APPLICATION

`ctmm`: an R package for analyzing animal relocation data as a continuous-time stochastic process

Justin M. Calabrese[1,2]*, Chris H. Fleming[1,2] and Eliezer Gurarie[2]

[1]*Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Rd., Front Royal, VA 22630, USA; and*
[2]*Department of Biology, University of Maryland, College Park, MD 20742, USA*

Summary

1. Movement ecology has developed rapidly over the past decade, driven by advances in tracking technology that have largely removed data limitations. Development of rigorous analytical tools has lagged behind empirical progress, and as a result, relocation data sets have been underutilized.

2. Discrete-time correlated random walk models (CRW) have long served as the foundation for analyzing relocation data. Unfortunately, CRWs confound the sampling and movement processes. CRW parameter estimates thus depend sensitively on the sampling schedule, which makes it difficult to draw sampling-independent inferences about the underlying movement process. Furthermore, CRWs cannot accommodate the multiscale autocorrelations that typify modern, finely sampled relocation data sets.

3. Recent developments in modelling movement as a continuous-time stochastic process (CTSP) solve these problems, but the mathematical difficulty of using CTSPs has limited their adoption in ecology. To remove this roadblock, we introduce the `ctmm` package for the R statistical computing environment. `ctmm` implements all of the CTSPs currently in use in the ecological literature and couples them with powerful statistical methods for autocorrelated data adapted from geostatistics and signal processing, including variograms, periodograms and non-Markovian maximum likelihood estimation.

4. `ctmm` is built around a standard workflow that begins with visual diagnostics, proceeds to candidate model identification, and then to maximum likelihood fitting and AIC-based model selection. Once an accurate CTSP for the data has been fitted and selected, analyses that require such a model, such as quantifying home range areas via autocorrelated kernel density estimation or estimating occurrence distributions via time-series Kriging, can then be performed.

5. We use a case study with African buffalo to demonstrate the capabilities of `ctmm` and highlight the steps of a typical CTSP movement analysis workflow.

Key-words: autocorrelated kernel density estimation, home range estimation, non-Markovian maximum likelihood, periodogram analysis, tracking data, variogram analysis

Introduction

An understanding of animal movement can inform a wide range of biological topics including population and community ecology, animal physiology, disease spread, gene flow, and wildlife management and conservation (Nathan *et al.* 2008; Kays *et al.* 2015). Historically, a lack of movement data limited progress, but technological advances have fuelled an explosion of tracking studies (Kays *et al.* 2015), with the Movebank.org repository alone now featuring 1300+ data sets on 400+ species (Wikelski & Kays 2015). Now, the key bottleneck is the dearth of accurate and reliable analytical tools for extracting information from these accumulating data sources.

The discrete-time correlated random walk (CRW; Kareiva & Shigesada 1983; Turchin 1998; Codling, Plank & Benhamou 2008) and its extensions (Morales *et al.* 2004; Jonsen,

Flemming & Myers 2005; McClintock *et al.* 2012) have long been the standard approach to modelling movement data. However, the CRW confounds the sampling and movement processes and thus yields different results even when the exact same movement path is sampled at different resolutions (Bovet & Benhamou 1988; Codling & Hill 2005; Nams 2013; Fleming *et al.* 2014a). In other words, inferences based on the CRW frequently reflect more about the sampling schedule than about the underlying movement process. Secondly, real movement data often feature rich, multiscale autocorrelation structure (Boyce *et al.* 2010; Dray, Royer-Carenzi & Calenge 2010; Fleming *et al.* 2014a,b). The CRW, which is a first-order Markov chain model, cannot accommodate multiscale autocorrelations and therefore returns biased results when fitted to such data. Extending the CRW paradigm to higher-order (e.g. 2nd or 3rd order) discrete-time Markov processes is not sufficient because the autocorrelation time scales in the data can span orders of magnitude and can be completely unrelated to the data sampling interval (Fleming *et al.* 2014a,b).

*Correspondence author. E-mail: CalabreseJ@si.edu

Continuous-time stochastic process (CTSP) models solve these problems by separating the continuous-time movement process from the discrete-time sampling process (Blackwell 1997, 2003; Johnson *et al.* 2008; Gurarie & Ovaskainen 2011; Fleming *et al.* 2014a,b), which also renders CTSPs robust to irregular sampling. CTSPs span a range from no autocorrelation, through Markovian (position) autocorrelation, to multiscale non-Markovian autocorrelation (position and velocity autocorrelation, etc.), and thus can account for whatever autocorrelations are in the data (Fleming, Subaşı & Calabrese 2015). Adoption of CTSP models has, unfortunately, been hindered by a lack of statistical methods and comprehensive software. Though several highly specialized software packages implementing particular CTSP models exist (Appendix S1, Supporting information), none of these provides a general environment for working with CTSPs. The continuous-time movement modelling (ctmm) package for the R environment for statistical computing (R Core Team, 2015) fills this gap. ctmm integrates a broad range of CTSPs with statistical methods for autocorrelated data such as variogram (Fleming *et al.* 2014a) and periodogram analyses (Péron *et al.* In Review), and non-Markovian maximum likelihood estimation (Fleming *et al.* 2014b). ctmm facilitates detailed analyses of movement data including: (i) visual diagnostics for CTSPs, (ii) fitting CTSPs via maximum likelihood, (iii) AIC-based model selection and (iv) analyses that require a fitted, selected CTSP. Currently, this latter category includes home range estimation via autocorrelated kernel density estimation (AKDE; Fleming *et al.* 2015a) and occurrence distribution estimation via time-series Kriging (Fleming *et al.* 2016).

Models

Relocation data feature observations that are ordered in time. Consequently, position observations that are closer together in time will tend to be more similar than those farther apart in time. Relocation data sets therefore frequently feature autocorrelated positions. For finely sampled data sets, directional persistence in motion is often observable. Directional persistence is a consequence of autocorrelated velocities, which means that an individual's direction and speed at one point in time tend to be correlated with those quantities at adjacent times. Finally, at larger scales, most animals will exhibit a tendency to remain in a defined region or 'home range'. These three features, position autocorrelation, velocity autocorrelation and range residency, can be used to classify CTSP models (Table 1). ctmm includes all single-component CTSP models currently in use in the ecological literature (Table 1).

The independent identically distributed (IID) process assumes uncorrelated positions and velocities and is thus not typically recognized as a movement model. However, the IID assumption is implicit when performing conventional home range estimation (e.g. KDE). Brownian motion (BM) is a standard model of regular diffusion, where velocities are uncorrelated and space use is not limited (Einstein 1905; Horne *et al.* 2007). BM can be appropriate for data that are sampled both too coarsely to reveal velocity autocorrelation, and too briefly

Table 1. A summary of the movement models implemented in ctmm. We classify the models by whether or not they feature position autocorrelation (Pos. AC), velocity autocorrelation (Vel. AC) and a home range (H. Range). These models can be specified via the ctmm function by their autocorrelation time scales. The IID model is assumed when no autocorrelation time scales are given. Furthermore, BM is obtained as a limit of the OU process, when its position autocorrelation time scale goes to infinity, while IOU is obtained in the same way as a limit of the OUF process. It is important to realize that the range-resident models neither assume nor impose any movement barriers. Instead, these models assume that position correlations eventually 'die off' over some period of time τ_r, and throughout the paper, we refer to this phenomenon as 'restricted' or 'constrained' space use

Movement Model	Pos. AC	Vel. AC	H. Range	Parameterization
Ind. Ident. Distr. (IID)	No	No	Yes	$\tau = $ NULL
Brownian Motion (BM)	Yes	No	No	$\tau = \infty$
Ornstein–Uhlenbeck (OU)	Yes	No	Yes	$\tau = \tau_r$
Integrated OU (IOU)	Yes	Yes	No	$\tau = \{\infty, \tau_v\}$
Ornstein-Uhlenbeck F (OUF)	Yes	Yes	Yes	$\tau = \{\tau_r, \tau_v\}$

to show range residency. BM is not appropriate for use as a home range model as it is, by design, an endlessly diffusing process (Fleming *et al.* 2015a). The Ornstein–Uhlenbeck (OU) process combines BM with a tendency to remain in a particular area (Uhlenbeck & Ornstein 1930; Dunn & Gipson 1977). The OU process is therefore appropriate for data that lack evidence of directional persistence, but where restricted space use is apparent. The integrated OU (IOU) process diffuses endlessly like BM, but unlike BM, features autocorrelated velocities at short time scales (Johnson *et al.* 2008; Gurarie & Ovaskainen, 2011). IOU is therefore appropriate for finely sampled data sets that show directional persistence, but where sampling duration was too brief to identify range residency. Finally, the recently introduced Ornstein–Uhlenbeck Foraging (OUF) process is a hybrid of the OU and IOU processes that features both correlated velocities and restricted space use (Fleming *et al.* 2014a, b). The OUF process will be appropriate for many modern data sets where sampling is both fine enough to reveal velocity autocorrelation and long enough to show range residency.

WORKFLOW

ctmm is designed around a particular workflow. Firstly, scatter plots of the relocation data should be made to identify problems in the data and coarse aspects of the movement. Next, empirical variograms and periodograms should be plotted to reveal more features of the movement behaviour and judge the appropriateness of the different models for the data. Suitable models thus identified can then be fit to the data via maximum likelihood and compared via AIC. Selecting an AIC-best model then facilitates subsequent analyses that condition on an appropriate movement model. Specifically, ctmm 0.3.0 supports autocorrelated kernel density estimation of home ranges as well as estimation of occurrence

Table 2. A description of the main functions in the `ctmm` package

Function	Description
`as.telemetry`	Coerces dataframes that follow the Movebank.org format to `telemetry` objects. Also coerces `move` objects created by the `move` package to `telemetry` objects. Data passed to `as.telemetry` must contain at least the following columns: `individual.local.identifier`, `timestamp`, `location.long`, and `location.lat`.
`plot`	When passed a `telemetry` object, produces simple plots of location data and can optionally overlay a Gaussian `ctmm` or `akde` object. When passed a `variogram` object, plots the empirical variogram, optionally overlaying the SVF from a fitted `ctmm` model. When passed a `periodogram` object, plots the empirical periodogram.
`variogram`	Calculates the empirical variogram for a `telemetry` object.
`variogram.fit`	Allows visually fitting model SVFs to an empirical variogram via interactive parameter sliders. This approach produces initial parameter guesses for `ctmm`'s model fitting functions. These guesses can be saved to a variable by clicking the button on the slider palette, and that variable can then be passed to `ctmm.fit` or `ctmm.select`.
`periodogram`	Calculates the Lomb–Scargle periodogram for a `telemetry` object.
`ctmm`	Creates a `ctmm` object defining a CTSP model based on the user-specified parameter vector.
`ctmm.fit`	Performs maximum likelihood parameter and interval estimation of `ctmm` models using each model's exact likelihood function. Accepts a `ctmm` object with parameter guesses (e.g. as generated by `variogram.fit`).
`ctmm.select`	A wrapper for `ctmm.fit` that automatically fits the range-resident models (IID, OU and OUF) and ranks them via AICc. Accepts a `ctmm` object with parameter guesses (e.g. as generated by `variogram.fit`).
`akde`	Estimates the range distribution for autocorrelated telemetry data via autocorrelated kernel density estimation. Requires both a `telemetry` object and a fitted, selected model as input.
`occurrence`	Estimates the occurrence distribution for autocorrelated telemetry data via Kriging. Requires both a `telemetry` object and a fitted, selected model as input.
`summary`	When passed a `telemetry` object, gives an overview of the data including sampling interval and sampling duration. When passed a `ctmm.fit` object, gives point estimates and 95% confidence intervals for model parameters. When passed an `akde` object, displays the estimated 95% home range area, with 95% confidence intervals on the area estimate.

distributions and trajectories via time-series Kriging. The core functions of `ctmm` are summarized in Table 2.

DATA IMPORT AND BASIC PLOTTING

`ctmm` defines a `telemetry` object for representing movement data and a geographical projection associated with the data. Data imported into `ctmm` must conform to a minimal version of the Movebank data format, which is a tabular format with observations on the rows and at least the columns indicated in Table 2. Data files downloaded directly from Movebank will have the correct format, but users can also manually create appropriately formatted data files. Additionally, `as.telemetry` can import `move` objects from the `move` R package (Kranstauber & Smolla 2015). `ctmm` users can therefore leverage the `move` package's well-developed functionality for interacting with Movebank programmatically. Once represented as a `telemetry` object, data can be easily visualized using `plot`, which we highly recommend to identify outliers, obvious migration events and other coarse features of the data.

VISUAL DIAGNOSTICS FOR CTSPS

Empirical variograms, which are plots of the semi-variance in positions as a function of the time lag separating observations (Cressie 1993; Diggle & Ribeiro 2007), provide an unbiased means of visualizing the autocorrelation structure of a relocation data set (Fleming *et al.* 2014a). A variogram can be calculated in `ctmm` via the `variogram` function and plotted with confidence intervals via `plot`. The behaviour of the variogram near the origin, its shape for intermediate lags and its long-lag behaviour can

help diagnose problems of model fit (Fleming *et al.* 2014a). We describe each of these features in turn.

Firstly, a linear increase in semi-variance from zero for short lags indicates uncorrelated velocities and suggests models such as BM or OU. Upward curvature at short lags indicates velocity autocorrelation and suggests models such as the IOU or OUF processes. When location error is large enough to be a concern, the semi-variance will approach some constant greater than zero as lag goes to zero. This so-called nugget effect also provides a crude estimate of the error magnitude. Secondly, the variogram will typically increase linearly over some range of intermediate lags, indicating autocorrelated positions consistent with most diffusion models including BM, OU, IOU and OUF (but not IID). Finally, the long-lag behaviour of the variogram contains information about space use. A range-resident species should have a variogram that eventually reaches an asymptote, which is proportional to home range size. If the variogram does not approach an asymptote with increasing lag, the data would not be appropriate for a home range analysis of any kind. For a range-resident animal, this usually means that the individual was not tracked long enough to reveal its home range extent, but it can also be a consequence of the individual migrating or shifting its range.

The semi-variance function (SVF) is the theoretical analogue of the variogram (Fig. 1). `ctmm` implements SVFs for the supported models and allows the user to get rough initial parameter guesses by visually matching a theoretical SVF to an empirical variogram. For RSTUDIO users, the function `variogram.fit` provides interactive sliders that manipulate parameter values and display the corresponding SVF against the variogram (Fig. 2). The SVF of a good movement model

for the data should simultaneously match the short-, intermediate-, and long-lag features of the variogram.

LIKELIHOOD-BASED PARAMETER ESTIMATION AND MODEL SELECTION

After identifying reasonable candidate models, the next analysis step is fitting the models via maximum likelihood (Fleming *et al.* 2014b). ctmm includes the exact likelihood functions for all models in Table 1. The ctmm.fit function takes a telemetry object and a model specification as input and computes the maximum likelihood parameter and interval estimates. The ctmm function defines a CTSP model via its parameter vector as described in Table 1. The initial guesses obtained by using variogram.fit can be given as starting values in ctmm.fit either by clicking the save to GUESS button on the slider palette, or by using the ctmm function (see Appendix S3, Supporting information). Alternatively, if there is evidence of range residence in the variogram, one can automatically fit and rank (via AICc) both isotropic and anisotropic versions of the three models with home ranges with the ctmm.select function. The objects returned by ctmm.fit and ctmm.select contain a range of fit-related quantities including point estimates, confidence intervals and AICc. Finally, fitted models can be compared to the data via plot.

QUANTIFYING HOME RANGE AREA VIA AUTOCORRELATED KERNEL DENSITY ESTIMATION

Understanding the amount and type(s) of habitat animals require to survive is essential both for effective conservation planning and for basic ecological understanding. Home range estimation is, consequently, one of the key uses of relocation data. However, recent work has shown that conventional home range estimation methods will systematically, and sometimes grossly, underestimate home range areas for autocorrelated relocation data (Fleming

et al. 2014a,b, 2015a). Underestimation occurs because a given number of IID observations contain far more information about the home range than an equivalent number of highly autocorrelated observations (Fleming *et al.* 2015a). Home range estimation procedures that assume IID data will therefore overestimate the amount of information the data actually contain, which typically results in an underestimated home range area. In contrast, the recently introduced AKDE estimator (Fleming *et al.* 2015a) produces a more accurate estimate by appropriately discounting the information content of autocorrelated data.

The essential task in AKDE, which ctmm automates, is calculating the appropriate smoothing bandwidth given the data and a previously fitted and selected movement model. This means that the above-described steps to select a good model for the focal data set are required prior to an AKDE analysis. Furthermore, both the data and selected model must feature a home range (see Table 1). If there is no evidence of constrained space use in the data (i.e. no asymptote in the variogram), then a home range analysis of any kind is not appropriate. The akde function takes both a telemetry object and a ctmm object as input and returns an akde object. plot can then be used to show the akde estimate, the 95% home range contour and the 95% confidence intervals around this contour, while summary gives the 95% home range area point estimate and corresponding 95% confidence intervals. Different home range percentages or different confidence interval limits can be optionally specified.

CASE STUDY

We demonstrate the capabilities of ctmm on an example with African buffalo (*Syncerus caffer*) tracked in Kruger National Park, South Africa (Getz *et al.* 2007; Cross *et al.* 2009). The African buffalo data set consists of hourly GPS position locations for six individuals over a periods ranging from 2·5 to 8 months. Some individuals feature many missing observations.
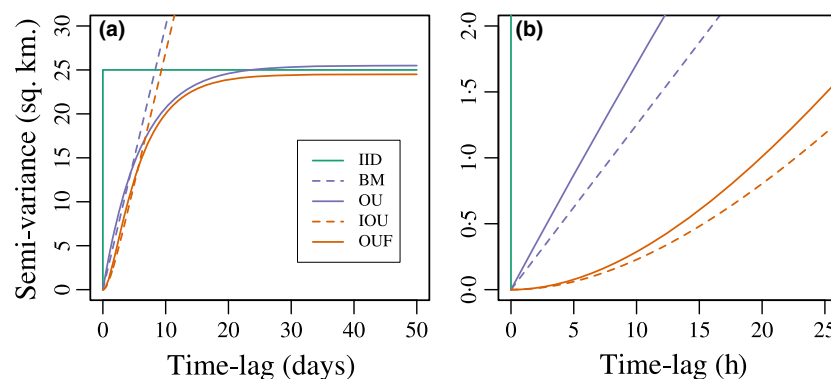

Fig. 1. The SVFs of all CTSP models currently implemented in ctmm. Parameter values are chosen such that the SVFs can all be displayed on the same axes. Panel A shows the intermediate- and large-lag behaviour of the SVFs while panel B shows the fine-scale features. Models that feature home ranges are shown with solid curves, while those lacking home ranges are displayed with dashed curves. A shared colour indicates model pairs where one model (dashed) is parameterized as a limit of another (solid; i.e. IOU and OUF). This colour and line-type scheme is carried throughout all subsequent figures that show theoretical SVFs.
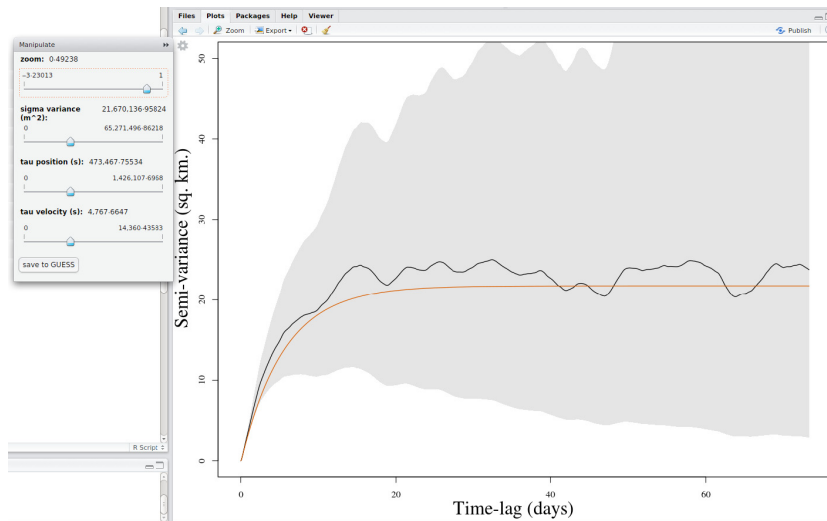
Fig. 2. A screenshot of `variogram.fit` and its parameter and zoom slider palette.

A COMPLETE ANALYSIS

We now walk through the main analysis steps with a single individual, Cilla, from the buffalo data. Cilla was tracked for 147 days between July and December 2005, yielding 3528 location fixes with no missing observations. We begin by loading the package and then extracting and plotting the data for Cilla. All analyses were performed using R version 3.2.3 (R Core Team, 2015), RStudio version 0.99.489 (RStudio Team, 2016) and ctmm version 0.3.0 (Fleming & Calabrese 2015).

```
#Load package
library(ctmm)

#Load example buffalo data
data("buffalo")

#Extract data for buffalo 1, Cilla
cilla<-buffalo[[1]]

#Plot the positions
plot(cilla)
```

Figure 3 suggests that Cilla is range resident with no obvious migrations. We then calculate and plot Cilla's variogram:

```
#Calculate variogram
vg.cilla<-variogram(cilla)

#Plot up to 50% of the maximum lag in the data
plot(vg.cilla)

#Zoom in on the shortest lags
plot(vg.cilla,fraction=0·005)
```

The asymptote of the variogram at longer lags (Fig. 4a) suggests that Cilla is indeed range resident, while the slight upward curvature at short lags (Fig. 4b) provides evidence of directional persistence in the data. We therefore consider the IID, OU and OUF models (Table 1) as reasonable candidates for the data. All three models feature a home

range, while the OU process also has autocorrelated positions and the OUF process has both position and velocity autocorrelation. We then use `variogram.fit` to get starting values for model parameters:

```
#Use the sliders provided by variogram.fit to specify
starting values.
#The default choices are usually acceptable.
variogram.fit(vg.cilla)
```

Once reasonable parameter values have been found by manipulating the sliders provided by `variogram.fit`, the save button allows the user to save the parameter values to a variable named GUESS. This variable can be passed to either `ctmm.fit` or `ctmm.select`, which will then refine the initial guesses via maximum likelihood estimation:

```
#Automatically fit the range-resident models via
maximum likelihood
```
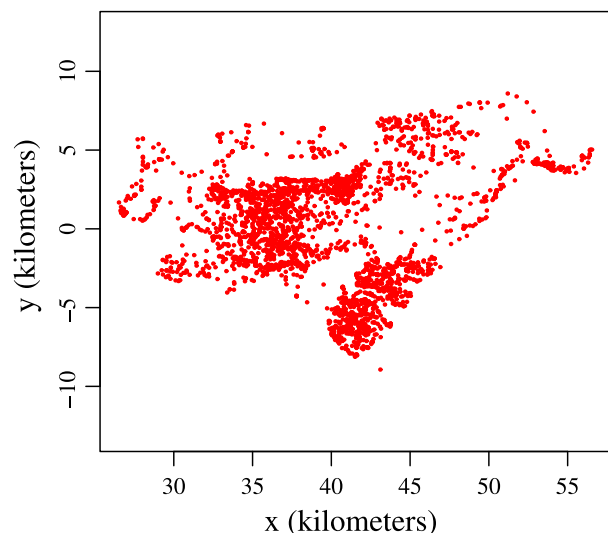

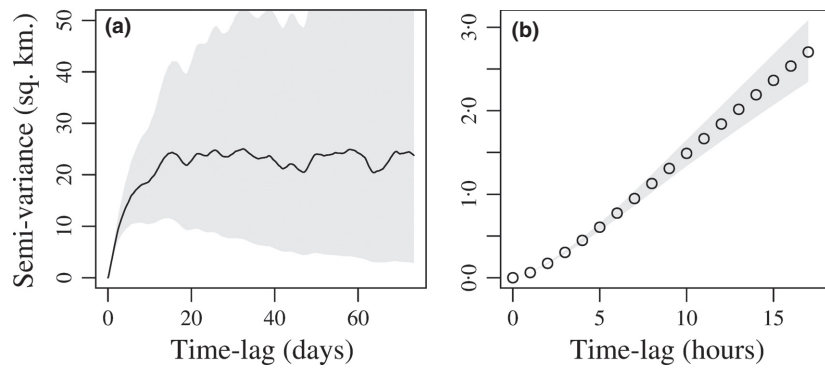**Fig. 3.** A scatterplot of the position observations for Cilla.

Fig. 4. (a) zoomed-out view of the variogram, which allows inspection of the large-scale features. The linear increase at intermediate time lags (~1–5 days) is evidence of autocorrelated positions, while the asymptote suggests restricted space use. (b) zoomed-in view focusing on the fine-scale features. The upward curvature at short lags is evidence of velocity autocorrelation in the data.

```
#using the initial parameter values obtained from
variogram.fit()
fitted.mods <- ctmm.select(cilla,CTMM=GUESS,
verbose=TRUE)
```

After fitting, we examine the model selection results:

```
summary(fitted.mods)
                                    dAICc
OUF anisotropic                     0·0000
OUF isotropic                       113·9392
OU anisotropic                      1458·2961
OU isotropic                        1643·9289
IID anisotropic                     36617·4880
IID isotropic                       37197·1365
```

Overall, we see that the anisotropic version of OUF is the clear winner. As the anisotropic version of each model was favoured over its isotropic counterpart, we visually examine the fits of the anisotropic versions of IID, OU and OUF.

```
#Extract the fitted anisotropic version of IID, OU,
and OUF.
iid <- fitted.mods[[5]]
ou <- fitted.mods[[3]]
ouf <- fitted.mods[[1]]
```

The SVFs of the fitted models can then be plotted against the empirical variogram to visually check the model selection results. Here, we only show the code for the first model, as the code for the others is similar:

```
plot(vg.cilla,CTMM=iid,col.CTMM="#1b9e77")
plot(vg.cilla,CTMM=iid,col.CTMM="#1b9e77",
fraction=0·005)
```

The results are shown in Fig. 5. It is clear from visual inspection that the OUF model is far superior to the others because it accounts for the three main features in the data: restricted space use, autocorrelated positions and autocorrelated velocities. The fit of the OU model is severely biased because it cannot account for the velocity autocorrelation apparent at short lags. While the upward curvature of the

variograms for short lags appears visually subtle, this feature is statistically very important and thus has a strong effect on how models fit the data (Fleming *et al.* 2014a). Next, we inspect the parameter and interval estimates for the three models.

```
summary(iid)
                         low          ML           high
area (square kilometers) 360·5731     369·1364     377·7989

summary(ou)
                         low          ML           high
area (square kilometers) 206·178715   397·23732    649·89803
tau position (days)      5·669666     12·98879     29·75639

summary(ouf)
                         low          ML           high
area (square kilometers) 280·312770   439·330756   633·50335
tau position (days)      3·512759     5·930834     10·01344
tau velocity (minutes)   43·749371    46·734973    49·92432
speed (kilometers/day)   10·922975    11·105522    11·29112
```

We conclude that the OUF model is the best for Cilla due to both the overwhelming AIC difference in its favour and its ability to accommodate all of the main features of the variogram.

We can now condition on the selected OUF model to obtain accurate home range estimates via autocorrelated kernel density estimation (Fleming *et al.* 2015a). We also calculate the conventional (uncorrelated) KDE on the same data for reference. The `akde` function calculates the appropriate smoothing bandwidth given the data and a model describing the autocorrelations in the data. The AKDE reduces to the conventional reference function KDE when an uncorrelated model is given as input. Thus, the conventional KDE is obtained by passing the data and fitted IID model to the `akde` function.

```
#Conventional KDE estimate
kde.cilla <- akde(cilla,CTMM=iid)
```

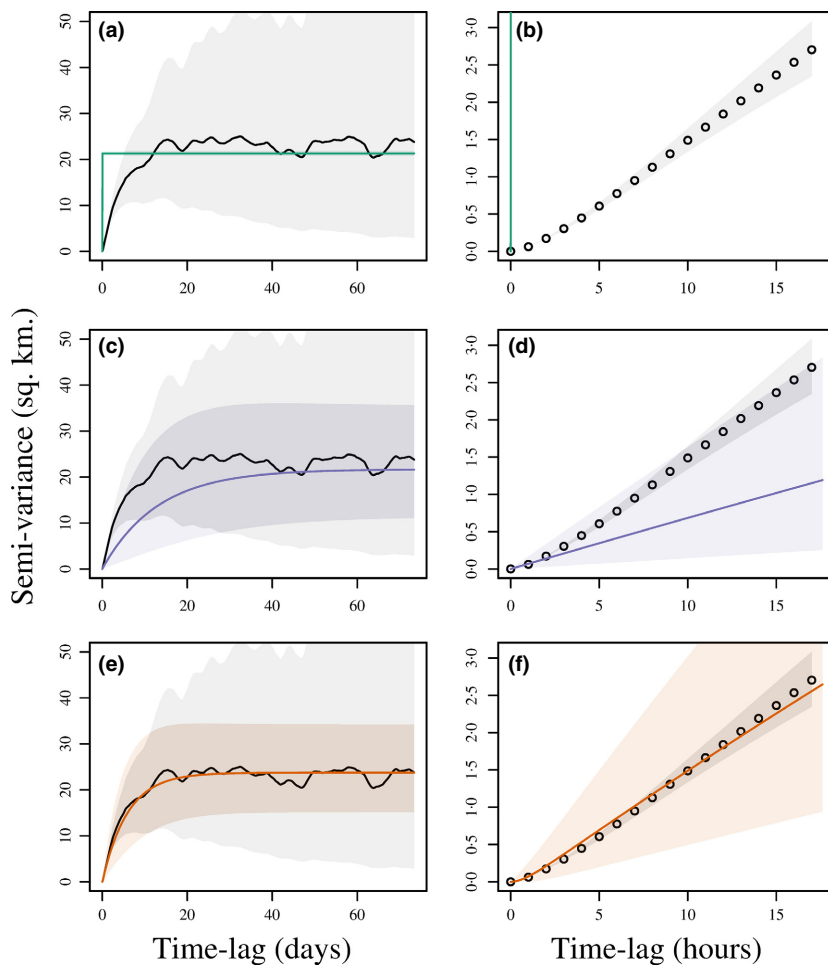The autocorrelated home estimate is then obtained by passing `akde` the selected OUF model

Fig. 5. Visual assessment of the fitted models against the empirical variogram. The left column presents a large-scale view of the fit, while the right column shows zoomed-in views. Each row represents a model, with row one corresponding to IID, row two to OU and row three to OUF. The colours representing the different models are as in Fig. 1.

```
#Autocorrelated KDE estimate
akde.cilla <- akde(cilla,CTMM=ouf)
```

Comparing the uncorrelated and autocorrelated estimates

```
summary(kde.cilla)
                            low        ML         high
area (square kilometers)    284·0115   285·3876   286·7635

summary(akde.cilla)
                            low        ML         high
area (square kilometers)    380·1221   493·3808   628·2644
```

we see that the AKDE predicts substantial additional space use relative to the conventional KDE. This is consistent with recent results proving that, in general, home range estimates that ignore autocorrelation in the data will be too small (Fleming *et al.* 2014a,b, 2015a). Finally, `plot` can be used to visualize the home range density estimates against the position data:

```
plot(cilla,UD=kde.cilla)
plot(cilla,UD=akde.cilla)
```

In addition to the appropriately larger home range estimate, the AKDE also provides a more honest accounting of uncertainty, as evidenced by the width of the confidence intervals around the 95% contour (cf. Fig. 6a, b). The corresponding intervals for the KDE are much narrower, which gives the false impression that the conventional estimate is more precise. For a given sample size, however, uncorrelated data provide more information than autocorrelated data. The conventional KDE, which assumes independent data, will thus yield confidence intervals that are far too narrow when input data are autocorrelated. The AKDE, in contrast, gives realistic confidence intervals even for very strongly autocorrelated data. However, these results will only be as good as the model used to describe autocorrelation in the data, which underlines the importance of visual diagnostics and model selection.

The supplementary material provides further examples of `ctmm`'s capabilities. Appendix S2 (Supporting information) shows an example with both apparently periodic movement and many missing observations. Appendix S3 (Supporting information) highlights the BM and IOU models for cases where evidence of range residency is lacking. Appendix S4 (Supporting information) demonstrates the estimation of occurrence distributions via time-series Kriging. Finally, Appendix D in Fleming *et al.* (2016) gives examples of importing `move` objects and incorporating location errors in `ctmm`.
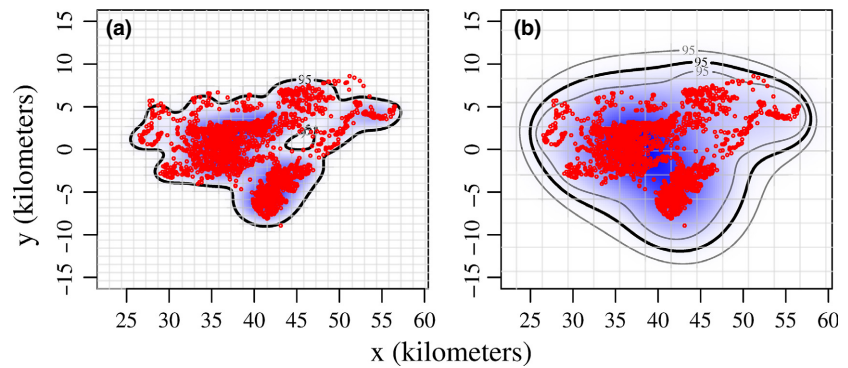
Fig. 6. (a) Conventional KDE home range estimate, which ignores autocorrelation. (b) The AKDE estimate, which conditions on the selected OUF model for Cilla. Note that the gridlines in each panel reflect the spatial resolution of the estimate. Spatial details in the density estimates that are smaller than the grid size should generally be ignored.

Discussion

Recently, movement ecology has reached a point where rapid advances in tracking technology have begun to expose the weaknesses of existing analytical frameworks. Continuous-time stochastic process models, which overcome many of the limitations suffered by their discrete-time counterparts, provide a robust foundation for the future of movement ecology. However, mainstream adoption of these next-generation tools has been hampered by the level of mathematical expertise required to use them by hand. The ctmm package for R helps to solve this problem by coupling user-friendly implementations of all major CTSPs movement models with a suite of powerful statistical tools for autocorrelated data adapted from geostatistics and signal processing.

ctmm is built around a standard workflow that progresses from exploratory visual diagnostics, to candidate model identification, to formal model fitting and selection, and finishes with estimating home ranges and occurrence distributions. The tools currently in ctmm allow a very thorough analysis of movement data. However, ctmm will continue to grow as new CTSP-based techniques and new models enter the literature. For example, our research group is currently developing methods to quantify correlated movements among multiple simultaneously tracked individuals (Calabrese *et al*. In Review), and to detect range shifts and migrations in individual relocation data (Gurarie *et al*. In Review). These and other new techniques will allow future versions of ctmm to extract even more biological detail from movement data and will open a broader range of relocation data sets to analysis with CTSP-based tools.

Acknowledgements

Data accessibility

The African buffalo data used in this paper can be accessed on Movebank.org as 'Study-Kruger African Buffalo, GPS Tracking, South Africa (Movebank ID 1764627).

References

Blackwell, P. (1997) Random diffusion models for animal movement. *Ecological Modelling*, **100**, 87–102.

Blackwell, P. (2003) Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika*, **90**, 613–627.

Bovet, P. & Benhamou, S. (1988) Spatial analysis of animals' movements using a correlated random walk model. *Journal of Theoretical Biology*, **131**, 419–433.

Boyce, M.S., Pitt, J., Northrup, J.M., Morehouse, A.T., Knopff, K.H., Cristescu, B. & Stenhouse, G.B. (2010) Temporal autocorrelation functions for movement rates from global positioning system radiotelemetry data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2213–2219.

Calabrese, J.M., Fleming, C.H., Fagan, W.F., Rimmler, M., Kaczensky, P., Bewick, S., Leimgruber, P. & Mueller, T. (In Review). A model-based approach to quantifying wildlife movement correlation. *Ecology*.

Codling, E.A. & Hill, N.A. (2005) Sampling rate effects on measurements of correlated and biased random walks. *Journal of Theoretical Biology*, **233**, 573–588.

Codling, E.A., Plank, M.J. & Benhamou, S. (2008) Random walk models in biology. *Journal of the Royal Society Interface*, **5**, 813–834.

Cressie, N. (1993) *Statistics for Spatial Data.* Wiley, New York.

Cross, P., Heisey, D., Bowers, J., Hay, C., Wolhuter, J., Buss, P. *et al*. (2009) Disease, predation and demography: assessing the impacts of bovine tuberculosis on African buffalo by monitoring at individual and population levels. *Journal of Applied Ecology*, **46**, 467–475.

Diggle, P.J. & Ribeiro, P.J. (2007) *Model-Based Geostatistics.* Springer, New York, NY, USA.

Dray, S., Royer-Carenzi, M. & Calenge, C. (2010) The exploratory analysis of autocorrelation in animal-movement studies. *Ecological Research*, **25**, 673–681.

Dunn, J.E. & Gipson, P.S. (1977) Analysis of radio telemetry data in studies of home range. *Biometrics*, **33**, 85–101.

Einstein, A. (1905) Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, **322**, 549–560.

Fleming, C.H. & Calabrese, J.M. (2015) ctmm: Continuous-Time Movement Modeling. R package version 0.3.0.

Fleming, C.H., Subaşı, Y. & Calabrese, J.M. (2015b) Maximum-entropy description of animal movement. *Physical Review E*, **91**, 032107.

Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2014a) From fine-scale foraging to home ranges: a semi-variance approach to identifying movement modes across spatiotemporal scales. *American Naturalist*, **183**, E154–E167.

Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2014b) Non-Markovian maximum likelihood estimation of autocorrelated movement processes. *Methods in Ecology and Evolution*, **5**, 462–472.

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

Fleming, C.H., Fagan, W.F., Mueller, T., Olson, K.A., Leimgruber, P. & Calabrese, J.M. (2016) Estimating where and how animals travel: an optimal framework for path reconstruction from autocorrelated tracking data. *Ecology*, **97**, 576—582.

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

Gurarie, E. & Ovaskainen, O. (2011) Characteristic spatial and temporal scales unify models of animal movement. *American Naturalist*, **178**, 113–123.

Gurarie, E., Peters, W., Fleming, C.H., Cagnacci, F., Calabrese, J.M., Mueller, T. & Fagan, W.F. (In Review). Whether, when, whither… and will it return? A statistical framework for modeling animal range shifts and migrations. *Journal of Animal Ecology*.

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

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.

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

Kareiva, P.M. & Shigesada, N. (1983) Analyzing insect movement as a correlated random walk. *Oecologia*, **56**, 234–238.

Kays, R., Crofoot, M.C., Jetz, W. & Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, **348**, aaa2478.

Kranstauber, B. & Smolla, M. (2015). Move: Visualizing and analyzing animal track data. R package version 1.5.514. http://cran.r-project.org/package=move.

McClintock, B.T., King, R., Thomas, L., 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.

Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E. & Fryxell, J.M. (2004) Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, **85**, 2436–2445.

Nams, V.O. (2013) Sampling animal movement paths causes turn autocorrelation. *Acta Biotheoretica*, **61**, 269–284.

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse, P.E. (2008) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences USA*, **105**, 19052–19059.

Péron, G., Fleming, C.H., dePaula, R.C. & Calabrese, J.M. (In Review). Periodograms for movement ecologists: a non-parametric method to uncover periodic patterns of space use in animal tracking data. *Movement Ecology*.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

RStudio Team (2016) *RStudio: Integrated Development Environment for R*. RStudio Inc, Boston, MA.

Turchin, P. (1998) *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. Sinauer Associates, Sunderland, MA, USA.

Uhlenbeck, G.E. & Ornstein, L.S. (1930) On the theory of the Brownian motion. *Physical Review*, **36**, 823–841.

Wikelski, M. & Kays, R. (2015). Movebank: archive, analysis and sharing of animal movement data. World Wide Web electronic publication, https://www.movebank.org/ (accessed 06 January 2016).

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Appendix S1. Overview of other software package for continuous-time movement models.

Appendix S2. Periodic behaviors.

Appendix S3. Data without restricted space use.

Appendix S4. Time-series Kriging with `ctmm`.