`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], and Toby A. Patterson[5]

[1]Department of Biological Sciences, Macquarie University, Sydney, NSW, Australia
[2]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]CSIRO Ocean and Atmosphere Research, Hobart, TAS, Australia
[*]corresponding author, ian.jonsen@mq.edu.au

Abstract

1.
2.
3.
4.

Keywords:

1 | Introduction

Animal biotelemetry as a discipline has matured, with telemetry data now virtually essential for understanding behaviour and social interactions, foraging ecology, habitat use and population dynamics of mobile and/or cryptic species. Additionally, the sophistication of current telemetry devices 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, animal biotelemetry requires rigourous quality control procedures to account for common, though not universally present, data issues such as irregularly timed measurements, sensor biases and time-varying errors in location. Some of these issues may be handled by a manufacturer's on-board or subsequent processing and some must be dealt with by researchers using the data.

transition paragraph focusing on analytical methods here…

transition paragraph focusing on related R packages & defining the foieGras niche here…

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

Here, we describe the main features of `foieGras` and illustrate its use through examples using both real and simulated data. Full R code for each of the examples 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's vignettes.

2 | `foieGras` overview

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

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 track 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). Other types of track data can be considered in a similar manner (see the package vignette for further details).

2.2 | State-space model fitting - `fit_ssm`

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

- `data` the input data structured as described in **2.1**
- `vmax` a maximum threshold speed (ms^{-1}) to help identify potential outlier locations
- `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) datetimes are converted to POSIXt format, chronological order is ensured, and duplicate datetime records are removed; 3) observations occurring less than `min.dt` seconds after a prior observation are removed; 4) a speed filter [`sda` from the `trip` R package; Sumner et al. (2009)] is used to identify potential outlier locations; 5) locations are projected from spherical lon-lat coordinates to planar x,y coordinates in km.

The function then fits a state-space model to the processed data, where the process model (currently, either a continuous-time `rw` or a continuous-time `crw`) is specified by the user and the measurement model(s) are selected automatically (see Jonsen et al., 2020 for model details). The model is fit by numerical optimization of the likelihood using either the `optim` or `nlminb` R function. 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

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

Function	Description
`fit_mpm`	Fit a Move Persistence Model to location data
`fit_ssm`	Fit a State-Space Model to location data
`fmap`	Plot fitted/predicted locations on a map with or without a defined projection
`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
`sim`	Simulate individual animal tracks with Argos LS or KF errors
`simfit`	Simulate animal tracks from ‘fG_ssm‘ fit objects
`sim_filter`	Filter tracks simulated with ‘simfit‘ according to similarity criteria
`plot.fG_ssm`	Plot the fit of a foieGras SSM to data
`plot.fG_osar`	Plot One-Step-Ahead Residuals from a foieGras SSM
`plot.fG_mpm`	Plot move persistence estimates as 1-D or 2-D (along track) time-series
`plot.fG_sim`	Plot simulated animal tracks

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 `foieGras` fit object (a nested data frame with class `fG_ssm`). 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 model parameter and state estimates, 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, typically irregular, observation times) or predicted values (locations predicted at regular `time.step` intervals) using the `grab` function.

2.3 | Model checking and visualisation - `osar`, `plot`, `fmap`

Before using fitted or predicted locations, a model fit should be checked and visualised to confirm that the model adequately describes the data. In linear regression and a variety of 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. 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 infer habitat availability, e.g., a null model of the distribution of foraging animals in the absence of external drivers, in habitat utilization studies (Hindell et al., 2020; Raymond et al., 2015). The `simfit` function extracts movement parameters from a `fG_ssm` fit object and simulates an arbitrary number of random tracks of the same duration from these parameters. The argument `cpf = TRUE` ensures that the simulated tracks start and end at approximately the same location, thereby simulating a central place forager. Something about `sim_filter` here...

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 as supplements.

3.1 | Southern Elephant seal - SSM validation with prediction residuals

We use a subadult male southern elephant seal track included in `foieGras` (`sese1`), 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, 73 % 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 suggests both models yield similar fits (Fig. 1a,b), however, there are marked trends in the time-series of residuals for the `rw` model fit (Fig. 1c) and the `rw` ACF's reveal consistent positive autocorrelation in the prediction residuals (Fig. 1e). The corresponding `crw` prediction residuals show no apparent trends through time and have relatively little autocorrelation (Fig. 1d,f), implying that the `crw` provides a better fit to the data.

3.2 | Assessing SSM robustness with simulated data

Using the `sim` function, we simulate animal movement tracks with a variety of plausible movement patterns. We use these simulated data to examine the accuracy of SSM fits to data generated by processes that differ from the `rw` and `crw` process models use in `fit_ssm`. While we regard this example as an atypical use of the track simulation function, it nonetheless illustrates one of the many possible uses of such a simulation tool. Another, more common application of track simulation as a preparatory step for a habitat usage analysis is highlighted in example 3.4.

We used `sim` to generate 50 animal tracks from each of the following 3 process models: 1) the same `crw` process used in `fit_ssm`; 2) a 2-state `crw` with stochastic switching between movement states; 3) the movement persistence model, where persistence in directionality and speed varies as a random walk. All tracks were simulated for 300 regular, 6-h time steps and a randomly selected Argos Kalman Filter error ellipse was assigned to each simulated location independent of the movement process. Further details on the track simulations are in Supplement xx. Representative simulated tracks for each movement process are shown in (Fig. 2a-c). We then used `fit_ssm` to fit both the `rw` and `crw` SSM's to these simulation tracks and calculated the Root Mean Squared Error of the SSM fits using the Euclidean distance between each SSM-estimated location and the corresponding simulated location (without Argos error). As the spatial scale of the simulated tracks differed among the 3 movement processes, we normalized RMSE values by the mean step length calculated across all tracks within each movement process type.

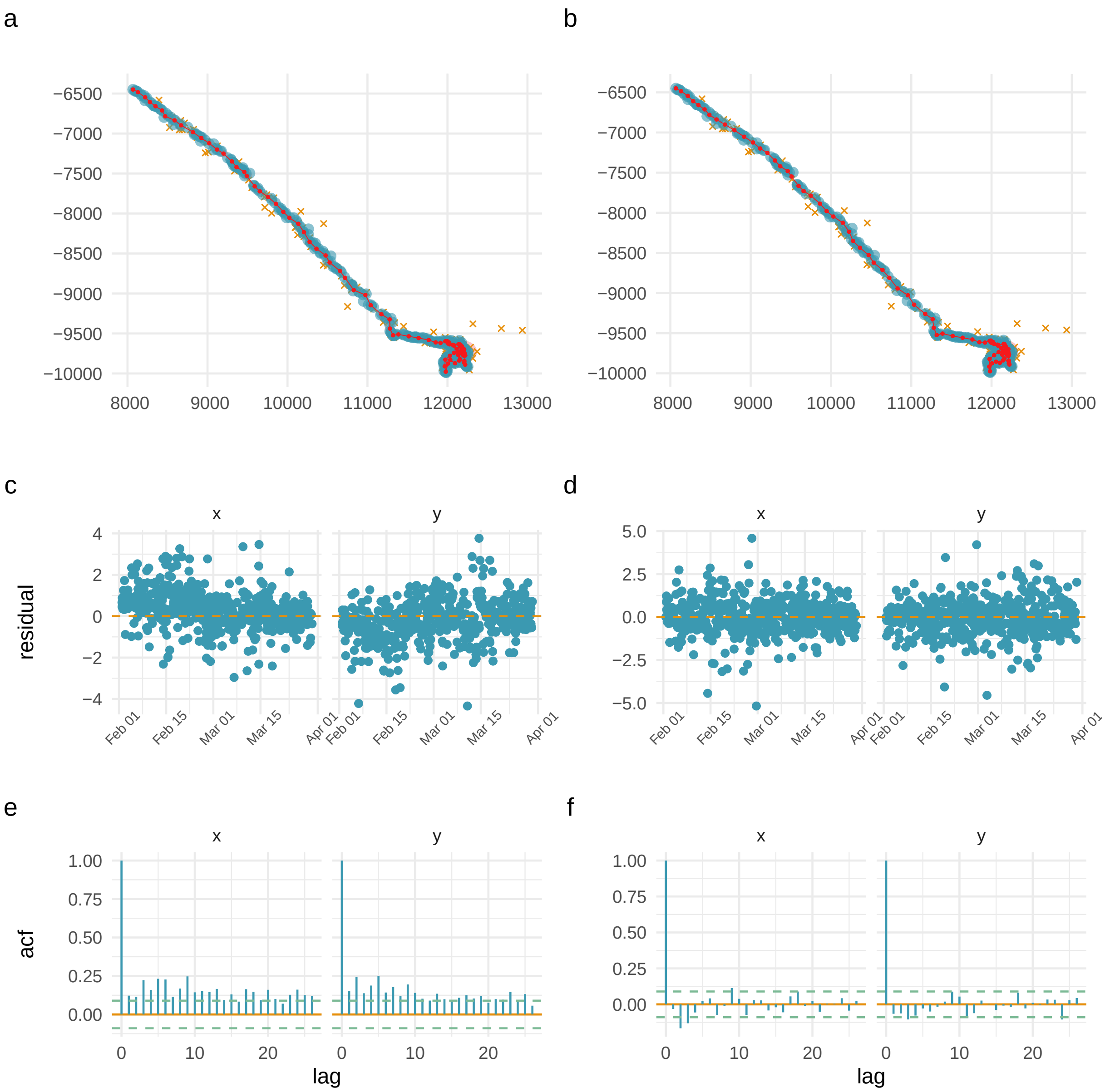

Figure 1: Selected diagnostic plots for assessing `rw` (a,c,e) and `crw` (b,d,f) state-space model fits to a southern elephant seal track. Top panels (a,b) are plots of predicted states (red; regular 12-h time intervals) and observations (blue) with pre-filtered observations (orange; ignored by the SSM), 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.

Regardless of the simulated movement process, the `crw` SSM consistently provided more accurate fits than the `rw` SSM (Fig. 2d-f). Although fits to the simulated 2-state `crw` and `mpm` tracks were less accurate than to the `crw` tracks, the differences were not large and goodness-of-fit based on prediction residuals were reasonable (Supplement xx). We advocate that in most cases the `crw` SSM should be preferred for quality-control of Argos tracking data, however the `rw` SSM may converge more reliably when fitting to problematic data (e.g., tracks with frequent and relatively large temporal data gaps).

3.3 | Inferring movement persistence as an index of behaviour from Argos and GPS data

Drawing on an expanded version of the data used in 3.1, we quality control and infer movement persistence, γ_t, along five southern elephant seal tracks. We fitted the `crw` SSM with a 24-h prediction interval using `fit_ssm` and assuming bivariate normal location measurement errors consistent with Argos Least-Squares-derived locations (Jonsen et al., 2020). We used the SSM-predicted locations to estimate movement persistence jointly among the 5 seals using `fit_mpm` and visualise the behavioural index along the seals' tracks. The data can be accessed in `foieGras` via `data(sese, package = 'foieGras')`. As the estimation of γ_t is sensitive to choice of time scale, we examined the influence of different prediction intervals (1 - 20 min) on the ability of the movement persistence model to resolve changes in movement pattern along the penguin tracks.

To illustrate how the method can accommodate other types of animal tracking data, we also infer γ_t along six little penguin (*Eudyptula minor*) GPS tracks from Montague Island, NSW Australia, described in Phillips et al. (2021). The data are temporally irregular GPS locations that are assumed to have minimal measurement error. We fitted the `crw` SSM to predict temporally regular locations at 5-min intervals, assuming consistently small bivariate normal location measurement errors (ie. $\pm$ 10 m sd).

Movement persistence estimates along the quality-controlled southern elephant seal tracks highlight some fundamental differences in movement pattern among the seals. The two seals engaging in pelagic foraging trips (Fig. 3a,c and f) had less contrast in their movements with consistently higher γ_t estimates compared to the three seals engaging in trips to the fast-ice on the Antarctic shelf (Fig. 3b,d-e and f). Although γ_t's were higher overall for the pelagically foraging seals, they both spent little time making fast, highly directed movements ($\gamma_t \rightarrow 1$) relative to the shelf-foraging seals (3a,c vs b,d-e). This suggests the pelagically-foraging seals may spend considerable time searching for suitable foraging habitat in the highly variable eddy fields between the Subantarctic and Polar Fronts (Jonsen et al., 2019), whereas foraging habitat may be more predictable for seals travelling rapidly and directly to the Antarctic shelf region. These seals may also haulout periodically on available fast-ice to rest. This behaviour could also contribute to the higher contrast in movement persistence, relative to pelagically-foraging seals who would not have access to fast-ice.

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. 4a-e). 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. 4f).

3.4 | Simulating tracks from `foieGras` model fits

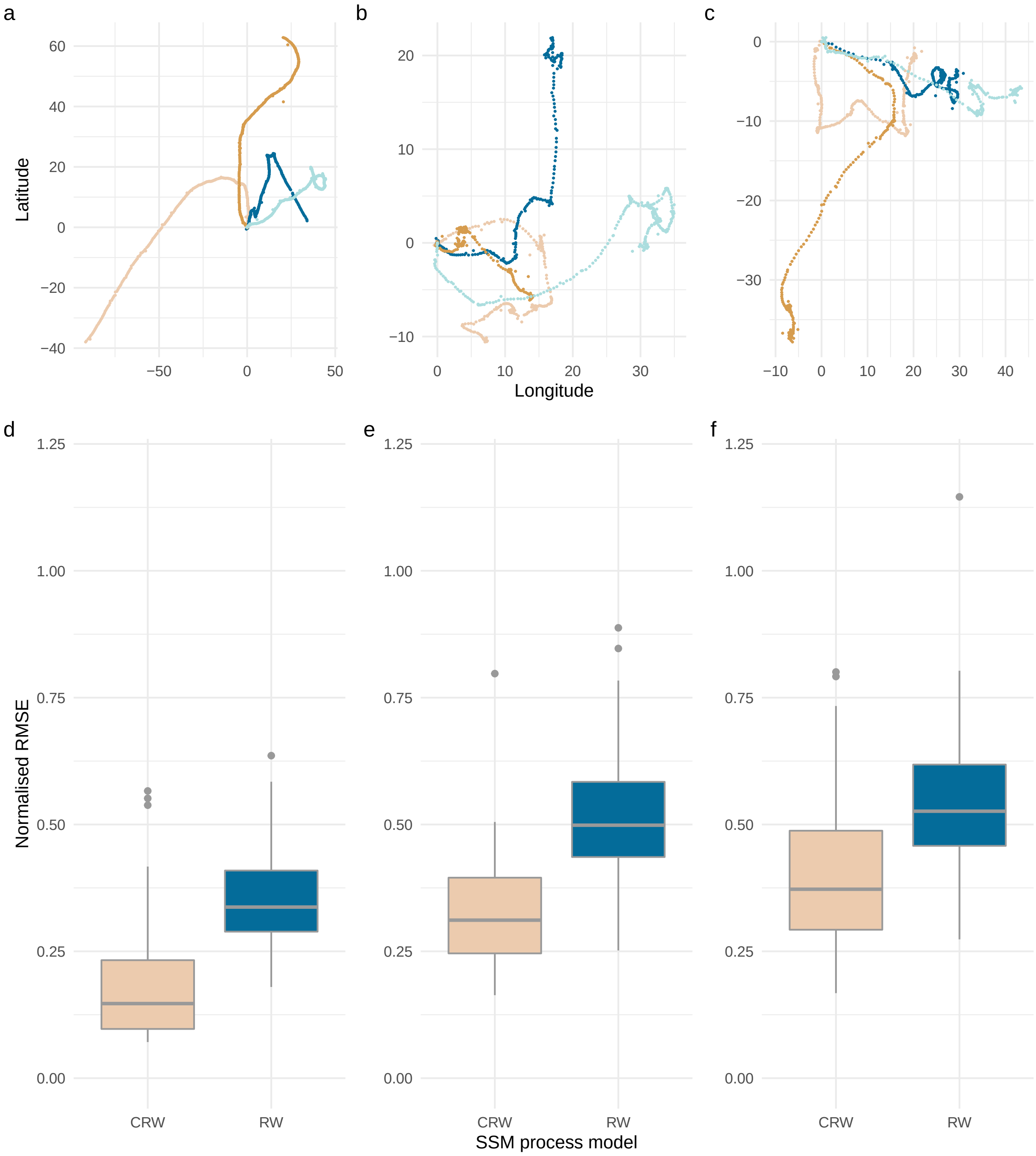

Figure 2: Four example tracks simulated from the correlated random walk process model (`crw`, a), a 2-state `crw` model (b), and the movement persistence model (`mpm`, c). Normalised Root Mean Squared Errors of state-space models fit with either the `crw` or random walk (`rw`) process model to 50 simulated `crw` tracks (d), 50 simulated `2-state crw` tracks (e), and 50 simulated `mpm` tracks (f). The SSM fits using the `crw` process model were consistently more accurate than those using the `rw` process model, regardless of the type of movements simulated.

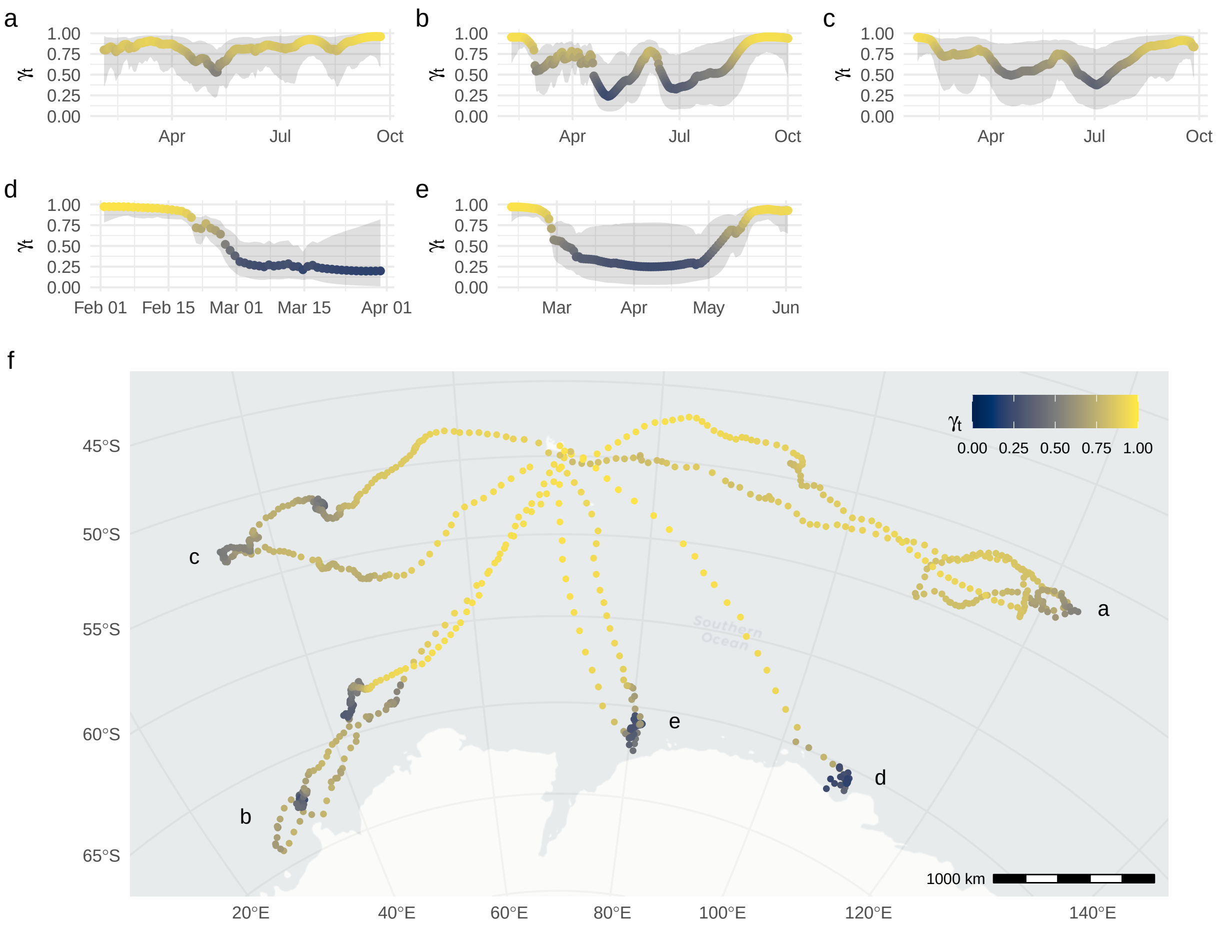

Figure 3: Inferred move persistence, γ_t, 1-D time-series for five southern elephant seals (a-e; grey envelopes are 95 % CI's) and along their 2-D tracks (f; track labels, a-e, correspond to the 1-D time-series plots). Locations associated with low move persistence (blue) are indicative of slow, undirected 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.

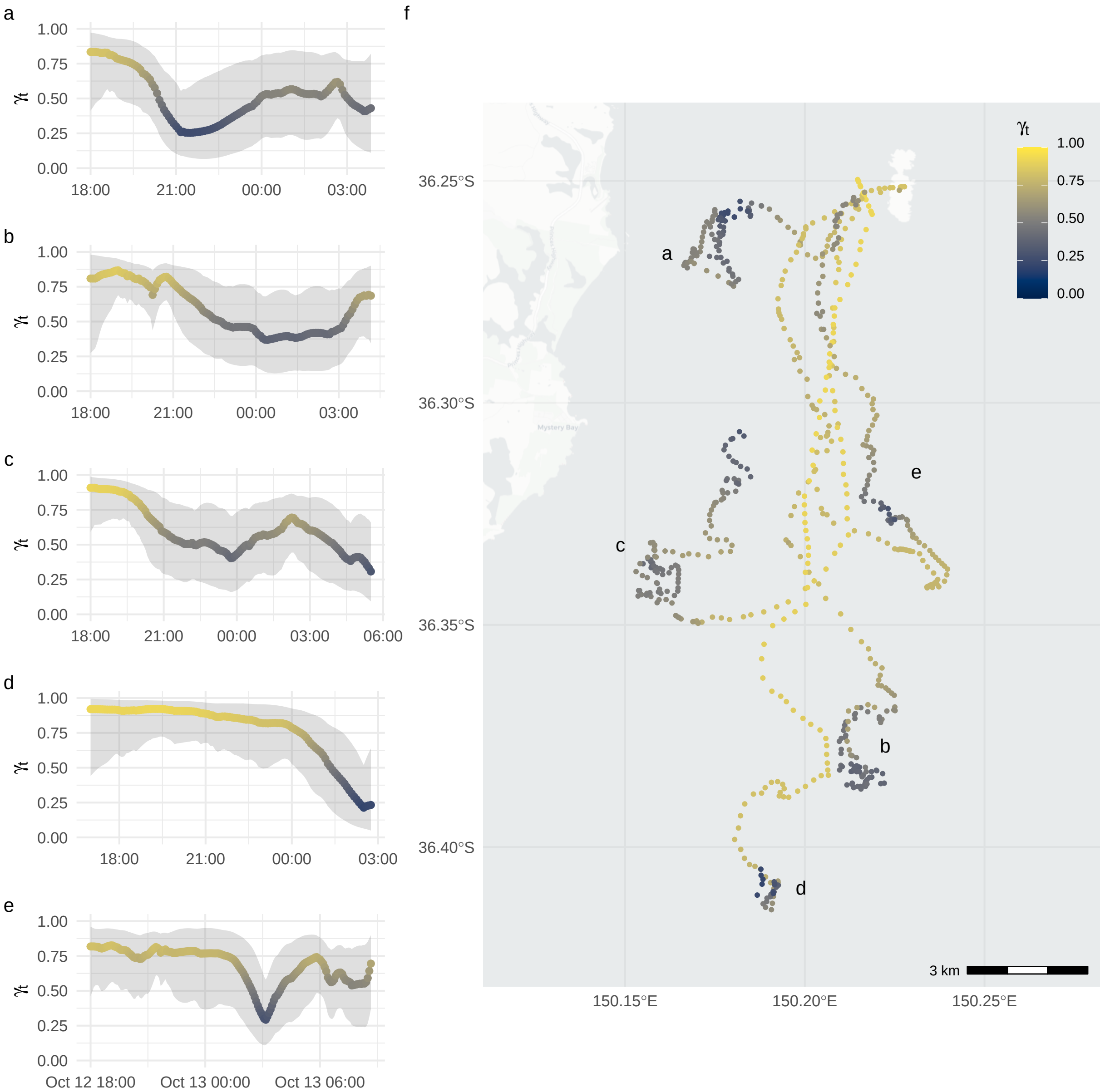

Figure 4: Inferred move persistence, γ_t, 1-D time-series (a-f; grey envelopes are 95 % CI's) and along little penguin GPS tracks (g). Colour palette as in 3. Movement persistence was estimated from SSM-predicted locations with a regular 5-min interval.

4 | Discussion

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

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^{o} 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 (species).

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; IDJ and TAP developed the state-space models; IDJ wrote an initial draft of the manuscript with a contribution from WJG; all authors edited the manuscript.

Data Accessibility

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

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
Robert G Harcourt https://orcid.org/0000-0003-4666-2934
Toby A Patterson https://orcid.org/0000-0002-7150-9205

References

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.

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.

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.

Jonsen, I. D. (2016). Joint estimation over multiple individuals improves behavioural state inference from animal movement data. *Scientific Reports*, *6*, 20625.

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.

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.

Lopez, R., Malardé, J.-P., Danès, P., & Gaspar, P. (2015). Improving Argos Doppler location using multiple-model smoothing. *Animal Biotelemetry*, *3*, 32.

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.

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*.

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.

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

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:%2010.1007/s10651-017-0372-4