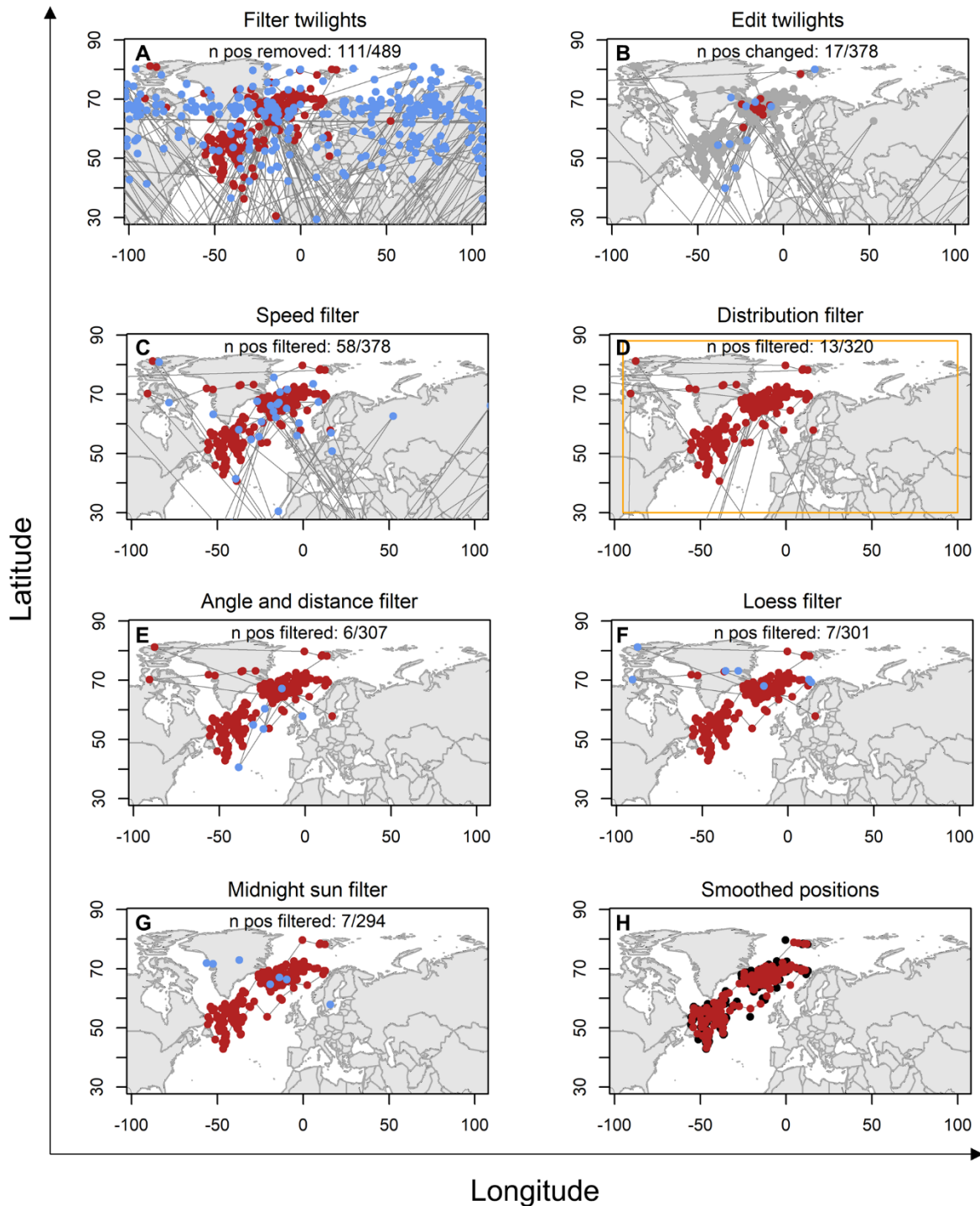


The large number of redundant twilights removed by the 'Filter twilights' shows its ability to clean and improve the quality of the dataset. The second step of editing twilights moved many fewer twilights than the first filter removed, but presuming that moving one twilight affects the location of two calculated positions, the 'Edit twilights' filter is still an important filter in our procedure. However, their contribution is probably greatest during the breeding season, as illustrated in the previous examples (**Figure 2.4**).

Among filters that remove unrealistic positions, the 'Speed filter' had the largest effect, 'Angle and distance filter' the second largest, 'Loess filter' the third, the 'Midnight sun filter' the fourth and the 'distribution filter' the lowest. However, the effect of each filter is largely dependent on their order in the procedure, and their implications on the final data quality vary through the annual cycle.

When comparing the median number of removed positions to each species' winter habitat (**Table 3.1**), typical inshore (coastal) species seem to record more erroneous twilights than offshore (pelagic) species, with coastal European shags having by far the most positions removed. A notable exception when comparing pelagic and coastal species was the Atlantic puffin, which required a high degree of filtering. Some of the reasons can be that they, being a burrowing species, are prone to record erroneous positions through most of the summer, that they may tuck leg with logger inside their plumage when resting, and a large proportion reside in the southern range of SEATRACK where fewer positions are lost to the midnight sun or polar night.

Finally, the number of final positions obtained from the light-level data seems to reflect a species' latitudinal range more than whether the species remains coastal or pelagic during the non-breeding season. In other words, it reflects the number of positions lost to periods of polar night or midnight sun. For example, the number of positions obtained is much lower in typical high-arctic species like glaucous gull and little auks, than for a species mostly concentrated in the southern range of the dataset, like the European shag. Considering that data recorded from European shags required the most filtering of positions, while data from the little auk required the second least, the combined effects of the automated filtering do not seem to be exaggerated, but instead reflect the typical quality of the recorded light-level we can expect from each species.



**Figure 3.1.** Examples of filtering and smoothing of data from a northern fulmar breeding in Iceland. Red positions are those retained or changed ("Edit twilights") in each step. Blue positions are the original positions that are removed or changed ("Edit twilights"). A grey line is drawn between consecutive positions. In panel H, smoothed positions are shown in red, while the unsmoothed positions are shown in black. Both sets of positions from panel H are available in our resulting dataset. Positions affected by the equinox periods have been removed from plots and from the displayed number of filtered positions.

## 4 Comparing SEATRACK procedures; manual versus automatic

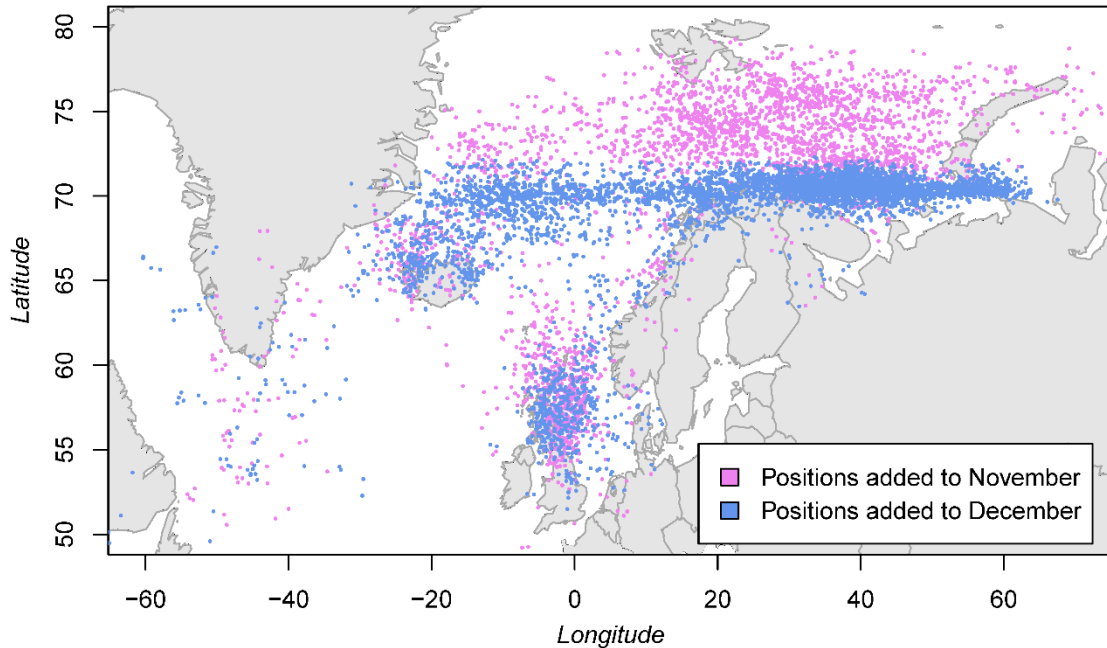
In this chapter, we compare positions obtained with the present automated procedure and those obtained with the previous manual procedure used in SEATRACK from 2014 to 2018/2019. In the latter, filtering and editing of unrealistic twilights were mainly done manually by a person with expert knowledge, after initially having used the same thresholds as in the present procedure to aid the identification of twilight events. The manual process also included two subsequent steps of automated filtering, which consisted of a minimum day- and night length-filter (very helpful in reducing influence of e.g. shorter events of shading) and a *Loess filter*. A satisfying result from comparing these datasets should not, in theory, yield a better result for the automatically processed dataset but one close to that of the manually processed, since the automated procedure does not involve any advanced probabilistic modelling nor provides as extensive fine-tuning of twilights like the previous manual procedure could. However, there are two exceptions where the new method should or could outperform the old. First, the shortest day- and night length allowed has been reduced from two to one hour in the automated procedure, which should expand the northern positional range of the automatically processed dataset during summer and winter. Second, while a standard *k*-value of 3 was applied for the *Loess filter* in the manual procedure, the *k*-value has been made species-specific in the present automated procedure, which could include more realistic positions for some species or filter out more unrealistic positions in others.

For comparisons, light-level data were processed by both procedures using the same sun angles for each individual year tracked. Paired t-tests were used to identify mean differences between the two procedures. We chose to compare data from November and December, since comparing short timespans increase interpretability, these months are part of the non-breeding season and are unaffected by the equinoxes. However, the two months also differ in some aspects. November is characterized as a month with enough light to obtain positions far north (in theory: 1 Nov at approx. 80°N, 30 Nov at approx. 72°N) and where individuals still cross large distances to reach their wintering grounds. December is the darkest month of the year, where individuals can experience polar night (northern range, in theory: 1 Dec. at approx. 72°N, 21 Dec at approx. 69°N) and fewer individuals undertake large scale movements. We assessed four variables that were calculated for each unique monthly track; the difference in size (km<sup>2</sup>) of 50% kernels, the average latitude and longitude, and the average number of positions obtained per monthly track. Kernels UDs were computed using a Lambert azimuthal equal-area projection, selecting “href” as a smoothing factor and a grid of 500. An overview of the results is presented in **Table 4.1** and examples plotted in **Figure 4.1**. Complete tables with species specific comparisons are presented in the appendix (appendix, **Table 7.1 & 7.2**).

**Table 4.1.** The table present differences in the positional data after using the same light-level data to estimate positions, first by SEATRACK’s manual procedure, and then with the new automatic procedure (this report). We compare 50% kernel HR size (km<sup>2</sup>), latitudinal mean position, longitudinal mean position and number of positions per individual in November and December. The column « $\Delta$  from man» show how a track produced by the automated procedure is different from a track produced with the manual procedure, along with results from paired t-tests.

Variable		$\Delta$ from man	paired t-test	Month
Area KHR 50%		+5.67 %	t = -5.261, df = 2358, p-value < 0.001	November
Latitude	mean	0.07°N	t = 7.357, df = 2358, p-value < 0.001	November
Longitude	mean	0.18°E	t = 7.127, df = 2358, p-value < 0.001	November
N positions	mean	+1.20	t = -14.191, df = 2358, p-value < 0.001	November
Area KHR 50%		+2.31 %	t = -1.815, df = 2096, p-value = 0.070	December
Latitude	mean	0.08°N	t = 8.371, df = 2096, p-value < 0.001	December
Longitude	mean	0.18°E	t = 4.524, df = 2096, p-value < 0.001	December
N positions	mean	+2.43	t = -18.470, df = 2096, p-value < 0.001	December

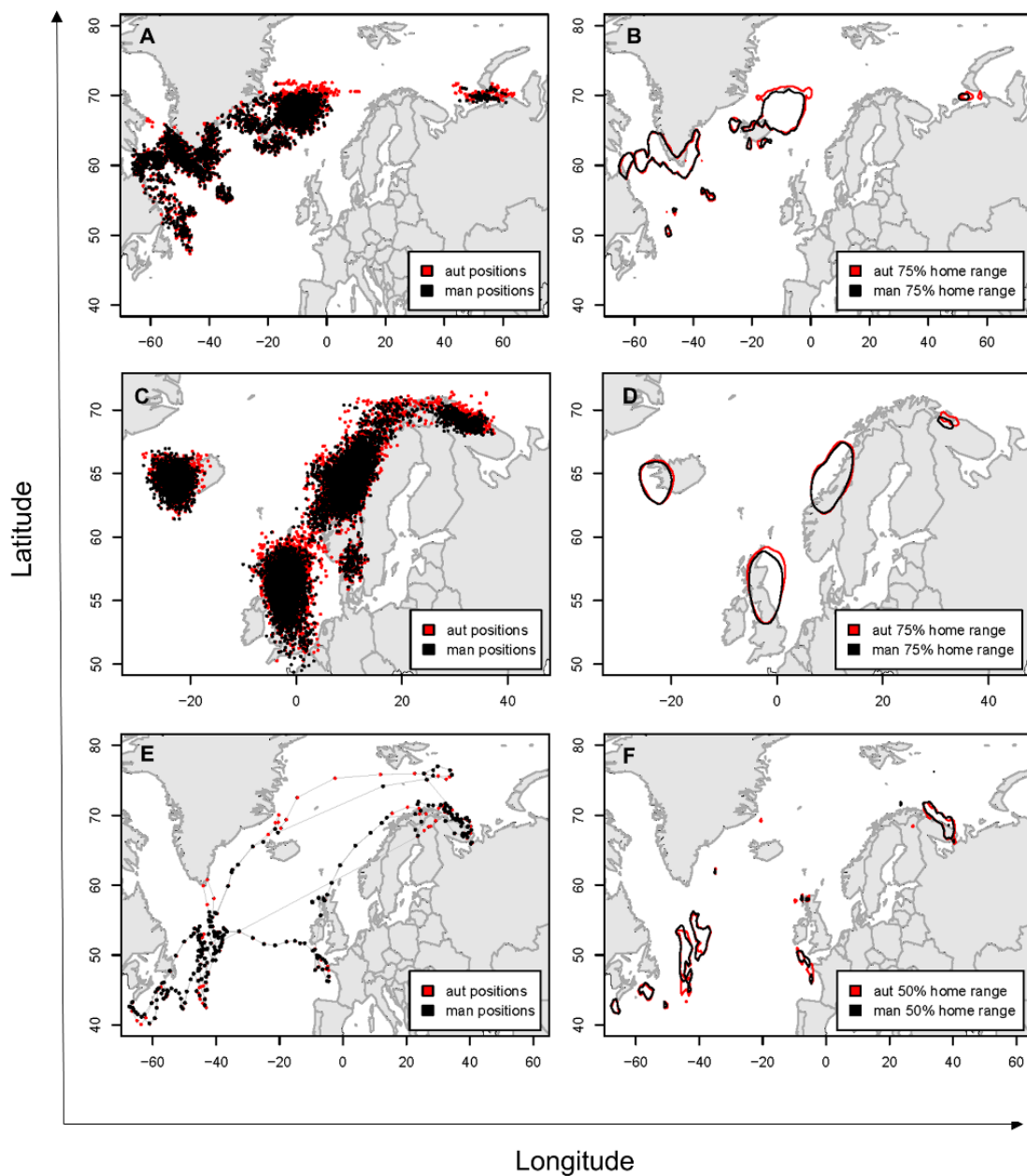
Overall, the average monthly kernel size was slightly larger when geolocator data was processed with the automated procedure, as more positions were added to the dataset and the average monthly location were shifted slightly east and north (**Table 4.1**). When plotting automatically obtained positions from dates that were not present in the manually processed dataset (**Figure 4.1**), added positions were mainly present in the northern and eastern range of the dataset. When plotting kernels on a map, the southern distribution seems to be less or not at all affected by the applied processing routine (e.g. little auks in panel A & B, **Figure 4.2**).



**Figure 4.1.** Automatically obtained positions for all species from dates that were not present in the previous manual processed dataset are plotted in violet for November and blue for December. Points concentrated in the north of the British Isles and Iceland reflect the large numbers of individuals from our dataset aggregated in these areas all year through.

The difference in size of kernels was slightly less for December than in November, as fewer positions could be estimated at higher latitudes in December due to polar night, whereas the difference in latitude and longitude was the same and the difference in number of positions larger. Species by species, it is clear that differences in kernel size increased in December for species like alcids and glaucous gulls that were still aggregated in large numbers at the northern positional range of the dataset, where most new positions also were added to the dataset (**Figure 4.1** and panel A & B, **Figure 4.2**), while the differences were smaller for species that are well-known to reside in the southern range of the dataset in December, like the black-legged kittiwake (appendix, **Table 7.1 & 7.2**).

For the black-legged kittiwake, the differences in November compared to December can also be related to a less restrictive `Loess` filtering in the automated versus the manually processed data, which can have a positive effect on the numerous large-scale movements towards wintering grounds from east to west in November. Here, the low  $k$ -value used for the `LoessFilter` in the manual procedure could filter twilights during rapid movements that we otherwise would like to keep, as illustrated in example E & F in **Figure 4.2**, and further push the mean distribution slightly north-east when processing data with the automated routine.



**Figure 4.2.** Examples of positions from SEATRACK's previous manual (black) and current automated (red) procedure (v2.0). Left panels (A, C, E) show positions and right panels (B, D) the 75% and (F) the 50% kernel density. Panels A and B show little auks in December and illustrate that the automated procedure has determined more twilights (and positions) around 70°N, i.e. at the northern range of apparent daylight in December (light threshold/twilight correspond to when the sun is approximately 3.5 below the horizon). Panels C and D show European shags in December and illustrates the limited difference for a species that is mostly south of this northern light limit. Panels E and F show a track of a black-legged kittiwake from October to February which illustrate the effect of less restrictive *k*-value for *Loess* filtering in the automated procedure.

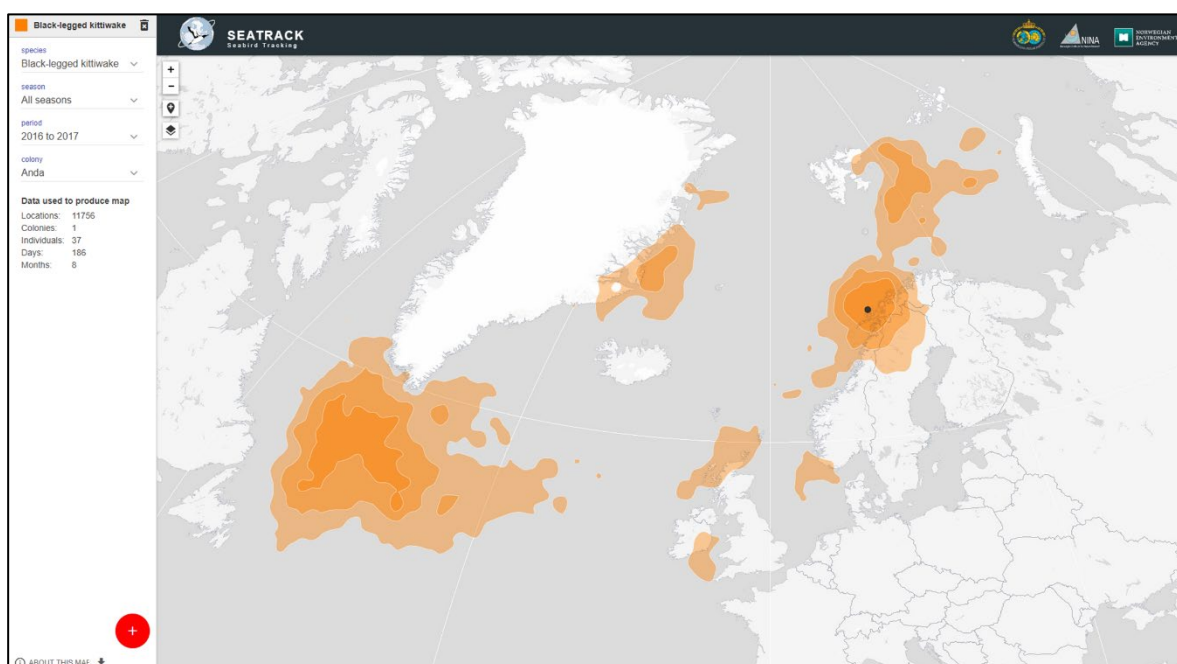
However, the results are not all in favour of the automated procedure. It is likely that automatically processed data sometimes offer less accurate positions that could increase kernel size and the number of positions added to the dataset. For example, the European shag is distributed along the coast throughout the year and is not expected to cross large stretches of water, but when

positions were plotted on top of each other, a larger proportion of positions obtained with the automated procedure were spread away from the coast (panel C, **Figure 4.2**). Indeed, a lower k-value was assigned for this species in the automated procedure, which should result in more twilights being removed by the `LoessFilter`. Yet, the automated procedure produced larger individual kernels in both months, which in this case shows that the accuracy can drop slightly with the automated procedure (**Table 7.1 & 7.2**, panel D, **Figure 4.2**).

We did not attempt to make a thorough comparison of the time required for each procedure to process data. However, we estimate that a manual inspection of twilight events from an individual tracked for one year would require 10 – 40 minutes, depending on the level of influence from artificial lights or shading, after which running scripts and sun angle calibrations would require an additional 1-3 minutes, disregarding the time to prepare files or folders in between. The present automated procedure requires roughly 2-3 minutes to process a year tracked, including sun angle calibrations and running scripts a second time for estimating and filter coordinates with a calibrated sun angle.

## 5 Applications and limitations

The present automated procedure offers a faster and more consistent and reproducible way of obtaining positions from geolocators than a procedure which includes manually inspection of twilights. The positional dataset is a key deliverable for the SEATRACK project to be used for management and environmental impact assessments as well as in scientific publications. In the appendix, we provide a readme-file and a basic script for reading, handling, calculating home ranges and plotting data in the open source R Statistical Software (R Core Team, 2019). The positional dataset is used to produce the kernel maps which are available in our web application (**Figure 5.1**, <http://seatrack.seapop.no/map/>). Ownership of the data is shared between the project and partners under our Agreement of Understanding, and access to SEATRACK data can be granted on formal request to the project leader (application form can be downloaded from the project's webpage), which in turn will ask for approval from the respective data owners.



**Figure 5.1.** Example of a kernel distribution map presented on SEATRACK's web app.

However, there are several limitations to keep in mind. The positional error is considered high when positions are obtained from light-level data (Edwards et al. 2016, Philips et al. 2014, Fudickar et al. 2012), and we do not recommend using the data in studying movements of less than 200 km (Lisovski et al. 2019). As an example of the limited precision of our dataset, one can often find that half of the positions for an individual are over land if it remains at the coast. Further, the positional dataset has gaps; 1) during the equinox periods, where latitudes are unreliable and removed by our filter (see readme in appendix 7 for how to retain longitudes and removed latitudes during these periods), 2) at high latitudes because of apparent midnight sun or polar night and 3) due to filtering of twilights and unrealistic positions.

A tailored informed random movement algorithm (IRMA, Fauchald et al. 2019) has been developed for filling data gaps and reducing statistical bias of the positional dataset. Currently this IRMA dataset includes six pelagic species and is available by submitting a Data Request Form (see above).

## 6 References

Biotrack 2013. M-Series Geolocator User Manual. Biotrack Ltd, Cambridge.

Bustnes, J.O., Moe, B., Helberg, M., Phillips, R.A. 2013. Rapid long-distance migration in Norwegian Lesser Black-backed Gulls *Larus fuscus fuscus* along their eastern flyway. *Ibis* 155, 2: 402-406.

Dias, M.P., Martin, R., Pearmain, E.J., Burfield, I.J., Small, C., Phillips, R.A., Yates, O., Lascelles, B., Borborogluuf, P.G. & Croxall, J.P. 2019. Threats to seabirds: A global assessment. *Biological Conservation*, 237: 525–537.

Dupuis, B., Amélineau, F., Tarroux, A., Bjørnstad, O., Bråthen, V.S., Danielsen, J., Descamps, S., Fauchald, P., Hallgrímsson, G.T., Hansen, E. S., Helberg, M., Helgason, H.H., Jónsson, J.E., Kolbeinsson, Y., Lorentzen, E., Thompson, P., Thórarinnsson, T.L. & Strøm, H. 2021. Light-level geolocators reveal spatial variations in interactions between northern fulmars and fisheries. *Marine Ecology Progress Series*. In press.

Edwards, E.W.J., Quinn, L.R., Wakefield, E.D., Miller, P.I. & Thompson, P.M. 2013. Tracking a northern fulmar from a Scottish nesting site to the Charlie-Gibbs fracture zone: evidence of linkage between coastal breeding seabirds and Mid-Atlantic ridge feeding sites. *Deep Sea Research Part II*, 98, 438–444.

Edwards, E.W.J., Quinn, L.R. & Thompson, P.M. 2016. State-space modelling of geolocation data reveals sex differences in the use of management areas by breeding northern fulmars. *Journal of Applied Ecology*, 53: 1880-1889.

Elliott, K.H. & Gaston, A.J. 2014. Dive behaviour and daily energy expenditure in thick-billed Murres *Uria lomvia* after leaving the breeding colony. *Marine Ornithology*. 42: 183–189.

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 from year-round modelling of environmental niche and abundance using light-logger data. NINA Report 1657. Norwegian Institute for Nature Research.

Fayet, A.L., Freeman, R., Shoji, A., Boyle, D., Kirk, H.L., Dean, B.J., Perrins, C.M. & Guilford, T. 2016. Drivers and fitness consequences of dispersive migration in a pelagic seabird. *Behavioral Ecology*, 27: 1061-1072.

Frederiksen, M., Descamps, S., Erikstad, K.E., Gaston, A.J., Gilchrist, H.G., Grémillet, D., Johansen, K.L., Kolbeinsson, Y., Linnebjerg, J.F., Mallory, M.L., McFarlane Tranquilla, L.A., Merkel, F.R., Montevecchi, W. A., Mosbech, A., Reiertsen, T.K., Robertson, G.J., Steen, H., Strøm H. & Thórarinnsson T.L. 2016. Migration and wintering of a declining seabird, the thick-billed murre *Uria lomvia*, on an ocean basin scale: Conservation implications. *Biological Conservation*, 200: 26–35

Frederiksen, M., Moe, B., Daunt, F., Philips, R.A., Barret, R.A., Bogdanova, M.I., Boulinier, T., Chardine, J.W., Chastel, O., Chivers, L.S., Christensen-Dalsgaard, S., Clément-Chastel, C., Colhoun, K., Freeman, R., Gaston, A.J., González-Solís, J., Goutte, A., Grémillet, D., Guilford, T., Jensen, G.H., Krasnov, Y., Lorentsen, S-H., Mallory, M.L., Newell, M., Olsen, B., Shaw, D., Steen, H., Strøm, H., Systad, G.H., Thórarinnsson, T.L. & Anker-Nilssen, T. 2012. Multicolony tracking reveals the winter distribution of a pelagic seabird on an ocean basin scale. *Diversity and Distributions*, 18: 530-542.



- Freitas, C., Lydersen, C., Fedak, M.A. & Kovacs K.M. 2007. A simple new algorithm to filter marine mammal Argos locations. *Marine Mammal Science*, 24: 315-325.
- Freitas, C. 2012. *argosfilter*: Argos locations filter. R package version 0.63. <https://CRAN.R-project.org/package=argosfilter>
- Foster, J. 2020. *roll*: Rolling and Expanding Statistics. R package version 1.1.6. <https://CRAN.R-project.org/package=roll>
- Fox, J. 2010. *Geolocator manual v8*. British Antarctic Survey, Cambridge.
- Fox, J. 2015. *Intiproc geolocation processing software. User reference manual Intiproc v1.03* (January 2015). Migrate Technology Ltd, Cambridge.
- Fox, J. 2018. *Intigeo® series geolocator*. Migrate Technology Ltd, Cambridge.
- Fudickar, A.M., Wikelski, M. & Partecke, J. 2012. Tracking migratory songbirds: accuracy of light-level loggers (geolocators) in forest habitats. *Methods in Ecology and Evolution*. 3: 47-52.
- Hanssen, S.A., Gabrielsen, G.W., Bustnes, J.O., Bråthen, V.S., Skottene, E., Fenstad, A.A., Strøm, H., Bakken, V., Phillips, R.A. & Moe, B. 2016. Migration strategies of common eiders from Svalbard: Implications for bilateral conservation management. *Polar Biology* 39: 2179–2188.
- Hedenström, A. 1993. Migration by soaring or flapping flight in birds: the relative importance of energy cost and speed. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 342:353–361.
- Grissot, A., Graham, I.M., Quinn, L., Bråthen, V.S. & Thompson, P.M. 2020. Breeding status influences timing but not duration of moult in the Northern Fulmar *Fulmarus glacialis*. *Ibis*. 162. 446-459.
- Jakubas, D., Iliszko, L.M., Strøm, H., Darecki, M., Jerstad, K. & Stempniewicz, L. 2016. Foraging behavior of a high-Arctic zooplanktivorous alcid, the little auk, at the southern edge of its breeding range. *Journal of Experimental Marine Biology and Ecology*. 475: 89–99
- Klaassen, R.H.G., Ens, B.J., Shamoun-Baranes, J., Exo, K.-M. & Bairlein, F. 2012. Migration strategy of a flight generalist, the Lesser Black-backed Gull *Larus fuscus*. *Behaviour Ecology*. 23: 58–68.
- Lascelles, B.G., Taylor, P., Miller, M., Dias, M., Oppel, S., Torres, L., Hedd, A., Le Corre, M., Phillips, R. & Shaffer, S.A. 2016. Applying global criteria to tracking data to define important areas for marine conservation. *Diversity and Distributions*, 22: 422–431.
- Lisovski, S. & Hahn, S. 2012. GeoLight—processing and analysing light-based geolocator data in R. *Methods in Ecology and Evolution*. 3: 1055–1059.
- Lisovski, S., Hewson, C. M., Klaassen, R.H.G., Korner-Nievergelt, F., Kristensen, M.W. & Hahn, S. 2012. Geolocation by light: accuracy and precision affected by environmental factors. *Methods in Ecology and Evolution*. 3: 603-612
- Lisovski, S., Bauer, S., Briedis, M., Davidson, S.C., Dhanjal-Adams, K.L., Hallworth, M.T., Karagicheva, J., Meier, C.M., Merkel, B., Ouwehand, J., Pedersen, L., Rakhimberdiev, E., Roberto-Charron A., Seavy N.E., Sumner M.D., Taylor, C.M., Wotherspoon, M.D. & Bridge, E.S. 2019. Light-level geolocator analyses: A user's guide. *Journal of Animal Ecology* 89: 221-236.

Merkel, B., Descamps, S., Yoccoz, N.G., Grémillet, D., Daunt, F., Erikstad, K.E., Ezhov, A.V., Harris, M.P., Gavrilov, M., Lorentsen, S.H., Reiertsen, T.K., Steen, H., Systad, G.H., Þórarinnsson, Þ.L., Wanless, S. & Strøm, H. 2020. Individual migration strategy fidelity but no habitat specialization in two congeneric seabirds. *Journal of Biogeography*, 48: 263-275.

Murphy, R., Perry, E., Keisman, J., Harcum, J. & Leppo, E.W. 2020. baytrends: Long Term Water Quality Trend Analysis. R package version 1.2.1. <https://CRAN.R-project.org/package=baytrends>

Pennycuik, C. 1997. Actual and 'optimum' flight speeds: field data reassessed. *Journal of Experimental Biology*. 200: 2355–2361.

Pennycuik, C.J., Fuller, M.R., Oar, J.J. & Kirkpatrick, S.J. 1994. Falcon versus grouse: flight adaptations of a predator and its prey. *Journal of Avian Biology*. 25: 39–49.

Phillips, R.A., Silk, J.R.D., Croxall, J.P., Afanasyev, V. & Briggs, D.R. 2004. Accuracy of geolocation estimates for flying seabirds. *Marine Ecology Progress Series*, 266: 265–272.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Stoffel, M.A., Nakagawa, S. & Schielzeth, H. 2017. rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution*. 8: 1639-1644.

Strøm, H., Bråthen, V.S., Descamps, S., Ekker, M., Fauchald, P., Helgason, H.H., Merkel, B., Moe, B., Tarroux, A. & Åstrøm, J. 2021. Large-scale tracking of seabirds in the Northeast-Atlantic 2014-2018. Unpublished manuscript. Final report from SEATRACK phase I. Norsk Polarinstitutt. Rapportserie.

Thieurmél, B. & Elmarhraoui, A. 2019. suncalc: Compute Sun Position, Sunlight Phases, Moon Position and Lunar Phase. R package version 0.5.0. <https://CRAN.R-project.org/package=suncalc>

van Bemmelen, R.S.A., Moe, B., Hanssen, S.A., Schmidt, N.M., Hansen, J., Lang, J., Sittler, B., Bollache, L., Tulp, I., Klaassen, R. & Gilg O. 2017. Flexibility in otherwise consistent non-breeding movements of a long-distance migratory seabird, the long-tail skua. *Marine Ecology Progress Series* 578: 197-211

van Bemmelen, R.S.A., Kolbeinsson, Y., Ramos, R., Gilg, O., Alves, J., Smith, M., Schekkerman, H., Lehikoinen, A., Petersen, I. K., Thorisson, B., Sokolov, A.A., Välimäki, K., van der Meer, T., Okill, J.D., Bolton, M., Moe, B., Hanssen, S.A., Bollache, L., Petersen, A., Thorstensen, S., González-Solís, J., Klaassen, R.H.G. & Tulp, I. 2019. A Migratory Divide Among Red-Necked Phalaropes in the Western Palearctic Reveals Contrasting Migration and Wintering Movement Strategies. *Frontiers in Ecology and Evolution*. 7: 86.

Watanabe, Y.Y. 2016. Flight mode affects allometry of migration range in birds. *Ecology letters* 19: 907–914.

Wilson, R.P., Ducamp, J.-J., Rees, W.G., Culik, B.M., & Niekamp, K. 1992. Estimation of location: Global coverage using light intensity. In: Priede, I.G. & Swift, S.M. eds. Remote monitoring and tracking of animals. Ellis Horwood, England, UK. Pp. 131– 134.

## 7 Appendix

**Table 7.1.** The table present differences in the positional data after using the same light-level data to estimate positions, first by SEATRACK's manual procedure (Strøm et al. 2021), and then with the new automatic procedure (this report). We compare 50% kernel HR size (km<sup>2</sup>), latitudinal mean position, longitudinal mean position and number of positions per individual in November. The column « $\Delta$  from man» show how a track produced by the automated procedure is different from a track produced with the manual procedure, along with results from paired t-tests.

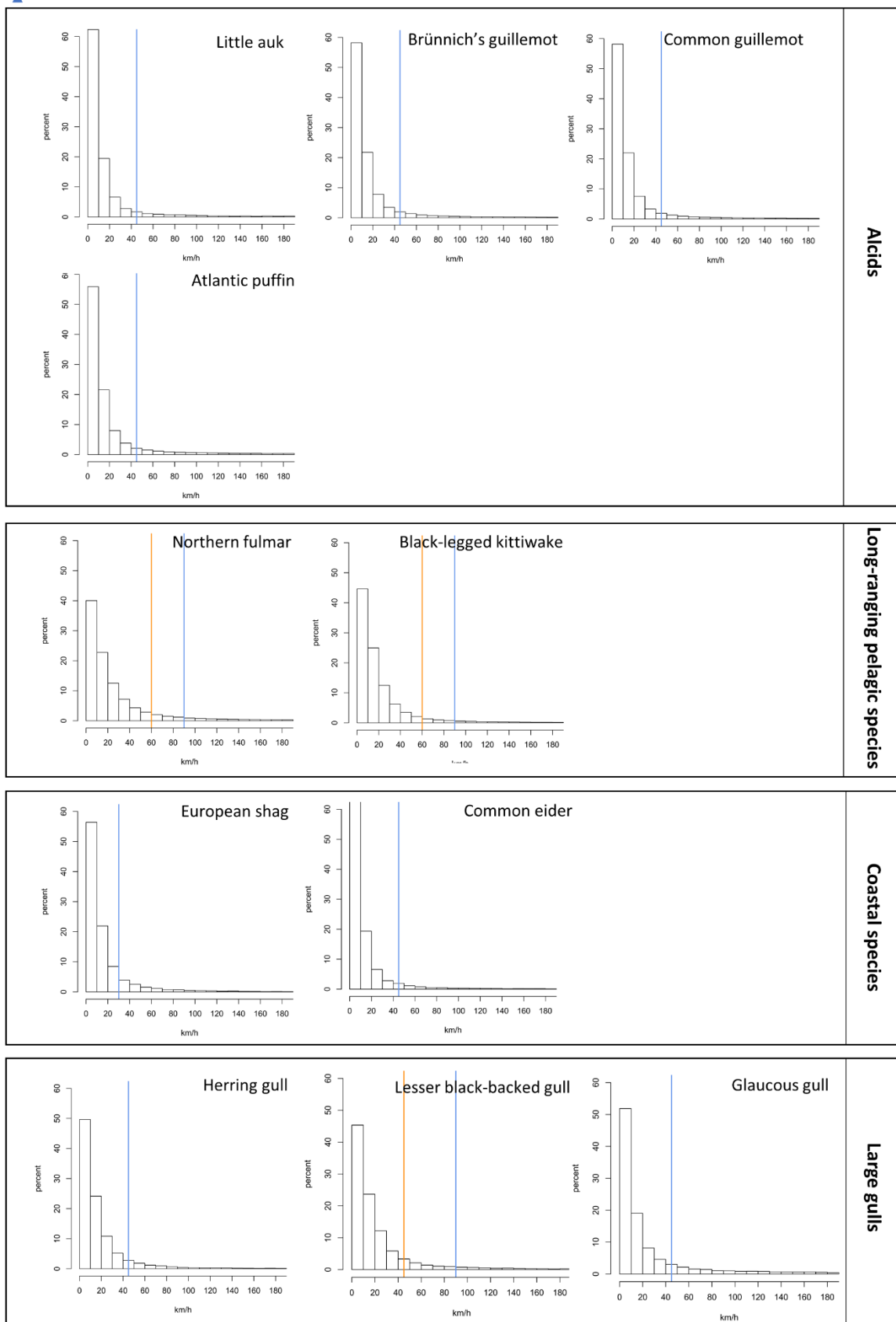
Species	Variable		$\Delta$ from man	paired t-test	period
Atlantic puffin	area KHR 50%		-0.61 %	t = 0.336, df = 360, p-value = 0.737	November
Black-legged kittiwake	area KHR 50%		+7.23 %	t = -4.924, df = 515, p-value < <b>0.001</b>	November
Brünnich's guillemot	area KHR 50%		+1.86 %	t = -1.470, df = 364, p-value = 0.143	November
Common eider	area KHR 50%		+1.67 %	t = -1.094, df = 310, p-value = 0.275	November
Common guillemot	area KHR 50%		+2.87 %	t = -4.015, df = 320, p-value < <b>0.001</b>	November
European shag	area KHR 50%		+17.66 %	t = -6.815, df = 189, p-value < <b>0.001</b>	November
Herring gull	area KHR 50%		+3.48 %	t = -0.540, df = 36, p-value = 0.593	November
Glaucous gull	area KHR 50%		+4.69 %	t = -0.758, df = 58, p-value = 0.452	November
Northern fulmar	area KHR 50%		+3.81 %	t = -1.422, df = 253, p-value = 0.156	November
Lesser black-backed gull	area KHR 50%		-0.81 %	t = 0.245, df = 16, p-value = 0.810	November
Little auk	area KHR 50%		+3.55 %	t = -2.187, df = 91, p-value = <b>0.031</b>	November
Atlantic puffin	latitude	mean	+0.101°	t = 6.752, df = 360, p-value < <b>0.001</b>	November
Black-legged kittiwake	latitude	mean	+0.034°	t = 1.351, df = 515, p-value = 0.177	November
Brünnich's guillemot	latitude	mean	+0.083°	t = 3.695, df = 368, p-value < <b>0.001</b>	November
Common eider	latitude	mean	-0.008°	t = -0.912, df = 314, p-value = 0.362	November
Common guillemot	latitude	mean	+0.064°	t = 6.745, df = 321, p-value < <b>0.001</b>	November
European shag	latitude	mean	+0.305°	t = 10.012, df = 289, p-value < <b>0.001</b>	November
Herring gull	latitude	mean	-0.006°	t = -0.111, df = 36, p-value = 0.913	November
Glaucous gull	latitude	mean	+0.167°	t = 2.737, df = 60, p-value = <b>0.008</b>	November
Northern fulmar	latitude	mean	+0.071°	t = 1.794, df = 253, p-value = 0.074	November
Lesser black-backed gull	latitude	mean	-0.899°	t = -1.886, df = 16, p-value = 0.078	November
Little auk	latitude	mean	0.060°	t = 3.021, df = 97, p-value = <b>0.003</b>	November
Atlantic puffin	longitude	mean	+0.027°	t = 0.690, df = 360, p-value = 0.491	November
Black-legged kittiwake	longitude	mean	+0.559°	t = 7.471, df = 515, p-value < <b>0.001</b>	November
Brünnich's guillemot	longitude	mean	+0.079°	t = 1.516, df = 368, p-value = 0.130	November
Common eider	longitude	mean	+0.057°	t = 3.560, df = 314, p-value < <b>0.001</b>	November
Common guillemot	longitude	mean	+0.159°	t = 7.998, df = 321, p-value < <b>0.001</b>	November
European shag	longitude	mean	+0.166°	t = 6.598, df = 189, p-value < <b>0.001</b>	November
Herring gull	longitude	mean	-0.038°	t = -0.573, df = 36, p-value = 0.571	November
Glaucous gull	longitude	mean	-0.117°	t = -0.545, df = 60, p-value = 0.588	November
Northern fulmar	longitude	mean	+0.141°	t = 1.065, df = 253, p-value = 0.288	November
Lesser black-backed gull	longitude	mean	-0.099°	t = -1.179, df = 16, p-value = 0.256	November
Little auk	longitude	mean	-0.070°	t = -0.554, df = 97, p-value = 0.581	November
Atlantic puffin	N positions	mean	+1.28	t = -4.823, df = 360, p-value < <b>0.001</b>	November
Black-legged kittiwake	N positions	mean	+1.46	t = -11.133, df = 515, p-value < <b>0.001</b>	November
Brünnich's guillemot	N positions	mean	+1.53	t = -9.278, df = 364, p-value < <b>0.001</b>	November
Common eider	N positions	mean	+0.27	t = -1.507, df = 310, p-value = 0.133	November
Common guillemot	N positions	mean	+1.54	t = -7.758, df = 320, p-value < <b>0.001</b>	November
European shag	N positions	mean	+1.65	t = -2.745, df = 189, p-value = <b>0.007</b>	November
Herring gull	N positions	mean	-1.65	t = 4.530, df = 36, p-value < <b>0.001</b>	November
Glaucous gull	N positions	mean	-0.95	t = 1.078, df = 58, p-value = 0.285	November
Northern fulmar	N positions	mean	+1.34	t = -4.005, df = 253, p-value < <b>0.001</b>	November
Lesser black-backed gull	N positions	mean	+2.29	t = -1.220, df = 16, p-value = 0.240	November
Little auk	N positions	mean	+2.20	t = -5.096, df = 91, p-value < <b>0.001</b>	November

**Table 7.2.** The table present differences in the positional data after using the same light-level data to estimate positions, first by SEATRACK's manual procedure (Strøm et al. 2021), and then with the new automatic procedure (this report). We compare 50% kernel HR size (km<sup>2</sup>), latitudinal mean position, longitudinal mean position and number of positions per individual in December. The column « $\Delta$  from man» show how a track produced by the automated procedure is different from a track produced with the manual procedure, along with results from paired t-tests.

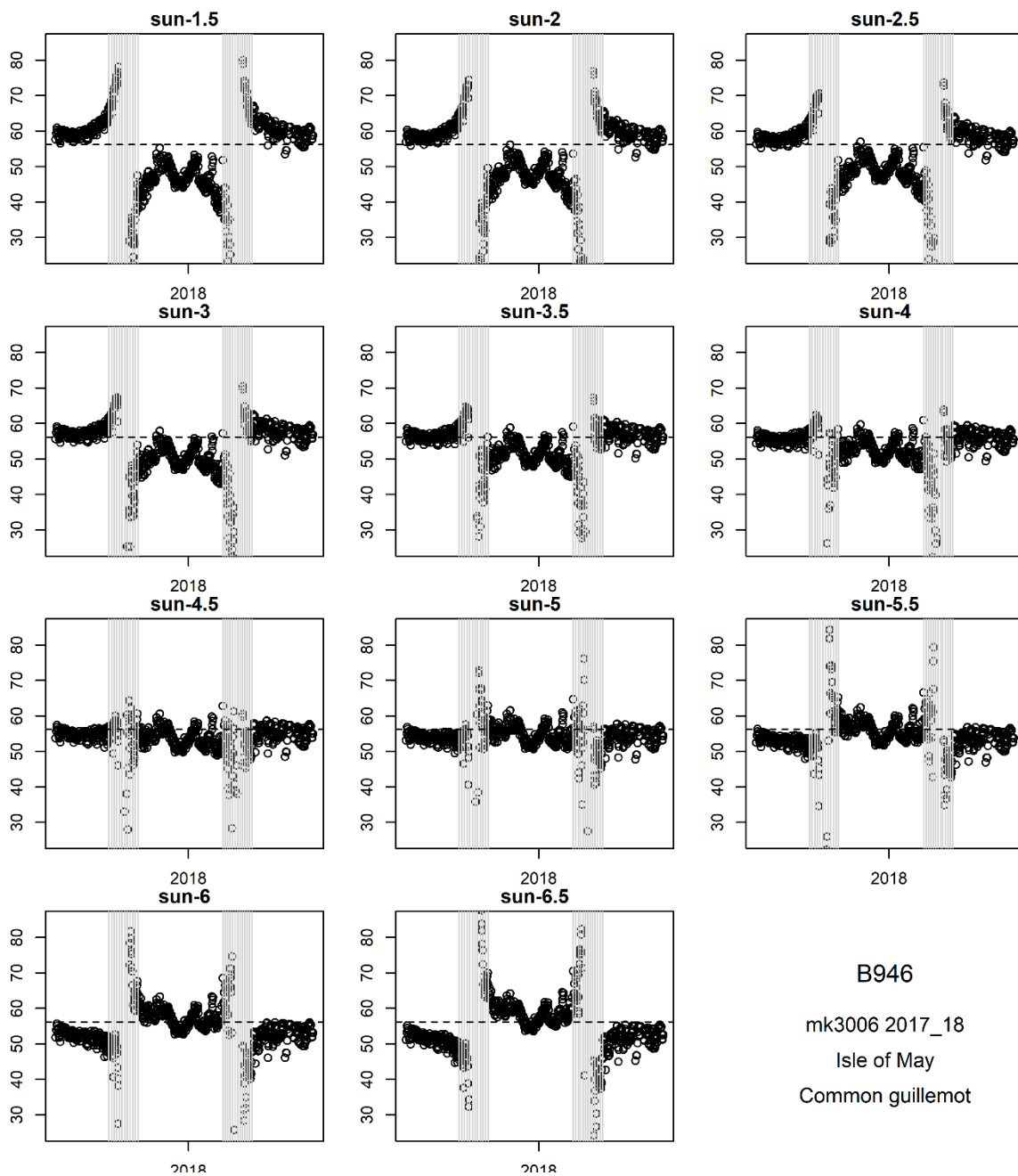
Species	Variable		$\Delta$ from man	paired t-test	period
Atlantic puffin	area KHR 50%		+2.65 %	t = -1.566, df = 284, p-value = 0.118	December
Black-legged kittiwake	area KHR 50%		+1.98 %	t = -2.070, df = 513, p-value = <b>0.039</b>	December
Brünnich's guillemot	area KHR 50%		+10.81 %	t = -5.893, df = 308, p-value < <b>0.001</b>	December
Common eider	area KHR 50%		-0.10 %	t = 0.859, df = 305, p-value = 0.391	December
Common guillemot	area KHR 50%		+9.64 %	t = -4.819, df = 308, p-value < <b>0.001</b>	December
European shag	area KHR 50%		+17.32 %	t = -8.148, df = 189, p-value < <b>0.001</b>	December
Herring gull	area KHR 50%		-0.05 %	t = 0.015, df = 35, p-value = 0.988	December
Glaucous gull	area KHR 50%		+13.71 %	t = -0.668, df = 23, p-value = 0.511	December
Northern fulmar	area KHR 50%		+0.55 %	t = -0.180, df = 196, p-value = 0.857	December
Lesser black-backed gull	area KHR 50%		+11.26 %	t = -1.462, df = 15, p-value = 0.164	December
Little auk	area KHR 50%		+13.67 %	t = -3.546, df = 66, p-value < <b>0.001</b>	December
Atlantic puffin	latitude	mean	+0.113°	t = 7.143, df = 303, p-value < <b>0.001</b>	December
Black-legged kittiwake	latitude	mean	-0.108°	t = -5.013, df = 514, p-value < <b>0.001</b>	December
Brünnich's guillemot	latitude	mean	+0.124°	t = 10.915, df = 317, p-value < <b>0.001</b>	December
Common eider	latitude	mean	-0.006°	t = -0.493, df = 310, p-value = 0.623	December
Common guillemot	latitude	mean	+0.097°	t = 10.282, df = 318, p-value < <b>0.001</b>	December
European shag	latitude	mean	+0.202°	t = 11.255, df = 189, p-value < <b>0.001</b>	December
Herring gull	latitude	mean	+0.035°	t = 1.235, df = 35, p-value = 0.225	December
Glaucous gull	latitude	mean	+0.441°	t = 6.054, df = 29, p-value < <b>0.001</b>	December
Northern fulmar	latitude	mean	+0.308°	t = 5.227, df = 206, p-value < <b>0.001</b>	December
Lesser black-backed gull	latitude	mean	-0.317°	t = -4.166, df = 15, p-value < <b>0.001</b>	December
Little auk	latitude	mean	+0.167°	t = 5.678, df = 67, p-value < <b>0.001</b>	December
Atlantic puffin	longitude	mean	+0.175°	t = 2.277, df = 303, p-value = <b>0.024</b>	December
Black-legged kittiwake	longitude	mean	+0.271°	t = 5.217, df = 514, p-value < <b>0.001</b>	December
Brünnich's guillemot	longitude	mean	+0.211°	t = 2.053, df = 317, p-value = <b>0.041</b>	December
Common eider	longitude	mean	+0.065°	t = 2.864, df = 310, p-value = <b>0.004</b>	December
Common guillemot	longitude	mean	+0.043°	t = 0.877, df = 318, p-value = 0.381	December
European shag	longitude	mean	+0.212°	t = 8.368, df = 189, p-value < <b>0.001</b>	December
Herring gull	longitude	mean	-0.012°	t = -0.266, df = 35, p-value = 0.792	December
Glaucous gull	longitude	mean	+0.884°	t = 1.511, df = 29, p-value = 0.142	December
Northern fulmar	longitude	mean	+0.209°	t = 0.675, df = 206, p-value = 0.500	December
Lesser black-backed gull	longitude	mean	+0.057°	t = 1.656, df = 15, p-value = 0.118	December
Little auk	longitude	mean	+0.233°	t = 2.754, df = 67, p-value = <b>0.008</b>	December
Atlantic puffin	N positions	mean	+2.19	t = -5.621, df = 284, p-value < <b>0.001</b>	December
Black-legged kittiwake	N positions	mean	+0.37	t = -5.363, df = 513, p-value < <b>0.001</b>	December
Brünnich's guillemot	N positions	mean	+5.16	t = -12.321, df = 308, p-value < <b>0.001</b>	December
Common eider	N positions	mean	-0.40	t = 1.731, df = 305, p-value = 0.085	December
Common guillemot	N positions	mean	+6.02	t = -14.083, df = 308, p-value < <b>0.001</b>	December
European shag	N positions	mean	+1.85	t = -5.004, df = 189, p-value < <b>0.001</b>	December
Herring gull	N positions	mean	-0.17	t = 0.354, df = 35, p-value = 0.725	December
Glaucous gull	N positions	mean	+6.75	t = -3.600, df = 23, p-value = <b>0.002</b>	December
Northern fulmar	N positions	mean	+2.17	t = -4.593, df = 196, p-value < <b>0.001</b>	December
Lesser black-backed gull	N positions	mean	+0.25	t = -1.732, df = 15, p-value = 0.104	December
Little auk	N positions	mean	+5.76	t = -5.838, df = 66, p-value < <b>0.001</b>	December

logger_id_retrieved	logger_model_retrieved	year_tracked	logger_producer	ring_number	euring_code	species	colony	col_lat	col_lon	date_deployed	date_retrieved	light_treshold	analyzer	data_responsible	sun1_db
V2016089	mk4083	2014_15	Biotrack	6199024	NOS	Black-legged kittiwake	Anda	69.065	15.17	10/07/2014	15/07/2016	1	Vegard Sandøy Bråthen	Signe Christensen-Dalsgaard	-3.25
V2016089	mk4083	2015_16	Biotrack	6199024	NOS	Black-legged kittiwake	Anda	69.065	15.17	10/07/2014	15/07/2016	1	Vegard Sandøy Bråthen	Signe Christensen-Dalsgaard	-3.75

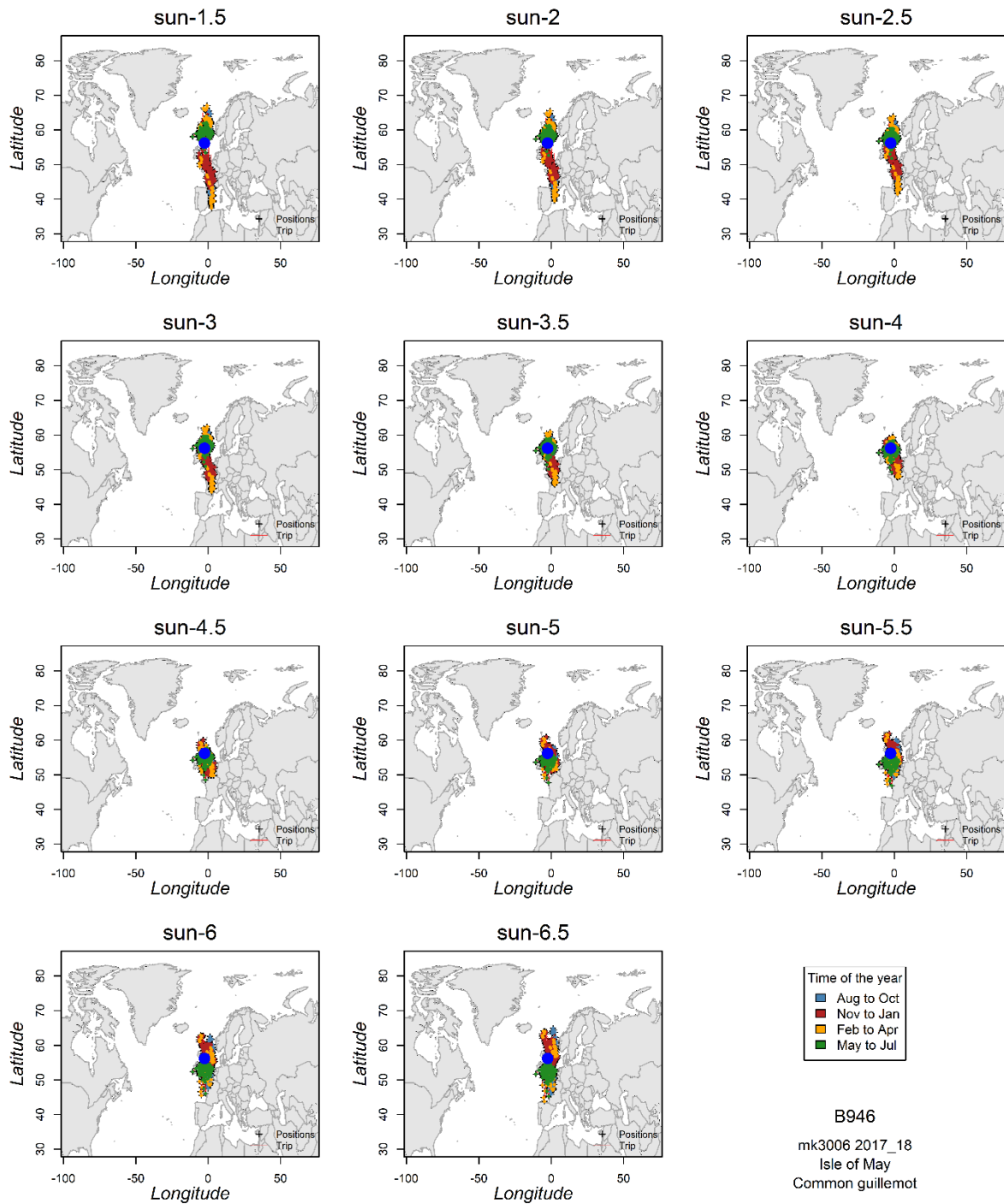
**Table 7.3.** Columns and two rows with example data used as complementary information in the automated processing of light-level data into positions. The tables with complementary information contain information about the individual (bird id), logger (logger id, model, producer), deployment (dates), colony (name, location), persons to contact (analyser, data responsible) and the final calibrated sun elevation angles. The tables consist of one row for each logger and year tracked.



**Figure 7.1.** Proportion (%) of distances covered in km/h. Blue vertical lines are the chosen speed thresholds for each species. Orange vertical lines are the thresholds that were used in the first version of the automated procedure (2018-2019).

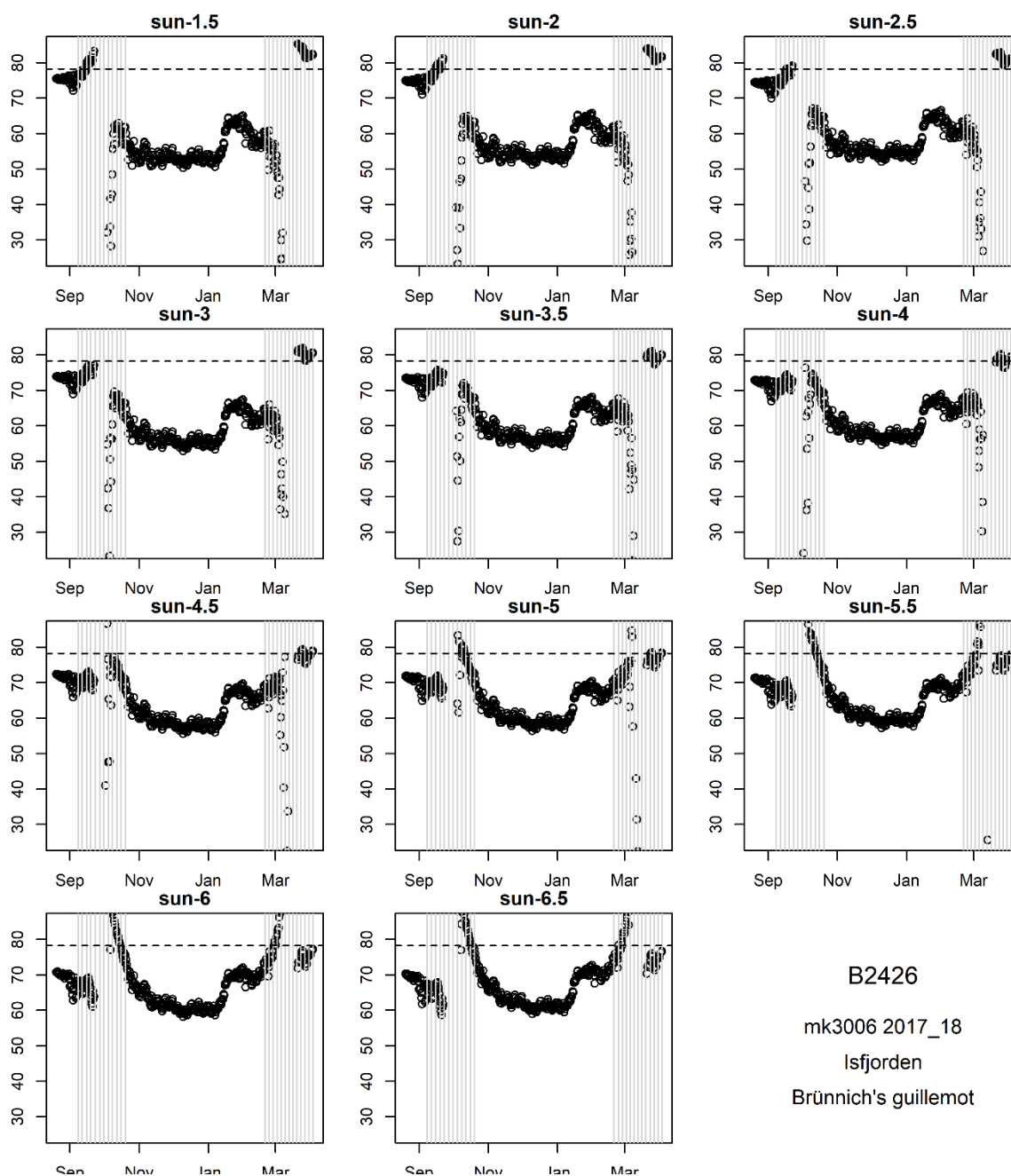


**Figure 7.2.** Latitude versus time plotted for different sun elevation angles, between  $-1.5$  and  $-6.5^{\circ}$ , for a common guillemot tracked from Isle of May. The plots span one non-breeding season from June 2017 to June 2018. The horizontal dotted line shows the latitude of the colony, and the vertical grey lines indicate the periods around the autumn and spring equinoxes. Here, we selected  $-4.5^{\circ}$  as the appropriate sun elevation angle. In this example we used all three criteria and chose the sun elevation angle that 1) minimized the amplification of latitudinal error close to the equinoxes, 2) resulted in matching latitudes at both sides of the equinox, and 3) resulted in positions that fitted the latitude of the colony in the beginning and the end of the track.

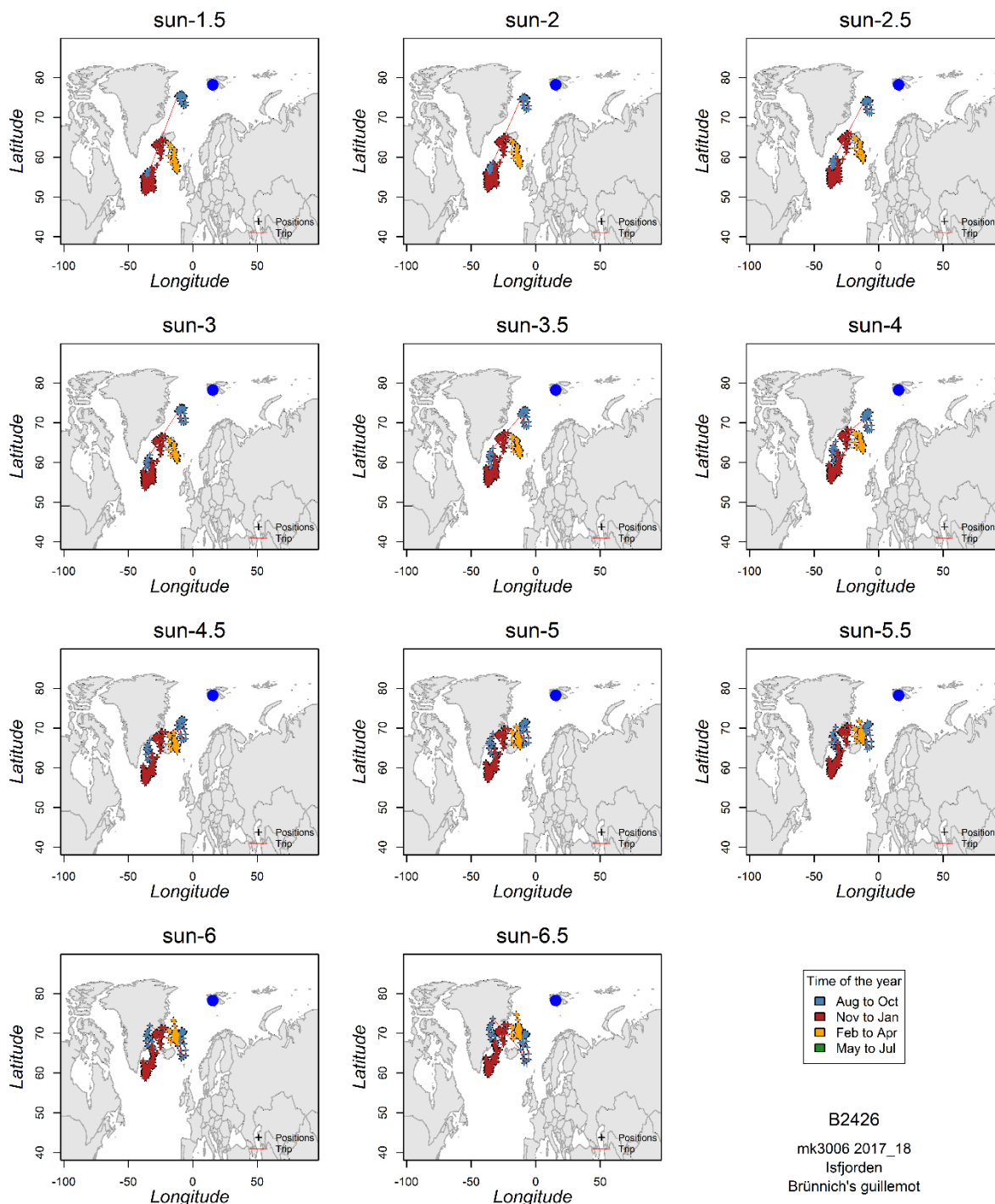


**Figure 7.3.** Smoothed and filtered positions calculated with different sun elevation angles for a common guillemot tracked from Isle of May in 2017/18. These maps supported the selection of  $-4.5^\circ$  as sun elevation angle. The location of the colony is marked with a filled blue symbol, and map limits is limited to the species-specific distribution range from the Distribution filter. Positions from the equinox periods has been excluded from the map. Positions are coloured by month of the year





**Figure 7.4.** Latitude versus time plotted for different sun elevation angles for a Brünnich's guillemot tracked from Isfjorden. Time spans one non-breeding season from first realistic twilight at the end of August 2017 to last one in early April 2018. The horizontal dotted line shows the latitude of the colony, and the vertical grey lines indicate the periods around the autumn and spring equinoxes. In this example, we selected  $-4.0^\circ$  as the appropriate sun elevation angle. We mainly used criteria 1) and 2). It was less useful to use criteria 3 since the bird was not at the colony at the beginning of the track. The colony is located far north and there is still apparent constant daylight at the time of departure from the colony. First latitude appears south of the colony. The latitudes at the end of the track may indicate that it was at the colony in the beginning of April, but this is inside the equinox period and should be given low weight.



**Figure 7.5.** Smoothed and filtered positions calculated with different sun elevation angles for a Brünnich's guillemot tracked from Isfjorden in 2017/18. These maps supported the selection of  $-4.0^\circ$  as sun elevation angle. The location of the colony is marked with a filled blue symbol, and map limits is limited to the species-specific distribution range from the Distribution filter. Positions from the equinox periods has been excluded from the map. Positions are coloured by month of the year

# Readme

- use of positional data from SEATRACK (IRMA positions are not included)

The dataset is usually tab separated.

We recommend to plot the positions with 'lon\_smooth2' and 'lat\_smooth2' and positions where 'eqFilter3' = 1. By doing so, you are using positions that are smoothed twice and filter out periods where latitudes are affected by equinox.

#####

A brief description of each column in the positional data file:

'date\_time' = date and time of the calculated position in GMT (DD.MM.YYYY HH:MM)  
 'logger' = specific logger ID and logger model ( e.g N032\_c65)  
 'logger\_id' = specific logger ID for each GLS logger (e.g N032).  
 'logger\_model' = GLS logger model (e.g c65, mk4083 etc)  
 'year\_tracked' = 'year tracked' specify the tracking period from one breeding season to the next (e.g. '2013\_14').  
 'session\_id' = This value will always be unique for each time a logger was deployed  
 'individ\_id' = This value will always be unique for each individual  
 'year\_deployed' = year logger was deployed (YYYY)  
 'year\_retrieved' = year logger was retrieved (YYYY)  
 'ring\_number' = id-number on the metal ring (e.g CA443314)  
 'euring\_code' = what bird ringing scheme the metal ring belongs to (e.g. 'NOS' for Norway/Stavanger museum)  
 'species' = english name of the species (e.g 'Black-legged kittiwake')  
 'colony' = name of the colony the logger was deployed and retrieved in.  
 Colony name can often cover multiple breeding locations when in close proximity.  
 'lon\_raw' = longitude calculated from time (GMT) of noon or midnight and date.  
 'lat\_raw' = latitude calculated from the length of day or night and date.  
 'lon\_smooth1' = averaging consecutive noon-midnight or midnight-noon longitudes from 'lon\_raw'.  
 'lat\_smooth1' = averaging consecutive noon-midnight or midnight-noon latitudes from 'lat\_raw'.  
 'lon\_smooth2' and 'lat\_smooth2' = two point moving average over the lon\_smooth1-lat\_smooth1 coordinates using spherical trigonometry. We recommend using these double smoothed positions for plotting and interpretations.  
 'disttocol\_s2' = distances from double smoothed positions (lon\_smooth2, lat\_smooth2) to the breeding colony calculated as great circle distances assuming a spherical earth.  
 'eqfilter3' = this filter identifies fixed periods where the latitudes are assumed to be unreliable during equinox (1 = reliable/outside fixed equinox period, 0 = unreliable/inside equinox). Value 0 represent the periods from 8 Sep to 20 Oct and 22 Feb to 3 Apr.  
 'lat\_smooth2\_eqfilt3' = contain data from 'lat\_smooth2', when 'eqfilter3'= 1.  
 'sex' = sex of the bird  
 'morph' = morph of the bird if morph has been assigned.  
 'subspecies' = subspecies of the bird if assigned  
 'age' = identifies the bird as adult or juvenile, or the specific age of the bird if known.  
 'col\_lon' = longitude of the colony.  
 'col\_lat' = latitude of the colony.  
 'tfirst' = time of first sun event (DD.MM.YYYY HH:MM) for a pair of sun events (sunrise-sunset or sunset-sunrise) used for calculating positions  
 'tsecond' = time of second sun event (DD.MM.YYYY HH:MM) for a pair of sun events (sunrise-sunset or sunset-sunrise) used for calculating positions  
 'twl\_type' = identify whether 'tfirst' is sunrise (1) or sunset (2)  
 'conf' = if manual inspection of twilight events: 9 repr a high confidence, 2 low conf, 4 high conf but moved after manual inspection, 3 low conf still after moving it. If automatically processed: always 9.  
 'sun' = the sun's angle to the horizon used to calculate the positions from twilight events.  
 'software' = identifies the script version (e.g.'DEC\_2019') when automated procedure has been used to identify twilights, or the name of the software used to identify twilights (e.g. Transedit, Intiproc) if the manual procedure has been used to identify twilights  
 'light\_threshold' = selected threshold that defines a sunset/sunrise  
 'analyzer' = name of person inspecting the light data and identifying the twilights (if manual procedure) and performing the sun elevation angle calibration  
 'data\_responsible' = name of data responsible  
 'logger\_yeartracked' = 'logger\_id' +'logger\_model' + 'year\_tracked'(e.g N032\_c65\_2014\_15) This value will always be unique for each year tracked.  
 'posdata\_file' = identifies the original datafile from where the tracking data was imported into the database

```
#####
## Example of plotting positions and kernels (50% homerange) with R ##

#Load packages
library(GeoLight) #install.packages("GeoLight")
library(lubridate) #install.packages("lubridate")
library(sp) #install.packages("sp")
library(adehabitatHR) #install.packages("adehabitatHR")
library(maps)# install.packages("maps")

#Load dataset
posdata <- read.csv("https://www.dropbox.com/s/kiftqlatmvgsz0s/plotting_posdata_example.csv?dl=1",
row.names=NULL)
posdata$colony <- gsub("Ã", "ø", posdata$colony)

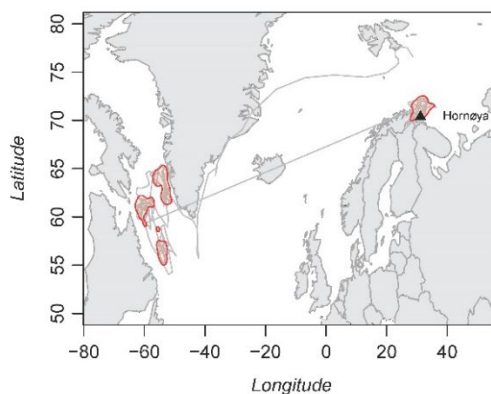
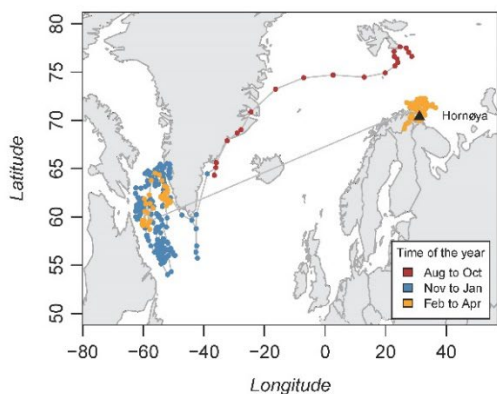
par(mfrow=c(1,2))

#remove equinox periods:
example<- posdata[posdata$eqfilter3==1,]

#Map with positions
tripMap(cbind(0,0), equinox=F,xlim=c(-75,52), ylim=c(50,80),col="grey",legend=F, cex=0.2,add=F)
lines(example$lon_smooth2, example$lat_smooth2,col="grey") # Plot our coordinates on it.
points(example$lon_smooth2[month(example$date_time)%in%c(8,9,10)],example$lat_smooth2[month(exam-
ple$date_time)%in%c(8,9,10)],cex=0.6,pch=16,col="firebrick")
points(example$lon_smooth2[month(example$date_time)%in%c(11,12,1)],example$lat_smooth2[month(exam-
ple$date_time)%in%c(11,12,1)],cex=0.6,pch=16,col="steelblue")
points(example$lon_smooth2[month(example$date_time)%in%c(2,3,4)],example$lat_smooth2[month(exam-
ple$date_time)%in%c(2,3,4)],cex=0.6,pch=16,col="orange")
#points(example$lon_smooth2[month(example$date_time)%in%c(5,6,7)],example$lat_smooth2[month(exam-
ple$date_time)%in%c(5,6,7)],cex=0.6,pch=16,col="forestgreen")
points(example$col_lon[1],example$col_lat[1],cex=1,pch=17)
text(example$col_lon[1]+15,example$col_lat[1], example$colony,cex=0.7)
legend("bottomright", inset=.02, title="Time of the year",
c("Aug to Oct","Nov to Jan","Feb to Apr"), fill=c("firebrick","steelblue","orange"), horiz=FALSE,
cex=0.7)

#Map with 50% kernels:
proj.azeqd <- "+proj=aeqd +lat_0=0 +lon_0=-10 +units=km"
proj.cc <- '+init=epsg:4326'
test.grid <- GridTopology(cellcentre.offset=c(-10000,-
10000),cellsize=c(20,20),cells.dim=c(1000,1500))
test.point <- SpatialPoints(cbind(c(0),c(0)))
test.pixel <- SpatialPixels(test.point,proj4string = CRS(proj.azeqd), round = NULL, grid =
test.grid)
cg <- example[!is.na(example$lon_smooth2) & !is.na(example$lat_smooth2_eqfilt3),]
coordinates(cg) <- cbind(cg$lon_smooth2,cg$lat_smooth2_eqfilt3)
proj4string(cg) <- CRS(proj.cc)
cg1 <- spTransform(cg,CRS(proj.azeqd))
kus <- kernelUD(cg1[,38],grid=test.pixel,h=44)
c50 <- getverticeshr(kus,50)
wgs<"+proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0"
c50 <- spTransform(c50, wgs) #project to wgs

tripMap(cbind(0,0), equinox=F,xlim=c(-75,52), ylim=c(50,80),col="grey",legend=F, cex=0.2,add=F)
color.transparent <- adjustcolor("red", alpha.f = 0.10)
lines(example$lon_smooth2, example$lat_smooth2,col="grey") # Plot our coordinates on it.
plot(c50,col=color.transparent, lwd=1,add=T, border="red") # Plot our coordinates on it.
points(example$col_lon[1],example$col_lat[1],cex=1,pch=17)
text(example$col_lon[1]+15,example$col_lat[1], example$colony,cex=0.7)
#example plot:
```









*The Norwegian Institute for Nature Research, NINA, is 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: 978-82-426-4666-8

## 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>



Cooperation and expertise for a sustainable future