

¹ **foieGras** an R package for animal movement data:
² rapid quality control, behavioural estimation and
³ simulation

⁴ **Ian D. Jonsen^{1,*}, W. James Grecian², Lachlan Phillips¹, Gemma Carroll³, Clive
⁶ McMahon⁴, Robert G. Harcourt¹, and Toby A. Patterson⁵**

⁷ ¹School of Natural Sciences, Macquarie University, Sydney, NSW, Australia

⁸ ²Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St
⁹ Andrews, Fife, United Kingdom

¹⁰ ³Environmental Defense Fund, Seattle, WA, United States

¹¹ ⁴Sydney Institute of Marine Science, Mosman, NSW, Australia

¹² ⁵CSIRO Ocean and Atmosphere Research, Hobart, TAS, Australia

¹³ *corresponding author, ian.jonsen@mq.edu.au

¹⁴ **Abstract**

- ¹⁵ 1.
- ¹⁶ 2.
- ¹⁷ 3.
- ¹⁸ 4.

¹⁹ **Keywords:**

²⁰ 1 | Introduction

²¹ Animal biotelemetry as a discipline has matured, with telemetry data now virtually essential for
²² understanding behaviour and social interactions, foraging ecology, habitat use and population
²³ dynamics of mobile and/or cryptic species. The sophistication of current telemetry devices
²⁴ also enables the use of animal-borne sensors as a cost-effective approach for observing our
²⁵ planet that complements more traditional observing platforms (Harcourt et al., 2019; Kays
²⁶ et al., 2015; McMahon et al., 2021). In all these applications, animal biotelemetry requires
²⁷ rigorous quality control procedures to account for common, though not universally present,
²⁸ data issues such as irregularly timed measurements, sensor biases and location measurement
²⁹ error. Some of these issues may be handled by a manufacturer's on-board or post-processing
³⁰ algorithms and some must be addressed by researchers using the data.

³¹ State-space models (SSMs) and hidden Markov models (HMMs) are powerful tools for
³² conducting quality control of and making behavioural inference from animal biotelemetry
³³ data (Jonsen et al., 2013; Patterson et al., 2008). These are time-series models used across
³⁴ a wide range of research disciplines that estimate the state of an unobserved process from
³⁵ an observed data set. Here, we view an animal's true location and/or behaviour as the
³⁶ unobserved state(s), though many other types of states are possible (Hooten et al., 2019;
³⁷ e.g., Schick et al., 2013), and measurements recorded by biotelemetry devices provide the
³⁸ observations. In practical yet simplistic terms, SSMs are usually preferred when the goal is to
³⁹ quality control error-prone location data and/or make inference directly from the parameters
⁴⁰ of their underlying movement models. HMMs are usually preferred when the goal is to infer
⁴¹ behavioural states hidden within biotelemetry data and their drivers, where measurements
⁴² have negligible error and occur at regular time intervals (but see McClintock & Michelot,
⁴³ 2018). Other more technical distinctions and reasons for preferring one of these methods exist
⁴⁴ (Jonsen et al., 2013; Patterson et al., 2017). Our primary focus here is on SSMs as tools for
⁴⁵ quality control of error-prone location data.

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

54 The `foieGras`, pronounced “*fwah grah*”, package for R (R Core Team, 2021) was developed
55 to be simple and fast to use, providing SSMs for quality control of error-prone (re)location
56 data collected via the Argos satellite [Service Argos (2016);Jonsen:2020] and other telemetry
57 systems, and for inference of changes in behaviour along movement tracks (Jonsen et al., 2019).
58 The simplicity of use sets `foieGras` apart from many of the related SSM R packages listed
59 above, yet users can exert control over many aspects of the package functions via optional
60 arguments. This design accommodates both novice and experienced users, and facilitates use
61 in automated, operational quality-assurance/quality-control workflows (Jonsen et al., 2020).

62 Here, we describe the main features of `foieGras` and illustrate its use through a set of
63 examples using Argos and GPS biotelemetry data. Full R code and data for each of the
64 examples is provided in the Supporting Information. Additional details on package functions
65 and their use can be found in their help files and in the package’s vignettes.

66 2 | `foieGras` overview

67 The workflow for `foieGras` is deliberately simple, with many of the usual track data processing
68 checks and formatting handled automatically. Here we outline the main aspects of the
69 `foieGras` package. Further details on usage are provided in the package’s help documentation
70 and vignettes. The package’s main functions are listed in Table 1.

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

Function	Description
<code>fit_ssm</code>	Fit a State-Space Model to location data
<code>fit_mpm</code>	Fit a Move Persistence Model to location data
<code>grab</code>	Extract fitted/predicted/observed locations from a foieGras model, with or without projection information
<code>osar</code>	Estimate One-Step-Ahead Residuals from a foieGras SSM
<code>map</code>	Map fitted/predicted locations with or without a defined projection
<code>sim</code>	Simulate individual animal tracks with Argos LS or KF errors
<code>simfit</code>	Simulate animal tracks from ‘ssm_df’ fit objects
<code>sim_filter</code>	Filter tracks simulated with ‘simfit’ according to similarity criteria
<code>route_path</code>	Reroute path so estimated locations are off land
<code>plot.ssm_df</code>	Plot the fit of a foieGras SSM to data
<code>plot.osar</code>	Plot One-Step-Ahead Residuals from a foieGras SSM
<code>plot.mpm_df</code>	Plot move persistence estimates as 1-D or 2-D (along track) time-series
<code>plot.sim</code>	Plot simulated animal tracks

71 **2.1 | Data preparation**

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

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

83 **2.2 | State-space model fitting - `fit_ssm`**

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

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

- 94
 - `data` the input data structured as described in **2.1**
 - `model` the process model to be used

- 96 • `time.step` the prediction time interval (h)
- 97 The function first invokes an automated data processing stage where the following occurs:
- 98 1) data type (Argos Least-Squares, Argos Kalman Filter/Smoother, GPS, or General (e.g.,
99 processed light-level geolocations, acoustic telemetry, coded VHF telemetry) is determined; 2)
100 date-times are converted to POSIXt format, chronological order is ensured, and duplicate
101 date-time records are removed; 3) observations occurring less than `min.dt` seconds after a
102 prior observation are removed (default: 60 s); 4) a speed filter [`sda` from the `trip` R package;
103 Sumner et al. (2009)] is used to identify potential extreme locations to be ignored by the
104 SSM; 5) locations are projected from spherical lon,lat coordinates to global Mercator x,y
105 coordinates in km.
- 106 The function then fits a state-space model to the processed data, where the process model
107 (one of `rw`, `crw`, or `mp`) is specified by the user via the `model` argument, and the measurement
108 model(s) are selected automatically. The model is fit by numerical optimization of the
109 likelihood using either of the standard R optimizers, `optim` or `nlinmib`. The R package TMB,
110 Template Model Builder (Kristensen et al., 2016), is used to compute the gradient function
111 in C++ via reverse-mode auto-differentiation and the Laplace Approximation is used to
112 integrate out the latent states (random effects). Fits to a single versus multiple individuals are
113 handled automatically, with sequential SSM fits occurring in the latter case. No hierarchical
114 or pooled estimation among individuals is currently available.
- 115 `fit_ssm` returns a `ssm_df` fit object (a nested data frame with class `ssm_df`). The outer data
116 frame lists the individual id(s), basic convergence information and a list with class `ssm`. This
117 list contains dense information on the estimated parameters and states, predictions, processed
118 data, optimizer results, and other diagnostic and contextual information. Users can extract a
119 simple data frame of SSM fitted (location estimates corresponding to the observation times) or
120 predicted values (locations predicted at regular `time.step` intervals) using the `grab` function.
121 Parameter estimates, AIC and other model fit information can be viewed in tabular form

122 using the `summary` function.

123 2.3 | Behavioural estimation - `fit_ssm`, `fit_mpm`

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

136 2.4 | Model checking and visualization - `osar`, `plot`, `fmap`

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

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

154 Additionally, `ssm_df` fit objects can be mapped using the `map` function for single or multiple
155 individuals. By default, `map` uses the coastline data from the `rnatuRearth` R package
156 (South, 2022a) at medium or high resolution, if the `rnatuRearthhires` (South, 2022b) R
157 package is installed, but can also use tiled maps for finer-scale detail, if the `rosm` (Dunnington,
158 2019) and `ggspatial` (Dunnington, 2021) R packages are installed. Mapping aesthetics (e.g.,
159 plot symbols, sizes, colours, fills) can be customized via the `aes` argument and use of the
160 `aes_lst` function. See code in SI for examples.

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

164 2.5 | Simulation - `sim`, `simfit`, `sim_filter`

165 Track simulation can be a helpful, yet informal, way of evaluating the degree to which
166 statistical movement models capture essential features of animal movement data (Michelot et
167 al., 2017). The `sim` function can simulate a variety of movement process, including the `rw`,
168 `crw`, and `mp` process models, as well as simple multiple movement state switching processes.
169 These simulations can be used to evaluate the robustness of location and move persistence
170 estimates from SSMs to model mis-specification. An example is provided in SI.

171 Simulation is also used frequently in habitat usage modelling to provide a measure of habitat

172 availability (Aarts et al., 2012) by generating a source of ‘background’ points representing a
173 null model of the distribution of foraging animals in the absence of external drivers (Hindell
174 et al., 2020; S. J. Phillips et al., 2009; Raymond et al., 2015). The `simfit` function extracts
175 movement parameters from a `ssm_df` fit object and simulates a user defined number of random
176 tracks of the same duration from these parameters. The argument `cpf = TRUE` allows the
177 user to simulate central place foragers by ensuring that the simulated tracks start and end
178 at approximately the same location. It is also possible to constrain movements to remain
179 mostly in water via a potential function, using included gradient rasters and the `grad` and
180 `beta` arguments. These are illustrated in the code for Example 3.3.

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

190 **2.6 | Path rerouting - route_path**

191 As the SSMS implemented in `foieGras` have no information about potential barriers to animal
192 movement, it is possible for locations to be estimated in implausible locations such as on
193 land for entirely or mostly aquatic species. To overcome this, `foieGras` makes use of the
194 `pathroutr` R package (London, 2020) to re-route locations from land back to water.
195 In principle, the reverse could be implemented for land-based species

¹⁹⁶ **3 | Examples**

¹⁹⁷ We illustrate the main capabilities of `foieGras` through a series of examples that are for
¹⁹⁸ demonstration purposes and not intended as a comprehensive guide for conducting analyses
¹⁹⁹ with `foieGras`. Complete code and data for reproducing the examples and for gaining a
²⁰⁰ deeper understanding of `foieGras` functions are provided as supplements.

²⁰¹ **3.1 | Southern Elephant seal - SSM validation with prediction residuals**

²⁰² We use a sub-adult male southern elephant seal track included in `foieGras` (`sese2`, id:
²⁰³ `ct36-E-09`), sourced from from the Australian Integrated Marine Observing System (IMOS;
²⁰⁴ data publicly available via imos.aodn.org.au) deployments at Iles Kerguelen in collaboration
²⁰⁵ with the French IPEV and SNO-MEMO programmes. The data are temporally irregular
²⁰⁶ Argos Least-Squares based locations, 74 % of which are in the poorest location quality classes:
²⁰⁷ A and B. We fit both the `rw` and `crw` models using `fit_ssm` with a speed filter threshold
²⁰⁸ (`vmax`) of 4 ms⁻¹ and a 12-h time step. We calculate prediction residuals using `osar`, and then
²⁰⁹ use the generic `plot` method for `osar` residuals to assess and compare the model fits (Fig. 1).

²¹⁰ The plots of predicted states on top of the observations suggest both models yield similar
²¹¹ fits (Fig. 1a; orange vs red lines). However, corresponding predicted locations from the two
²¹² models differ by a median 6.62 km (range: 0.02, 53.02 km), and there are marked trends in the
²¹³ time-series of residuals for the `rw` model fit (Fig. 1b) with significantly positive autocorrelation
²¹⁴ in both the x and y directions (Fig. 1d). The `crw` prediction residuals show little trend
²¹⁵ through time and have relatively little autocorrelation (Fig. 1c,e), implying that the `crw`
²¹⁶ process model provides a better fit to the data.

²¹⁷ **3.2 | Inferring movement persistence from Argos and GPS data**

²¹⁸ Drawing on an expanded version of the data used in 3.1, we quality control and infer movement
²¹⁹ persistence, γ_t , along four southern elephant seal tracks. We fitted the `mp` SSM with a 24-h

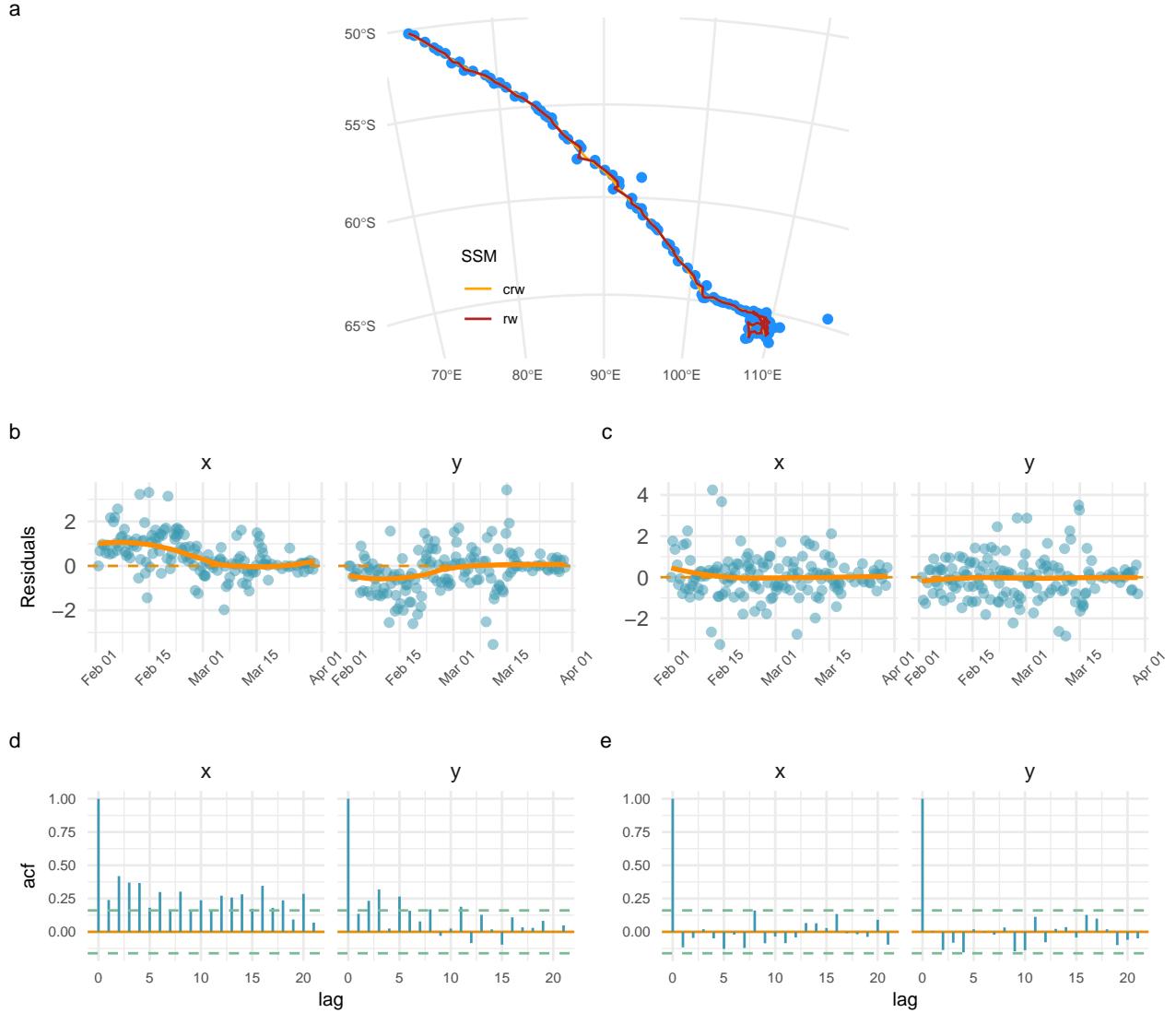


Figure 1: State-space model fits to a southern elephant seal track (a), and diagnostic plots for assessing **rw** (c - time-series of prediction residuals; e - autocorrelation of prediction residuals) and **crw** (d,f) state-space model goodness-of-fit. All residual plots generated using the `plot.fG_osar`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

220 prediction interval using `fit_ssm`, assuming bivariate normal measurement errors consistent
221 with Argos Least-Squares-derived locations (Jonsen et al., 2020). The `mp` SSM simultaneously
222 estimates locations and γ_t , and their uncertainties. We then assess how γ_t changes along the
223 seals' tracks to infer regions where the seals spend disproportionately more or less time during
224 their foraging trips.

225 To illustrate how the method can accommodate other types of animal tracking data, we also
226 infer γ_t along four little penguin (*Eudyptula minor*) GPS tracks from Montague Island, NSW
227 Australia, described in L. Phillips et al. (2021). The data are temporally irregular GPS
228 locations, with high frequency sampling (1 - 2s) periodically disrupted by the birds' diving
229 behaviour, and are assumed to have minimal measurement error. We consequently chose to
230 fit the `crw` SSM to the GPS data and predict temporally regular locations at 5-min intervals,
231 assuming consistently small bivariate normal location measurement errors (ie. ± 10 m sd).
232 We then used `fit_mpm` to estimate γ_t from the SSM-predicted locations.

233 Movement persistence estimates along the quality-controlled southern elephant seal tracks
234 highlight some fundamental differences in movement pattern among the seals. The two seals
235 engaging in pelagic foraging trips (Fig. 2a,c and e) had less contrast in their movements with
236 consistently higher γ_t estimates compared to the two seals engaging in trips to the fast-ice on
237 the Antarctic shelf (Fig. 2b,d and e). Although γ_t 's were higher overall for the pelagically
238 foraging seals, they both spent little time making fast, highly directed movements ($\gamma_t \rightarrow 1$)
239 relative to the shelf-foraging seals (2a,c vs b,d). This suggests the pelagically-foraging seals
240 may spend considerable time searching for suitable foraging habitat in the highly variable
241 eddy fields between the Subantarctic and Polar Fronts (Jonsen et al., 2019), whereas foraging
242 habitat may be more predictable for seals travelling rapidly and directly to the Antarctic
243 shelf region. These seals may also haulout periodically on available fast-ice to rest. This
244 behaviour could also contribute to the higher contrast in movement persistence, relative to
245 pelagically-foraging seals who would not have access to fast-ice.

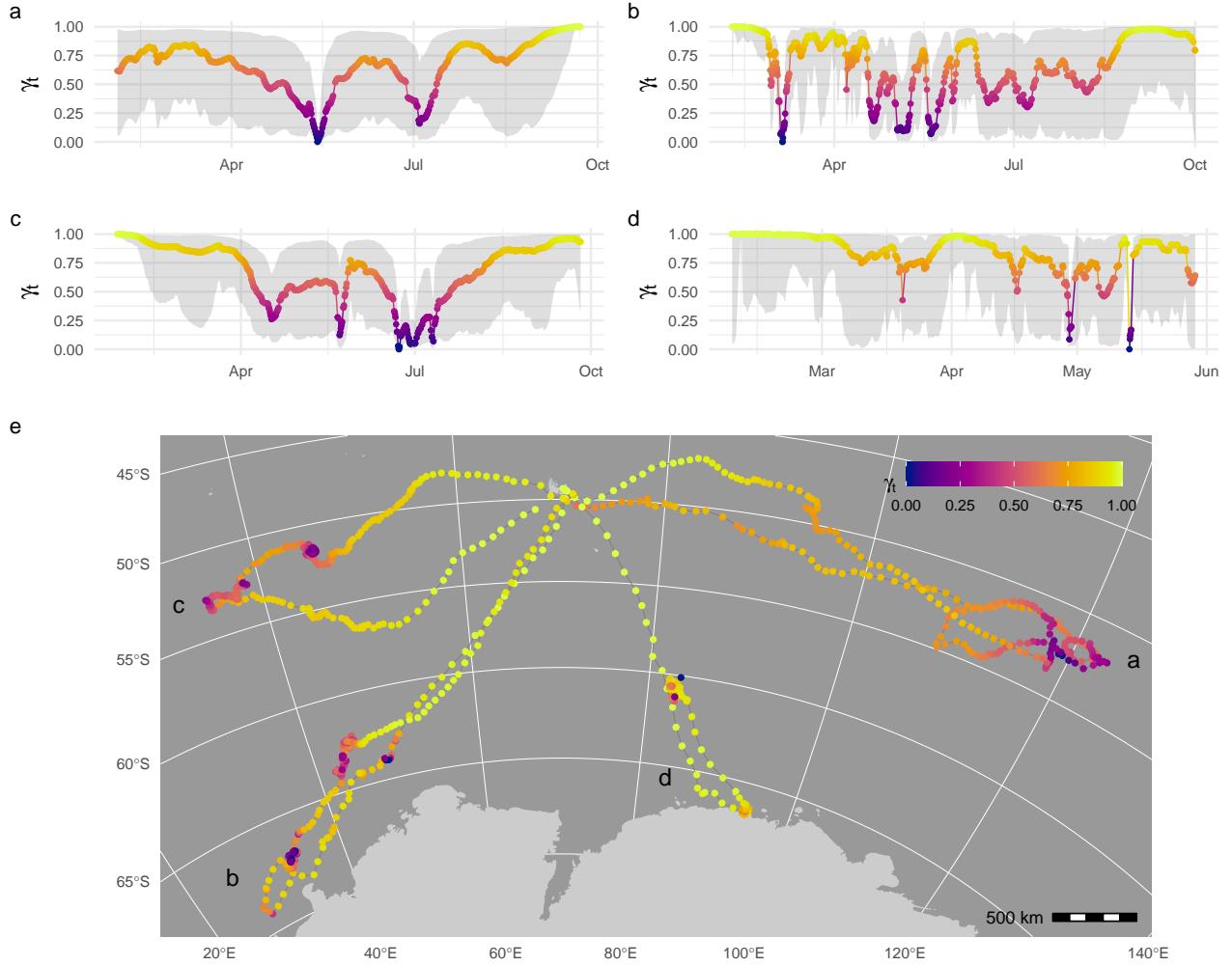


Figure 2: Inferred move persistence, γ_t , 1-D time-series for four southern elephant seals (a-d; grey envelopes are 95 % CI's) and along their 2-D tracks (e; track labels, a-d, correspond to the 1-D time-series plots). Locations associated with low move persistence (purple) are indicative of slow, undirected movements, whereas high move persistence (yellow) is indicative of faster, directed movements.

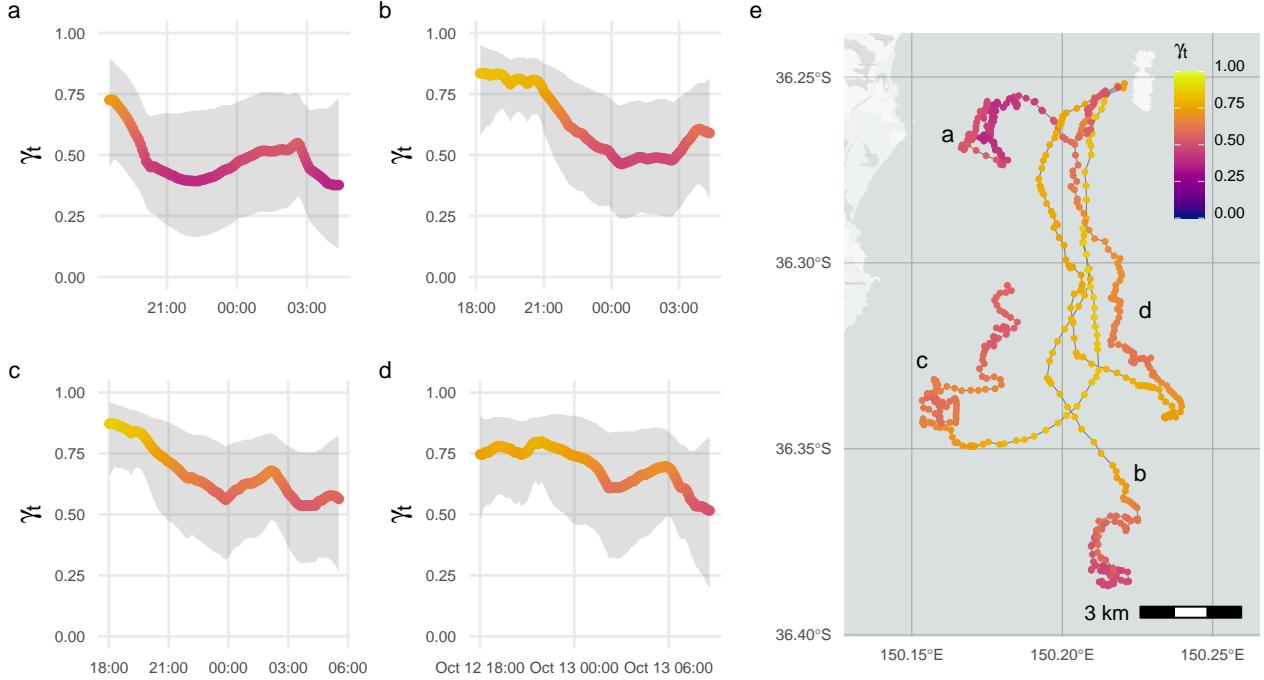


Figure 3: Inferred move persistence, γ_t , 1-D time-series (a-d; grey envelopes are 95 % CI's) and along little penguin GPS tracks (e).

Despite vastly different scales of movement, the time series of little penguin movement persistence estimates were broadly similar to those of the southern elephant seals (Fig. 3a-d). The little penguin foraging trips likely reflect the underlying spatial distribution of their forage-fish prey, with spatially diffuse bouts of lower movement persistence potentially indicative of foraging both within and among neighbouring discrete prey patches (Carroll et al., 2017) (Fig. 3e).

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`).

258 We simulate 50 animal movement paths from the `crw` process model using `simfit`, and apply
259 a potential function using the `grad` and `beta` arguments to constrain the simulated paths
260 to largely remain in water. These tracks are then filtered based on their similarity to the
261 original path using `sim_filter` and the top 10% retained (`keep = 0.1`) (Fig. 4a,b).

262 As the potential function does not guarantee all locations remain off land, we re-route any
263 remaining simulated locations from land back to water using `route_path` (Fig. 4c).

264 In combination, these functions provide a simple, pragmatic method to generate and objectively
265 filter pseudo-tracks for use in movement or habitat modelling applications.

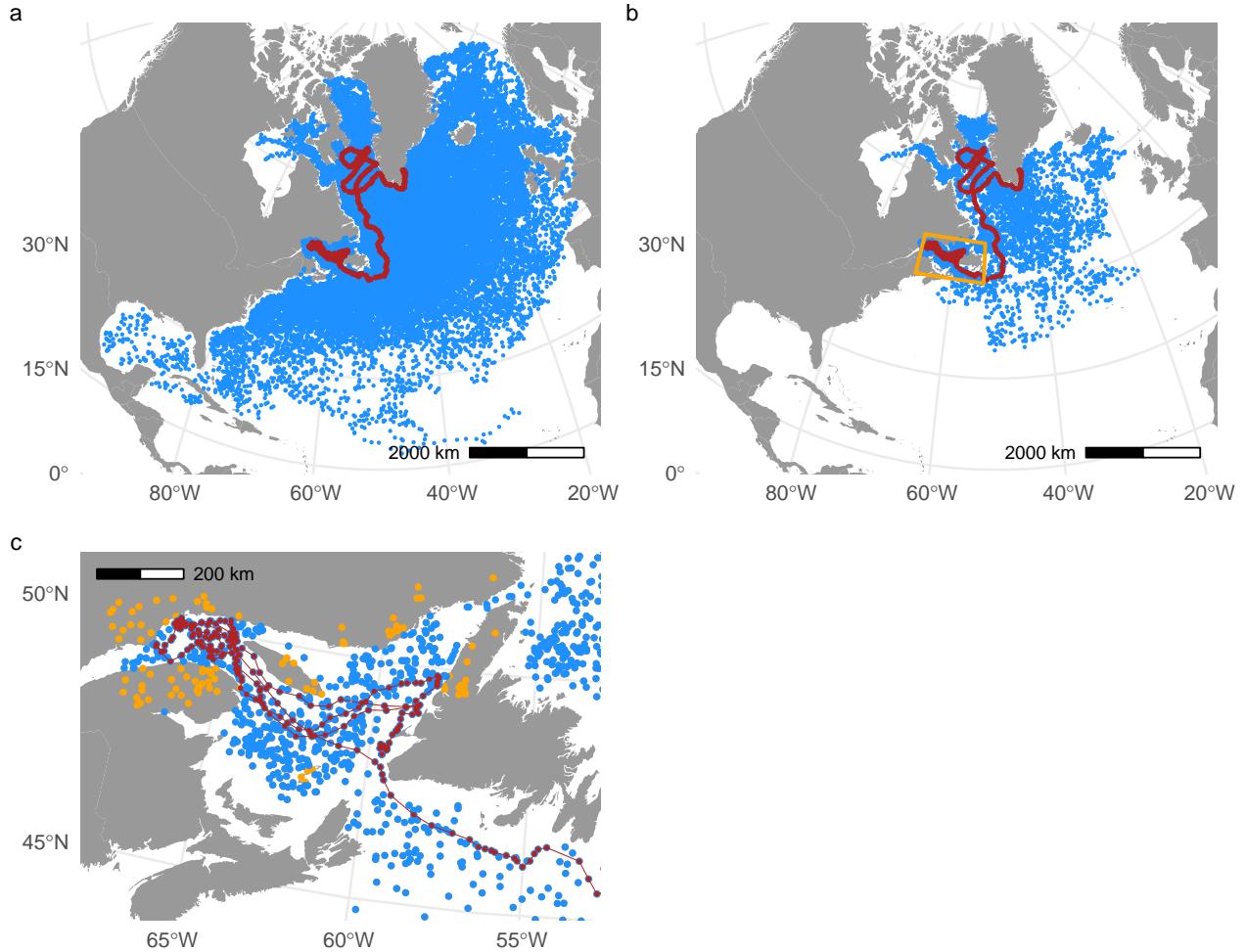


Figure 4: Simulating (a) 100 movement paths from a correlated random walk process model; (b) filtering those tracks to select the top 10% based on their similarity to the original SSM-predicted track (red); and (c) re-routing simulated locations on land (orange) back to ocean (blue). The orange box in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

²⁶⁶ **4 | Discussion**

²⁶⁷ The estimation speed and simplicity of the SSMs allows them to be fit as part of a fully
²⁶⁸ automated, real-time quality control/quality assurances process (Jonsen et al., 2020).

- ²⁶⁹ • re-routing paths is pragmatic, not a statistical solution... discuss pitfalls.
- ²⁷⁰ • discuss nuances of `fit_ssm` vs `fit_mpm` for move persistence estimation & relation of
²⁷¹ move persistence idea to discrete movement states, eg via HMM's
- ²⁷² • touch on applicability of fitting SSM to GLS data ... possibly ok-ish for seabirds but
²⁷³ not recommended for fish that spend majority of time at depth. Link is more about how
²⁷⁴ light-level data is converted to location, model-based methods eg WC can yield location
²⁷⁵ likelihoods that are highly non-Gaussian and therefore not amenable to subsequent
²⁷⁶ modelling with `foieGras`. Also touch on fitting a model to output from a model...

²⁷⁷ **Acknowledgements**

²⁷⁸ We thank Marie Auger-Méthé for contributing original code to the movement persistence
²⁷⁹ models. IDJ acknowledges support from a Macquarie University co-Funded Fellowship and
²⁸⁰ from partners: the US Office of Naval Research, Marine Mammal Program (grant N00014-18-
²⁸¹ 1-2405); the Integrated Marine Observing System (IMOS); Taronga Conservation Society;
²⁸² the Ocean Tracking Network; Birds Canada; and Innovasea/VEMCO. TAP was supported
²⁸³ by CSIRO Oceans & Atmosphere internal research funding scheme. The Integrated Marine
²⁸⁴ Observing System (IMOS) supported seal fieldwork. IMOS is a national collaborative research
²⁸⁵ infrastructure, supported by the Australian Government and operated by a consortium of
²⁸⁶ institutions as an unincorporated joint venture, with the University of Tasmania as Lead
²⁸⁷ Agent. Field work at Illes Kerguelen was conducted as part of the IPEV programme N°
²⁸⁸ 109 (PI H. WEIMERSKIRCH) and of the SNO-MEMO programme (PI C. GUINET) in
²⁸⁹ collaboration with IMOS. CTD tags were partly funded by CNES-TOSCA and IMOS. Little

290 penguin fieldwork was supported by an Australian Research Council Linkage grant to IDJ,
291 GC and RGH (LP160100162). All animal tagging procedures approved and executed under
292 the Animal Ethics Committee guidelines of the University of Tasmania (elephant seals),
293 Macquarie University (little penguins), and . . . University (harp seals).

294 **Author's Contributions**

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

299 **Data Accessibility**

300 All code and data used here are provided in the `foieGras` package for R or in the Supplemen-
301 tary Information. The latest stable and cross-platform tested version of the package (currently,
302 1.0-6) 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
303 GitHub repository: <https://github.com/ianjonsen/foieGras>. An older version of `foieGras`
304 (0.7-6) remains on CRAN at <https://CRAN.R-project.org/package=foieGras>, however, we
305 recommend users upgrade to the latest R-universe version for full access to the functionality
306 presented here.
307

308 **ORCID**

309 *Ian D Jonsen* <https://orcid.org/0000-0001-5423-6076>
310 *W James Grecian* <https://orcid.org/0000-0002-6428-719X>
311 *Lachlan Phillips* <https://orcid.org/0000-0002-7635-2817>
312 *Gemma Carroll* <https://orcid.org/0000-0001-7776-0946>

³¹³ Robert G Harcourt <https://orcid.org/0000-0003-4666-2934>

³¹⁴ Clive R McMahon <https://orcid.org/0000-0001-5241-8917>

³¹⁵ Toby A Patterson <https://orcid.org/0000-0002-7150-9205>

³¹⁶ References

- ³¹⁷ Aarts, G., Fieberg, J., & Matthiopoulos, J. (2012). Comparative interpretation of count,
³¹⁸ 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>
- ³²⁰ Albertsen, C. M., Whoriskey, K., Yurkowski, D., Nielsen, A., & Mills Flemming, J. (2015).
³²¹ Fast fitting of non-Gaussian state-space models to animal movement data via Template
³²² Model Builder. *Ecology*, 96, 2598–2604.
- ³²³ Auger-Méthé, M., Albertsen, C. M., Jonsen, I. D., Derocher, A. E., Lidgard, D. C., Studholme,
³²⁴ K. R., Bowen, W. D., Crossin, G. T., & Flemming, J. M. (2017). Spatiotemporal modelling
³²⁵ of marine movement data using Template Model Builder (TMB). *Marine Ecology Progress Series
Series*, 565, 237–249.
- ³²⁷ Baktoft, H., Gjelland, K. Ø., Økland, F., & Thygesen, U. H. (2017). Positioning of aquatic
³²⁸ animals based on time-of-arrival and random walk models using YAPS (yet another
³²⁹ positioning solver). *Scientific Reports*, 7, 14294. [https://doi.org/10.1038/s41598-017-14278-z](https://doi.org/10.1038/s41598-017-
330 14278-z)
- ³³¹ Calabrese, J. M., Fleming, C. H., & Gurarie, E. (2016). Ctmm: An R package for analyzing
³³² animal relocation data as a continuous-time stochastic process. *Methods in Ecology and
Evolution*, 7, 1124–1132.
- ³³⁴ Carroll, G., Cox, M., Harcourt, R., Pitcher, B., Slip, D., & Jonsen, I. (2017). Hierarchical
³³⁵ influences of prey distribution on patterns of prey capture by a marine predator. *Functional
Ecology*, 31, 1750–1760.
- ³³⁷ Dunnington, D. (2019). *Rosm: Plot raster map tiles from open street map and other sources*.
³³⁸ <https://CRAN.R-project.org/package=rosm>

- 339 Dunnington, D. (2021). *Ggspatial: Spatial data framework for ggplot2*. <https://CRAN.R-project.org/package=ggspatial>
- 340
- 341 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>
- 342
- 343
- 344
- 345 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.
- 346
- 347
- 348
- 349
- 350 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>
- 351
- 352
- 353
- 354 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.
- 355
- 356
- 357
- 358 Hooten, M. B., Scharf, H. R., & Morales, J. M. (2019). Running on empty: Recharge dynamics from animal movement data. *Ecology Letters*, 22, 377–389.
- 359
- 360 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.
- 361
- 362 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.
- 363
- 364
- 365 Jonsen, I. D., McMahon, C. R., Patterson, T. A., Auger-Méthé, M., Harcourt, R., Hindell, M.

- 366 A., & Bestley, S. (2019). Movement responses to environment: Fast inference of variation
367 among southern elephant seals with a mixed effects model. *Ecology*, *100*, e02566.
- 368 Jonsen, I. D., Patterson, T. A., Costa, D. P., Doherty, P. D., Godley, B. J., Grecian, W. J.,
369 Guinet, C., Hoenner, X., Kienle, S. S., Robinson, P. W., Votier, S. C., Whiting, S., Witt,
370 M. J., Hindell, M. A., Harcourt, R. G., & McMahon, C. R. (2020). A continuous-time
371 state-space model for rapid quality control of Argos locations from animal-borne tags.
372 *Movement Ecology*, *8*, 31.
- 373 Jonsen, I., Flemming, J., & Myers, R. (2005). Robust state-space modeling of animal
374 movement data. *Ecology*, *86*(11), 2874–2880.
- 375 Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an
376 eye on life and planet. *Science*, *348*(6240).
- 377 Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic
378 differentiation and Laplace approximation. *Journal of Statistical Software*, *70*, 1–21.
- 379 Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences
380 and their importance in species distribution modelling. *Ecography*, *33*(1), 103–114.
381 <https://doi.org/https://doi.org/10.1111/j.1600-0587.2009.06039.x>
- 382 London, J. M. (2020). *Pathroutr: An r package for (re-)routing paths around barriers*
383 (v0.1.1-beta) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.4321827>
- 384 McClintock, B. T., & Michelot, T. (2018). MomentuHMM: R package for generalized hidden
385 Markov models of animal movement. *Methods in Ecology and Evolution*, *9*, 1518–1530.
- 386 McMahon, C. R., Roquet, F., Baudel, S., Belbeoch, M., Bestley, S., Blight, C., Boehme, L.,
387 Carse, F., Costa, D. P., Fedak, M. A., Guinet, C., Harcourt, R., Heslop, E., Hindell, M.
388 A., Hoenner, X., Holland, K., Holland, M., Jaine, F. R. A., Jeanniard du Dot, T., ...
389 Woodward, B. (2021). Animal borne ocean sensors – AniBOS – an essential component of
390 the global ocean observing system. *Frontiers in Marine Science*, *8*, 751840.
- 391 Michelot, T., Langrock, R., Bestley, S., Jonsen, I. D., Photopoulou, T., & Patterson, T.
392 A. (2017). Estimation and simulation of foraging trips in land-based marine predators.

- 393 *Ecology*, 98(7), 1932–1944.
- 394 Michelot, T., Langrock, R., & Patterson, T. A. (2016). MoveHMM: An R package for the
395 statistical modelling of animal movement data using hidden Markov models. *Methods in
396 Ecology and Evolution*, 7, 1308–1315.
- 397 Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L., & King, R. (2017).
398 Statistical modelling of individual animal movement: An overview of key methods and a
399 discussion of practical challenges. *AStA Advances in Statistical Analysis*, 101, 399–438.
- 400 Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., & Matthiopoulos, J. (2008).
401 State-space models of individual animal movement. *Trends in Ecology and Evolution*, 23,
402 87–94.
- 403 Pedersen, T. L. (2020). *Patchwork: The composer of plots*. <https://CRAN.R-project.org/package=patchwork>
- 404 Phillips, L., Carroll, G., Jonsen, I. D., Harcourt, R. G., Brierley, A., Wilkins, A., & Cox, M.
405 (2021). Variability in prey field structure drives inter-annual differences in prey encounter
406 by a marine predator, the little penguin. *Proceedings of the Royal Society B, In Review*.
- 407 Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier,
408 S. (2009). Sample selection bias and presence-only distribution models: implications for
409 background and pseudo-absence data. *Ecological Applications*, 19(1), 181–197. <https://doi.org/10.1890/07-2153.1>
- 410 R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation
411 for Statistical Computing. <https://www.R-project.org/>
- 412 Raymond, B., Lea, M.-A., Patterson, T., Andrews-Goff, V., Sharples, R., Charrassin, J.-B.,
413 Cottin, M., Emmerson, L., Gales, N., Gales, R., Goldsworthy, S. D., Harcourt, R., Kato,
414 A., Kirkwood, R., Lawton, K., Ropert-Coudert, Y., Southwell, C., Hoff, J. van den,
415 Wienecke, B., ... Hindell, M. A. (2015). Important marine habitat off east Antarctica
416 revealed by two decades of multi-species predator tracking. *Ecography*, 38, 121–129.
- 417 Schick, R. S., New, L. F., Thomas, L., Costa, D. P., Hindell, M. A., McMahon, C. R., Robinson,

- 420 P. W., Simmons, S. E., Thums, M., J., Harwood., & Clark, J. S. (2013). Estimating
421 resource acquisition and at-sea body condition of a marine predator. *Journal of Animal*
422 *Ecology*, 82, 1300–3015.
- 423 Service Argos. (2016). *Argos Users' Manual*. CLS. <http://www.argos-system.org/manual>
- 424 South, A. (2022a). *Rnaturalearth: World map data from natural earth*.
- 425 South, A. (2022b). *Rnaturalearthhires: High resolution world vector map data from natural*
426 *earth used in rnaturalearth*.
- 427 Sumner, M. D., Wotherspoon, S. J., & Hindell, M. A. (2009). Bayesian estimation of animal
428 movement from archival and satellite tags. *PLoS ONE*, 4(10). <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0007324>
- 430 Thygesen, U. H., Albertsen, C. M., Berg, C. W., Kristensen, K., & Nielsen, A. (2017).
431 Validation of ecological state space models using the Laplace approximation. *Environmental*
432 *and Ecological Statistics*. <https://doi.org/DOI: 10.1007/s10651-017-0372-4>
- 433 Whoriskey, K., Auger-Méthé, M., Albertsen, C. M., Whoriskey, F. G., Binder, T. R., Kreuger,
434 C. C., & Mills Flemming, J. (2017). A hidden markov movement model for rapidly
435 identifying behavioral states from animal tracks. *Ecology and Evolution*, 7, 2112–2121.
- 436 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.
437 <https://ggplot2.tidyverse.org>