Running title: R tracking packages review
Number of words: ~XXXX
Number of tables: 2
Number of figures: 4
Number of references: XX

Navigating through the R packages for movement

Rocio Joo[1*] et al.

1. *Department of Wildlife Ecology and Conservation, Fort Lauderdale Research and Education Center, University of Florida, Fort Lauderdale, FL, USA*

*Correspondence author. `rocio.joo@ufl.edu`

Date of submission: 8th October 2018

Summary

1. The advent of miniaturized biologging devices has provided ecologists with unparalleled opportunities to record animal movement across scales. Technological advancements, including improvements to battery life and data storage, have lead to ever-increasing quantities of data. There have also been concurrent advances in the abundance and sophistication of tools used to process, visualize and analyze biologging data.

2. In recent years, there has been greater emphasis on the standardization of methods, but these efforts often occur in parallel, such that there has been a proliferation of programming tools available, but little consensus or advice on their use. Within the R software alone, there are as many as 57 packages for the processing or analysis of tracking data.

3. We present a review of these R packages, called here tracking packages, which aims to enable researchers to access the appropriate tools and to provide package developers criteria on what needs to be improved, from a user perspective. Since the packages respond to needs of tools for data processing and analysis, we first divide these aspects into pre-processing, post-processing, data visualization, track description, path reconstruction, behavioral pattern identification, space use characterization, trajectory simulation and others. We describe each of these aspects and, for each of them, we assess which packages offer suitable functionalities and summarize them.

4. Supporting documentation is key to render a package accessible to new users. Based on a survey of users, we review the quality of the supporting documentation provided in conjunction with the packages, and identify 12 packages considered as having good or excellent documentation.

5. Compatibility and connectivity between packages is assessed through a network graph analysis. Though a large group of packages has some degree of connectivity (depending on functions or suggesting the use of another tracking package), a third of the packages work on isolation, reflecting a fragmentation in the R movement-ecology programming community.

6. At the end of this review, we provide some recommendations to users for choosing packages and to developers for maximizing the usefulness of individual packages and strengthening the links between the programming community.

Keywords: biologging, R packages, movement ecology, tracking data

General Introduction

Animal movement plays a crucial role in ecological and evolutionary processes, from the individual to ecosystem level (Nathan et al. 2008, Kays et al. 2015, other refs). However, studying animal movement has presented challenges to researchers, as individuals are often difficult to follow for extended time periods and over large distances. Over recent decades, decreases in the size and cost of animal-borne sensors or biologging devices have led to an exponential increase in their use. This has substantially improved our understanding of how and why animals move (Nathan et al. 2008, other refs). Technological advancements have also enabled a wide range of sensors to be used by ecologists, which can be integrated to remotely record a suite of metrics, including x, y and z (i.e. altitude or depth) locations, acceleration, prey capture attempts, as well as in-situ environmental conditions (refs). From these multiple sensors, fine-scale behaviors and physiological states can be inferred (Wilmers et al. 2015, other refs).

The increase in quantity and complexity of biologging data requires appropriate analytical and software tools that aid processing and interpretation of data. Those tools should be standardized to allow for reproducibility of results and computation time optimization (Reichman et al. 2011, Stewart Lowndes et al. 2017). Mainly in the last decade, many of these tools have been made available for the scientific community in the form of R packages, which has facilitate their widespread use (refs). Nonetheless, many packages have been made in isolation, and there is no formal appraisal and comparison of the tools provided by the packages. This limits their use as researchers are required to review each package to identify, within a package, the most appropriate function for their analysis, and between packages with similar objectives, the most appropriate package for their analysis.

The aim of this study is to review the available packages within the R platform, for movement ecology. Movement of an organism is defined as a change in the spatial location of an individual in time, so movement data is defined by a space and a time component. For the purpose of this review, we focus on a specific type of movement data: tracking data; i.e. data composed by at least 2-dimensional coordinates (x, y) and a time index (t), and can be seen as the geometric representation (the trajectory) of an individual's path. Since most movement data is collected using tagging devices, some R packages focus on extracting or analyzing data from these devices, dealing with the limitations related to the device they focus on; for instance, some packages provide tools for extracting locations from the light level information collected with Global Location Sensors (GLS). Some other packages have been created to process or analyze any dataset in a tracking data format (x,y,t) regardless of the way the data was collected. All of these packages, that are either for transforming data into a tracking format or to analyze tracking data, are reviewed here and will be henceforth called tracking packages. To our knowledge, there are 57 of them.

The next section will summarize the packages by stage of data processing and analysis (Fig. 1). In some cases, biologging devices do not provide raw data in the form of tracking data, e.g. for GLS loggers, for the most part just light intensity is provided. The process by which data is transformed into the (x,y,t) format we refer to as pre-processing. After this, the data

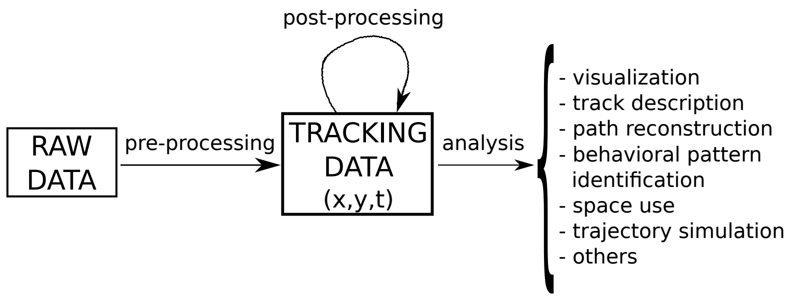

Fig. 1 Stages of data processing and analysis.

may not be immediately usable, e.g. errors or outliers need to be identified, or other second or third order variables need to be derived for the dataset to be ready for analysis; we defined this type of data processing as post-processing. We divide any further analyses into data visualization, track description, path reconstruction, behavioral pattern identification, space use characterization, trajectory simulation and others (e.g. population parameter estimation, interaction between individuals). In each of these subsections, we describe the tools provided by tracking packages to achieve these goals. An additional subsection will briefly describe some R packages that did not deal with tracking data (as defined above), but were developed to process and analyze data from biologging devices such as accelerometers and time-depth recorders.

Since the documentation provided in conjunction with the packages are key for rendering them accessible, we also review this supporting documentation and, based on a survey, show how useful these documents are to package users. The links between packages, showing how much they rely on each other and the compatibility between them are also assessed.

This review is aimed at both users and developers of tracking packages, since any user can potentially become a developer at a given time, and any developer can use other packages. This work is a step forward for a complete knowledge on the existing tracking packages; for users, offering criteria to select packages to perform movement ecology analysis, and for developers, discussing issues that could help maximize the usefulness of individual packages and strengthen the links between the community.

Data processing, analysis and the R packages

Multiple sources were used to identify the tracking packages (as defined in the introduction); mainly, 1) the spatio-temporal task view on CRAN (`https://cran.r-project.org/web/views/SpatioTemporal.html`, 2) an updated list of this task on GitHub (`https://gist.github.com/mdsumner/0a3cb0e58bf9d37b782943ac269e1eff`), 3) packages suggested in the description files of other packages, 4) Google search engine and 5) e-mail/Twitter exchanges with ecologists. The package search was done between March and August 2018. Tracking packages that were either removed from CRAN or described as in a 'very early version' on their GitHub repositories were discarded.

Fifty seven packages assist with processing and analysis of tracking data (Fig. 1). Some R packages have been developed to tackle several of these stages of data processing and analysis, while others focus on only one. The number of packages regarding each type of processing/analysis is showed in Table 1. When appropriate, the type of biologging devices from which the tracking data originates will be described, so that readers that are not familiar with these devices have a basic idea of the advantages and limitations of the devices, and why some packages need to focus on specific issues related to them.

Type of processing/analysis	Count
Pre-processing	9
Post-processing	17
Visualization*	2
Track description	5
Path reconstruction	8
Behavioral patterns identification	9
Space use	17
Trajectory simulation	10
Others	8
Total	57

Table 1 Number of packages dealing with each type of data processing and analysis. Some packages may correspond to more than one category, except for data visualization (*), where only packages created for that purpose are counted.

The description of packages in this section will also include information on the year each package was publicly available (Fig.2), the main repository where the package is stored and whether it is actively maintained (hereafter referred to as 'active'). The official repository for R packages is the Comprehensive R Archive Network (CRAN) repository. CRAN enforces technical consistency, with a set of rules such as the inclusion of ownership information, cross-platform portable code (i.e. to work with Windows, Mac OS and UNIX platforms), minimum and maximum sizes for package components, among others. The great majority of the reviewed packages are on CRAN; the ones that are not on CRAN are mostly on GitHub and a few others are in other repositories (e.g. r-forge or independent websites). Regarding package maintenance, we consider that a GitHub package is actively maintained if

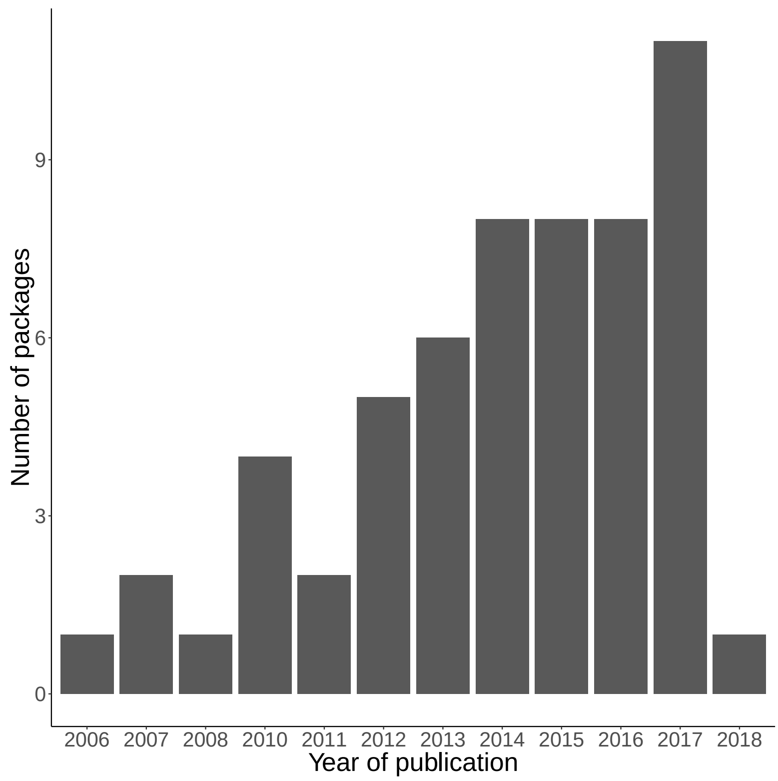

Fig. 2 Number of packages per year of publication (assessment performed in August 2018).

a 'commit' has been made in the last year, and for the others (if they are not also on GitHub), that the most recent version of the package is no older than one year (analysis conducted in August 2018). Links to each package repository along with a summary of their main characteristics are included in Supplementary File 1.

Pre-processing

When raw biologging data are not in a tracking data format, some pre-processing is required. The methods used for pre-processing depend heavily on the type of biologging device used. Among the tracking packages, 6 are focused on global location sensors (GLS), one on radio telemetry, one uses accelerometry and magnetometry, and another one uses GPS data in addition to accelerometry and magnetometry data.

GLS data pre-processing

GLS are electronic archival tracking devices which record ambient light intensity and elapsed time. The timings of sunset and sunrise are estimated, and latitude is calculated from day length, and longitude from the time of local midday relative to Greenwich Mean time (Afanasyev, 2004). GLS can record data for several years and their small size and low mass ($<$1 g) make them suitable for studying long-distance movements in a wide range of species. Several methodologies have been developed to reduce errors in geographic locations generated from the light data, which is reflected by the large number of packages for pre-processing GLS data. We classified these methods in three categories: threshold, template-fitting and twilight-free.

- Threshold methods. Arbitrary threshold levels of solar irradiance are fixed to identify the timing of sunrise and sunset. The packages that use threshold methods are `GeoLight` (2012, CRAN, inactive) and `probGLS` (2016, GitHub, inactive). `GeoLight` uses astronomical equations from Montenbruck & Pfleger (2013) to calculate locations from timings of sunrise and sunset, and from sun elevation angles. `probGLS` implements a probabilistic method that takes into account uncertainty in sun elevation angle and twilight events to estimate locations. Starting with the first known location (where the individual was tagged), it estimates the location of the subsequent twilight event which is replicated several times adding an error term, repeating this process for the whole data series; it then computes probabilities for each location based on the plausibility of the estimated speed or on environmental conditions (e.g. sea surface temperature or SST) (Merkel *et al.*, 2016).

- Template-fitting methods. The observed light irradiance levels for each twilight are modeled in function of theoretical light levels (i.e. the template). Then, parameters from the model (e.g. a slope in a linear regression) are used for estimating the locations. The formulation of the model and the parameters used for location estimation vary from method to method (Ekstrom, 2004). The packages that use template-fitting methods are `FLightR` (2015, CRAN, active), `TripEstimation` (2007, GitHub, inactive) and `trackit` (2012, GitHub, active). `FLightR` was particularly developed for bird movement. In its state-space modeling framework (Patterson *et al.*, 2008), the locations are hidden states and the observation model is a physical model of light level changes as a function of position and time. A detailed description of the model and the package functions can be found in Rakhimberdiev *et al.* (2015) and Rakhimberdiev *et al.* (2017), respectively. `trackit` was developed mainly for fish movement. Light intensity around sunrise and sunset are used as inputs in a state-space model that includes solar altitude and SST as covariates (Lam *et al.*, 2010). `TripEstimation` was developed for marine organisms. It uses a Bayesian approach modeling light level as a function of sun elevation at each plausible location, prior knowledge of the animal's movement, and complementary environmental information (e.g. SST, depth of the water column) (Sumner *et al.*, 2009). The package is still available on CRAN, but in its GitHub repository it is indicated that the package was deprecated for `SGAT` (the authors of `TripEstimation` are main authors of `SGAT` and `GeoLight`). `SGAT` contains functions to implement both threshold and template-fitting methods. As one has been deprecated by the other, we consider both packages as one `TripEstimation/SGAT`. As auxiliary packages, `TwGeos` and `BAStag` contain functions to process GLS data such as detecting the timing of twilight periods from light data. The estimated twilight periods can be later used as inputs in the above mentioned packages for location estimation.

- Twilight-free methods. It is possible to estimate locations without depending on twilight event identification. `TwilightFree` (2017, GitHub, active) uses a Hidden Markov Model (HMM) where the hidden states are the daily geographic locations (the spatial domain is discretized as gridded cells) and the observed variable is the observed pattern of light and dark over the day (Bindoff *et al.*, 2017). SST and land/sea marks can be used as covariates. Parameter estimation is performed using functions from the SGAT package.

Radio tagging data pre-processing

Radio tagging involves the attachment of a radio transmitter to an animal. The radio signals transmitted are picked up by an antenna and transformed into a beeping sound by a receiver. As the receiver gets closer to the transmitter, the beeps get louder. Radio tagging can be used to study animal movement by either placing a fixed antenna at a location or through a method called homing, where the researcher moves towards the loudest beeps until the animal has been located. If antennae are located in different locations, triangulation in a map of the direction vectors of the received signals can give an estimate of the animal's location. RFID (radio-frequency identification data) tags can also be used to record when an individual passes through a receiver. With RFID, the researcher does not need to search for beeping signals, but the individual must be adjacent to the receiver to be detected. Radio tags are comparable in price and mass to GLS.

`telemetr` (2012, GitHub, inactive) implements several triangulation methods as well as a maximum likelihood procedure to estimate locations from bearing data (triangulation information). Since there are no reference to the methods in the package documentation, it is aimed at users that are already familiar with the methods.

Accelerometry, magnetometry and GPS data pre-processing

Magnetometers measure magnetic fields. Accelerometers measure non-gravitational acceleration, quantifying movement through time by way of changes in velocity. Acceleration is used by ecologists to measure three dimensional movement to classify behaviors such as flight and prey capture, and estimate overall dynamic body acceleration, a measure of energy expenditure. A combined use of magnetometer and accelerometer data (and optionally, other sensors such as gyroscopes) allow obtaining tracks using dead-reckoning (DR)

(Wilson *et al.*, 2007; Bidder *et al.*, 2015; Williams *et al.*, 2017). Typically, data from magnetometers and accelerometers are used to calculate travel vectors (i.e. vectors representing distance covered and direction) for each given time interval, and, since the initial tagging point is known, the three dimensional movement path can be reconstructed by integrating the vectors in sequence. `animalTrack` (2013, CRAN, inactive) and `TrackReconstruction` (2014, CRAN, inactive) implement DR to obtain tracks.

When GPS data are available, `TrackReconstruction` takes DR outputs and forces them to go through known GPS points via space transformation. GPS loggers are perhaps the most widely used type of biologging device. They are very cheap and easy to obtain, and location information can be downloaded directly without any post processing. GPS receivers collect but do not transmit information, and infer their own location based on the location of GPS satellites and the time of transmission. Four or more satellites should be visible by the receiver to obtain an accurate result (< 100 m). GPS receivers can collect precise location data at short time intervals (in the order of minutes or seconds). Combined use of DR and GPS data allows obtaining very high resolution tracking data.

Post-processing

Post-processing of tracking data comprises data cleaning (e.g. identification of outliers or errors), compressing (i.e. reducing data resolution which is sometimes called resampling) and computation of metrics based on tracking data, which are useful for posterior analyses.

Data cleaning

`argosfilter` (2007, CRAN, inactive) and `SDLfilter` (2014, CRAN, active) implement functions to filter implausible platform terminal transmitter (PTT) data. Platform terminal (or Argos) transmitters send signals to polar-orbital Argos satellites, which geographically locate the source of the data. They preserve battery life by only needing to transmit signals, leading them to be used for tracking of large-scale migrations, particularly marine mammals and turtles. When the tracked animals are under water, the chances of a satellite receiving PTT signals decrease, so fewer locations can be estimated, and they are likely estimated with fewer satellites, so their accuracy also diminishes. PTTs are particularly useful for individuals that cannot be recaptured, and hence a device recovered. Along with locations, Argos provide accuracy classes (1, 2, 3, 0, A, B, Z) which are associated with different degrees of spatial error (Costa *et al.*, 2010). `argosfilter`'s algorithm is described in Freitas *et al.* (2007). It essentially removes records where a location was not estimated as well as locations that required unrealistic travel speeds. SDLfilter, which can also handle GPS data, allow the removal of duplicates, locations estimated with a low number of satellites, biologically unrealistic locations based on speed thresholds or turning angles and locations above high tide lines. The filtering methods are described in Shimada *et al.* (2012, 2016).

Other packages with functions for cleaning tracking data are `T-LoCoH` (2013, R-forge, active), `ctmm` (2015, CRAN, active), `TrajDataMining` (2017, CRAN, active) and `trip` (2006, CRAN, active).

Data compression

Concerning data compression, rediscretization (equal step lengths) can be achieved with `adehabitatLT` (2010, CRAN, active), `trajectories` (2014, CRAN, active) or `trajr` (2018, CRAN, active). Regular time-step interpolation can be performed using `adehabitatLT`, `trajectories` or `amt` (2016, CRAN, active). Other compression methods include Douglas-Peucker (`trajectories` and `TrajDataMining`), opening window (`TrajDataMining`) or Savitzky-Golay (`trajr`). For a brief review on compression methods, see Meratnia & de By (2004).

`rsMove` (2017, CRAN, active) provides functions to explore and transform tracking data for a posterior linkage with remote sensing data. Location fixes are transformed into pixels and grouped into regions. The spatial or temporal resolution of the tracking data can be changed to match the resolution of the remote sensing data.

Computation of metrics

Some packages automatically derive second or third order movement variables (e.g. distance and angles between consecutive fixes) when transforming the tracking data into the package's data-class (most packages define their own data classes, see Supplementary File 1). These packages are `adehabitatLT`, `trajectories`, `moveHMM` (2015, CRAN active), `momentuHMM` (2017, CRAN, active) and `rhr` (2014, GitHub, inactive). `trip`, `amt`, `trajr` and `move` (2012, CRAN, active) also contain functions for computing those metrics, but the user needs to specify which metric they need to compute.

`feedr` (2016, GitHub, active) works specifically with RFID data (described above in the radio telemetry subsection). Raw RFID data typically contain an individual line of data for each read event made by each RFID logger. `feedr` contains functions to read raw data from several RFID loggers, and to transform the data of logger detection into movement data for each individual, computing statistics such as the time of arrival and departure from each logger station, and how much time was spent near a station at each visitation.

`VTrack` (2015, CRAN, active) handles acoustic telemetry data. Acoustic telemetry uses high frequency sound (between 30 and 300 kHz) to transmit information through water. Tags (transmitters) emit a pulse of sound, which is detected by a hydrophone (or an array of hydrophones) with an acoustic receiver. The distance at which a transmitter can be detected depends on the power and frequency of the tag, and the characteristics of the surrounding environment (e.g. background noise, water turbidity and temperature) (DeCelles & Zemeckis, 2014). `VTrack` was created to deal with VEMCO©(a tagging equipment provider) data, which has a similar structure than RFID; it is composed of transmitter ID, receiver ID, datetime stamps, location of receivers information. Like `feedr` for RFID, `VTrack` can compute statistics such as the time of arrival and departure from each receiver, and how much time was spent near a receiver at each visitation.

Visualization

Most of the tracking packages contain functions to visualize the data they analyze and we encourage users to explore these functions. In this section, we focus on the packages mainly developed for visualization purposes. Those are `anipaths` (2017, CRAN, active) and `moveVis` (2017, CRAN, active).

They were both conceived for producing animations of tracks. `anipaths` relies on the `animation` package. Users can specify time-steps and seconds per frame for animation, add a background map (e.g. Google Maps) and an individual-level covariate (e.g. migrant, stationary), among others. Consecutive fixes are joined via a spline-based interpolation and a confidence interval for the interpolation of the path for animation can be shown.

`moveVis` is based on a `ggplot2` plotting architecture and works with `move`-class objects. Users can choose between 'true time' which displays the animation respecting the timestamps provided, or 'simple' animations where time is not taken into account and all individuals are displayed together as if their tracks started at time 0. Consecutive fixes are joined via linear interpolation. As in `anipaths`, users can specify the number of frames per second and personalize the background map. Statistics related to the background layer (e.g. temperature, land cover) can also be shown as animated lines or bar plots.

For both packages, animations can be saved in many different formats such as mpeg, mp4 and gif.

Track description

`amt`, `movementAnalysis` (2013, GitHub, inactive) and `trajr` compute summary metrics of tracks, such as total distance covered, straightness index, sinuosity and others related to net squared displacement. It should be noted that `movementAnalysis` depends on the `adehabitat` package, was which was officially removed from CRAN in 2018, as it was superseeded by `adehabitatLT`, `adehabitatHR`, `adehabitatHS` and `adehabitatMA` in 2010. `marcher` (2017, CRAN, active), which is focused on migration analysis, also computes net squared displacement as well as a range shift index; i.e. the ratio of the distance between successive circular ranges and their diameter.

`trackeR` (2015, CRAN, active), which was created to analyze running, cycling and swimming data from GPS-tracking devices for humans, computes metrics summarizing movement effort during each track (or workout effort per session). Those metrics include total distance covered, total duration, time spent moving, work to rest ratio, averages of speed, pace and heart rate. The functionality of this package could be adapted to non-human tracking data.

Path reconstruction

Whether it is for correcting for sampling errors, obtain finer data resolutions or regular time steps, path reconstruction is a common goal in movement analysis. Here we mention methods available, however, before choosing a method, users should be aware that every method is constructed under unique movement assumptions (either inherent to the mathematical model or constructed for a particular species or type of data), and users should refer to the literature on the methods first. Packages available for path reconstruction are `HMMoce` (2017, CRAN, active), `kftrack` (2011, GitHub, active), `ukfsst/kfsst` (2012, GitHub, active), `bsam` (2016, CRAN, active), `argosTrack` (2014, GitHub, active), `BayesianAnimalTracker` (2014, CRAN, inactive), `crawl` (2008, CRAN, active) and `ctmcmove` (2015, CRAN, active).

While the first three are focused on GLS data, `bsam` is intended for PTT data, `BayesianAnimalTracker` combines GPS data and dead-reckoning, and the other two could be used with any tracking data.

Improving location estimation from GLS data

`kftrack`, `kfsst` and `ukfsst` were developed by the same team of `trackit`, described in the pre-processing section. As `trackit`, they are mainly focused on fish movement. `kftrack`, `ukfsst` and `kfsst` use already estimated positions, either by the threshold method or given by the provider, and improve those estimations using a 2-dimensional random walk model (Sibert *et al.*, 2003). Because of the generality of this modeling framework, `kftrack` could actually be used for any tracking data. In addition to the random walk model, `kfsst` includes SST as a covariate in the model (Nielsen *et al.*, 2006), but it has been superseded by `ukfsst`, which implements an optimized parameter estimation. For that reason, we count `kfsst` and `ukfsst` as one.

`HMMoce`, also adapted to fish movement and working with already estimated/provided locations, uses HMMs (like `TwilightFree`) and incorporates depth-temperature profiles and SST as covariates in the observed model (Braun *et al.*, 2017).

Improving location estimation from PTT data

`bsam` estimates locations by fitting Bayesian state-space models to the data. They offer possibility of accounting for different movement patterns using 'switching' or HMMs; if this is opted out, first-difference correlated random walk models (DCRWs) are used. It is possible to estimate some of the model parameters for each individual and others at the population level (see Jonsen *et al.* (2013); Jonsen (2016) for more details). The `argosTrack` package fits several types of movement models to PTT data (Albertsen *et al.*, 2015), such as correlated random walks (CRWs) in discrete and continuous versions, and Ornstein-Uhlenbeck (OU) models, using Laplace approximation via Template Model Builder.

Combining dead-reckoning and GPS data

`BayesianAnimalTracker` takes an already estimated DR path and combines it with GPS data via a Bayesian approach (Liu *et al.*, 2016): it maximizes a likelihood of a model where it is assumed that the true points come from a Brownian Bridge, both the GPS and DR points are linearly dependent on the true path, and random and measurement error parameters are added to the model.

Modeling movement of general tracking data

`crawl` reconstructs paths by fitting continuous-time correlated random walk (CTCRW) models (Johnson *et al.*, 2008) to tracking data. Though it can be used for any tracking data, it can account for the accuracy classes of PTT data in the models to model the error in location. `ctmcmove` fits a functional movement model (Buderman *et al.*, 2016) to the data and a set of probable true paths can be generated.

Behavioral pattern identification

Another common goal in movement ecology is to get a proxy of the individual's behavior through the observed movement patterns, based on either the locations themselves or second/third order variables such as distance, speed or turning angles. Covariates, mainly related to the environment, are frequently used for behavioral pattern identification.

We classify the methods in this section as: non-sequential classification or clustering techniques, segmentation and hidden Markov models.

Non-sequential classification or clustering techniques

Each fix in the track is classified as a given type of behavior, independently of the classification of the preceding or following fixes. `EMbC` (2015, CRAN, active) and `m2b` (2017, CRAN, inactive) present tools that fall in the first category. `EMbC` implements the Expectation-maximization binary clustering method (Garriga *et al.*, 2016). `m2b` implements a random forest (a wrapper for the randomForest package functions) to classify behaviors using a supervised training dataset; i.e. a dataset of tracking data and known behaviors is needed to train the model.

Segmentation methods

A track (typically a time series of movement patterns) is cut into several segments; the edges of each segment represent a change in behavior. `adehabitatLT`, `bcpa` (2013, CRAN, inactive), `marcher` and `migrateR` (2016, GitHub, active) implement segmentation methods. `adehabitatLT` presents two of these methods: Gueguen and Lavielle. `bcpa` implements the behavioral change point analysis (Gurarie *et al.*, 2009). Both `marcher` and `migrateR` are suited for migrant individuals. `marcher` enables a mechanistic range shift analysis (Gurarie *et al.*, 2017) that identifies changes in locations of focal ranges, so migration and resident behaviors can be distinguished. The ranging models available in the package can take into account autocorrelation in location and in velocity. `migrateR` uses net displacement models to identify migrant, resident and nomad behavior (Spitz *et al.*, 2017). The models can incorporate factors such as elevation, sensitivity to starting date in the series, minimum time out of residence zone, among other features.

Hidden Markov models

The main idea is that there is a hidden state process (representing the sequence of non-observed behaviors) conditioning the observed movement patterns, and that the states follow a Markov process (Langrock *et al.*, 2012). In this category we consider standard as well as more complex versions of these models; e.g. adding hierarchical structures, a second observation process for locations (state-space modeling), covariates affecting different components in the model, autoregressive processes or a spatial covariance structure. `bsam`, `moveHMM` and `momentuHMM` implement methods that fall in the HMM category. `bsam`, for PTT data, implements Bayesian state-space models as described in the path reconstruction section, and may incorporate a a layer of two switching states into the model: one state representing directed fast movement, and the other representing relatively undirected slow movement (Jonsen *et al.*, 2013). `moveHMM` and `momentuHMM` are not restricted to two states. moveHMM implements HMMs incorporating covariates and allowing for state sequence reconstruction, i.e. sequences of the behavioral proxies, via the Viterbi algorithm. In moveHMM, the variables modeled in the observed process are step length and turning angles, or two variables that statistically behave as step length and turning angles. momentuHMM implements generalized Hidden Markov models (McClintock *et al.*, 2012) with great flexibility for the choice of observed variables and their probability distributions, and covariate incorporation in the models. Since HMMs require regular time steps, momentuHMM offers a multiple imputation method (McClintock, 2017): it fits a CTCRW (from crawl) to the data obtaining regular (time-step) realizations and then fits an HMM to those realizations; all of this is done multiple times. Even if the data classes and model formulation in the package differ from moveHMM, many of the HMM-related functions are based on moveHMM. moveHMM is more user-friendly than momentuHMM, but momentuHMM offers greater modeling possibilities.

Space and habitat use characterization

Spatial ecology precedes movement ecology as a research field, which is why a main interest in movement ecology is to use tracking data to answer questions related to space and habitat use, such as: where do individuals spend their time, how long do they stay in different places and what role environmental conditions play in these choices? Multiple packages implement functions to help answering these questions, which are typically split into two categories: home range calculation and habitat selection (both including true tracking data, i.e. serial locations).

Home range

Several packages allow the estimation of home ranges: `adehabitatHR`, `rhr`, `T-LoCoH`, `BBMM`, `mkde`, `MovementAnalysis` and `move`. They provide a variety of methods, from simple Minimum convex polygons (MCP) (Mohr, 1947) to more complex probabilistic Utilization distributions (UD) (Van Winkle, 1975), potentially accounting for the temporal autocorrelation in tracking data, as detailed below.

- `adehabitatHR` (2010, CRAN, active) contains a comprehensive list of methods to estimate home ranges: convex hull methods like MCP, clustering techniques, Local convex hulls (LoCoH) (Getz *et al.*, 2007) and the characteristic hull method Downs & Horner (2009); UD methods like kernel home ranges, also with the modification from Benhamou & Cornélis (2010) to account for boundaries, and methods to account for temporal autocorrelation between locations (Brownian bridge kernel method) (Bullard, 1991); biased random bridge kernel method also known as movement-based kernel estimation (Benhamou & Cornélis, 2010; Benhamou, 2011); and product-kernel algorithm, Horne *et al.* (2007).

- `rhr` (Signer & Balkenhol, 2015) provides a graphical user interface to estimate home ranges using several non-movement based methods, such as parametric home ranges, MCP, kernel UD, or local convex hulls, as well as the Brownian Bridge kernel method (as a wrapper to the `adehabitatHR` function). Complementary analyses include time to statistical independence, site fidelity test (against random permutation of step lengths and angles), among others.

- `T-LoCoH` is focused on constructing home-range hulls (Lyons *et al.*, 2013). A time-scale distance metric and a set of different nearest-neighbor criteria are available to choose which points to consider in a same hull. Hull metrics for space use, such as number of revisitations (repeated visits of an individual to the same hull) and their durations are also computed. Although the package was originally implemented for GPS data, it can be used for tracking data in general.

- `BBMM` (2010, CRAN, inactive), `MovementAnalysis` and `mkde` (2014, CRAN, inactive) use Brownian bridge movement models to obtain utilization distributions. `mkde` allows for a 3D extension of the Brownian bridges (Tracey *et al.*, 2014).

- `move`, in turn, calculates UDs of tracking data via dynamic Brownian Bridge modeling (Kranstauber *et al.*, 2012) or uses MCP for home range estimation; for the latter, it imports functions from `adehabitatHR`.

Habitat use

The role of habitat features on animal space use, or habitat selection, can be investigated with any of the following four packages.

- `adehabitatHS` (2010, CRAN, active) provides several tools for exploratory habitat selection analyses, from simple univariate analyses, such as resource selection ratios (Manly *et al.*, 2007) or compositional analysis (Aebischer *et al.*, 1993), to a family of multivariate analyses based on the geometric concept of ecological niche (Hutchinson, 1957), or the Outlying Mean Index (OMI) (Dolédec *et al.*, 2000) and the K-select (Calenge *et al.*, 2005) at the individual level.

- `hab` (2015, GitHub, inactive) enhances several utility functions of `adehabitatHS`, `adehabitatHR` and `adehabitatLT`, and provides core functions to prepare, fit and evaluate Step Selection Functions (SSFs) (Fortin *et al.*, 2005) while relying on `adehabitatLT` classes to handle trajectories. SSFs essentially investigate habitat selection along the trajectory, by comparing habitat features at observed step locations with those at alternative random steps taken from the same starting point (Thurfjell *et al.*, 2014).

- `amt` contains functions and wrappers to streamline the process of fitting SSFs from pairs of coordinates defining locations, to the conditional logistic regression model.

- In `ctmcmove`, the role of habitat features is investigated through a glm-framework, for which these features are rasterized, and the animal track is first imputed via functional movement modeling, then discretized in a grilled space (more details in Hanks *et al.* (2015)).

Non-conventional approaches for space use

Other non-conventional approaches for investigating space use from tracking data can be found in `ctmm`, `moveNT` (2017, GitHub, active), `recurse`, `rsMove`, `feedr` and `VTrack`.

- `ctmm` fits several candidate movement models via a variogram regression approach (Fleming *et al.*, 2014); those models can account for spatial autocorrelation in locations and periodicity in space use if required (Péron *et al.*, 2016). Space utilization is computed via an autocorrelated kernel estimator, where the autocorrelation term comes from the movement model previously fitted (Fleming *et al.*, 2015).

- `moveNT` tackles space use analysis via network graph theory (Bastille-Rousseau *et al.*, 2018). We summarize the procedure here: 1) Tracking data is represented over a gridded map and the number of transitions between pixels are counted. 2) The adjacency matrix, i.e. the counts of transitions, are then used to compute some network metrics at the pixel level. 3) A Gaussian mixture model is fitted to one of the metrics (user choice) to cluster values in two groups potentially representing patches and interpatch movement.

- `rsMove` implements a procedure to identify feeding sites from tracking data as a function of environmental variables (remote sensing data). It uses a random forest classification model; however, there is no information about how to fix the parameters of the model, so users should be careful when using this method. An application of the method can be found in Remelgado *et al.* (2017), but the parametrization is not described in the manuscript.

- `recurse` (2017, CRAN, active) aims at computing number of revisitations to pre-defined areas and their duration. These areas can be defined by the user by entering their center of gravity (by default, the fixes in the track) and a radius. The vignette gives important criteria to use the functions and interpret the results, though no there are no citations of scientific publications. `feedr` and `VTrack`, for radio and acoustic telemetry data, respectively, provide statistics on animal visits to given logger stations/receivers.

Trajectory simulation

Simulating trajectories can be useful to test hypotheses concerning movement, by comparing the patterns of simulated movement from several alternative theoretical models, or the patterns in the simulated movement to those of real observed tracks. In addition, simulation allows the quantification of estimator uncertainty by parametric bootstrapping (e.g. Michelot *et al.* (2016)). As with other types of data analysis, simulations highly depend on the model used by the researcher. The trajectory simulation with the existing packages are mainly based on Hidden Markov models, correlated random walks, Brownian motions, Lévy walks or Ornstein-Uhlenbeck processes.

Packages that allow simulation of trajectories from movement models fitted to tracking data are `moveHMM`, `momentuHMM` (HMMs), `bsam` (DCRWs), `crawl` (CTCRWs), `argosTrack` (discrete and continuous CRWs, and OU processes) and `ctmm` (several continuous time movement models). These packages have been described in the previous sections, and the simulations are presented as additional features after model fitting in their documentation. Another package for model fitting and simulation is `smam` (2013, CRAN, inactive). It can fit and simulate two types of movement models: Brownian motions with measurement error (Pozdnyakov *et al.*, 2014) and moving-resting processes with Brownian motion for the moving stage (Yan *et al.*, 2014).

Other packages implemented simulation functions when movement parameters are known; i.e. there is no previous fitting to tracking data. `adehabitatLT` proposes trajectory simulation using Brownian motion-based models, Lévy walks, CRWs and bivariate OU motion. `trajr` allows for CRWs, directed random walks (direction is equal to a constant plus a small noise), Brownian motion and Lévy walks. `moveNT` enables simulation of movement within and between patches. Movement within patches can follow an OU process (wrapping functions from `adehabitatLT`) or two-states movement model (wrapping functions from `moveHMM`). Movement between patches is simulated via a Brownian bridge movement model (from `adehabitatLT`).

`SiMRiv` (2016, CRAN, active) is another package created for simulation and it can take into account environmental constraints. It allows simulating random walks, correlated random walks, multi-state movement and constraining the area by an environmental resistance variable – defined by the user – that conditions the direction of the movement. The available documentation gives a detailed explanation of the simulation process.

Other analyses of tracking data

Interactions

Interactions between individuals can be assessed using metrics from `wildlifeDI` (2014, CRAN, active), which quantifies the dynamic interaction between two tracks of distinct individuals through several metrics (see Long *et al.* (2014) for details). The package relies on ltraj objects (the `adehabitatLT` data class). Other packages that include functions investigating interaction are `TrajDataMining` and `movementAnalysis`: `TrajDataMining` can identify potential partners based on distance and time thresholds fixed by the user and `MovementAnalysis` computes the expected duration of encounters at each location for every pair of IDs, based on a Brownian Bridge movement model fitted to the tracking data.

Movement similarity

`SimilarityMeasures` (2015, CRAN, inactive) assesses similarity between trajectories using metrics such as the longest common subsequence (LCSS), Fréchet distance, edit distance and dynamic time warping (DTW). Curious readers can refer to Magdy *et al.* (2015) for a brief review on trajectory similarity measures. `trajectories` also computes the Fréchet distance for two trajectories.

Population size

`caribou` was specifically created to estimate population size from Caribou tracking data, but can also be used for wildlife populations with similar home-range behavior. The methods implemented here are described in Rivest *et al.* (1998). The user needs to specify parameters concerning the size of each detected group, the number of collars in each of these groups and the detection model to use.

Inferring environmental variables from tracking data

Using tracking data to infer an environmental variable is the objective of `moveWindSpeed` (2016, CRAN, active). It uses bird tracking data to estimate wind speed via a maximum likelihood approach (Weinzierl *et al.*, 2016). The estimation is only performed for segments where the bird is circling in a thermal, so a function in the package identifies those segments. Speed is modeled as a mean with an autocorrelated drift.

Database management

Finally, `rpostgisLT` handles database management for trajectory data by integrating R and the 'PostgreSQL/PostGIS' database system. The package relies on `adehabitatLT`, and allows users to run the analyses on their database that can be usually done with an ltraj object in `adehabitatLT`.

Analysis of biologging but not tracking data

Time-depth recorders collect data on depth, velocity and other parameters as animals move through the water. These biologging data by itself does not allow obtaining tracking data (x,y,t) and thus comparable analyses to the ones presented above. `diveMove` and `rbl`, the latter also for accelerometer data, are the two packages implementing TDR data analysis. `diveMove` contains functions to identify wet and dry periods in the series, calibrate depth and speed sensor readings, identify individual dives and their phases, summarize statistics per dive and plot the data. With `rbl`, accelerometry data are used for identifying prey catch attempts (Viviant *et al.*, 2010) and swimming effort from frequency and magnitude of tail movement (Bras *et al.*, 2016). Other functions allow the extraction of summary statistics from dives (e.g. maximum depth), fitting broken stick models (i.e. piecewise linear regression) to dive series and identifying dive phases.

Accelerometry data is also used in human studies, primarily to assess levels of physical activity. Six R packages focus on the analysis of human accelerometry data, mainly to describe periodicity and levels of activity. `accelerometry`, `GGIR` and `PhysicalActivity` identify wear and non-wear time of the accelerometers. `nparACT` Computes descriptive statistics such as interdaily stability, intradaily variability and relative amplitude of activity (Blume *et al.*, 2016). `acc`, `GGIR` and `pawacc` classify wear data into different levels of activity (e.g. sedentary, moderate and vigorous) using thresholds given by the user, and offer some functions for visual representation of the data and descriptive statistics on the types of activities. Additionally, `acc` allows for activity simulation via Hidden Markov modeling.

The packages described in this section are not tracking packages and will not be discussed in the next sections, but readers should take them into consideration when analyzing TDR and accelerometer data. The packages focused on human data can be used for animal data as well.

Packages documentation

Documentation in the form of manuals, vignettes, tutorials or published articles is key to understanding how to use a package's features for the first time, especially if the package contains a large number of functions and tools. Without proper user testing and peer editing, package documentation can lead to large gaps of understanding and lower usefulness for users. If functions and work flows are not expressly defined, a packages capacity to help users is undermined. Vignettes can act as road maps for the user, and published articles expressly pertaining to the package help provide context and guidance on the internal workings of functions. Moreover, since packages make specific methods available for R users, the documentation should not only explain how to use the packages but also explain or provide references for the methods.

To assess package documentation, an online survey was conducted between August and October 2018. Questions in the survey regarded helpfulness of package documentation and the frequency of package use. The survey was posted on Twitter and sent to several email lists of ecology and R related groups, and completed by 225 people. The exact formulation of each question in the survey, summarized results and a discussion on the representativity of the survey are shown in SupplementaryFile. All of the packages in this review were considered in the survey except for `trajr`, which was added to the review after the survey started.

We identified 12 packages (for which we had at least 10 respondents) as having 'great documentation', meaning that more than 75% of the respondents expressed that the documentation was either good (allowing the user to do everything they wanted and needed to do with the package) or excellent (allowing users to do even more than what they initially planned because of the excellent quality of the information). These are: `momentuHMM` (93.8%), `moveHMM` (89.5%), `adehabitatLT` (88.6%), `adehabitatHS` (86.1%), `adehabitatHR` (83.1%), `EMbC` (81.8%), `wildlifeDI` (81.3), `ctmm` (80.0%), `GeoLight` (76.9%), `move` (76.6%), `recurse` (76.5%), and `bsam` (76.2%) (see Fig. 3). From these group of packages, `momentuHMM` and `move` offer manuals and vignettes, while all the others offer in addition scientific articles centered on the package. Also, if we look at the packages used by more than 50 participants, all except for `crawl` had 'great documentation'.

Package developers could use the results of this survey (see Supplementary File for more details) as a guidance to decide on whether to improve the documentation of their packages so more researchers can use them.

Links between the packages

We analyzed the links between tracking packages. If a package needs functions that have already created by another package, they can use those functions by declaring this dependency in the description file of the package under 'Depends on ', 'Imports' or 'Linking to' categories. Theoretically there are some differences between the three, but in practice developers mix those groups, so we consider them as part of the same concept: dependency. A package can also suggest using other packages, for instance, a package focused on data analysis can suggest, in case data has to be cleaned first, the use of a package that allows post-processing. Since most packages define their own data classes, packages suggesting others often offer functions that enable working with data classes with the other packages.

Suggestions can also be declared in the description file, but we also took into account the packages suggested in vignette examples; these analyses

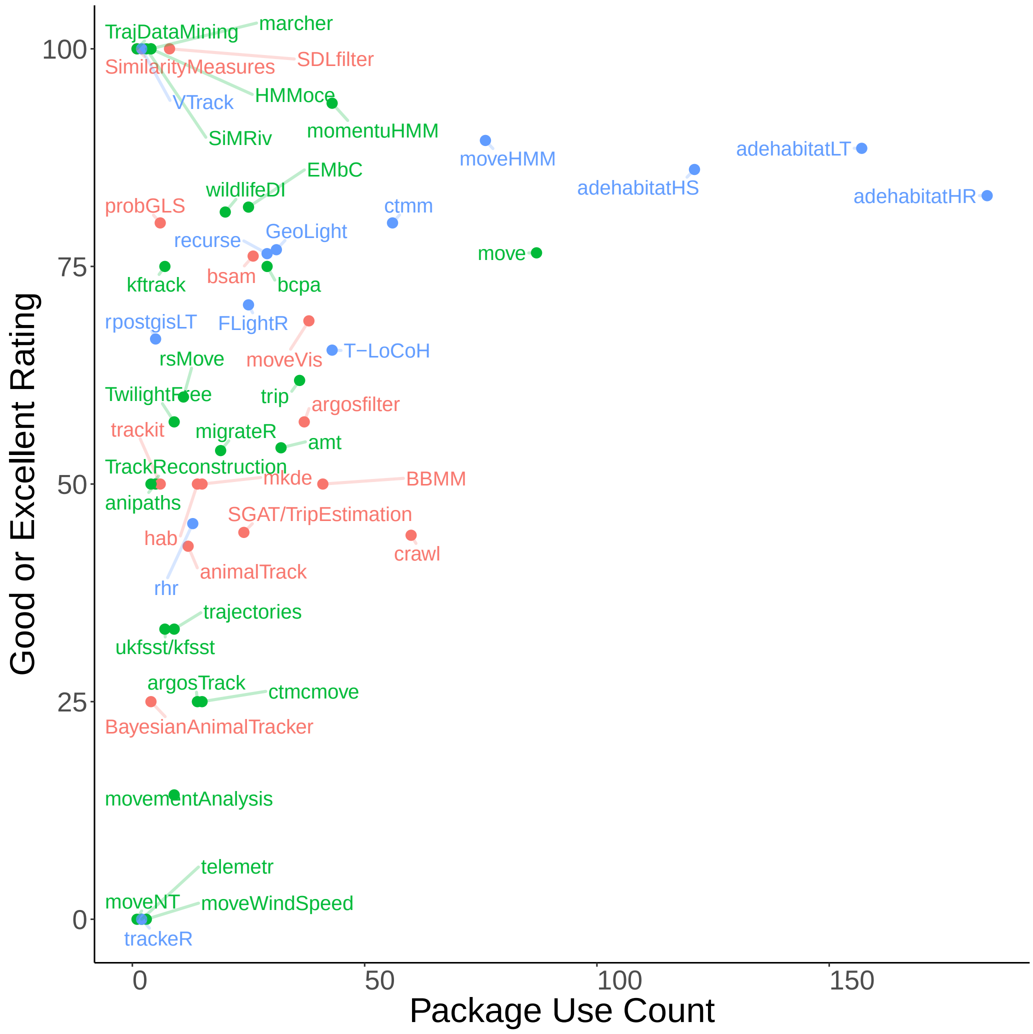

Fig. 3 Packages with good and excellent documentation (survey results). x axis: Number of survey participants that used each package; y axis: percentage of users of each package that considered the documentation good or excellent. Text color in red corresponds to packages with standard documentation only, green is for packages with vignettes, and blue is for packages that also released other types of documentation.

were performed in August 2018. With the dependency and suggestion information, we performed a graph analysis (Fig. 4). 39 packages in total showed some level of connections among them (30 in the form of one large group and three other small ones), while 18 (32%) of the packages worked in isolation. `adehabitatLT` and `move` were the most suggested/depended-on packages with 14 and 7 links to them, respectively. Indeed, many packages use functions compatible with the ltraj data-class from `adehabitatLT`, and some others with the move class from `move`. `amt` suggests most packages than any other (6), and it provides coercion methods for data-classes from the packages it suggests.

Discussion

Conclusions from the review

As the quantity and diversity of biologging data increases, so does the need for suitable statistical techniques and software resources. These tools are essential to convert data into ecologically meaningful measures and analyze outputs to test hypotheses. Through a systematic search we identified 57 R packages aimed at processing or analyzing tracking data. The packages offer tools for data processing, visualization, computation of statistics for track description, path reconstruction, behavioral pattern identification, space use characterization and trajectory simulation, among others. The main issues for extracting and processing tracking data from biologging devices are already covered by the reviewed packages, and the main types of analyses are covered as well. In some cases, there is more than one package imple-

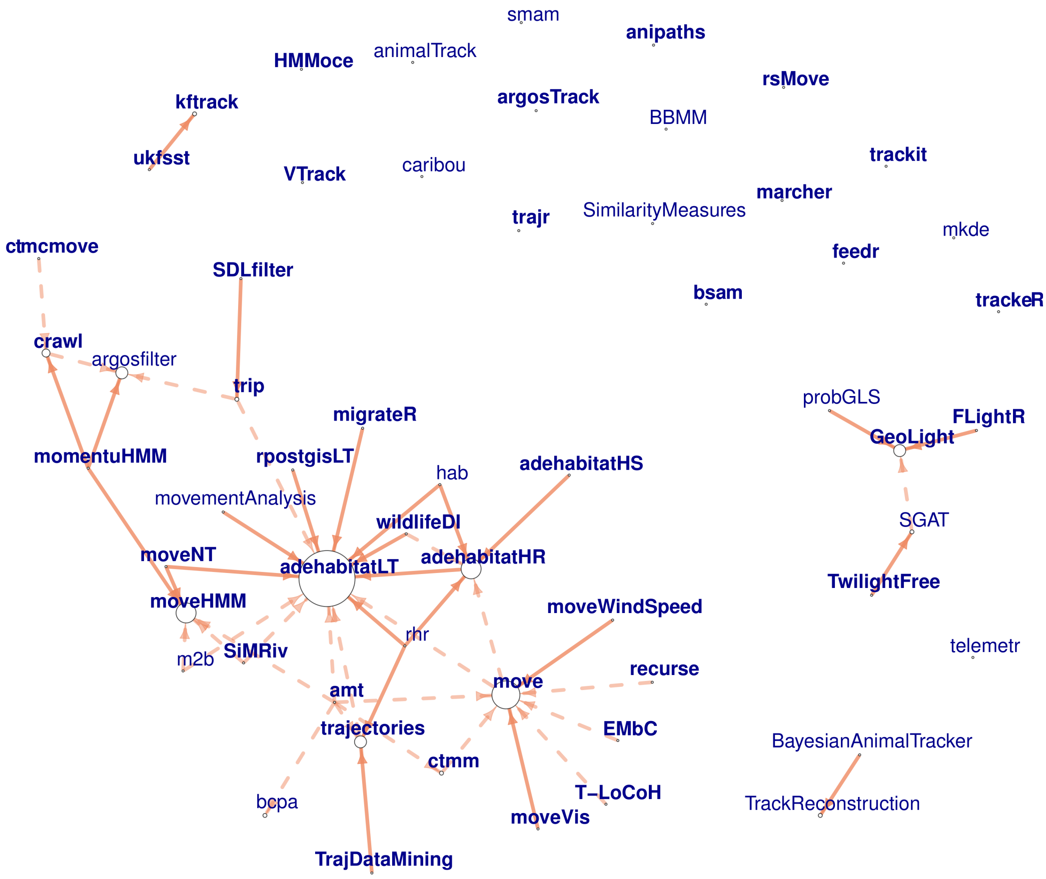

Fig. 4 Social network representation of the dependency and suggestion between tracking packages. The arrows go towards the package the others suggest (dashed arrows) or depend on (solid arrows). Bold font corresponds to active packages. The size of the circle is proportional to the number of packages that suggest or depend on this one.

menting the same type of analysis with the same or very similar approaches, such as `animalTrack` and `TrackReconstruction` for dead-reckoning, `BBMM`, `MovementAnalysis` and `mkde` for Brownian bridge movement models, to cite some examples. Few packages focus on collective motion: mainly `wildlifeDI` and, to a lesser degree, `TrajDataMining` and `movementAnalysis` allow computing descriptive metrics on encounters between individuals, periods of proximity or other metrics of interaction. The lack of tools to analyze collective movement beyond descriptive statistics may be the reflection of a remaining challenge in ecology, and not in computational programming. Overall, the review highlighted the abundance and depth of analytical tools available, and identified a need to improve accessibility to existing packages rather than developing new packages.

Integration over proliferation

Transparency in science is facilitated by the sharing of data and analysis tools, including code. This has result in a general tendency in the scientific community to convert functions into publicly available packages. In movement ecology, this has translated into a proliferation of R tracking packages, many of them, isolated from the rest (Fig. 4), and with similar goals and methods, as described before. While it is promising that there is a large amount of code available to the scientific community, it is hard to maintain an overview of their functionality and availability. Here we presented a list of 57 packages but the number is expected to keep increasing, with possibilities of task repetitions and disconnected from each other. Due to the already overwhelming number of tracking packages, we suggest developers only create new packages in the future when they represent a substantial contribution to the scientific and programming community. This is difficult to implement however, as package necessity is not assessed through any repository. We recommend future developers reflect on how their package fits into the landscape of current movement ecology packages, in conjunction with users. Methodological journals, which often publish R packages, should ensure that users are involved in the review process and that necessity is a component of the publication decision.

Recommendations

This work is not intended to tell ecologists exactly which packages to use, but to provide them the catalog of tracking packages, a description of their function and show the similarities and differences between them. We suggest researchers use packages with good documentation, that are actively maintained and that have a large number of users. Good documentation facilitates the initial use of a package, and offers developers an opportunity to link this with pre-existing packages. A regularly maintained package means that there is a person or team behind it, and that, in case of an error in the package, it will likely be fixed rapidly and a new version will be available. A package that has a large number of users means 1) more chances to spot bugs in the package, calling the attention of the maintainer for a rapid repair, and thus making it better and 2) more chances of getting additional guidance on package use from other users. Regarding the methods available in the packages, we previously stated the importance of describing them and citing references. On the other hand, it is the responsibility of the researchers to solely apply a method if they correctly understand it, and not only because it is available in a package.

When developers are working on new packages, we recommend they consider the following questions:

- Does your package fill a gap or need? Does a function of the package perform a novel task that does not already exist in another published package? Can those functions be instead added to an existing package? Developers should contemplate the possibility (and appropriateness) of contacting authors of existing and actively maintained packages to add functions into them. We also suggest the authors of the existing packages to be open to considering the integration of new functions (and new collaborators) to their package.

- Does the package handle commonly used data classes (e.g. sp, ltraj), so that it is compatible with the use of other packages?

- Is the documentation clear, exhaustive on the functions, with methods description or references available? The latter is even more important if the package implements a new method of analysis. Citing papers rather than explaining the methods is the easiest way to back their procedures up, but authors should consider that not all scientific articles are open access, which means that some potential package users could have free access to the package but not to its methodological support. Worked examples and vignettes can enable researchers to learn the package more easily, minimizing the need for additional support.

- Who will maintain the package over time? If a PhD student or a postdoctoral researcher creates a package and after a while is no longer invested in its maintenance, then the lab's PI could take responsibility for the package or delegate responsibility to someone else.

Taking these questions into account can help maximizing the usefulness of individual packages and strengthen the links between package developers. A stronger community of developers highly benefits the users: a limited number of strongly-related packages, that are continuously maintained with new functions (and sound documentation) added, are easy to follow and use. Communication between movement ecologists is essential to foster this community.

Summary

The ability to analyze biologging data is essential to answer ecological questions. While the abundance of devices is enabling researchers to collect ever increasing amounts of data, without the necessary tools to interpret this data, their contribution to the field of ecology is unlikely to be realized. Programmers have responded to this need developing up to 57 R packages, 19% (11) of those in the last year; this review serves as a map of the tools implemented by the packages for data analysis in movement ecology. An increased accessibility and understanding of existing packages will help the advancement of research in this field, allowing researchers to continue to address novel and exciting questions.

Acknowledgments

A lot of people are to thank here. HFSP Seabird Sound. Alcohol. So on.

References

Aebischer, N.J., Robertson, P.A. & Kenward, R.E. (1993) Compositional analysis of habitat use from animal radio-tracking data. *Ecology*, **74**, 1313–1325.

Afanasyev, V. (2004) A miniature daylight level and activity data recorder for tracking animals over long periods. *Memoirs of National Institute of Polar Research, Special Issue*, **58**, 227–233.

Albertsen, C.M., Whoriskey, K., Yurkowski, D., Nielsen, A. & Flemming, J.M. (2015) Fast fitting of non-Gaussian state-space models to animal movement data via Template Model Builder. *Ecology*, **96**, 2598–2604. ISSN 0012-9658.

Bastille-Rousseau, G., Douglas-Hamilton, I., Blake, S., Northrup, J.M. & Wittemyer, G. (2018) Applying network theory to animal movements to identify properties of landscape space use. *Ecological Applications*, **28**, 854–864. ISSN 1939-5582.

Benhamou, S. (2011) Dynamic Approach to Space and Habitat Use Based on Biased Random Bridges. *PLOS ONE*, **6**, e14592. ISSN 1932-6203.

Benhamou, S. & Cornélis, D. (2010) Incorporating movement behavior and barriers to improve kernel home range space use estimates. *Journal of Wildlife Management*, **74**, 1353–1360. ISSN 0022-541X.

Bidder, O.R., Walker, J.S., Jones, M.W., Holton, M.D., Urge, P., Scantlebury, D.M., Marks, N.J., Magowan, E.A., Maguire, I.E. & Wilson, R.P. (2015) Step by step: Reconstruction of terrestrial animal movement paths by dead-reckoning. *Movement Ecology*, **3**, 1–16. ISSN 20513933.

Bindoff, A.D., Wotherspoon, S.J., Guinet, C., Hindell, M.A. & Orme, D. (2017) Twilight-free geolocation from noisy light data. *Methods in Ecology and Evolution*, **0**. ISSN 2041-210X.

Blume, C., Santhi, N. & Schabus, M. (2016) 'nparACT' package for R: A free software tool for the non-parametric analysis of actigraphy data. *MethodsX*, **3**, 430–435. ISSN 2215-0161. URL `http://www.sciencedirect.com/science/article/pii/S221501611630022X`

Bras, Y.L., Jouma'a, J., Picard, B. & Guinet, C. (2016) How Elephant Seals (Mirounga leonina) Adjust Their Fine Scale Horizontal Movement and Diving Behaviour in Relation to Prey Encounter Rate. *PLOS ONE*, **11**, e0167226. ISSN 1932-6203.

Braun, C.D., Galuardi, B., Thorrold, S.R. & Parrini, F. (2017) HMMoce: An R package for improved geolocation of archival-tagged fishes using a hidden Markov method. *Methods in Ecology and Evolution*, **0**. ISSN 2041-210X.

Buderman, F.E., Hooten, M.B., Ivan, J.S. & Shenk, T.M. (2016) A functional model for characterizing long-distance movement behaviour. *Methods in Ecology and Evolution*, **7**, 264–273. ISSN 2041210X.

Bullard, F. (1991) *Estimating the home range of an animal: a Brownian bridge approach*. University of North Carolina Chapel Hill.

Calenge, C., Dufour, A. & Maillard, D. (2005) K-select analysis: a new method to analyse habitat selection in radio-tracking studies. *Ecological Modelling*, **186**, 143–153.

Costa, D.P., Robinson, P.W., Arnould, J.P.Y., Harrison, A.L., Simmons, S.E., Hassrick, J.L., Hoskins, A.J., Kirkman, S.P., Oosthuizen, H., Villegas-Amtmann, S. & Crocker, D.E. (2010) Accuracy of ARGOS Locations of Pinnipeds at-Sea Estimated Using Fastloc GPS. *PLOS ONE*, **5**, e8677. ISSN 1932-6203.

DeCelles, G. & Zemeckis, D. (2014) Chapter Seventeen - Acoustic and Radio Telemetry. *Stock Identification Methods (Second Edition)* (eds. S.X. Cadrin, L.A. Kerr & S. Mariani), pp. 397–428. Academic Press. ISBN 978-0-12-397003-9.

Dolédec, S., Chessel, D. & Gimaret-Carpentier, C. (2000) Niche Separation in Community Analysis: A New Method. *Ecology*, **81**, 2914–2927. ISSN 1939-9170.

Downs, J.A. & Horner, M.W. (2009) A characteristic-hull based method for home range estimation. *Transactions in GIS*, **13**, 527–537.

Ekstrom, P.A. (2004) An advance in geolocation by light. *Memoirs of National Institute of Polar Research. Special issue*, **58**, 210–226. ISSN 03860744.

Fleming, C., Fagan, W., Mueller, T., Olson, K., Leimgruber, P. & Calabrese, J. (2015) Rigorous home range estimation with movement data: a new autocorrelated kernel density estimator. *Ecology*, **96**, 1182–1188.

Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2014) From Fine-Scale Foraging to Home Ranges: A Semivariance Approach to Identifying Movement Modes across Spatiotemporal Scales. *The American Naturalist*, **183**, E154–E167. ISSN 0003-0147.

Fortin, D., Beyer, H.L., Boyce, M.S., Smith, D.W., Duchesne, T. & Mao, J.S. (2005) Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park. *Ecology*, **86**, 1320–1330.

Freitas, C., Lydersen, C., Fedak, M.A. & Kovacs, K.M. (2007) A simple new algorithm to filter marine mammal Argos locations. *Marine Mammal Science*, **24**, 315–325. ISSN 0824-0469.

Garriga, J., Palmer, J.R.B., Oltra, A. & Bartumeus, F. (2016) Expectation-Maximization Binary Clustering for Behavioural Annotation. *PLOS ONE*, **11**, e0151984. ISSN 1932-6203.

Getz, W.M., Fortmann-Roe, S., Cross, P.C., Lyons, A.J., Ryan, S.J. & Wilmers, C.C. (2007) LoCoH: nonparameteric kernel methods for constructing home ranges and utilization distributions. *PloS one*, **2**, e207.

Gurarie, E., Andrews, R.D. & Laidre, K.L. (2009) A novel method for identifying behavioural changes in animal movement data. *Ecology letters*, **12**, 395–408. ISSN 1461-0248.

Gurarie, E., Cagnacci, F., Peters, W., Fleming, C.H., Calabrese, J.M., Mueller, T. & Fagan, W.F. (2017) A framework for modelling range shifts and migrations: asking when, whither, whether and will it return. *Journal of Animal Ecology*, **86**, 943–959. ISSN 1365-2656.

Hanks, E.M., Hooten, M.B. & Alldredge, M.W. (2015) Continuous-time discrete-space models for animal movement. *Annals of Applied Statistics*, **9**, 145–165. ISSN 19417330.

Horne, J.S., Garton, E.O., Krone, S.M. & Lewis, J.S. (2007) Analyzing animal movements using Brownian bridges. *Ecology*, **88**, 2354–2363.

Hutchinson, G. (1957) Concluding remarks Cold Spring Harbor Symposia on Quantitative Biology, 22: 415–427. *GS SEARCH.*

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**, 1208–1215. ISSN 0012-9658.

Jonsen, I. (2016) Joint estimation over multiple individuals improves behavioural state inference from animal movement data. *Scientific reports*, **6**, 20625–20625. ISSN 2045-2322.

Jonsen, I., Basson, M., Bestley, S., Bravington, M., Patterson, T., Pedersen, M., Thomson, R., Thygesen, U. & Wotherspoon, S. (2013) State-space models for bio-loggers: A methodological road map. *Deep Sea Research Part II: Topical Studies in Oceanography*, **88-89**, 34–46. ISSN 09670645.

Kranstauber, B., Kays, R., Lapoint, S.D., Wikelski, M. & Safi, K. (2012) A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *Journal of animal ecology*, **81**, 738–746. ISSN 1365-2656.

Lam, C.H., Nielsen, A. & Sibert, J.R. (2010) Incorporating sea-surface temperature to the light-based geolocation model TrackIt. *Marine Ecology Progress Series*, **419**, 71–84. ISSN 0171-8630, 1616-1599.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. & Morales, J.M. (2012) Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, **93**, 2336–2342. ISSN 0012-9658. URL `http://www.ncbi.nlm.nih.gov/pubmed/23236905`

Liu, Y., Zidek, J.V., Trites, A.W. & Battaile, B.C. (2016) Bayesian data fusion approaches to predicting spatial tracks: Application to marine mammals. *The Annals of Applied Statistics*, **10**, 1517–1546. ISSN 1932-6157, 1941-7330.

Long, J.A., Nelson, T.A., Webb, S.L. & Gee, K.L. (2014) A critical examination of indices of dynamic interaction for wildlife telemetry studies. *Journal of animal ecology*, **83**, 1216–1233. ISSN 1365-2656.

Lyons, A.J., Turner, W.C. & Getz, W.M. (2013) Home range plus: a space-time characterization of movement over real landscapes. *Movement Ecology*, **1**, 2. ISSN 2051-3933.

Magdy, N., Sakr, M.A., Mostafa, T. & El-Bahnasy, K. (2015) Review on trajectory similarity measures. *Intelligent Computing and Information Systems (ICICIS), 2015 IEEE Seventh International Conference on*, pp. 613–619. IEEE.

Manly, B.F.L., McDonald, L., Thomas, D.L., McDonald, T.L. & Erickson, W.P. (2007) *Resource selection by animals: statistical design and analysis for field studies*. Springer Science & Business Media.

McClintock, B.T. (2017) Incorporating Telemetry Error into Hidden Markov Models of Animal Movement Using Multiple Imputation. *Journal of Agricultural, Biological and Environmental Statistics*, **22**, 249–269. ISSN 1085-7117, 1537-2693.

McClintock, B.T., King, R., Len, T., Matthiopoulos, J., McConnell, B.J. & Morales, J.M. (2012) A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs*, **82**, 335–349. ISSN 0012-9615.

Meratnia, N. & de By, R.A. (2004) Spatiotemporal Compression Techniques for Moving Point Objects. *Advances in Database Technology - EDBT 2004* (eds. E. Bertino, S. Christodoulakis, D. Plexousakis, V. Christophides, M. Koubarakis, K. Böhm & E. Ferrari), Lecture Notes in Computer Science, pp. 765–782. Springer Berlin Heidelberg. ISBN 978-3-540-24741-8.

Merkel, B., Phillips, R.A., Descamps, S., Yoccoz, N.G., Moe, B. & Strøm, H. (2016) A probabilistic algorithm to process geolocation data. *Movement Ecology*, **4**, 26. ISSN 2051-3933.

Michelot, T., Langrock, R., Patterson, T.A. & McInerny, G. (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. ISSN 2041-210X.

Mohr, C.O. (1947) Table of equivalent populations of North American small mammals. *The American Midland Naturalist*, **37**, 223–249.

Montenbruck, O. & Pfleger, T. (2013) *Astronomy on the personal computer*. Springer.

Nielsen, A., Bigelow, K.A., Musyl, M.K. & Sibert, J.R. (2006) Improving light-based geolocation by including sea surface temperature. *Fisheries Oceanography*, **15**, 314–325. ISSN 1365-2419.

Patterson, T.A., Thomas, L., Wilcox, C., Ovaskainen, O. & Matthiopoulos, J. (2008) State-space models of individual animal movement. *Trends in Ecology and Evolution*, **23**, 87–94.
URL `http://dx.doi.org/10.1016/j.tree.2007.10.009`

Pozdnyakov, V., Meyer, T., Wang, Y.B. & Yan, J. (2014) On modeling animal movements using Brownian motion with measurement error. *Ecology*, **95**, 247–253. ISSN 0012-9658.

Péron, G., Fleming, C.H., de Paula, R.C. & Calabrese, J.M. (2016) Uncovering periodic patterns of space use in animal tracking data with periodograms, including a new algorithm for the Lomb-Scargle periodogram and improved randomization tests. *Movement Ecology*, **4**, 19. ISSN 2051-3933.

Rakhimberdiev, E., Saveliev, A., Piersma, T. & Karagicheva, J. (2017) FLightR: an r package for reconstructing animal paths from solar geolocation loggers. *Methods in Ecology and Evolution*, **8**, 1482–1487. ISSN 2041-210X.

Rakhimberdiev, E., Winkler, D.W., Bridge, E., Seavy, N.E., Sheldon, D., Piersma, T. & Saveliev, A. (2015) A hidden Markov model for reconstructing animal paths from solar geolocation loggers using templates for light intensity. *Movement Ecology*, **3**, 25. ISSN 2051-3933.

Remelgado, R., Leutner, B., Safi, K., Sonnenschein, R., Kuebert, C. & Wegmann, M. (2017) Linking animal movement and remote sensing – mapping resource suitability from a remote sensing perspective. *Remote Sensing in Ecology and Conservation*, **0**. ISSN 2056-3485.

Rivest, L.P., Couturier, S. & Crépeau, H. (1998) Statistical Methods for Estimating Caribou Abundance Using Postcalving Aggregations Detected by Radio Telemetry. *Biometrics*, **54**, 865–876. ISSN 0006-341X.

Shimada, T., Jones, R., Limpus, C. & Hamann, M. (2012) Improving data retention and home range estimates by data-driven screening. *Marine Ecology Progress Series*, **457**, 171–180. ISSN 0171-8630, 1616-1599.

Shimada, T., Limpus, C., Jones, R., Hazel, J., Groom, R. & Hamann, M. (2016) Sea turtles return home after intentional displacement from coastal foraging areas. *Marine Biology*, **163**, 8. ISSN 0025-3162, 1432-1793.

Sibert, J.R., Musyl, M.K. & Brill, R.W. (2003) Horizontal movements of bigeye tuna (Thunnus obesus) near Hawaii determined by Kalman filter analysis of archival tagging data. *Fisheries Oceanography*, **12**, 141–151. ISSN 1365-2419.

Signer, J. & Balkenhol, N. (2015) Reproducible home ranges (rhr): A new, user-friendly R package for analyses of wildlife telemetry data. *Wildlife Society Bulletin*, **39**, 358–363. ISSN 19385463.

Spitz, D.B., Hebblewhite, M. & Stephenson, T.R. (2017) 'MigrateR': Extending model-driven methods for classifying and quantifying animal movement behavior. *Ecography*, pp. 788–799. ISSN 16000587.

Sumner, M.D., Wotherspoon, S.J. & Hindell, M.A. (2009) Bayesian Estimation of Animal Movement from Archival and Satellite Tags. *PLOS ONE*, **4**, e7324. ISSN 1932-6203.

Thurfjell, H., Ciuti, S. & Boyce, M.S. (2014) Applications of step-selection functions in ecology and conservation. *Movement Ecology*, **2**, 4. ISSN 2051-3933.

Tracey, J.A., Sheppard, J., Zhu, J., Wei, F., Swaisgood, R.R. & Fisher, R.N. (2014) Movement-Based Estimation and Visualization of Space Use in 3D for Wildlife Ecology and Conservation. *PLOS ONE*, **9**, e101205. ISSN 1932-6203.

Van Winkle, W. (1975) Comparison of several probabilistic home-range models. *The Journal of wildlife management*, pp. 118–123.

Viviant, M., Trites, A.W., Rosen, D.A.S., Monestiez, P. & Guinet, C. (2010) Prey capture attempts can be detected in Steller sea lions and other marine predators using accelerometers. *Polar Biology*, **33**, 713–719. ISSN 1432-2056.

Weinzierl, R., Bohrer, G., Kranstauber, B., Fiedler, W., Wikelski, M. & Flack, A. (2016) Wind estimation based on thermal soaring of birds. *Ecology and Evolution*, **6**, 8706–8718. ISSN 2045-7758.

Williams, H.J., Holton, M.D., Shepard, E.L.C., Largey, N., Norman, B., Ryan, P.G., Duriez, O., Scantlebury, M., Quintana, F., Magowan, E.A., Marks, N.J., Alagaili, A.N., Bennett, N.C. & Wilson, R.P. (2017) Identification of animal movement patterns using tri-axial magnetometry. *Movement Ecology*, **5**, 6. ISSN 2051-3933.

Wilson, R.P., Liebsch, N., Davies, I.M., Quintana, F., Weimerskirch, H., Storch, S., Lucke, K., Siebert, U., Zankl, S., Müller, G., Zimmer, I., Scolaro, A., Campagna, C., Plötz, J., Bornemann, H., Teilmann, J. & McMahon, C.R. (2007) All at sea with animal tracks; methodological and analytical solutions for the resolution of movement. *Deep Sea Research Part II: Topical Studies in Oceanography*, **54**, 193–210. ISSN 0967-0645.

Yan, J., Chen, Y.w., Lawrence-Apfel, K., Ortega, I.M., Pozdnyakov, V., Williams, S. & Meyer, T. (2014) A moving–resting process with an embedded Brownian motion for animal movements. *Population Ecology*, **56**, 401–415. ISSN 1438-3896, 1438-390X.