foieGras an R package for animal movement data: rapid quality control, behavioural estimation and simulation

Ian D. Jonsen[1,*], W. James Grecian[2], Lachlan Phillips[1], Gemma Carroll[3], Clive McMahon[4], Robert G. Harcourt[1], Mark A. Hindell[5], and Toby A. Patterson[6]

[1]School of Natural Sciences, Macquarie University, Sydney, NSW, Australia

[2]Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife, United Kingdom

[3]Environmental Defense Fund, Seattle, WA, United States

[4]Sydney Institute of Marine Science, Mosman, NSW, Australia

[5]Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS, Australia

[6]CSIRO Ocean and Atmosphere Research, Hobart, TAS, Australia

[*]corresponding author, ian.jonsen@mq.edu.au

Abstract

1.

2.

3.

4.

Keywords:

1 | Introduction

The use of animal-borne electronic sensors has matured, with tracking data now virtually essential for understanding behaviour and social interactions, foraging ecology, physiology, habitat use and population dynamics of mobile and/or cryptic species. The sophistication of current technology also enables the use of animal-borne sensors as a cost-effective approach for observing our planet that compliments more traditional observing platforms (Harcourt et al., 2019; Kays et al., 2015; McMahon et al., 2021). In all these applications, data from electronic tracking devices require rigourous quality control procedures to account for common, though not universally present, issues such as irregularly timed measurements, sensor biases and location measurement error. Some of these issues may be handled by a manufacturer's on-board or post-processing algorithms and some must be addressed by researchers using the data.

State-space models (SSMs) and hidden Markov models (HMMs) are powerful tools for conducting quality control of and making behavioural inference from animal tracking data (Jonsen et al., 2013; Patterson et al., 2008). These are time-series models used across a wide range of research disciplines that estimate the state of an unobserved process from an observed data set. Here, we view an animal's true location and/or behaviour as the unobserved state(s), though many other types of states are possible (Hooten et al., 2019; e.g., Schick et al., 2013), and measurements recorded by electronic devices provide the observations. In practical yet simplistic terms, SSMs are usually preferred when the goal is to quality control error-prone location data and/or make inference directly from the parameters of their underlying movement models. HMMs are usually preferred when measurements have negligible error and occur at regular time intervals (but see McClintock & Michelot, 2018), and when the goal is to infer behavioural states hidden within the data along with their potential external or internal drivers. Other more technical distinctions and reasons for preferring one of these methods exist (Jonsen et al., 2013; Patterson et al., 2017). Our primary focus here is on SSMs as tools for quality control of error-prone location data and for making fast inference of behavioural changes along animal tracks.

A number of R packages such as `moveHMM` (Michelot et al., 2016), `momentuHMM` (McClintock & Michelot, 2018), and `swim` (Whoriskey et al., 2017) provide highly accessible and flexible tools for fitting HMMs to animal tracking data, and facilitating general inference of animal movement behaviour and its drivers. Similarly, R packages such as `bsam` (I. Jonsen et al., 2005), `crawl` (Johnson et al., 2008), `argosTrack` (Albertsen et al., 2015), `ctmm` (Calabrese et al., 2016), and `yaps` (Baktoft et al., 2017) all provide tools for fitting movement process models in either discrete- or continuous-time, ranging from simple random walks to Ornstein-Uhlenbeck processes, in state-space form to various types of tracking data.

Here, we introduce the package `foieGras`, pronounced "*fwah grah*", developed for R (R Core Team, 2021). This package was developed with two aims: (1) to be a simple and fast implementation of SSMs for quality control of error-prone animal (re)location data (including via Argos satellite, Service Argos (2016);Jonsen:2020); and (2) for inference of changes in behaviour along animal movement tracks (Jonsen et al., 2019). The simplicity of use sets `foieGras` apart from many of the related SSM R packages listed above, yet users can exert control over many aspects of the package functions via optional arguments. This design accommodates both novice and experienced users.

Here, we describe the main features of `foieGras` and illustrate its use through a set of applications drawing on Argos and GPS tracking data. Full R code and data for each of the applications is provided in the Supporting Information. Additional details on package functions and their use can be found in their help files and in the package vignettes.

2 | `foieGras` overview

The workflow for `foieGras` is deliberately simple, with many of the usual track data processing checks and formatting handled automatically. Here we outline the main aspects of the

Table 1: Main functions for the R package `foieGras`

Function	Description
`fit_ssm`	Fit a State-Space Model to location data
`fit_mpm`	Fit a Move Persistence Model to location data
`grab`	Extract fitted/predicted/observed locations from a foieGras model, with or without projection information
`osar`	Estimate One-Step-Ahead Residuals from a foieGras SSM
`map`	Map fitted/predicted locations with or without a defined projection
`sim`	Simulate individual animal tracks with Argos LS or KF errors
`simfit`	Simulate animal tracks from 'ssm_df' fit objects
`sim_filter`	Filter tracks simulated with 'simfit' according to similarity criteria
`route_path`	Reroute path so estimated locations are off land
`plot.ssm_df`	Plot the fit of a foieGras SSM to data
`plot.osar`	Plot One-Step-Ahead Residuals from a foieGras SSM

`foieGras` package. The packages's main functions are listed in Table 1.

2.1 | Data preparation

Animal tracking data, consisting of a time-series of location coordinates, can be read into R as a data frame using standard functions such as `read.csv`. The canonical data format for Argos tracks consists of a data frame with 5 columns corresponding to the following named variables: `id` (individual id), `date` (date and time), `lc` (location class), `lon` (longitude), `lat` (latitude). Optionally, an additional 3 columns, `smaj` (semi-major axis), `smin` (semi-minor axis), `eor` (ellipse orientation), providing Argos error ellipse information may be included.

Other types of location data can be accommodated, for example, by including the `lc` column where all `lc = "G"` for GPS data. In this case, measurement error in the GPS locations is assumed to have a standard deviation of 0.1 x Argos class 3 locations (approximately 30 m). If location standard errors exist, these can be added by appending the columns `lonerr` and `laterr` to the data (see the Overview vignette for further details).

2.2 | State-space model fitting - `fit_ssm`

State-space models are fit using `fit_ssm`. When fitting a SSM to location data, the type of data is automatically detected from the location quality class designations that are typical of Argos data and that can be added to the data by the researcher for other types of location data. Based on the location quality classes and optional information on measurement errors contained in the data, an appropriate measurement error model is selected for each observation (Jonsen et al., 2020). This capability can allow different tracking data types, such as Argos and GPS, to be combined in a single input data frame and to be fit in a single state-space model.

There are a large number of arguments that can be set in `fit_ssm`, and these are explained in the documentation. We focus only the essential arguments here:

- `data` the input data structured as described in **2.1**
- `model` the process model to be used
- `time.step` the prediction time interval (h)

The function first invokes an automated data processing stage where the following occurs: 1) data type (Argos Least-Squares, Argos Kalman Filter/Smoother, GPS, or General (e.g., processed light-level geolocations, acoustic telemetry, coded VHF telemetry) is determined; 2) date-times are converted to POSIXt format, chronological order is ensured, and duplicate date-time records are removed; 3) observations occurring less than `min.dt` seconds after a prior observation are removed (default: 60 s); 4) a speed filter [`sda` from the `trip` R package; Sumner et al. (2009)] is used to identify potential extreme locations to be ignored by the SSM; 5) locations are projected from spherical lon,lat coordinates to global Mercator x,y coordinates in km.

The function then fits a state-space model to the processed data, where the process model (one of `rw`, `crw`, or `mp`) is specified by the user via the `model` argument, and the measurement model(s) are selected automatically. The model is fit by numerical optimization of the likelihood using either of the standard R optimizers, `optim` or `nlminb`. The R package `TMB`, Template Model Builder (Kristensen et al., 2016), is used to compute the gradient function in C++ via reverse-mode auto-differentiation and the Laplace Approximation is used to integrate out the latent states (random effects). Fits to a single versus multiple individuals are handled automatically, with sequential SSM fits occurring in the latter case. No hierarchical or pooled estimation among individuals is currently available.

`fit_ssm` returns a `ssm_df` fit object (a nested data frame with class `ssm_df`). The outer data frame lists the individual id(s), basic convergence information and a list with class `ssm`. This list contains dense information on the estimated parameters and states, predictions, processed data, optimizer results, and other diagnostic and contextual information. Users can extract a simple data frame of SSM fitted (location estimates corresponding to the observation times) or predicted values (locations predicted at regular `time.step` intervals) using the `grab` function. Parameter estimates, AIC and other model fit information can be viewed in tabular form using the `summary` function.

2.3 | Behavioural estimation - `fit_ssm`, `fit_mpm`

Move persistence, an index of along-track movement behaviour, can be estimated as a continuous-valued (0 - 1), time-varying latent variable that represents changes in movement pattern based on autocorrelation in speed and direction Jonsen et al. (2019). There are two approaches in `foieGras` for estimating move persistence. The first is to use `fit_ssm` with `model = 'mp'`, which fits a continuous-time move persistence model in state-space form and thereby simultaneously estimates true locations and move persistence from the error-prone telemetry data. This approach is most appropriate for fitting to irregularly-timed and error-prone Argos data as both aspects are taken into account explicitly. The second is to use `fit_mpm`, which can take as input either location data or SSM-estimated locations from an ssm_df fit object. This approach is generally more appropriate when the data have minimal measurement error (e.g., GPS locations), or when time-regularization is desired. We illustrate both approaches in Application 3.2.

2.4 | Model checking and visualization - osar, plot, map

Before using fitted or predicted locations, a fit_ssm model fit should be checked and visualized to confirm that the model adequately describes the data. There is no simple way to calculate residuals for latent variable models that have non-finite state-spaces and that may be nonlinear, but they can be computed based on iterative forecasts of the model (Thygesen et al., 2017). The osar function computes one-step-ahead (prediction) residuals via the oneStepPredict function from the TMB R package to make this as efficient as possible. A set of residuals are calculated for the x and y values corresponding to the fitted values from the SSM and returned as an osar object. A generic plot (plot.osar) method provides an easy way to visualize the osar residuals as time-series plots, quantile-quantile plots, or autocorrelation functions.

State-space model fits to data can also be visualised by using the generic plot (plot.ssm_df) function on an ssm_df fit object. Options exist to plot fitted or predicted values along with observations as either paired, 1-D time-series, as 2-D tracks with 95% confidence intervals or ellipses, by using the argument type = 1 and type = 2, respectively. These plots provide a rapid check on SSM fits to data. Additionally, when the fitted SSM is the move persistence model (i.e., model = 'mp'), 1-D time-series (type = 3) or 2-D track plots (type = 4) of move persistence can be viewed.

Additionally, a fit_ssm model fit can be mapped using the map function for single or multiple individuals. By default, map uses the coastline data from the rnaturalearth R package (South, 2022a) at medium or high resolution, if the rnaturalearthhires (South, 2022b) R package is installed, but can also use tiled maps for finer-scale detail, if the rosm (Dunnington, 2019) and ggspatial (Dunnington, 2021) R packages are installed. Mapping aesthetics (e.g., plot symbols, sizes, colours, fills) can be customized via the aes argument and use of the aes_lst function. See code in SI for examples.

All foieGras visualizations draw on the ggplot2 R package (Wickham, 2016), with multi-panel plots also using the patchwork R package (Pedersen, 2020), and generally can be modified through additive calls in the usual ggplot2 manner. See code in SI for examples.

2.5 | Simulation - sim, simfit, sim_filter

Track simulation can be a helpful, yet informal, way of evaluating the degree to which statistical movement models capture essential features of animal movement data (Michelot et al., 2017). The sim function can simulate a variety of movement process, including the rw, crw, and mp process models, as well as simple multiple movement state switching processes. Simulations from different process models can be used to evaluate the robustness of SSM location and move persistence estimates to model misspecification. An example is provided in the SI.

Simulation is also used frequently in habitat usage modelling to provide a measure of habitat availability (Aarts et al., 2012) by generating a source of 'background' points representing a null model of the distribution of foraging animals in the absence of external drivers (Hindell et al., 2020; S. J. Phillips et al., 2009; Raymond et al., 2015). The simfit function extracts movement parameters from a ssm_df fit object and simulates a user defined number of random tracks of the same duration from these parameters. The argument cpf = TRUE allows the user to simulate central place foragers by ensuring that the simulated tracks start and end at approximately the same location. It is also possible to constrain movements to remain mostly in water via a potential function (Preisler et al., 2013), using included gradient rasters and the grad and beta arguments. These are illustrated in the code for Application 3.3.

The choice of null points can have a large impact on the performance of habitat suitability models (Lobo et al., 2010; S. J. Phillips et al., 2009), and so the sim_filter function provides a tool to filter the simulated tracks based on their similarity to the original path. The filtering is based on one of two metrics that capture the difference in the net displacement and bearing between the two paths (see `similarity_flag` for more detail). This metric is motivated by the 'flag value' described in Hazen et al. (2017). The user can also specify the quantile of flag values to retain; i.e. `keep = 0.25` (the default) will return a `simfit` object containing those simulated tracks with flag values in the top 25% of values calculated for the input `simfit` object.

2.6 | Path rerouting - `route_path`

As the SSMs implemented in `foieGras` have no information about potential barriers to animal movement it is possible for locations to be estimated in implausible locations, such as on land for marine species. To overcome this, `foieGras` makes use of the `pathroutr` R package (London, 2020) to efficiently re-route locations from land back to water by using visibility graphs (Jan et al., 2014). The `route_path` function can be applied to either a `fit_ssm` model fit (`ssm_df` object) or the simulations generated by `simfit`. When the input is an `ssm_df` object the re-routed path can be appended to the object for visualisation and use in subsequent analyses. When the input is a `simfit` object the locations within the simulation are replaced with the re-routed paths. We illustrate how the latter can be achieved in Application 3.3.

3 | Applications

We illustrate the main capabilities of `foieGras` through a set of applications that are for demonstration purposes and not intended as a comprehensive guide for conducting analyses with `foieGras`. Complete code and data for reproducing the applications and for gaining a deeper understanding of `foieGras` functions are provided in the Supplementary Information.

3.1 | SSM validation with prediction residuals

We use a sub-adult male southern elephant seal (*Mirounga leonina*) track included as example data in `foieGras` (`sese2`, id: ct36-E-09), sourced from from the Australian Integrated Marine Observing System (IMOS; data publicly available via imos.aodn.org.au) deployments at Iles Kerguelen in collaboration with the French IPEV and SNO-MEMO programmes. The data are temporally irregular Argos Least-Squares based locations, 74 % of which are in the poorest location quality classes: A and B. We fit both the `rw` and `crw` models using `fit_ssm` with a speed filter threshold (`vmax`) of 4 ms^{-1} and a 12-h time step. We calculate prediction residuals using `osar`, and then use the generic `plot` method for `osar` residuals to assess and compare the model fits (Fig. 1).

The plots of predicted states on top of the observations suggest both models yield similar fits (Fig. 1a; orange vs red lines). However, corresponding predicted locations from the two models differ by a median 6.62 km (range: 0.02, 53.02 km), and there are marked trends in the time-series of residuals for the `rw` model fit (Fig. 1b) with significantly positive autocorrelation in both the x and y directions (Fig. 1d). The `crw` prediction residuals show little trend through time and have relatively little autocorrelation (Fig. 1c,e), implying that the `crw` process model provides a better fit to the data.

3.2 | Inferring move persistence from Argos and GPS data

Argos data - southern elephant seals

Drawing on additional IMOS tracking data from Iles Kerguelen, we infer move persistence, γ_t, along four southern elephant seal tracks. We fitted the `mp` SSM with a 12-h prediction interval with `fit_ssm`, using the Argos Kalman filter-derived error ellipse information provided with each observation (Jonsen et al., 2020). The `mp` SSM simultaneously estimates locations and γ_t, and their uncertainties. We then assess how γ_t changes along the seals' tracks to infer regions where the seals spend disproportionately more or less time during their foraging trips.

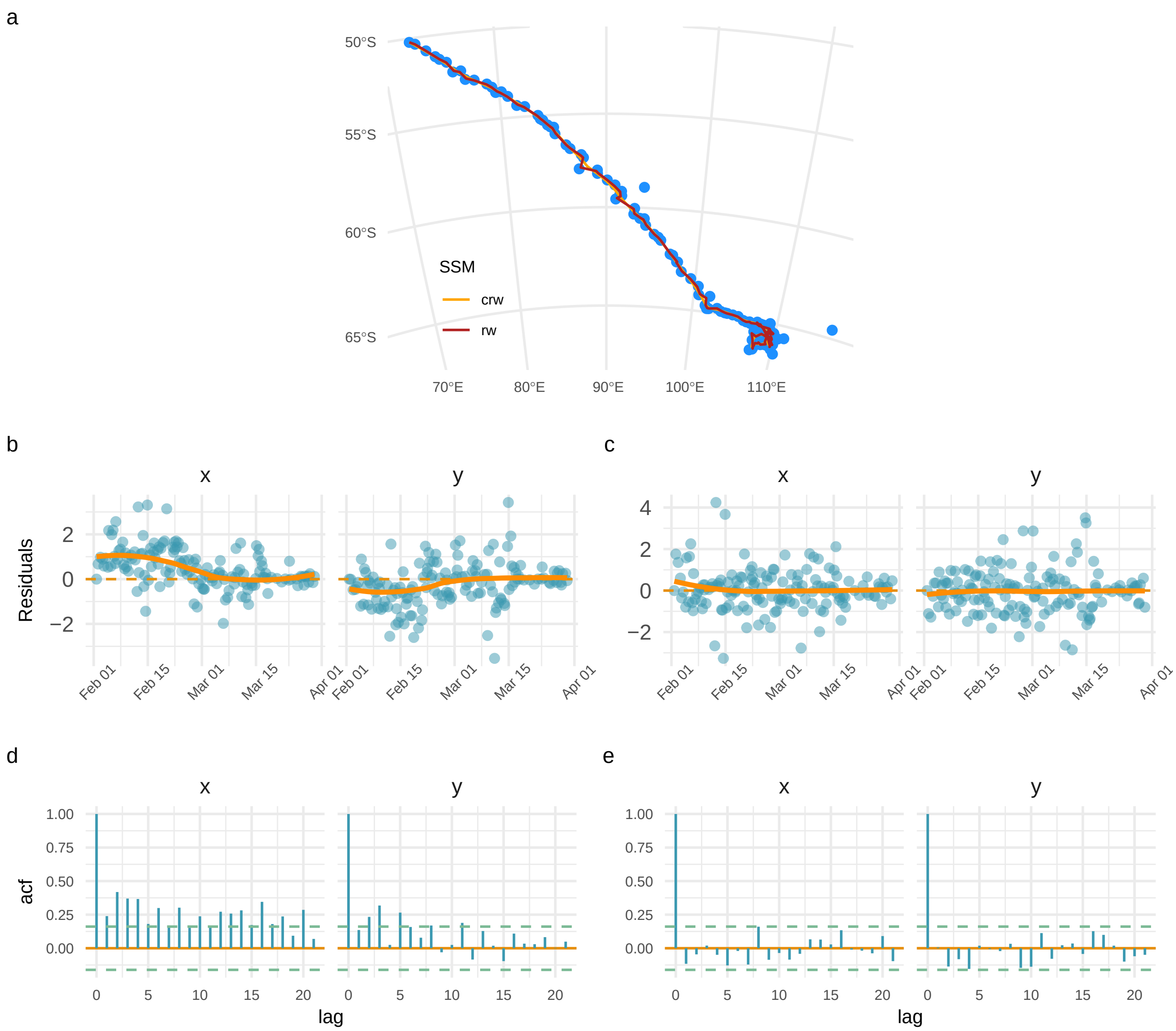

Figure 1: State-space model fits to a southern elephant seal track (a), and diagnostic plots for assessing goodness-of-fit of the `rw` (b - prediction residual time-series; d - prediction residual autocorrelation) and `crw` (c,e) state-space models. All residual plots generated using the `plot.osar` function.

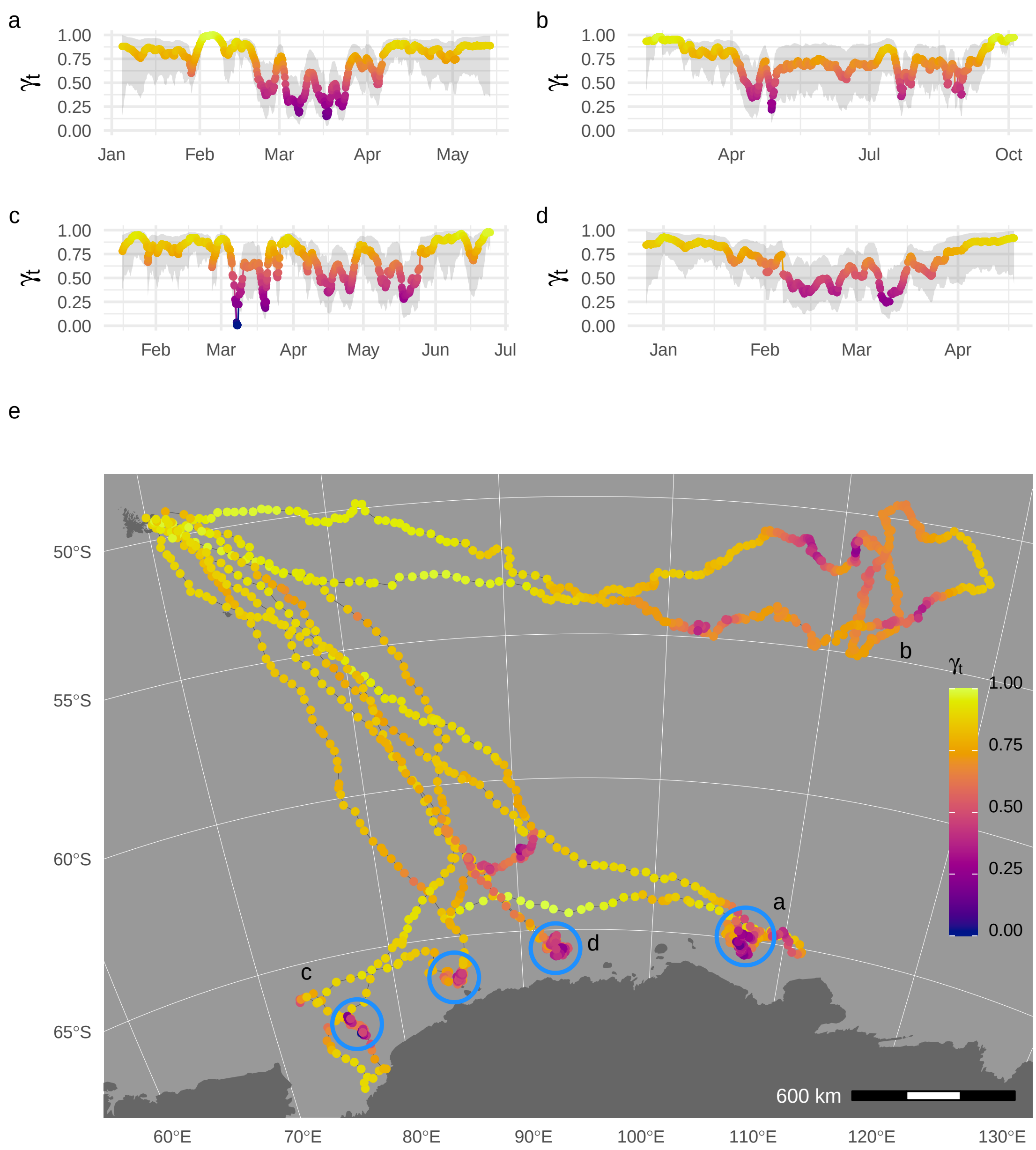

Figure 2: Inferred move persistence, γ_t, time-series for four southern elephant seals (a-d; grey envelopes are 95 % CI's, note differing date ranges on the x axes), and along their 2-D tracks (e; track labels, a-d, correspond to the time-series plots). Locations associated with low γ_t (purple) are indicative of slow, undirected movements, whereas high γ_t (yellow) is indicative of faster, directed movements. Blue circles highlight bouts of spatially constrained low γ_t.

The three southern elephant seals on foraging trips to the Antarctic shelf region all engaged in spatially constrained bouts of low move persistence while in the shelf region (Fig. 2 a,c,d; blue circles in e). Without additional data it is unclear exactly what these bouts of low horizontal move persistence represent. They could result from area-restricted search and foraging within dense prey aggregations, physical constraints of dense ice on horizontal movements, haulout (resting) on sea-ice, or some combination of these. Conversely, the seal on a pelagic foraging trip engaged in slower, more meandering movements with less spatially constrained bouts of lower move persistence (Fig. 2 b, e). This general movement pattern may be consistent with searching for suitable foraging resources within the highly variable eddy fields between the Subantarctic and Polar Fronts (Jonsen et al., 2019).

GPS data - little penguins

To illustrate how move persistence can be estimated from other types of animal tracking data, we use four little penguin (*Eudyptula minor*) GPS tracks from daily foraging trips during the chick-rearing period from Montague Island, NSW, Australia, and described in L. Phillips et al. (2021). The data are temporally irregular GPS locations, with high frequency sampling (15 s on average) intermittently disrupted by the birds' diving behaviour, and are assumed to have minimal measurement error. We fitted the `crw` SSM to the GPS data to predict temporally regular locations at 5-min intervals, and assumed consistently small bivariate normal location measurement errors (ie. $\pm$ 10 m sd). We then used `fit_mpm` to estimate γ_t from these regularised locations.

The little penguin GPS tracks did not exhibit strong contrast in move persistence, with γ_t declining below 0.5 for only two of the birds (Fig. 3a-d). Nonetheless, the move persistence estimates highlight change in movement pattern over the course of the penguins' daily foraging trips (Fig. 3e). The penguins departed Montague Is. with relatively fast movements, three directed southward (Fig. 3 tracks b-d in e) and one less directed and remaining close to the island (track a), before slowing down and engaging in meandering movements (orange - red

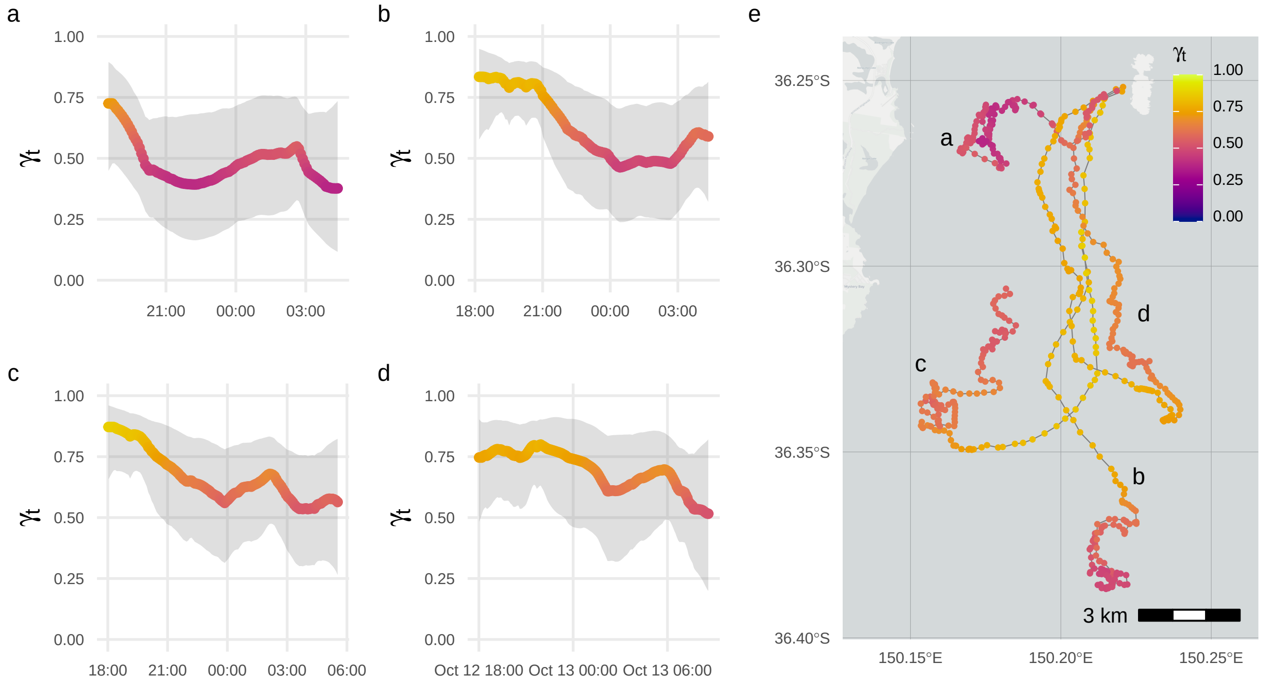

Figure 3: Inferred move persistence, γ_t, 1-D time-series (a-d; grey envelopes are 95 % CI's) and along little penguin GPS tracks (e).

in Fig. 3e). The spatially diffuse bouts of low move persistence within the penguin tracks may reflect the fine-scale patchiness of their forage-fish prey with search and prey-capture occurring both within and among discrete neighbouring prey aggregations (Carroll et al., 2017). **Lach/Gemma could we overlay SVM-inferred prey captures for these birds? (G123f02, G124m10, G126m05, L013m01)**

3.3 | Simulating tracks from `foieGras` model fits

To illustrate how to simulate tracks from `foieGras` model fits we use a juvenile harp seal (*Pagophilus groenlandicus*) tracked from the Gulf of St Lawrence, Canada, and described in Grecian et al. (2022). The data are temporally irregular Argos locations including error ellipse information. We fit the `crw` model using `fit_ssm` with a 4 ms^{-1} speed filter threshold (`vmax`) and a 12-h prediction interval (`time.step`).

We simulate 50 animal movement paths from the `crw` process model using `simfit`, and apply a potential function using the `grad` and `beta` arguments to constrain the simulated paths to largely remain in water. These tracks are then filtered based on their similarity to the original path using `sim_filter` and the top 10% retained (`keep = 0.1`)(Fig. 4a,b). As the potential function does not guarantee all locations remain off land, we re-route any remaining simulated locations from land back to water using `route_path` (Fig. 4c). In combination, these functions provide a pragmatic, non-statistical method to generate and objectively filter pseudo-tracks for use in movement or habitat modelling applications.

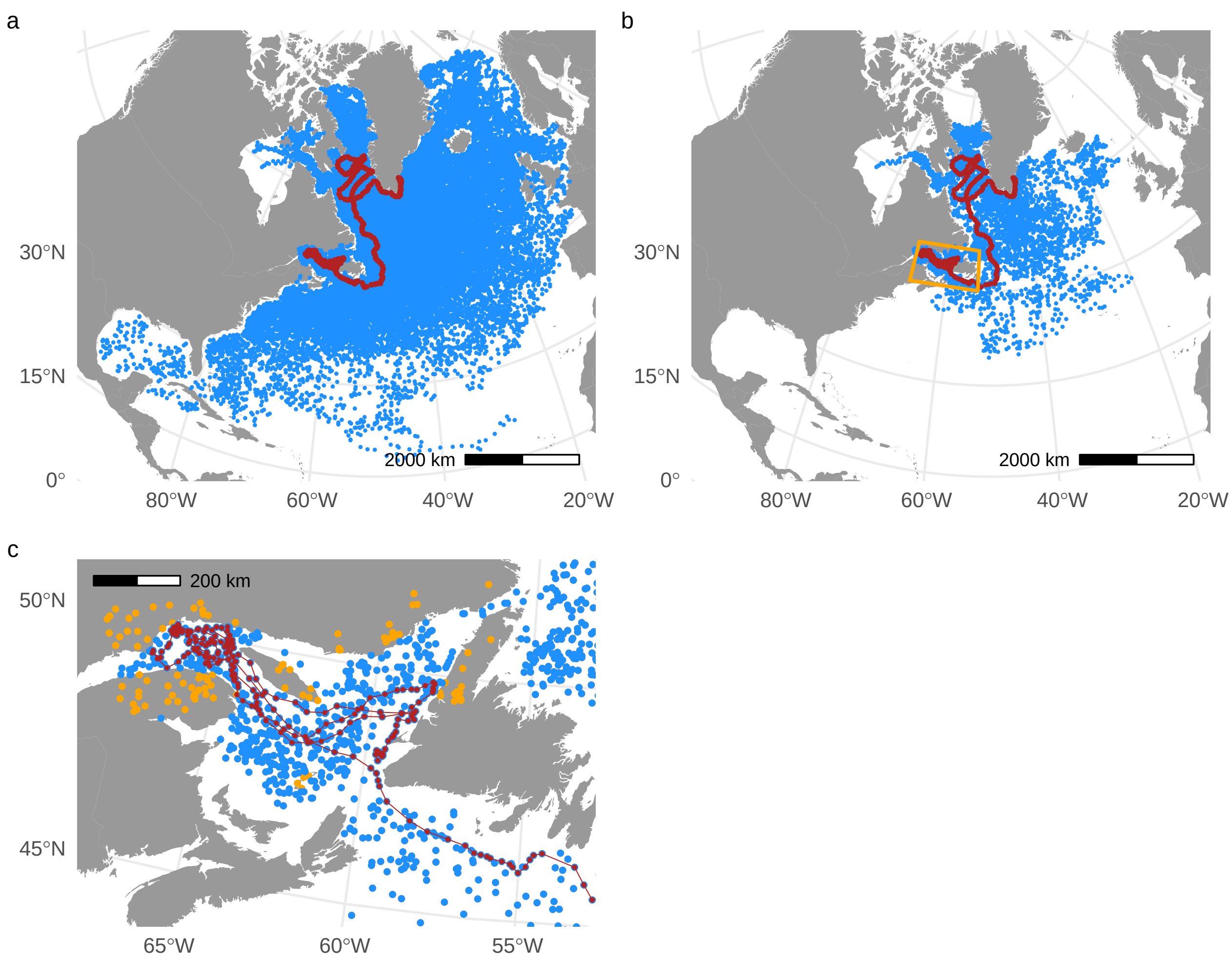

Figure 4: Simulating (a) 100 movement paths from a correlated random walk process model; (b) filtering those tracks to select the top 10% based on their similarity to the original SSM-predicted track (red); and (c) re-routing simulated locations on land (orange) back to ocean (blue). The orange box in (b) indicates region magnified in (c). SSM-predicted track (red) overlaid in all panels for context.

4 | Conclusions

The `foieGras` package was developed to ease fitting state-space models for quality control of animal location data and for inference of behavioural change along animal tracks. We achieve these primarily through a simple yet extensible workflow, model parsimony, and computational speed. Combined, these traits accommodate both novice or occasional and advanced users, and facilitate use in automated, operational quality-assurance/quality-control processes for animal-borne ocean observations (Jonsen et al., 2020; McMahon et al., 2021).

The `foieGras` package is an intermediate analysis toolbox where location quality control typically occurs after some initial data processing but prior to any comprehensive, final analysis. In this vein, the move persistence model tools provide a simple, rapid approach for objectively identifying changes in movement behaviour along animal tracks without any required *a priori* knowledge or decisions about the kind or number of behavioural states hidden within the data. Subsequent analysis could entail use of a hidden Markov model to infer relationships between behaviour and environmental or individual covariates (McClintock & Michelot, 2018; e.g., Michelot et al., 2016). Alternatively, move persistence - covariate relationships among multiple individuals can be inferred rapidly in a mixed-effect model framework using the `mpmm` package [Jonsen et al. (2019); https://github.com/ianjonsen/mpmm].

The package will undergo further development, expanding the range of available SSM's via new movement process models and/or enhanced measurement models for other electronic tracking data types. Additionally, we will seek to enhance integration with other R packages for processing and analysis of animal movement data, where this makes sense. Feedback from users is invaluable and highly encouraged. Users may submit bug reports and enhancement suggestions via the foieGras GitHub issues page (https://github.com/ianjonsen/foieGras/issues/new/choose). More general feedback is always welcome by contacting the lead author directly.

Acknowledgements

We thank Marie Auger-Méthé for contributing original code to the movement persistence models. IDJ acknowledges support from a Macquarie University co-Funded Fellowship and from partners: the US Office of Naval Research, Marine Mammal Program (grant N00014-18-1-2405); the Integrated Marine Observing System (IMOS); Taronga Conservation Society; the Ocean Tracking Network; Birds Canada; and Innovasea/VEMCO. TAP was supported by CSIRO Oceans & Atmosphere internal research funding scheme. The Integrated Marine Observing System (IMOS) supported seal fieldwork. IMOS is a national collaborative research infrastructure, supported by the Australian Government and operated by a consortium of institutions as an unincorporated joint venture, with the University of Tasmania as Lead Agent. Field work at Illes Kerguelen was conducted as part of the IPEV programme N° 109 (PI H. WEIMERSKIRCH) and of the SNO-MEMO programme (PI C. GUINET) in collaboration with IMOS. CTD tags were partly funded by CNES-TOSCA and IMOS. Little penguin fieldwork was supported by an Australian Research Council Linkage grant to IDJ, GC and RGH (LP160100162). All animal tagging procedures approved and executed under the Animal Ethics Committee guidelines of the University of Tasmania (elephant seals), Macquarie University (little penguins), and **... University (harp seals) - JAMES, need input here**.

Author's Contributions

IDJ developed the R package; WJG contributed harp seal data and to the R package; LP, GC, and RGH contributed little penguin data; CRM and RGH contributed Southern elephant seal data; IDJ and TAP developed the state-space models; IDJ wrote an initial draft of the manuscript with contributions from WJG; all authors edited the manuscript.

Data Accessibility

All code and data used here are provided in the `foieGras` package for R or in the Supplementary Information. The latest stable and cross-platform tested version of the package (currently, 1.0-7) is available via ROpenSci's R-universe, at https://ianjonsen.r-universe.dev/ui#package:foieGras. The latest partially tested stable and development versions are available on the GitHub repository: https://github.com/ianjonsen/foieGras. An older version of `foieGras` (0.7-6) remains on CRAN at https://CRAN.R-project.org/package=foieGras, however, we recommend users upgrade to the latest R-universe version for full access to the functionality presented here.

ORCID

Ian D Jonsen https://orcid.org/0000-0001-5423-6076

W James Grecian https://orcid.org/0000-0002-6428-719X

Lachlan Phillips https://orcid.org/0000-0002-7635-2817

Gemma Carroll https://orcid.org/0000-0001-7776-0946

Clive R McMahon https://orcid.org/0000-0001-5241-8917

Robert G Harcourt https://orcid.org/0000-0003-4666-2934

Mark A Hindell https://orcid.org/0000-0002-7823-7185

Toby A Patterson https://orcid.org/0000-0002-7150-9205

References

Aarts, G., Fieberg, J., & Matthiopoulos, J. (2012). Comparative interpretation of count, presence–absence and point methods for species distribution models. *Methods in Ecology and Evolution*, *3*(1), 177–187. https://doi.org/10.1111/j.2041-210x.2011.00141.x

Albertsen, C. M., Whoriskey, K., Yurkowski, D., Nielsen, A., & Mills Flemming, J. (2015). Fast fitting of non-Gaussian state-space models to animal movement data via Template

Model Builder. *Ecology*, *96*, 2598–2604.

Auger-Méthé, M., Albertsen, C. M., Jonsen, I. D., Derocher, A. E., Lidgard, D. C., Studholme, K. R., Bowen, W. D., Crossin, G. T., & Flemming, J. M. (2017). Spatiotemporal modelling of marine movement data using Template Model Builder (TMB). *Marine Ecology Progress Series*, *565*, 237–249.

Baktoft, H., Gjelland, K. Ø., Økland, F., & Thygesen, U. H. (2017). Positioning of aquatic animals based on time-of-arrival and random walk models using YAPS (yet another positioning solver). *Scientific Reports*, *7*, 14294. https://doi.org/10.1038/s41598-017-14278-z

Calabrese, J. M., Fleming, C. H., & Gurarie, E. (2016). Ctmm: An R package for analyzing animal relocation data as a continuous-time stochastic process. *Methods in Ecology and Evolution*, *7*, 1124–1132.

Carroll, G., Cox, M., Harcourt, R., Pitcher, B., Slip, D., & Jonsen, I. (2017). Hierarchical influences of prey distribution on patterns of prey capture by a marine predator. *Functional Ecology*, *31*, 1750–1760.

Dunnington, D. (2019). *Rosm: Plot raster map tiles from open street map and other sources*. https://CRAN.R-project.org/package=rosm

Dunnington, D. (2021). *Ggspatial: Spatial data framework for ggplot2*. https://CRAN.R-project.org/package=ggspatial

Grecian, W. J., Stenson, G. B., Biuw, M., Boehme, L., Folkow, L. P., Goulet, P. J., Jonsen, I. D., Malde, A., Nordøy, E. S., Rosing-Asvid, A., & Smout, S. (2022). Environmental drivers of population-level variation in the migratory and diving ontogeny of an arctic top predator. *Royal Society Open Science*, *9*(3). https://doi.org/10.1098/rsos.211042

Harcourt, R., Sequeira, A. M. M., Zhang, X., Roquet, F., Komatsu, K., Heupel, M., McMahon, C., Whoriskey, F., Meekan, M., Carroll, G., Brodie, S., Simpfendorfer, C., Hindell, M., Jonsen, I., Costa, D. P., Block, B., Muelbert, M., Woodward, B., Weise, M., … Fedak, M. A. (2019). Animal-borne telemetry: An integral component of the ocean observing toolkit. *Frontiers in Marine Science*, *6*, 326.

Hazen, E. L., Palacios, D. M., Forney, K. A., Howell, E. A., Becker, E., Hoover, A. L., Irvine, L., DeAngelis, M., Bograd, S. J., Mate, B. R., & Bailey, H. (2017). WhaleWatch: a dynamic management tool for predicting blue whale density in the California Current. *Journal of Applied Ecology*, *54*(5), 1415–1428. https://doi.org/10.1111/1365-2664.12820

Hindell, M. A., Reisinger, R. R., Ropert-Coudert, Y., Hückstädt, L. A., Trathan, P. N., Bornemann, H., Charrassin, J.-B., Chown, S. L., Costa, D. P., Danis, B.others. (2020). Tracking of marine predators to protect southern ocean ecosystems. *Nature*, *580*(7801), 87–92.

Hooten, M. B., Scharf, H. R., & Morales, J. M. (2019). Running on empty: Recharge dynamics from animal movement data. *Ecology Letters*, *22*, 377–389.

Jan, G. E., Sun, C.-C., Tsai, W. C., & Lin, T.-H. (2014). An o(n log n) shortest path algorithm based on delaunay triangulation. *IEEE/ASME Transactions on Mechatronics*, *19*(2), 660–666. https://doi.org/10.1109/TMECH.2013.2252076

Johnson, D. S., London, J. M., Lea, M.-A., & Durban, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, *89*(5), 1208–1215.

Jonsen, I. D., Basson, M., Bestley, S., Bravington, M. V., Patterson, T. A., Pedersen, M. W., Thomson, R., Thygesen, U. H., & Wotherspoon, S. J. (2013). State-space models for bio-loggers: A methodological road map. *Deep Sea Research II*, *88-89*, 34–46.

Jonsen, I. D., McMahon, C. R., Patterson, T. A., Auger-Méthé, M., Harcourt, R., Hindell, M. A., & Bestley, S. (2019). Movement responses to environment: Fast inference of variation among southern elephant seals with a mixed effects model. *Ecology*, *100*, e02566.

Jonsen, I. D., Patterson, T. A., Costa, D. P., Doherty, P. D., Godley, B. J., Grecian, W. J., Guinet, C., Hoenner, X., Kienle, S. S., Robinson, P. W., Votier, S. C., Whiting, S., Witt, M. J., Hindell, M. A., Harcourt, R. G., & McMahon, C. R. (2020). A continuous-time state-space model for rapid quality control of Argos locations from animal-borne tags. *Movement Ecology*, *8*, 31.

Jonsen, I., Flemming, J., & Myers, R. (2005). Robust state–space modeling of animal movement data. *Ecology*, *86*(11), 2874–2880.

Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*(6240).

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic differentiation and Laplace approximation. *Journal of Statistical Software*, *70*, 1–21.

Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, *33*(1), 103–114. https://doi.org/https://doi.org/10.1111/j.1600-0587.2009.06039.x

London, J. M. (2020). *Pathroutr: An r package for (re-)routing paths around barriers* (v0.1.1-beta) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.4321827

McClintock, B. T., & Michelot, T. (2018). MomentuHMM: R package for generalized hidden Markov models of animal movement. *Methods in Ecology and Evolution*, *9*, 1518–1530.

McMahon, C. R., Roquet, F., Baudel, S., Belbeoch, M., Bestley, S., Blight, C., Boehme, L., Carse, F., Costa, D. P., Fedak, M. A., Guinet, C., Harcourt, R., Heslop, E., Hindell, M. A., Hoenner, X., Holland, K., Holland, M., Jaine, F. R. A., Jeanniard du Dot, T., … Woodward, B. (2021). Animal borne ocean sensors – AniBOS – an essential component of the global ocean observing system. *Frontiers in Marine Science*, *8*, 751840.

Michelot, T., Langrock, R., Bestley, S., Jonsen, I. D., Photopoulou, T., & Patterson, T. A. (2017). Estimation and simulation of foraging trips in land-based marine predators. *Ecology*, *98*(7), 1932–1944.

Michelot, T., Langrock, R., & Patterson, T. A. (2016). MoveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. *Methods in Ecology and Evolution*, *7*, 1308–1315.

Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L., & King, R. (2017). Statistical modelling of individual animal movement: An overview of key methods and a discussion of practical challenges. *AStA Advances in Statistical Analysis*, *101*, 399–438.

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., & Matthiopoulos, J. (2008). State-space models of individual animal movement. *Trends in Ecology and Evolution*, *23*, 87–94.

Pedersen, T. L. (2020). *Patchwork: The composer of plots*. https://CRAN.R-project.org/package=patchwork

Phillips, L., Carroll, G., Jonsen, I. D., Harcourt, R. G., Brierley, A., Wilkins, A., & Cox, M. (2021). Variability in prey field structure drives inter-annual differences in prey encounter by a marine predator, the little penguin. *Proceedings of the Royal Society B*, *In Review*.

Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, *19*(1), 181–197. https://doi.org/10.1890/07-2153.1

Preisler, H. K., Ager, A. A., & Wisdom, M. J. (2013). Analyzing animal movement patterns using potential functions. *Ecosphere*, *4*, 1–13.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Raymond, B., Lea, M.-A., Patterson, T., Andrews-Goff, V., Sharples, R., Charrassin, J.-B., Cottin, M., Emmerson, L., Gales, N., Gales, R., Goldsworthy, S. D., Harcourt, R., Kato, A., Kirkwood, R., Lawton, K., Ropert-Coudert, Y., Southwell, C., Hoff, J. van den, Wienecke, B., … Hindell, M. A. (2015). Important marine habitat off east Antarctica revealed by two decades of multi-species predator tracking. *Ecography*, *38*, 121–129.

Schick, R. S., New, L. F., Thomas, L., Costa, D. P., Hindell, M. A., McMahon, C. R., Robinson, P. W., Simmons, S. E., Thums, M., J., Harwood., & Clark, J. S. (2013). Estimating resource acquisition and at-sea body condition of a marine predator. *Journal of Animal Ecology*, *82*, 1300–3015.

Service Argos. (2016). *Argos Users' Manual*. CLS. http://www.argos-system.org/manual

South, A. (2022a). *Rnaturalearth: World map data from natural earth*.

South, A. (2022b). *Rnaturalearthhires: High resolution world vector map data from natural earth used in rnaturalearth*.

Sumner, M. D., Wotherspoon, S. J., & Hindell, M. A. (2009). Bayesian estimation of animal movement from archival and satellite tags. *PLoS ONE*, *4*(10). https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0007324

Thygesen, U. H., Albertsen, C. M., Berg, C. W., Kristensen, K., & Neilsen, A. (2017). Validation of ecological state space models using the Laplace approximation. *Environmental and Ecological Statistics*. https://doi.org/DOI: 10.1007/s10651-017-0372-4

Whoriskey, K., Auger-Méthé, M., Albertsen, C. M., Whoriskey, F. G., Binder, T. R., Kreuger, C. C., & Mills Flemming, J. (2017). A hidden markov movement model for rapidly identifying behavioral states from animal tracks. *Ecology and Evolution*, *7*, 2112–2121.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org