

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

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

32 *transition paragraph focusing on analytical methods here*

33 *transition paragraph focusing on related R packages & defining the foieGras*

34 **niche here**

35 A diverse set of software tools now exist for processing and analysis of animal telemetry
36 data. The CRAN Task View: Processing and Analysis of Tracking Data (<https://CRAN.R-project.org/view=Tracking>) lists 42 R packages available on CRAN, including **foieGras**, and
37 several more available from other publicly accessible repositories (Joo et al., 2020). Of the R
38 packages available on the CRAN repository, **crawl** (Johnson et al., 2008), **ctmm** (Calabrese
39 et al., 2016), and **bsam** (Jonsen, 2016) fill similar data quality control and analysis roles, by
40 providing continuous- or discrete-time state-space modelling tools.

42 The **foieGras**, pronounced “*fwah grah*”, package for R (R Core Team, 2021) was developed
43 to be as simple and fast to use as possible. The package implements state-space models to
44 conduct quality control on animal (re)location data collected via the Argos satellite (Service
45 Argos, 2016) and other telemetry systems. The latest stable and fully cross-platform tested
46 version of the package (currently, 1.0.0) is available on the Comprehensive R Archive Network
47 (CRAN), at <https://cran.r-project.org/package=foieGras>. The latest partially tested
48 stable and development versions are available on the lead author’s GitHub repository: <https://github.com/ianjonsen/foieGras>.

50 Here, we describe the main features of **foieGras** and illustrate its use through examples
51 using both real and simulated data. Full R code for each of the examples is provided in the
52 Supporting Information. Additional details on package functions and their use can be found
53 in their help files and in the package’s vignettes.

54 **2 | foieGras overview**

55 The workflow for **foieGras** is deliberately simple, with many of the usual track data process-
56 ing checks and formatting handled automatically. The main functions are listed in Table 1.
57 When fitting a model, **foieGras** automatically detects the type of tracking data from the
58 location quality class designations that are typical of Argos data and that can be added to the
59 data by the researcher for other types of track data. Based on the location quality classes and
60 other, optional information on measurement errors contained in the data, **foieGras** chooses
61 an appropriate measurement error model for each observation. This capability allows for
62 combinations of different tracking data types, e.g., Argos and GPS, in a single input data
63 frame and to be fit in a single state-space model.

64 **2.1 | Data preparation**

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

71 Other types of track data can be accommodated, for example, by including the **lc** column
72 where all **lc = "G"** for GPS data. In this case, measurement error in the GPS locations is

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

| Function | Description |
|--------------------------|---|
| <code>fit_ssm</code> | Fit a State-Space Model to location data |
| <code>route_path</code> | Reroute path so estimated locations are off land |
| <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 ‘fG_ssm’ fit objects |
| <code>sim_filter</code> | Filter tracks simulated with ‘simfit’ according to similarity criteria |
| <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 |

73 assumed to have a standard deviation of 0.1 x Argos class 3 locations (approximately 30 m).
74 Other types of track data can be also be considered, including those derived from light-level
75 geolocation (see the package vignette for further details).

76 **2.2 | State-space model fitting - fit_ssm**

77 State-space models are fit using `fit_ssm`. There are a large number of options that can be
78 set in `fit_ssm` (see Suppl for details). We focus only the essential options here:

- 79 • `data` the input data structured as described in **2.1**
- 80 • `vmax` a maximum threshold speed (ms^{-1}) to help identify potential outlier locations
- 81 • `model` the process model to be used
- 82 • `time.step` the prediction time interval (h)

83 The function first invokes an automated data processing stage where the following occurs:
84 1) data type (Argos Least-Squares, Argos Kalman Filter/Smoother, GPS, or General (e.g.,
85 processed light-level geolocations, acoustic telemetry, coded VHF telemetry) is determined;
86 2) datetimes are converted to POSIXt format, chronological order is ensured, and duplicate
87 datetime records are removed; 3) observations occurring less than `min.dt` seconds after a
88 prior observation are removed; 4) a speed filter [`sda` from the `trip` R package; Sumner et al.
89 (2009)] is used to identify potential outlier locations; 5) locations are projected from spherical
90 lon-lat coordinates to planar x,y coordinates in km.

91 The function then fits a state-space model to the processed data, where the process model
92 (currently, either a continuous-time `rw` or a continuous-time `crw`) is specified by the user
93 and the measurement model(s) are selected automatically (see Jonsen et al., 2020 for model
94 details). The model is fit by numerical optimization of the likelihood using either the `optim`
95 or `nlminb` R function. The R package TMB, Template Model Builder (Kristensen et al.,
96 2016), is used to compute the gradient function in C++ via reverse-mode auto-differentiation
97 and the Laplace Approximation is used to integrate out the latent states (random effects).
98 Fits to a single versus multiple individuals are handled automatically, with sequential SSM
99 fits occurring in the latter case. No hierarchical or pooled estimation among individuals is
100 currently available.

101 `fit_ssm` returns a `foieGras` fit object (a nested data frame with class `fG_ssm`). The outer
102 data frame lists the individual id(s), basic convergence information and a list with class `ssm`.
103 This list contains dense information on the model parameter and state estimates, predictions,
104 processed data, optimizer results, and other diagnostic and contextual information. Users can
105 extract a simple data frame of SSM fitted (location estimates corresponding to the, typically
106 irregular, observation times) or predicted values (locations predicted at regular `time.step`
107 intervals) using the `grab` function.

108 **2.3 | Model checking and visualisation - osar, plot, fmap**

109 Before using fitted or predicted locations, a model fit should be checked and visualised to
110 confirm that the model adequately describes the data. In linear regression and a variety of
111 analogous methods, goodness-of-fit can be assessed by calculating standard residuals such

as Pearson or deviance residuals. 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 and uses 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 `fG_osar` object.

A generic `plot` method provides an easy way to visualise the `fG_osar` residuals. Time-series plots of the prediction residuals can be used to detect temporal changes in goodness-of-fit. Quantile-quantile plots of residuals against standard normal quantiles can be used to detect departures from normality. Sample autocorrelation function plots of the residuals are useful for detecting autocorrelation not accounted for by the model. Assessing residual autocorrelation can be particularly important as Argos locations, for example, are themselves derived from a time-series model (Lopez et al., 2015) which can introduce additional autocorrelation in the location errors.

State-space model fits to data can also be visualised by using the generic `plot` function on an `fG_ssm` data frame. Options exist to plot fitted or predicted values along with observations as either paired, 1-D time-series or as 2-D tracks with confidence intervals or ellipses, respectively. These plots provide a more intuitive and rapid method for assessing SSM fits to data, however, they do not replace the residual diagnostics. Fitted `fG_ssm` data frames can be mapped using the `fmap` function for single or multiple individuals. Estimated tracks can be displayed with or without confidence ellipses, observations, and/or a projection and maps of single tracks can be coloured by date.

2.4 | Behavioural estimation - `fit_mpm`

The `fit_mpm` function fits a simple move persistence model to estimate a continuous-valued, time-varying latent variable that indexes changes in movement behaviour (Jonsen et al., 2019). This variable measures the autocorrelation in speed and direction between consecutive pairs of movements such that high values correspond to fast, directed movements at one end of the continuum and low values correspond to slow, tortuous movements at the other end. It's important to note that this approach is unlike hidden Markov models (McClintock & Michelot, 2018; Michelot et al., 2016) and some state-space models (Jonsen, 2016) as there is no notion of discrete behavioural states that animals periodically switch between. Nonetheless, move persistence can be used to identify objectively places where animals spend disproportionately more or less time, and with extensions be correlated with environment or other covariates (See Examples 3.x).

The move persistence model assumes that locations are absent of measurement error and can occur either irregularly or regularly in time. `fit_mpm` takes either a `fG_ssm` data frame as input or a data frame with the follow variables: `id`, `date`, `x`, `y`, where `x` and `y` coordinates can be planar `x,y` or spherical `long,lat`. This latter input format allows the model to be fit easily to GPS or other tracking data with negligible measurement error. When the data contain multiple individuals, the default model is fit jointly by assuming all individuals share the same move persistence variance parameter (Jonsen, 2016). There is an option to fit the

model separately to each individual. The time-series of estimated move persistence with confidence intervals can be visualized by using the generic `plot` function with the resulting `fG_mpm` data frame. Visualization of move persistence along the 2-D tracks can be plotted or mapped by using the `plot` or `fmap` functions, respectively, and supplying both the `fG_mpm` and `fG_ssm` nested data frames. When using `fit_mpm` on, for example, GPS tracking data that do not require state-space filtering, the movement persistence estimates can be extracted from the `fG_mpm` data frame using the `grab` function and subsequently merged with the observed track data for visualization.

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). Michelot et al. (2016) advocate comparison of simulated tracks from fitted hidden Markov models to the observed tracks as a means of identifying potential weakness in the hidden Markov model formulation. Here, we suggest that the `rw` and `crw` state-space models and the `mpm` model can be fit to track data simulated from different movement processes to evaluate robustness of location and movement persistence estimates to model mis-specification. We illustrate this idea in section 3.x by drawing on flexibility in the `sim` function that allows a variety of movement processes to be simulated.

Simulation is also used frequently to provide a measure of habitat availability (Aarts et al., 2012) by providing a source of ‘pseudo-absence’ 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 `fG_ssm` 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 a central place forager by ensuring that the simulated tracks start and end at approximately the same location.

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.

3 | Examples

We illustrate the main capabilities of `foieGras` through a series of examples using real and simulated tracking data. These examples are for demonstration purposes and not intended as a comprehensive guide for conducting analyses with `foieGras`. Complete code for reproducing the examples and for gaining a deeper understanding of `foieGras` functions are provided

194 as supplements.

195 3.1 | Southern Elephant seal - SSM validation with prediction residuals

196 We use a subadult male southern elephant seal track included in `foieGras` (`sese2`), sourced
197 from the Australian Integrated Marine Observing System (IMOS; data publicly available
198 via imos.aodn.org.au) deployments at Iles Kerguelen in collaboration with the French IPEV
199 and SNO-MEMO programmes. The data are temporally irregular Argos Least-Squares based
200 locations, 73 % of which are in the poorest location quality classes: A and B. We fit both
201 the `rw` and `crw` models using `fit_ssm` with a speed filter threshold (`vmax`) of 4 ms⁻¹ and a
202 12-h time step. We calculate prediction residuals using `osar`, and then use the generic `plot`
203 method for `osar` residuals to assess and compare the model fits (Fig. 1).

204 The plots of predicted states on top of the observations suggest both models yield similar
205 fits (Fig. 1a). However, corresponding predicted locations from the two models differ by a
206 median 6.62 km (range: 0.02, 53.02 km; Fig. 1b), and there are marked trends in the time-
207 series of residuals for the `rw` model fit (Fig. 1c) and the `rw` ACF's reveal consistent positive
208 autocorrelation in the prediction residuals (Fig. 1e). The corresponding `crw` prediction
209 residuals show little trend through time and have relatively little autocorrelation (Fig. 1d,f),
210 implying that the `crw` process model provides a better fit to the data.

211 3.2 | Inferring movement persistence from Argos and GPS data

212 Drawing on an expanded version of the data used in 3.1, we quality control and infer move-
213 ment persistence, γ_t , along five southern elephant seal tracks. We fitted the `mp` SSM with
214 a 24-h prediction interval using `fit_ssm`, assuming bivariate normal location measurement
215 errors consistent with Argos Least-Squares-derived locations (Jonsen et al., 2020). The `mp`
216 SSM simultaneously estimates locations and γ_t , and their uncertainties. We then assess how
217 γ_t changes along the seals' tracks.

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

226 Movement persistence estimates along the quality-controlled southern elephant seal tracks
227 highlight some fundamental differences in movement pattern among the seals. The two seals
228 engaging in pelagic foraging trips (Fig. 2a,c and f) had less contrast in their movements
229 with consistently higher γ_t estimates compared to the three seals engaging in trips to the
230 fast-ice on the Antarctic shelf (Fig. 2b,d-e and f). Although γ_t 's were higher overall for the
231 pelagically foraging seals, they both spent little time making fast, highly directed movements
232 ($\gamma_t \rightarrow 1$) relative to the shelf-foraging seals (2a,c vs b,d-e). This suggests the pelagically-
233 foraging seals may spend considerable time searching for suitable foraging habitat in the

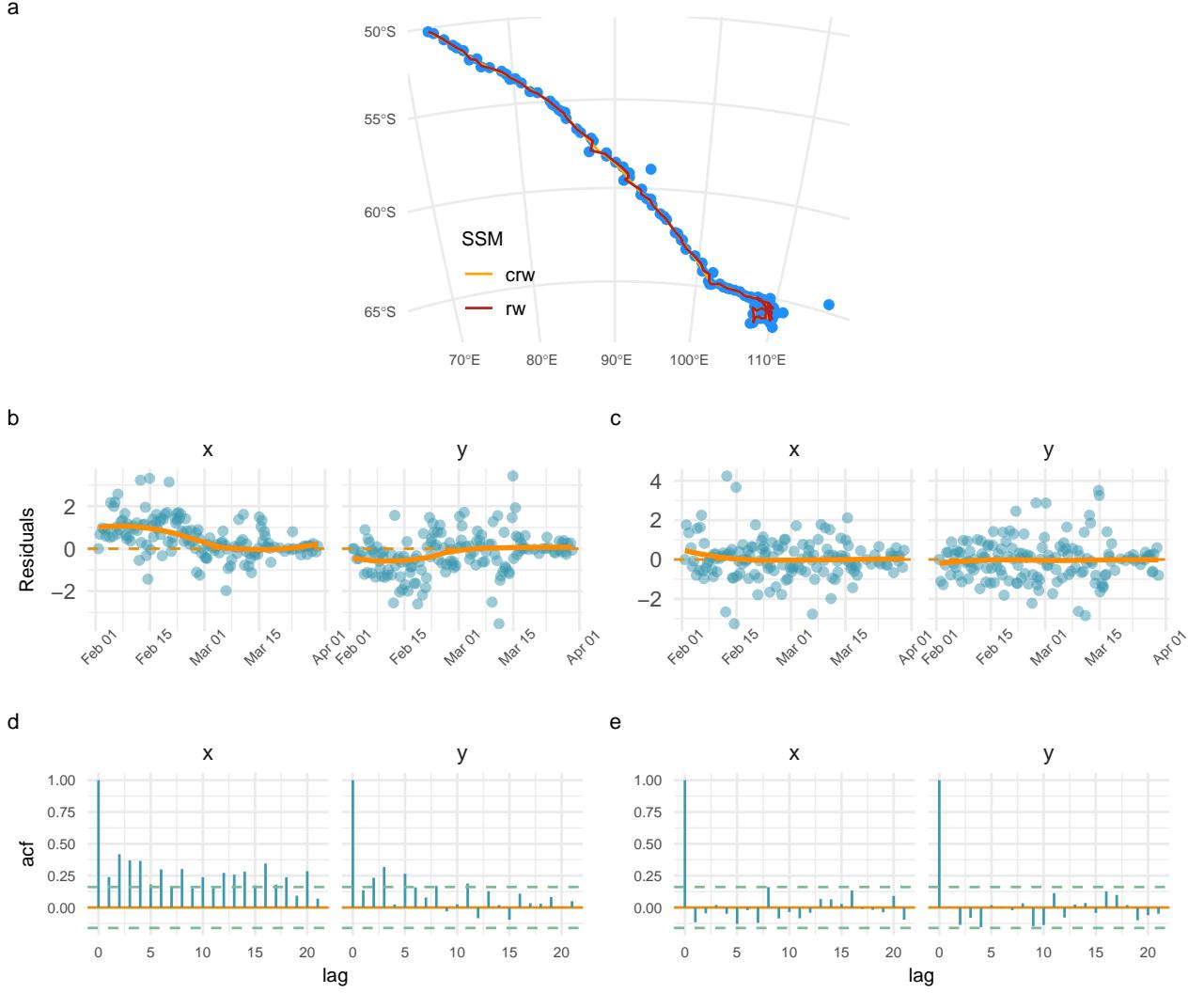


Figure 1: Selected diagnostic plots for assessing `rw` (a,c,e) and `crw` (b,d,f) state-space model fits to a southern elephant seal track. Top panels (a,b) are plots of predicted states (red; regular 12-h time intervals) and observations (blue) with pre-filtered observations (orange; ignored by the SSM), using the `plot.fG_ssm` function. Panels c,d are time-series plots of the prediction residuals for the x and y coordinates of each fitted state. Panels e,f are autocorrelation functions of the prediction residuals. All residual plots generated using the `plot.fG_osar` function.

234 highly variable eddy fields between the Subantarctic and Polar Fronts (Jonsen et al., 2019),
235 whereas foraging habitat may be more predictable for seals travelling rapidly and directly to
236 the Antarctic shelf region. These seals may also haulout periodically on available fast-ice to
237 rest. This behaviour could also contribute to the higher contrast in movement persistence,
238 relative to pelagically-foraging seals who would not have access to fast-ice.

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

245 3.3 | Simulating tracks from `foieGras` model fits

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

251 From this process model we simulate 50 animal movement paths using `simfit` and apply
252 a potential function using the `grad` and `beta` arguments to constrain the simulated paths
253 to water. These tracks are then filtered based on their similarity to the original path using
254 `sim_filter` and the top 10% retained (`keep = 0.1`) (Fig. 4).

255 In combination, these functions provide a user-friendly method to generate and objectively
256 filter pseudo-tracks for use in movement or habitat modelling applications.

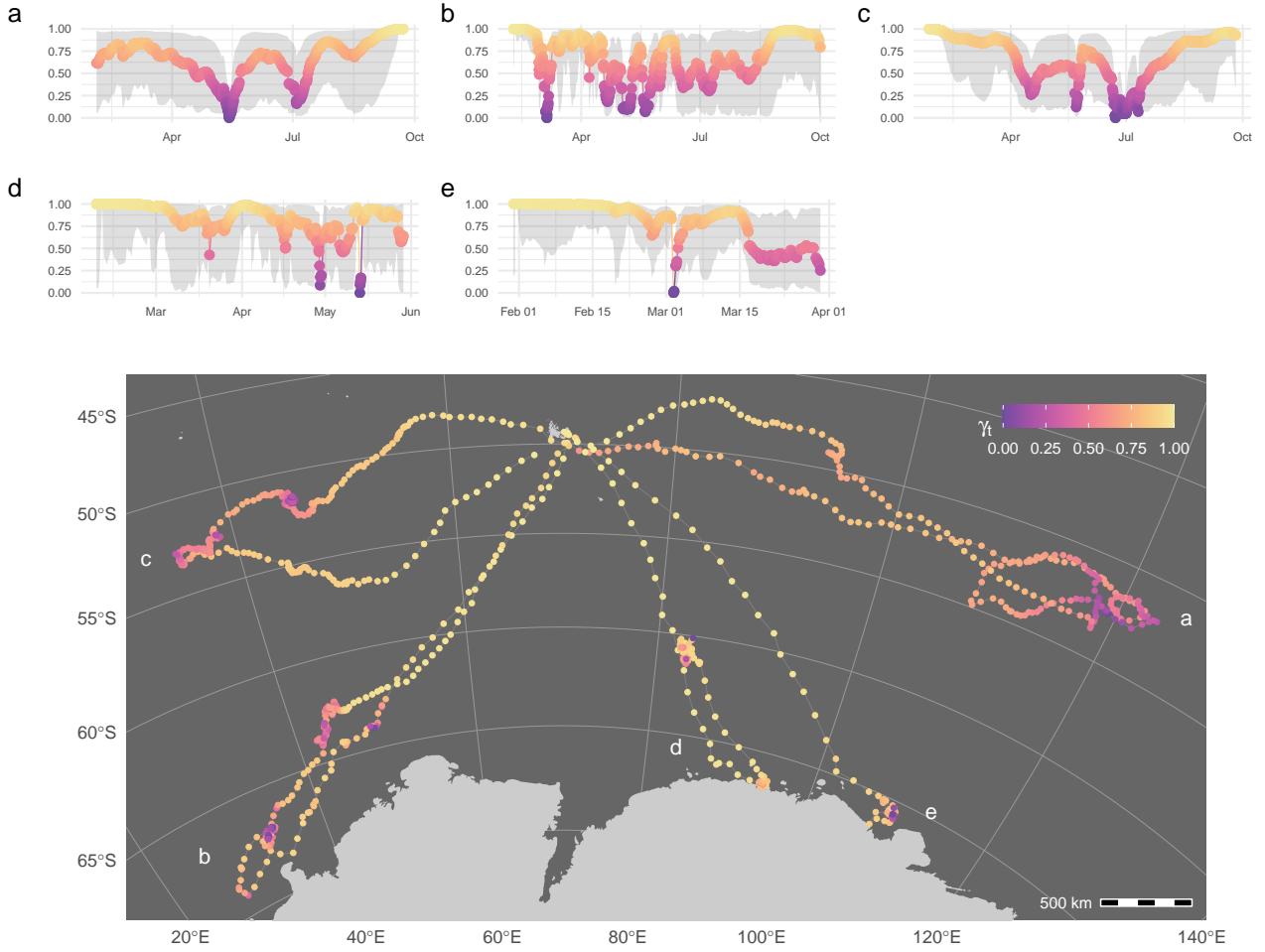


Figure 2: Inferred move persistence, γ_t , 1-D time-series for five southern elephant seals (a-e; grey envelopes are 95 % CI's) and along their 2-D tracks (f; track labels, a-e, correspond to the 1-D time-series plots). Locations associated with low move persistence (blue) are indicative of slow, undirected movements, whereas high move persistence (yellow) is indicative of faster, directed movements. The lowest move persistence tends to occur at the distal end of foraging trips from the colony at Iles Kerguelen, suggesting these bouts of low movement persistence are associated with foraging activity. Due to the stereographic projection used and huge area covered in (f), the scale bar is not accurate in all regions and is indicative only.

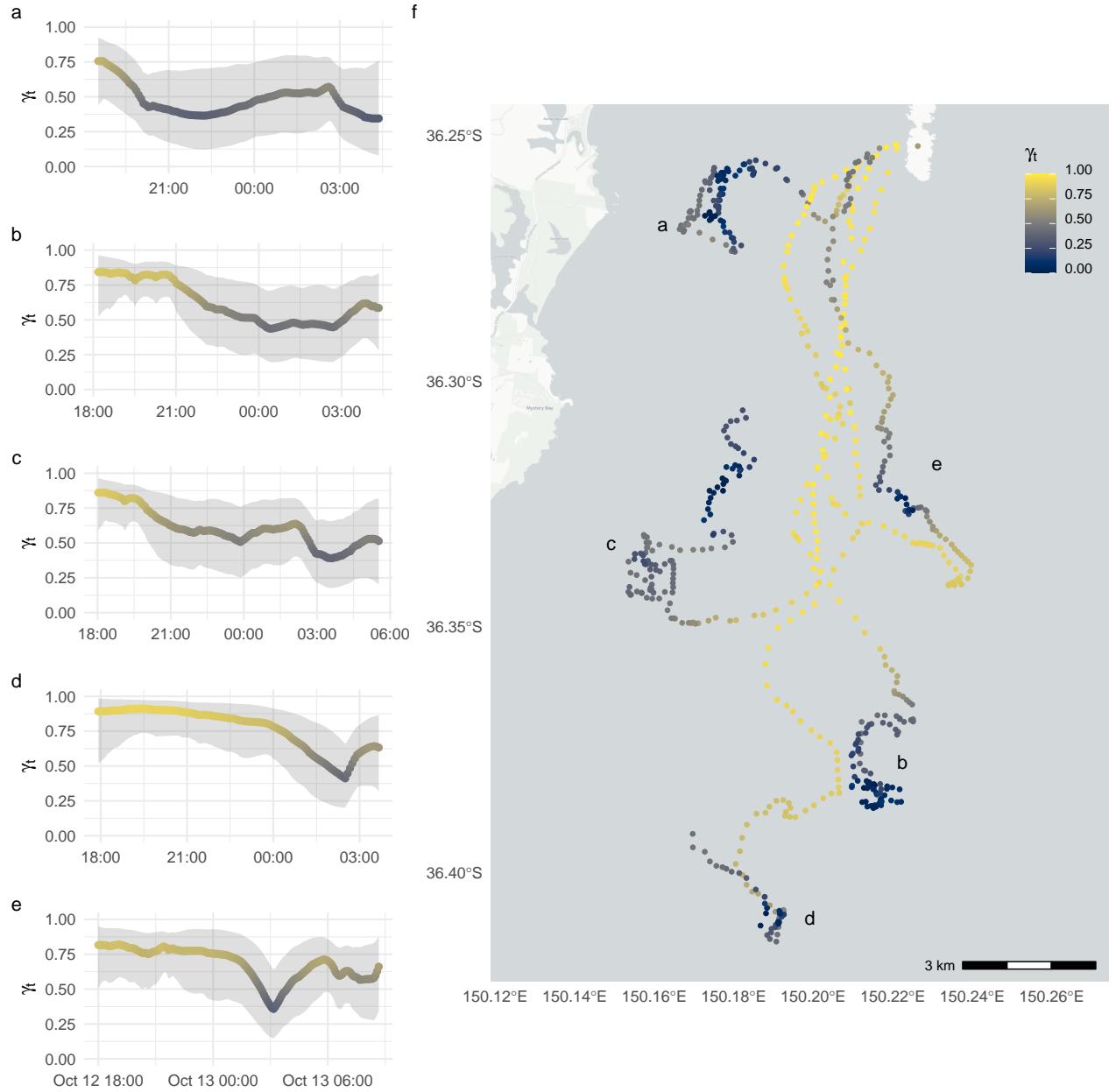


Figure 3: Inferred move persistence, γ_t , 1-D time-series (a-f; grey envelopes are 95 % CI's) and along little penguin GPS tracks (g). Colour palette as in 2. Movement persistence was estimated from SSM-predicted locations with a regular 5-min interval.

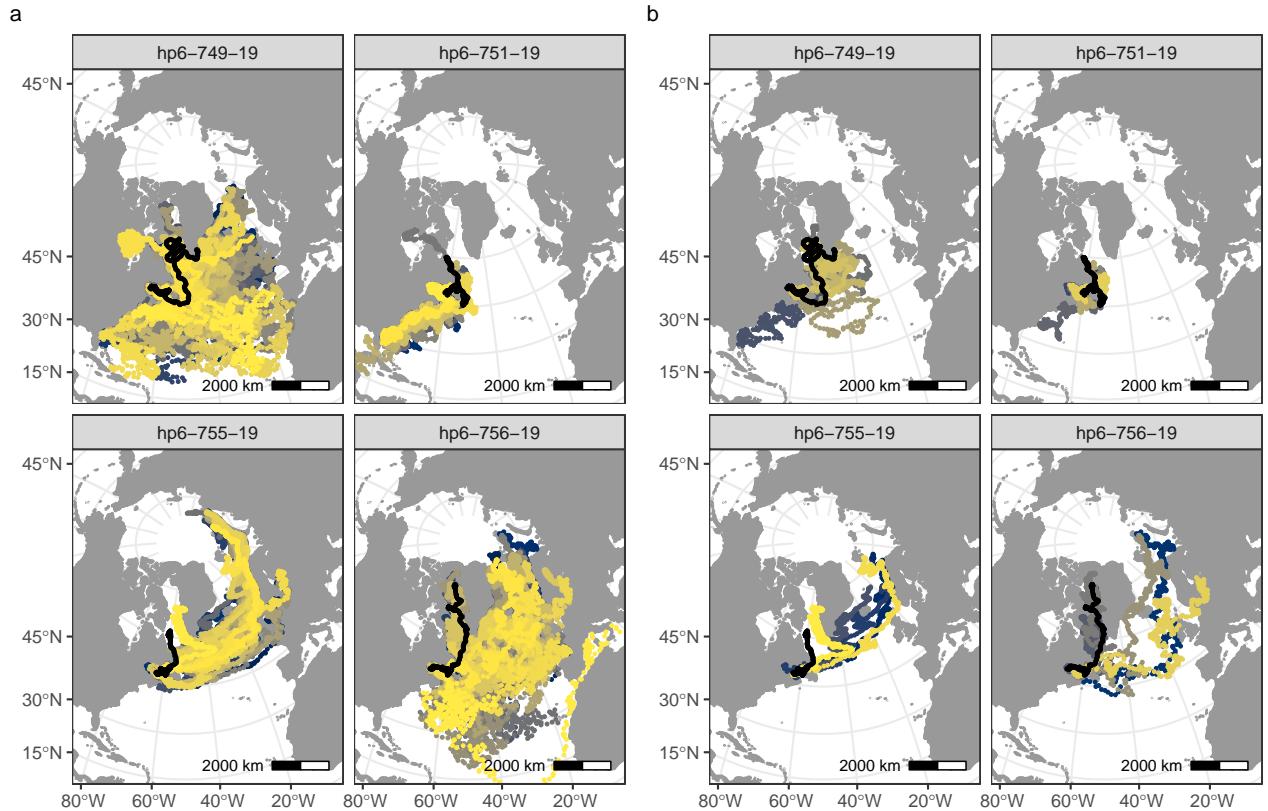


Figure 4: Simulating (a) 50 movement paths from a correlated random walk process model and then (b) filtering those tracks to select the top 10% based on their similarity to the original fitted track (shown in black).

257 **4 | Discussion**

258 Ex 3.2 In a limited way, this provides information on the robustness of the `foieGras` SSM's
259 to different kinds plausible animal movements

260 **Acknowledgements**

261 We thank Marie Auger-Méthé for contributing original code to the movement persistence
262 models. IDJ acknowledges support from a Macquarie University co-Funded Fellowship and
263 from partners: the US Office of Naval Research, Marine Mammal Program (grant N00014-
264 18-1-2405); the Integrated Marine Observing System (IMOS); Taronga Conservation Society;
265 the Ocean Tracking Network; Birds Canada; and Innovasea/VEMCO. TAP was supported
266 by CSIRO Oceans & Atmosphere internal research funding scheme. The Integrated Ma-
267 rine Observing System (IMOS) supported seal fieldwork. IMOS is a national collaborative
268 research infrastructure, supported by the Australian Government and operated by a consor-
269 tium of institutions as an unincorporated joint venture, with the University of Tasmania as
270 Lead Agent. Field work at Illes Kerguelen was conducted as part of the IPEV programme
271 N° 109 (PI H. WEIMERSKIRCH) and of the SNO-MEMO programme (PI C. GUINET) in
272 collaboration with IMOS. CTD tags were partly funded by CNES-TOSCA and IMOS. Little
273 penguin fieldwork was supported by an Australian Research Council Linkage grant to IDJ,
274 GC and RGH (LP160100162). All animal tagging procedures approved and executed un-
275 der the Animal Ethics Committee guidelines of the University of Tasmania (elephant seals),
276 Macquarie University (little penguins), and ... University (species).

277 **Author's Contributions**

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

282 **Data Accessibility**

283 All code mentioned here is provided in the `foieGras` package for R available on CRAN at
284 <https://CRAN.R-project.org/package=foieGras>. The development version of the package is
285 available on GitHub at <https://github.com/ianjonsen/foieGras>. Data used in the examples
286 are available at...

287 **ORCID**

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

- 292 Robert G Harcourt <https://orcid.org/0000-0003-4666-2934>
293 Clive R McMahon <https://orcid.org/0000-0001-5241-8917>
294 Toby A Patterson <https://orcid.org/0000-0002-7150-9205>

295 References

- 296 Aarts, G., Fieberg, J., & Matthiopoulos, J. (2012). Comparative interpretation of count,
297 presence-absence and point methods for species distribution models. *Methods in Ecology
298 and Evolution*, 3(1), 177–187. <https://doi.org/10.1111/j.2041-210x.2011.00141.x>
299 Calabrese, J. M., Fleming, C. H., & Gurarie, E. (2016). Ctmm: An R package for analyzing
300 animal relocation data as a continuous-time stochastic process. *Methods in Ecology and
301 Evolution*, 7, 1124–1132.
302 Carroll, G., Cox, M., Harcourt, R., Pitcher, B., Slip, D., & Jonsen, I. (2017). Hierarchical in-
303 fluences of prey distribution on patterns of prey capture by a marine predator. *Functional
304 Ecology*, 31, 1750–1760.
305 Grecian, W. J., Stenson, G. B., Biuw, M., Boehme, L., Folkow, L. P., Goulet, P. J., Jonsen,
306 I. D., Malde, A., Nordøy, E. S., Rosing-Asvid, A., & Smout, S. (2022). Environmental
307 drivers of population-level variation in the migratory and diving ontogeny of an arctic
308 top predator. *Royal Society Open Science*, 9(3). <https://doi.org/10.1098/rsos.211042>
309 Harcourt, R., Sequeira, A. M. M., Zhang, X., Roquet, F., Komatsu, K., Heupel, M., McMa-
310 hon, C., Whoriskey, F., Meekan, M., Carroll, G., Brodie, S., Simpfendorfer, C., Hindell,
311 M., Jonsen, I., Costa, D. P., Block, B., Muelbert, M., Woodward, B., Weise, M., ... Fedak,
312 M. A. (2019). Animal-borne telemetry: An integral component of the ocean observing
313 toolkit. *Frontiers in Marine Science*, 6, 326.
314 Hazen, E. L., Palacios, D. M., Forney, K. A., Howell, E. A., Becker, E., Hoover, A. L., Irvine,
315 L., DeAngelis, M., Bograd, S. J., Mate, B. R., & Bailey, H. (2017). WhaleWatch: a
316 dynamic management tool for predicting blue whale density in the California Current.
317 *Journal of Applied Ecology*, 54(5), 1415–1428. <https://doi.org/10.1111/1365-2664.12820>
318 Hindell, M. A., Reisinger, R. R., Ropert-Coudert, Y., Hückstädt, L. A., Trathan, P. N.,
319 Bornemann, H., Charrassin, J.-B., Chown, S. L., Costa, D. P., Danis, B.others. (2020).
320 Tracking of marine predators to protect southern ocean ecosystems. *Nature*, 580(7801),
321 87–92.
322 Johnson, D. S., London, J. M., Lea, M.-A., & Durban, J. W. (2008). Continuous-time
323 correlated random walk model for animal telemetry data. *Ecology*, 89(5), 1208–1215.
324 Jonsen, I. D. (2016). Joint estimation over multiple individuals improves behavioural state
325 inference from animal movement data. *Scientific Reports*, 6, 20625.
326 Jonsen, I. D., McMahon, C. R., Patterson, T. A., Auger-Méthé, M., Harcourt, R., Hindell, M.
327 A., & Bestley, S. (2019). Movement responses to environment: Fast inference of variation
328 among southern elephant seals with a mixed effects model. *Ecology*, 100, e02566.
329 Jonsen, I. D., Patterson, T. A., Costa, D. P., Doherty, P. D., Godley, B. J., Grecian, W. J.,
330 Guinet, C., Hoennner, X., Kienle, S. S., Robinson, P. W., Votier, S. C., Whiting, S., Witt,
331 M. J., Hindell, M. A., Harcourt, R. G., & McMahon, C. R. (2020). A continuous-time
332 state-space model for rapid quality control of Argos locations from animal-borne tags.
333 *Movement Ecology*, 8, 31.
334 Joo, R., Boone, M. E., Clay, T. A., Patrick, S. C., Clusella-Trullas, S., & Basille, M. (2020).

- 335 Navigating through the r packages for movement. *Journal of Animal Ecology*, 89, 248–
336 267.
- 337 Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as an
338 eye on life and planet. *Science*, 348(6240).
- 339 Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic
340 differentiation and Laplace approximation. *Journal of Statistical Software*, 70, 1–21.
- 341 Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of ab-
342 sences and their importance in species distribution modelling. *Ecography*, 33(1), 103–114.
343 <https://doi.org/https://doi.org/10.1111/j.1600-0587.2009.06039.x>
- 344 Lopez, R., Malardé, J.-P., Danès, P., & Gaspar, P. (2015). Improving Argos Doppler location
345 using multiple-model smoothing. *Animal Biotelemetry*, 3, 32.
- 346 McClintock, B. T., & Michelot, T. (2018). MomentuHMM: R package for generalized hidden
347 Markov models of animal movement. *Methods in Ecology and Evolution*, 9, 1518–1530.
- 348 McMahon, C. R., Roquet, F., Baudel, S., Belbeoch, M., Bestley, S., Blight, C., Boehme, L.,
349 Carse, F., Costa, D. P., Fedak, M. A., Guinet, C., Harcourt, R., Heslop, E., Hindell, M.
350 A., Hoenner, X., Holland, K., Holland, M., Jaine, F. R. A., Jeanniard du Dot, T., ...
351 Woodward, B. (2021). Animal borne ocean sensors – AniBOS – an essential component
352 of the global ocean observing system. *Frontiers in Marine Science*, 8, 751840.
- 353 Michelot, T., Langrock, R., Bestley, S., Jonsen, I. D., Photopoulou, T., & Patterson, T.
354 A. (2017). Estimation and simulation of foraging trips in land-based marine predators.
355 *Ecology*, 98(7), 1932–1944.
- 356 Michelot, T., Langrock, R., & Patterson, T. A. (2016). MoveHMM: An R package for the
357 statistical modelling of animal movement data using hidden Markov models. *Methods in
358 Ecology and Evolution*, 7, 1308–1315.
- 359 Phillips, L., Carroll, G., Jonsen, I. D., Harcourt, R. G., Brierley, A., Wilkins, A., & Cox, M.
360 (2021). Variability in prey field structure drives inter-annual differences in prey encounter
361 by a marine predator, the little penguin. *Proceedings of the Royal Society B, In Review*.
- 362 Phillips, S. J., Dudík, M., Elith, J., Graham, C. H., Lehmann, A., Leathwick, J., & Ferrier,
363 S. (2009). Sample selection bias and presence-only distribution models: implications for
364 background and pseudo-absence data. *Ecological Applications*, 19(1), 181–197. <https://doi.org/10.1890/07-2153.1>
- 366 R Core Team. (2021). *R: A language and environment for statistical computing*. R Founda-
367 tion for Statistical Computing. <https://www.R-project.org/>
- 368 Raymond, B., Lea, M.-A., Patterson, T., Andrews-Goff, V., Sharples, R., Charrassin, J.-
369 B., Cottin, M., Emmerson, L., Gales, N., Gales, R., Goldsworthy, S. D., Harcourt, R.,
370 Kato, A., Kirkwood, R., Lawton, K., Ropert-Coudert, Y., Southwell, C., Hoff, J. van den,
371 Wienecke, B., ... Hindell, M. A. (2015). Important marine habitat off east Antarctica
372 revealed by two decades of multi-species predator tracking. *Ecography*, 38, 121–129.
- 373 Service Argos. (2016). *Argos Users' Manual*. CLS. <http://www.argos-system.org/manual>
- 374 Sumner, M. D., Wotherspoon, S. J., & Hindell, M. A. (2009). Bayesian estimation of animal
375 movement from archival and satellite tags. *PLoS ONE*, 4(10). <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0007324>
- 377 Thygesen, U. H., Albertsen, C. M., Berg, C. W., Kristensen, K., & Nielsen, A. (2017). Vali-
378 dation of ecological state space models using the Laplace approximation. *Environmental
379 and Ecological Statistics*. <https://doi.org/DOI: 10.1007/s10651-017-0372-4>