Identifying stationary phases in multivariate time-series for highlighting behavioural modes and home range settlements

Rémi Patin[1], Marie-Pierre Étienne[2], Émilie Lebarbier[3],

Simon Chamaillé-Jammes[1,4,5] & Simon Benhamou[1*]

[1] Centre d'Écologie Fonctionnelle et Évolutive, CNRS et Université de Montpellier, France

[2] Institut de recherche mathématique de Rennes, Université de Rennes, AgroCampusOuest, Rennes, France

[3] Mathématiques et Informatique Appliquées, Agroparistech, Paris, France

[4] LTSER France, Zone Atelier "Hwange", Hwange National Park, Bag 62, Dete, Zimbabwe

[5] Mammal Research Institute, Department of Zoology & Entomology, University of Pretoria, Pretoria, South Africa

[*] Corresponding author: simon.benhamou@cefe.cnrs.fr

Authors' contributions.

RP analysed the data and contributed to the coding of the statistical model, which was developed by MPE and EL. SC provided the tracking data. SB led the project, performed computer simulations, and wrote the first draft of the manuscript, except the part describing the model which was first contributed to by MPE, EL and RP. All authors contributed significantly to the final manuscript.

Key-words: movement ecology, segmentation, clustering, foraging, home range, area-concentrated searching, transit, migration

Abstract

1. Recent advances in bio-logging open promising perspectives in the study animal movements at numerous scales. It is now possible to record time-series of animal locations and ancillary data (e.g. activity level derived from on-board accelerometers) over extended areas and long durations with a high spatial and temporal resolution. Such time-series are often piecewise stationary, as the animal may alternate between different stationary phases (i.e. characterised by a specific mean and variance of some key parameter for limited periods). Identifying when these phases start and end is a critical first step to understand the dynamics of the underlying movement processes.

2. We introduce a new segmentation-clustering method we called segclust2d. It can segment bi- (or more generally multi-) variate time-series and possibly cluster the various segments obtained, corresponding to phases assumed to be stationary. It is easy to use, as it only requires specifying the minimum length of a segment (to prevent over-segmentation) based on biological considerations.

3. Although this method can be applied to time-series of any nature, we focus here on two-dimensional piecewise time-series whose phases correspond at small scale to the expressions of different behavioural modes such as transit, feeding and resting, as characterised by two joint metrics such as speed and turning angles or, at larger scale, to temporary home ranges, characterised by stationary distributions of bivariate coordinates.

4. Using computer simulations, we show that segcust2d can rival and even outperform previous, more complex methods, which were specifically developed to highlight changes in movement modes or home range shifts (based on Hidden Markov or Ornstein-Uhlenbeck modelling, respectively), which, contrary to our method, require truly informative initial guesses to be efficient. Furthermore we demonstrate it on actual examples involving a zebra's small scale movements and an elephant's large scale movements, to illustrate the identification of various movement modes and of home range shifts, respectively.

1 Introduction

Landscapes are spatially and temporally variable at various scales (Levin, 1992), and animals are expected to adjust their movements to the characteristics of their local environment, so as to maximize the time spent in profitable (or safe) habitats and minimize time in adverse ones (Pyke, 1978). Recent advances in bio-logging have made it possible to acquire time-series of animal's locations, and possibly ancillary data such as activity level derived from on-board accelerometers, over extended areas and long durations with high spatial and temporal resolutions. Such locational time-series, and the other ones that can be derived from them to describe the movement behaviour (e.g. turning angle, speed), are therefore expected to be piecewise stationary, i.e. to present a specific mean and variance for limited periods corresponding to stationary phases, alternating with rapid transition phases corresponding to changes of area or behaviour. Identifying these phases is a prerequisite to determine the biologically relevant scales of movement (Benhamou, 2014). It is therefore of paramount importance in two types of movement studies:

Identifying behavioural modes. Foragers are generally expected to alternate intensive (area-concentrated) searching mode, characterized by high tortuosity and low speed, and extensive searching (transit) mode, characterized by low tortuosity and high speed (see Dias et al. 2009 for contrasting examples). This alternation of searching modes therefore results in piecewise "behavioural stationarity" when considering time-series of tortuosity and speed. Although different segmentation approaches have been developed to identify behavioural modes by looking at breakpoints (i.e. rapid transitions between stationary phases Barraquand and Benhamou, 2008; Gurarie et al., 2009; Nams, 2014), a more sophisticated approach based on Hidden Markov Models (HMM) has gained momentum in recent years. In this approach, the joint step lengths and turning angles calculated from successive relocations are categorized among a predefined number of different modes modelled as hidden states (Morales et al., 2004; Beyer et al., 2013; Langrock et al., 2012; McClintock et al., 2012; Michelot et al., 2016). However, the convergence of HMM may require specifying informative initial state-dependent probability distribution parameters (i.e. informative priors when HMM are designed in a Bayesian context), which can be difficult. Here, we aim at developing an alternative approach which does not require such a pre-specification.

Identifying home range shifts. The recently emerging question of piecewise "locational stationarity" at the home range scale has been addressed in terms of movement scales (Benhamou, 2014), migration characteristics (Naidoo et al., 2012; Cagnacci et al., 2016) and of within-season shifts (Couriot et al., 2018). Indeed, for an animal that exploits various temporary home ranges, the time-series of relocations coordinates can be assumed to be stationary for a relatively long time (when the animal exploited the area where it established its temporary home range), then non-stationary for a relatively short time (when the animal left its home range until it established a new one), and so on. It is worth noting that a shift in home range does not necessarily involve a shift in mean location. It may also correspond to a change in variance if the animal enlarged or shrank its home range, e.g. due to a change in season (Naidoo et al., 2012; Monsarrat et al., 2013) or in reproductive status. Various methods have been proposed to detect home range shifts. The simple univariate approach based on the change of the beeline distance from a starting point (Bunnefeld et al., 2011) appears to be convenient in some cases but fully ignores movements leading the animal at a similar distance from the starting point but in another direction. A more complex approach rests on multi-state Ornstein-Uhlenbeck modelling (OUM Breed et al., 2017; Gurarie et al., 2017). However, as it requires that all home range phases and shifts are explicitly modelled, this approach tends to become cumbersome when there are several shifts to consider. Furthermore, it may require truly informative initial guesses to correctly detect small shifts. We therefore aim at developing an alternative approach that could be more efficient than an OUM-based approach to detect home range shifts and simpler to use. Additionally, as detecting shifts in behavioural modes and in home ranges settlements are conceptually similar, we focused on a generic approach that can be applied to both types of studies.

Here we introduce a new method, called segclust2d, able to segment a bi- (and more generally multi-) variate time-series, and to cluster similar segments (corresponding to stationary phases) in a common class (corresponding to a given state) if desired. We demonstrate that this method, which is easy to use, can successfully identify stationary phases corresponding to temporary home ranges when based on bivariate locational time-series, as well as movement modes when based on bivariate time-series of metrics such as speed and tortuosity. It thus offers an efficient and user-friendly alternative to previous, more complex, approaches. Furthermore, as this model applies to multivariate piecewise stationary time-series based on

any kind of metrics, it can integrate additional time-series of ancillary data (e.g. activity level derived from on-board accelerometers) for a better segmentation-clustering of movement data.

2 Methods

2.1 Statistical model and parameter estimation

General principle. Consider a multivariate piecewise stationary time-series assumed to be regular (no gaps), composed of an unknown number of stationary phases. The C components of this signal (each corresponding to a univariate time-series) are assumed to be statistically independent (conditionally to the stationary phases) and should be normalized if they are of different nature, so as to have the same weight in subsequent procedures. One needs a reliable statistical model to find and characterize these phases and possibly to cluster them when they are assumed to be the expressions of a limited number of unobserved states of the underlying process (e.g. behavioural modes). Likelihood-based segmentation methods provide a suitable statistical framework to detect changes of phases but raise two main issues from a statistical and algorithmical point of view: (i) determining the optimal number of segments and (ii) for a given number of segments, finding the optimal segmentation, i.e. determining the locations of the starting/ending points of the segments (called breakpoints). The latter reduces to a well-known discrete optimization problem solved using a dynamic programming algorithm introduced by Bellman (1954; for a recent example, see Rigaill 2015). With n sites where the signal can be cut and K segments, the dynamic programming algorithm reaches the exact maximum likelihood solution with a complexity in $O(n^2K)$, drastically smaller than the complexity in $O(n^K)$ involved by a force brute algorithm when exploring the whole segmentation space. We will first introduce the models and the estimation procedure to optimally segment a multivariate signal for a predefined number of segments K and possibly (if clustering is required) a predefined number of states M. Afterwards, we will show how the optimal number of segments and possibly of clusters can be found based on a penalized likelihood criterion. Our approach is based on Lavielle (2005)'s segmentation method of univariate signals and its extension by Picard et al. (2007) to segment and cluster DNA sequences without assuming any kind of distribution for the segment lengths, such as a geometric distribution as HMM implicitly

do (Karlin and Taylor, 1975).

Optimal segmentation in K segments, with optional clustering in M states. Assume that there are K stationary phases in a bivariate time-series with total length n. A stationary phase corresponds to a segment. It is defined by a sequence of consecutive random variables sharing the exact same distribution, in particular the same mean $\boldsymbol{\mu}$ and variance matrix $\boldsymbol{\sigma}^2$. As soon as one of these parameters changes, a new segment starts. The C components within a given segment $k \in [1, ..., K]$ starting at time $t = t_{k-1} + 1$ and ending at time $t = t_k$ (with $t_0 = 0$ and $t_K = n$) are modelled as sequences of Gaussian independent random variables bY_t, for $t = 1, ..., n$. In the absence of clustering (segmentation-only model), bY_t (with C components labelled 1 to C) at time t is modelled simply as follows:

$$\boldsymbol{Y}_t \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\sigma}_k^2) \text{ with } \boldsymbol{\mu}_k = \begin{pmatrix} \mu_{k,1} \\ \mu_{k,...} \\ \mu_{k,C} \end{pmatrix} \text{ and } \boldsymbol{\sigma}_k^2 = \begin{pmatrix} \sigma_{k,1}^2 & 0 & 0 \\ 0 & \sigma_{k,...}^2 & 0 \\ 0 & 0 & \sigma_{k,C}^2 \end{pmatrix} \quad (1)$$

where $\boldsymbol{\mu}_k$ and $\boldsymbol{\sigma}_k^2$ are the mean and variance matrix for segment k. As the model parameters to be estimated vary independently between segments, dynamic programming can be used to segment the multivariate signal at best in K segments. Its application is straightforward in this case, as it relies on the log-likelihood of each segment, which is simply equal to the sum of the log-likelihoods of the C components.

In the segmentation-clustering model, a state m, among M possible states, is assigned to every segment, and random variables within a segment classified in state m are all assumed to share the same mean $\boldsymbol{\mu}_m$ and the same variance $\boldsymbol{\sigma}_m^2$, different from those involved in other states. More formally, let (S_k), with $k = 1, ..., K$, denote the state of the segment k: S_k is a latent random variable taking values in $[1, ..., M]$. It is modelled through a multinomial distribution of parameters $\pi = (\pi_m)$ with $m = 1, \ldots, M$, whereπ_m corresponds to the probability for a segment to belong to state m. Given $S_k = m$, the value of the bivariate

signal at time t, $\boldsymbol{Y}_t$, is therefore modelled as:

$$\boldsymbol{Y}_t|S_k = m \sim \mathcal{N}(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) \text{ with } \boldsymbol{\mu}_m = \begin{pmatrix} \mu_{m,1} \\ \mu_{m,...} \\ \mu_{m,C} \end{pmatrix} \text{ and } \boldsymbol{\sigma}^2_m = \begin{pmatrix} \sigma^2_{m,1} & 0 & 0 \\ 0 & \sigma^2_{m,...} & 0 \\ 0 & 0 & \sigma^2_{m,C} \end{pmatrix} \quad (2)$$

As parameters (π_m, $\boldsymbol{\mu}_m$, $\boldsymbol{\sigma}^2_m$) that characterise any state m are unknown and are to be estimated, resulting in a mixture distribution where segments are linked in terms of parameters, the optimal segmentation cannot anymore be obtained using dynamic programming alone. Following Picard et al. (2007), we designed the following two-step procedure, which is iterated up to convergence.

1. Given a set of parameters (π_m, $\boldsymbol{\mu}_m$, $\boldsymbol{\sigma}^2_m$) with $m = 1, ..., M$, the best segmentation in K segments is obtained using dynamic programming.

2. Given this segmentation, the values of parameters are estimated using an expectation-maximization algorithm which is commonly used in latent variable modelling (Dempster et al., 1977).

By mixing dynamic programming and expectation-maximization through this iterative procedure, segmentation and clustering processes work jointly (rather than the latter after the former) leading to the optimal segmentation given K and M. However, an additional smoothing procedure is required to solve possible convergence issues (see details in Supporting information 1).

Finding the optimal numbers of segments and states. For both methods (segmentation-only and segmentation-clustering), a minimum segment length L_{min} has to be set not only to speed up the algorithm, but also, more fundamentally, to prevent over-segmenting, based on biological considerations. For example, setting L_{min} to a value of a few weeks when analysing locational time-series will prevent the algorithm from considering an area exploited only for a few days, corresponding to foray outside the usual home range or to stop-over during migration, as a distinct home range. Similarly, setting L_{min} to a value of several minutes when looking for changes in behavioural modes will prevent force the algorithm to assign a given movement bout to a given behavioural mode only if this bout is long enough, even though it may be punctuated by ephemeral events related to another behaviour. To obtain the optimal solution, we calculated the likelihood

of all number of segments $K < K_{max} = \lfloor \frac{n}{L_{min}} \rfloor$, and for any number of states M ($< K$) one wishes to consider if clustering has to be involved. In this case, the optimal values of K and M are determined as those that maximize a Bayesian Information Criterion (BIC; Schwarz et al. 1978)-based penalised likelihood (Supporting information 1). However, as it will be shown in the Results section, it is usually preferable to consider a single value of M based on biologically relevant grounds than to let the model determine an optimal number of states based on a statistical basis (see also Pohle et al. 2017). When only segmentation is involved (no clustering), the optimal number of segments K is based in agreement with Lavielle (2005) on maximizing a K-penalized likelihood curve (Supporting Information 1).

2.2 Computer simulations

We run simulations to assess the ability of our approach to detect home range shifts and changes in behavioural modes from bivariate time-series, and to compare it with that of other (OUM-based and HMM-based) methods. For each set of parameters of each type of simulation, we simulated 100 replicates. Distances are expressed in arbitrary unit length u.

Home range shifts. For simplicity, the animal was assumed to behave as a central place forager. We simulated its fine-scale movement as a central-place biased correlated random walk, which results in a probability of presence decreasing exponentially with the distance D to the central place (Benhamou, 1989): at each time step, the animal turns by an angle α_i drawn from a wrapped Gaussian distribution with a null mean and standard deviation $\sigma_i = \sigma_0 \left[1 + \frac{D_i - D_{i-1}}{2}\right]$, and progresses by 1 unit length ($1u$) in the new direction. In a batch of simulations, σ_0 was set to 0.5 radians, and the central place was first set at a given location for the first 10,000 time steps (phase 1), then shifted to another location by $60u$ in both X and Y for 10,000 additional time steps (phase 2), resulting in disjoint home ranges, and then shifted to a third location by $20u$ in both X and Y for 10,000 additional time steps (phase 3), resulting in overlapping home ranges. In another batch of simulations, the central place remained at the same location for the 30,000 time steps, but σ_0 was set to 0.7 radians for time steps 10,001 to 20,000 (phase 2) and to 0.5 radians otherwise (phases 1 and 3), involving a transitory enlargement of the home range. We finally sub-sampled the data sets by keeping one location every 60. The home range phases were then defined by 166 locations each, with low

serial correlation, and were thus similar to actual datasets that are commonly used in home range studies. Note that in our approach that focuses on the contrast between the stationary phases, the actual lengths of these phases do not matter, provided they are longer than L_{min}.

Changes in behavioural modes. We simulated a random search movement as a correlated random walk where three types of activity – immobility (resting or standing), intensive (area-concentrated) searching and extensive searching (transit) – alternate, each one lasting 20 time steps, this 60-step sub-series being repeated 5 times. The step lengths L_i were drawn from a log-normal distribution with a mean equal to $0.5u$ in the intensive mode or $1.0u$ in the extensive mode, and with a standard deviation equal to 1/10th of the mean in both modes. Turning angles α_i were drawn from a wrapped Gaussian distribution with a null mean and a standard deviation equal to 0.4 radians in the intensive mode or 0.3 radians in the extensive mode. To mimic possible factors (e.g. GPS recording noise) that can blur the contrast between the modes, the locations obtained in this way, as well those obtained for immobility phases, were submitted to bivariate Gaussian random noise with a null mean and various standard deviations ζ. Note that in order to assess the ability of a method to segment a behavioural time-series, the precise movement rules used in simulations are not important. What really matters is the contrast between the different phases: with a high contrast, all methods should work well, whereas with a low contrast, all methods should fail, whatever the movement rules considered. In the results section, we will present the results obtained with a standard deviation of the noise ζ set to 0.3, involving a moderate contrast between the three modes. The results obtained with a lower noise ($\zeta = 0.2$; high contrast) or higher noise ($\zeta = 0.4$; low contrast) are provided in Supporting Information.

2.3 Metrics

For identifying home range shifts, the two signal components considered are orthonormal Cartesian coordinates (x_i, y_i)) of locations (GPS locations expressed in decimal degrees as longitude and latitude therefore require to be transformed in terms of easting and northing through a classical projection such as UTM). For identifying behavioural modes, the two components usually considered in HMM-based approaches are the classical metrics corresponding to the step lengths L_i and the turning angles α_i, computed from locations

recorded at constant time intervals Δt, and therefore acting as proxies for linear $\left(\frac{L_i}{\Delta t}\right)$ and angular $\left(\frac{\alpha_i}{\Delta t}\right)$ speeds, respectively. We used such metrics for comparative purpose, but we also tested some variants, assumed to improve the contrast between the different modes. We computed the linear speed as $\frac{L_i + L_{i+1}}{2\Delta t}$. Although this basic smoothing introduces some serial correlation which is not taken into account in our model, it should result in a less noisy signal. Furthermore, angular speed may show faded changes between searching modes because the intensive mode usually involves both a decrease in linear speed and an increase in path tortuosity but angular speed mechanically increases with both of them (Benhamou and Bovet, 1989; Barraquand and Benhamou, 2008).

We therefore computed turning angles α_i^* based on a constant step length r rather than at constant time interval. For this purpose, each location $\boldsymbol{X}_i = (X_i, Y_i)$ is considered the centre of a virtual circle with radius r, and the entrance and exit locations P_{en} and P_{ex} are determined through linear interpolation (Appendix 2). The turning angle α_i^* is then computed in $[-\pi, \pi]$ as the angular deviation between vectors $P_{en} \rightarrow X_i$ and $X_i \rightarrow P_{ex}$ (both with length r) rather than vectors $X_{i-1} \rightarrow X_i$ (with length L_i) and $X_i \rightarrow X_{i+1}$ (with length L_{i+1}) as done to compute α_i. When both L_i and L_{i+1} are larger than r, one gets $\alpha_i^* = \alpha_i$, whereas α_i^* tends, on average, to be larger (random search paths) or smaller (directed paths) than α_i when r is larger. We set r to the median of the step length distribution. In our simulations, we noticed that using a larger radius tends to improve the discrimination between the fast and slow movement modes but to worsen the discrimination between the slow movement mode and the immobility mode. We also tested the two orthogonal signals provided by the 'persistence speed' $\frac{L_{i+1}\cos(\alpha_i)}{\Delta t}$ and 'turning speed' $\frac{L_{i+1}\sin(\alpha_i)}{\Delta t}$ (Gurarie et al., 2009; Gloaguen et al., 2015).

2.4 Practical implementation of the method

Both segclust2d procedures (segmentation-only and segmentation-clustering) have been currently implemented to work on bivariate times-series (the only ones we considered in this paper) in a R package (https://cran.r-project.org/package=segclust2d), which optionally allows K_{max} to be set to a value lower than $\lfloor \frac{n}{L_{min}} \rfloor$ for preventing the algorithm from wasting time in looking for unlikely solutions. An integrated module makes it possible to derive the various movement variables mentioned in this paper from locations

data. Because our approach requires large amounts of computer memory, it cannot deal with too long time-series (¿ 10000 values) on small desktop computers. To circumvent this limit, some sub-sampling is automatically performed when necessary. It is worth noting however that, even in absence of any memory constraints, it is usually not a good practice to attempt to directly segment very long series, which encompass both very large scale phenomena thanks to their large extent and very small scales phenomena thanks to their high resolution. Indeed, small-scale data are usually not relevant for analysing large-scale patterns and therefore act more as noise than as information in this context. Thus, in home range simulations, subsampling by a factor 60 makes it possible to dramatically shorten the time-series by eliminating fine-scale movements (klinokinetic process), which are characterised by a high level of serial correlation in location and in direction. Such details are clearly not relevant for the question of home range shifts, where only the overall phase-dependent mean and/or the variance of locations matter (accordingly, our approach ignores serial correlations occurring in any stationary phase). Conversely, for fine-scale movement studies, the characteristics of the environment are liable to change (e.g. due to seasonal variations) when considering a time-series running over an extended duration, possibly leading to change in the characteristics of the behavioural classes expected. It appears therefore preferable to consider the various phases (e.g. seasons) separately rather than to attempt to deal with the long time-series as a whole.

Figure 1: **Examples of segmentation-only of simulated movements using segclust2d to highlight home range phases and shifts.** Top panels show the simulated paths (after 1/60 subsampling) corresponding to three home range phases (two shifts), either in mean location (a) or in variance (b). The corresponding time series for both location coordinates (x, y) are presented in panel (c) and (d), respectively. The horizontal colour bars running along the time axis show the true occurrences of the three phases, whereas the coloured bands appearing over the x and y signals show their occurrences as estimated using the segclust2d/segmentation-only method (with Lmin = 45 locations, i.e. 2700 unit times) and provide the estimated mean (plain horizontal line running in the middle of the band) ± standard deviation (band width) for each segment separately.

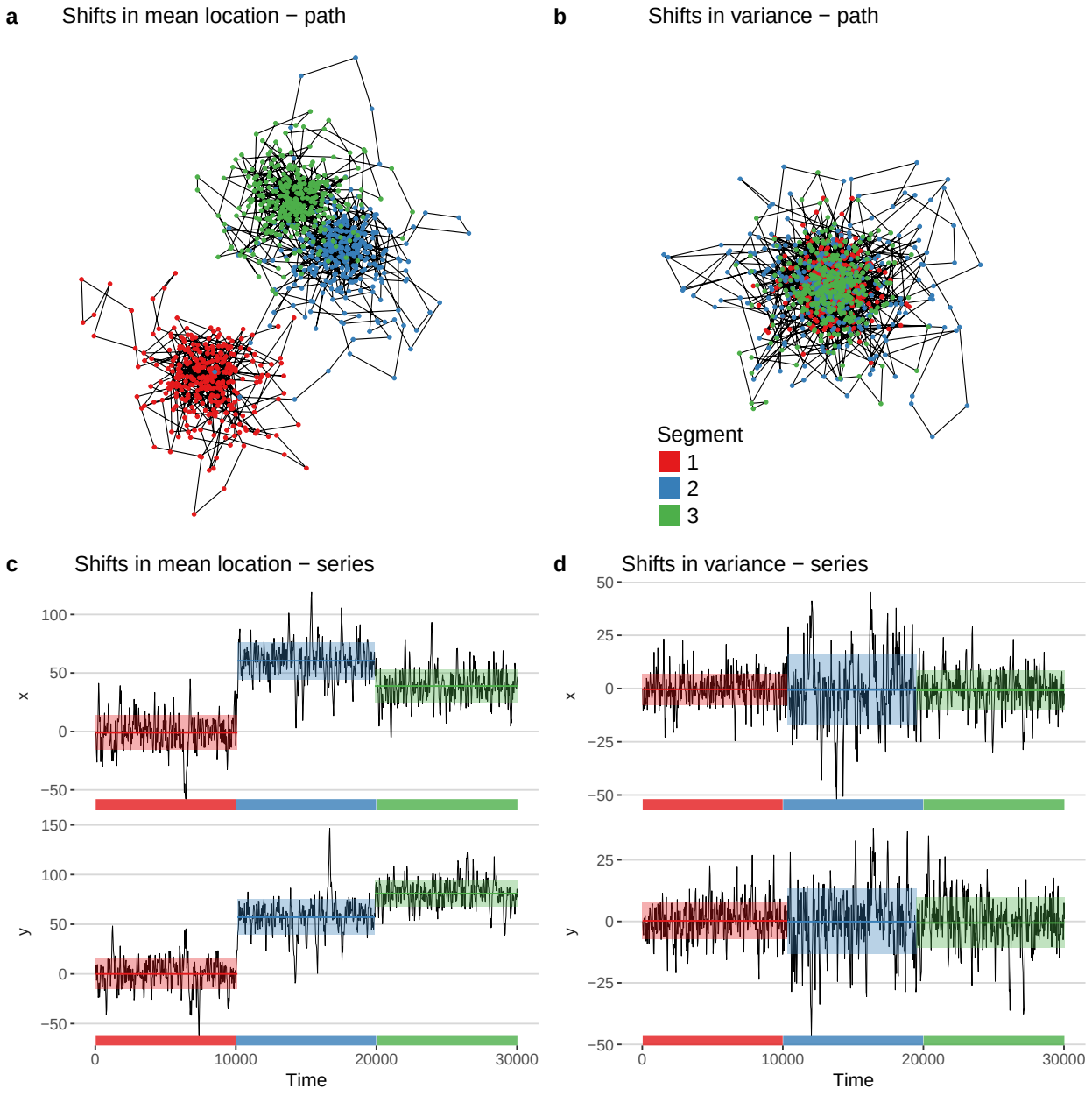
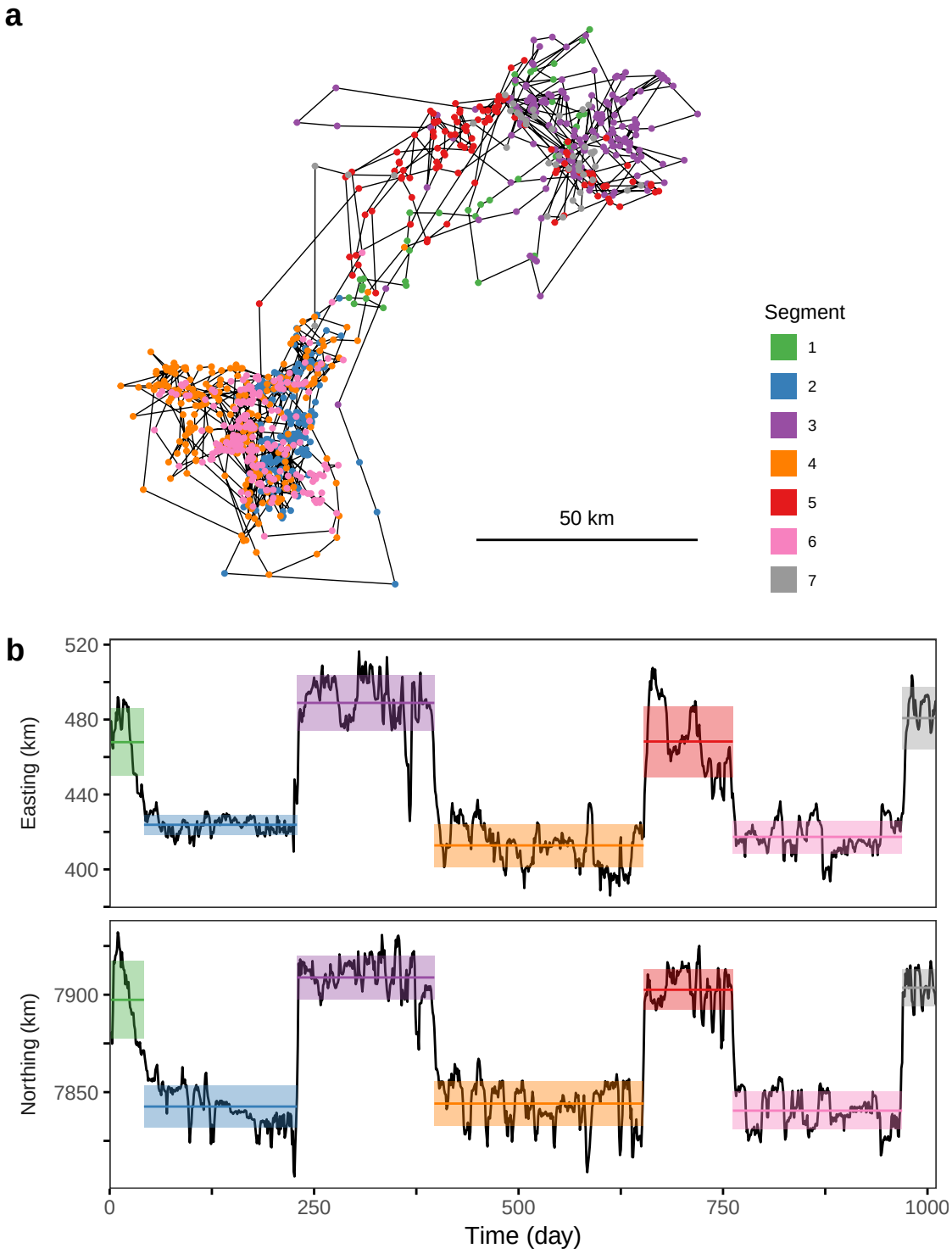


Figure 2: **Example of segmentation-only of an African elephant's movement recorded over 1000 days using segclust2d to highlight home range phases and shifts.** (a) Rough path representation obtained by linking the locations subsampled so as to keep a single GPS location per day; (b) Corresponding time series of locations coordinates (easting and northing). The coloured bands appearing over the time series show the estimated mean (plain horizontal line running in the middle of the band) ± standard deviation (band width) of each of the seven segments obtained using the segmentation-only method with Lmin = 30 days.



3 Results

3.1 Identifying home range shifts

Simulated movements. Figure 1 shows an example where the central place of the home range was shifted by 60 u between phases 1 and 2, and by $20u$ between phases 2 and 3, in both easting (X) and northing (Y), and an another example where the home range was enlarged during the phase 2 with respect to phases 1 and 3. The segclust2d/segmentation-only procedure (with Lmin set to 45 locations, corresponding to 2700 time steps because of the 1/60 subsampling) was able to correctly determine the true number of phases (3) in 98 out of the 100 replicates involving shifts in mean location (i.e. central place), and 88 out of the 100 replicates involving shifts in variance (i.e. change in home range size). In contrast, the OUM-based algorithm "marcher", which was specifically developed by Gurarie et al. (2017) to identify home range shifts in mean location, requires that the number of shifts has been specified, and is unable to detect shifts in variance. When the true number of phases (3) has been specified, our approach was able to correctly estimate the occurrences of the various shifts (mean$\pm$SD = 10152$\pm$79 and 20092$\pm$188 time units for the $60u$ and $20u$ shifts in mean location, respectively; mean$\pm$SD = 10035$\pm$1184 and 19942$\pm$1314 time units for the first and second shifts in variance of the same amplitude). When the actual home range centres and shift dates were provided as truly informative initial guesses to Gurarie et al. (2017)'s method, it was also able to correctly detect the two shifts in mean location (mean$\pm$SD = 10152$\pm$82 and 19967$\pm$239 time units for the $60u$ and $20u$ shifts in mean location, respectively). In contrast, when no information was provided, Gurarie et al. (2017)'s method, which then relies on a k-means procedure to get initial guesses, was still able to detect the large shift in location (mean$\pm$SD = 10152$\pm$937 time units) but was unable to detect the small one (the shift occurred at random between 15000 and 30000 time units, based only on 75 replicates, as the algorithm failed to provide any result for 25 replicates).

Illustrative example. We used the GPS track of an African Elephant (*Loxodonta africana*), recorded for > 2.5 years to illustrate the way the segclust2d/segmentation-only procedure can identify home ranging phases and shifts (Fig. 2). Considering that a time-series is (roughly) stationary when the partial means and variances obtained for its first and second halves or for its three thirds are not markedly different, the whole time-series of easting and northing coordinates can be said to be both stationary, corresponding to a large multiannual (possibly lifetime) home range, and piecewise stationary, corresponding to temporary (possibly seasonal) smaller home ranges. It can therefore be segmented to highlight these temporary home ranges and shifts between them. However, some of the phases so highlighted are clearly nonstationary. In particular, segments 1 and 5 correspond mainly to a slow south-westwards migration (rather than a temporary home range) between the two core areas of the multiannual home range. Segment 2 also corresponds to a nonstationary, migratory (southwards moving) phase, which went through an area used as a temporary home range during segments 4 and 6. This indicates that a same area can be used in different ways at different periods.

3.2 Identifying behavioural modes

Simulated movements. An example of path with three behavioural modes (extensive searching, intensive searching and resting) is shown in Fig. 3 with the corresponding time-series in terms of turning angle α_i^* and smoothed speed $\left(\frac{L_i + L_{i+1}}{2\Delta t}\right)$. In this example, the segclust2d/segmentation-clustering procedure appears able to detect the true number of modes ($M = 3$) and to attribute almost all locations to the right mode. We compared our method with a HMM-based method specifically designed to deal with movement data (Michelot et al., 2016; McClintock and Michelot, 2018) when the true number of modes has been specified. The results obtained from 100 replicates showed that our procedure rivals with the HMM-based method although the latter was helped by initial state-dependent probability distribution parameters which were tuned to their true values for each behavioural state (Fig. 4 with medium noise level $\zeta = 0.3$). With very low noise level ($\zeta =< 0.2$), an excellent fit was obtained with all methods and metrics considered, whereas with a very high nose level ($\zeta => 0.4$) the percentage of correct state assignment became closer to the value expected for a random assignment (33%; see Supporting information 3.1 results with $\zeta = 0.2$ and $\zeta = 0.4$).

It also appeared that, as expected, the angular $\left(\frac{\alpha_i}{\Delta t}\right)$ and linear $\left(\frac{L_i}{\Delta t}\right)$ speeds are not the most suitable the metrics for detecting behavioural changes. Thus, better results were obtained with both methods when using any other of the couples of metrics considered. The best fits were obtained with turning angle α_i^* or it absolute value $|\alpha_i^*|$ and smoothed speed. When the true number of modes is unknown, our method can also estimate this number as the most likely number of clusters, but the fraction of correct estimate is too low to consider the result as reliable (Supporting information 3.2).

Illustrative example. We used a 24-h GPS track of a plains zebra (*Equus quagga*) to illustrate the way the segclust2d/segmentation-clustering procedure can identify the occurrences of the various movement modes (Fig. 5). Although that, in this example, the most likely number of modes was estimated to be five, we present the segmentation obtained when setting this number to three, assuming that the biologically relevant modes should be resting (or any other non-moving behaviour such as standing), feeding and transiting (the other two modes detected by our procedure when using five modes were assumed to correspond to mixed behaviours).

Figure 3: **Comparative performances of segclust2d/segmentation-clustering vs. a HMM-based method for highlighting behavioural changes.** The boxplots show the proportion of correct state assignments, obtained for various bivariate signals when the true number of states is known (M = 3), as estimated from 100 replicates simulated with the same parameters as to the one illustrated in fig. 4 (noise $\zeta = 0.3$). The star (*) indicates turning angles computed with a constant step length, in terms of arithmetic (α_i^*) or absolute ($—\alpha_i^*—$) values. The white boxplots show the results obtained with HMM-based R package momentuHMM (McClintock and Michelot, 2018), with Gaussian priors set to the true means and variances of the various metrics in the different states. The grey boxplots shows the results obtained using segclust2d/segmentation-clustering with $L_min = 10$.

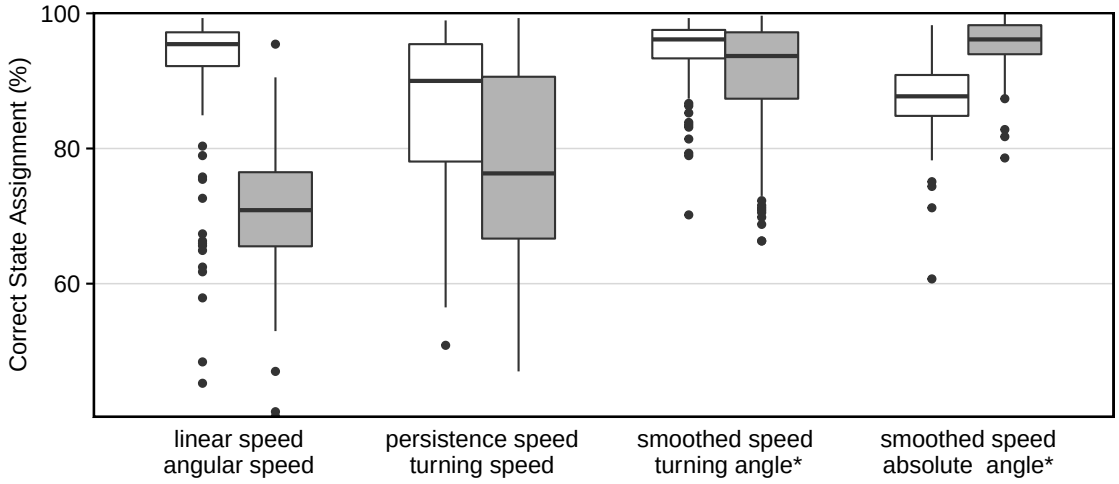


Figure 4: **Example of segmentation-clustering of a 24-h zebra's movement using segclust2d to highlight behavioural changes.** (a) Path representation obtained by linking GPS locations recorded every 8 minutes; (b) Determination using BIC-based penalised likelihood of the most likely numbers of states (M = 5) and segments (K = 20) (big orange dot), and most likely numbers of segments for the other number of states considered (large squares at the top of the curves), with Lmin = 5 (i.e. 40 min.). (c) Corresponding time-series in terms of absolute turning angle computed with a constant step length, $—\alpha_i^*—$, and smoothed speed, $\frac{L_i + L_{i+1}}{2\Delta t}$, segmented with M = 3 (leading to K = 15); the coloured bands appearing over the two time-series show the estimated occurrences and mean (plain horizontal line running in the middle of the band) +/- standard deviation (band width) for each of the three clusters considered.

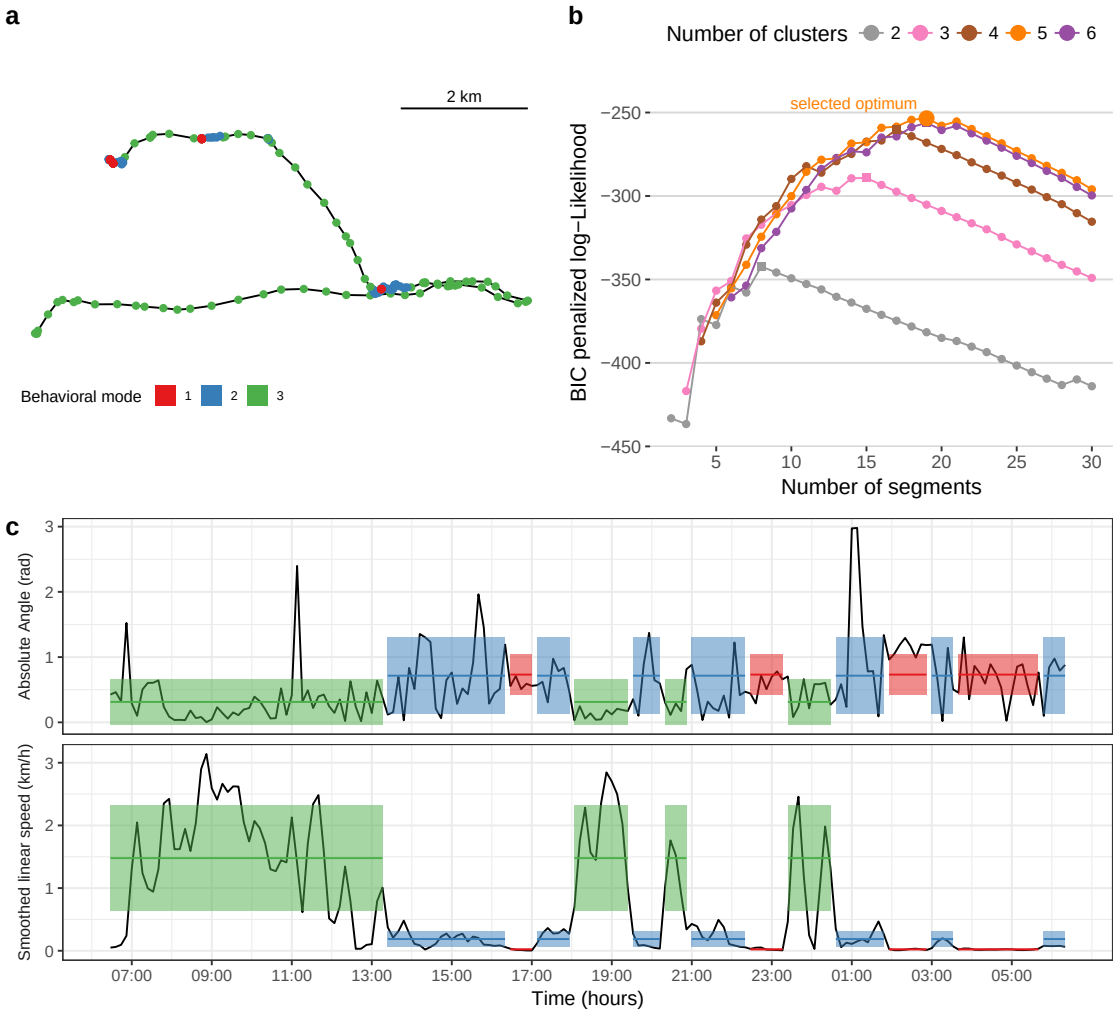
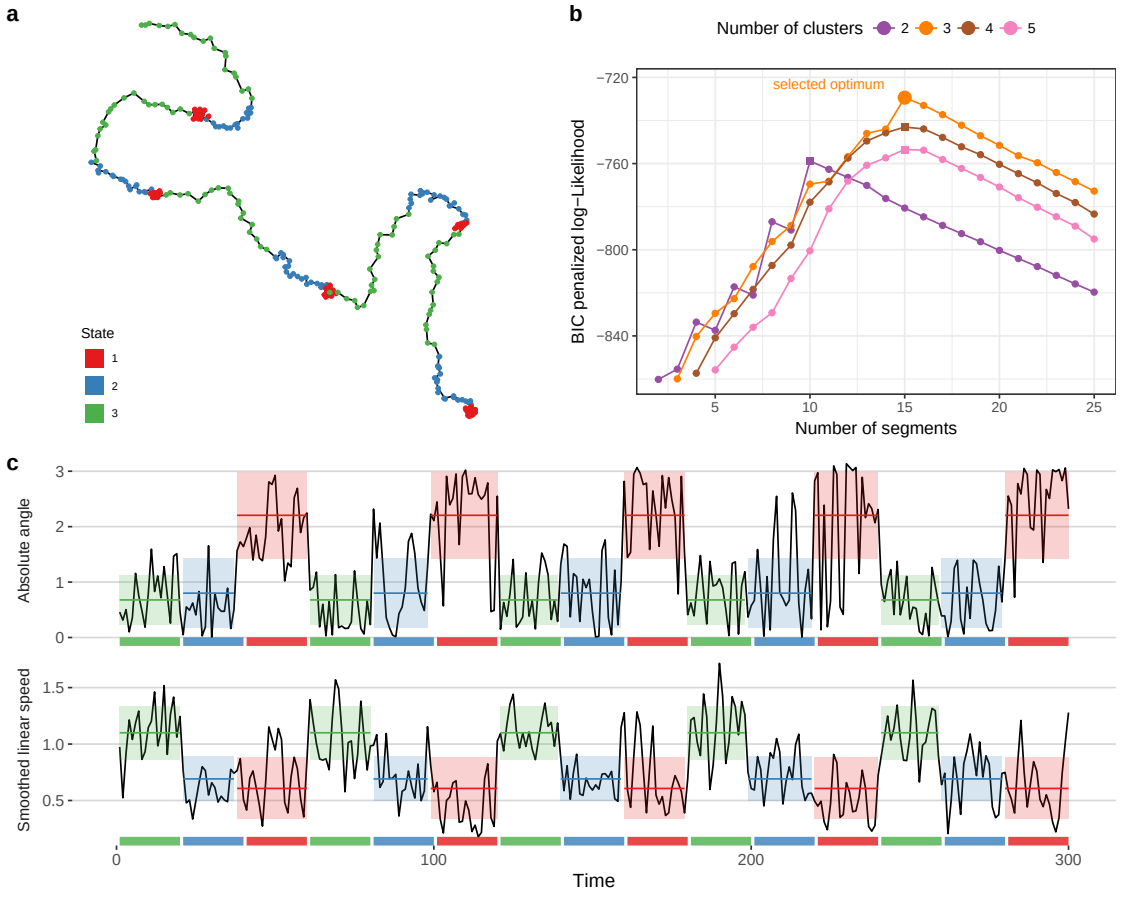


Figure 5: **Example of segmentation-clustering of a simulated movement using segclust2d to highlight behavioural changes.** (a) Simulated path as a composite correlated random walk, with additional noise $\zeta = 0.3$ u; (b) Determination using BIC-based penalised likelihood of the most likely numbers of states (M = 3) and segments (K = 15) (big orange dot), and of the most likely number of segments for the other three numbers of states considered (large squares at the top of the curves). (c) Corresponding time-series in terms of absolute turning angle computed with a constant step length, $—\alpha_i^*—$, and smoothed speed, $\frac{L_i + L_{i+1}}{2\Delta t}$, segmented with $L_{min} = 10$ and M = 3; the coloured bands appearing over the two time-series show the est +/- standard deviation (band width) for each of the three movement modes whereas the horizontal colour bars running along the time axis show the true occurrences of these modes.



4 Discussion

We showed that a generic method, segclust2d,that extends Lavielle (2005)'s and Picard et al. (2007)'s methods to multivariate time-series makes it possible to reliably detect two types of changes that are of key importance when studying free-ranging animal movements: home range shifts, based on bivariate time-series of location coordinates (segmentation-only procedure), and changes in behavioural modes, based on bivariate time-series of turning angles and speed (segmentation-clustering procedure). In any case, this new method is straightforward to parameterize: the user has just to set the minimum segment length (L_min) to a biologically relevant value so as to prevent the time-series from being over-segmented. Nevertheless, it proved to work at least as well as, and often better than, other recent methods specifically designed to deal with either home range shifts (Gurarie et al., 2017) or changes in behavioural modes (Michelot et al., 2016; McClintock and Michelot, 2018)

Breed et al. (2017) and Gurarie et al. (2017) independently developed an OUM-based method to identify home range shifts in mean location. Using computer simulations, we compared this approach, as implemented in Gurarie et al.'s "marcher" algorithm, with segclutst2d/segmentation-only. Both Gurarie et al. (2017)'s and our method are well able to detect large shifts in mean location. However, our method is also able to detect small shifts in mean location, whereas Gurarie et al. (2017)'s one requires a priori information on the actual mean locations and the shifts dates to correctly detect them, although this is precisely in this case that such information is usually lacking (i.e., they can hardly be guesstimated from visual inspection of the data). Furthermore, contrary to Gurarie et al. (2017)'s method, our method can work with any number of shifts, and is furthermore able to correctly estimate the number of shifts by itself in most cases, and it is also able to reveal changes in home range size. Yet, to be efficient, it does not require any more or less informative initial guess. It simply requires specifying a minimum length (L_min) for stationary phase to be called a temporary home range, shorter phases being assumed to correspond to transitory exploitations of restricted areas rather than to home ranges. However, contrarily to our method, which considers migrations as simple breakpoints, Gurarie et al. (2017)'s method can estimate the duration of migrations.

The elephant we considered in our illustrative example tended to move back and forth between two main areas. This kind of space use is common in migrating birds that commute between reproductive and

wintering home ranges. However, there are numerous studies showing more complex patterns, with an animal setting several distinct temporary home ranges successively (Naidoo et al., 2012; Benhamou, 2014; Cagnacci et al., 2016; Couriot et al., 2018). The segmentation of a long piecewise locational time-series in phases corresponding to temporary home ranges opens promising perspectives to understand how the occurrences and durations of home ranges are related to environmental co-variates, which is a prerequisite to infer long-term consequences for population distribution (Mueller and Fagan, 2008). The elephant illustrative example also shows that, although the model underlying segclust2d looks for stationary phases, there is no guaranty that all segments obtained are really stationary. This occurs because changes between stationary phases are modelled as breakpoints but may in fact correspond to slow progressive changes.

Since the pioneering paper by Morales et al. (2004), HMM-based methods have often been considered the best way to detect changes in behavioural modes of remotely tracked animals. An alternative approach was proposed by Barraquand and Benhamou (2008). It consisted in computing the series of residence time within a virtual circle running along the path and to search for the most likely breakpoints using Lavielle (2005)'s univariate segmentation method. However, although the residence time provides a simple and reliable univariate signal easy to segment and interpret, the values obtained depend not only on the type of behaviour that is performed but also on how long it is performed, preventing the segments corresponding to the same behaviour from being easily clustered. In the present study, we show that the segclust2/segmentation-clustering procedure rivals a HMM-based method when applied to the bivariate signal provided by the two classical metrics that are linear and angular speeds. Importantly, contrarily to what occurs with HMM-based methods, such performance is reached without the need to specify any informative initial state-dependent probability distribution parameters. With both types of method, better results were obtained when using other metrics such as smoothed speed and turning angle measured at constant step length, which were expected to improve the contrast between the intensive and extensive searching modes. Interestingly, using absolute rather than signed values of turning angles measured at constant step length works at best with our method whereas right and left turns were balanced in any mode in our simulated movements. Such metrics should be particularly useful to distinguish between intensive and extensive modes when the former involves turning systematically right or left, i.e. characterized by markedly either negative or positive mean turning

329 angles, whereas the latter involves balanced turning, as occurs in some species (e.g. Smith, 1974). As it
330 results also in a reliable identification when turns are balanced in both intensive and extensive searching, we
331 recommend using it systematically when using our method to distinguish between extensive and intensive
332 searching phases.

333 In the illustrative example on zebra's movements, five behavioural modes were detected by the segclust2d/segmentation-
334 clustering procedure in the time-series of smoothed speed and absolute value of the turning angles measured
335 at constant step length. Nevertheless, based on behavioural observations, we chose to segment the time-
336 series with only three modes assumed to correspond to immobility, feeding and transit. Indeed, although
337 our method can estimate the number of states on a statistical basis as the most likely number of clusters,
338 this estimate was not correct for a number of computer simulations whereas behavioural modes were clearly
339 defined. With actual data, there can be some mixing between modes, for instance transit and opportunis-
340 tic feeding at some times, so that the estimation of the number of relevant modes may become unreliable.
341 Thus, we recommend using the capacity of the segclust2d/ segmentation-clustering procedure to estimate
342 the number of states only when this number cannot be fixed a priori based on biological arguments. A similar
343 conclusion was reached by (Pohle et al., 2017) for HMM-based methods. It is also worth noting that feeding
344 and resting can be distinguished based on movement characteristics only in animals which have to move
345 significantly (with respect to the location recording noise) to feed. For animals which feed mainly without
346 markedly moving, such as some browser herbivores and carcass-eating carnivores, ancillary activity data
347 provided by on-board accelerometers are absolutely required to distinguish these two behavioural modes. As
348 our method can work conjointly on any number of time-series of any nature, future implementation could
349 integrate activity (accelerometer-based) time-series for a better identification of resting vs. active phases.

350 The segmentation of piecewise stationary time-series, possibly complemented by the clustering of the
351 resulting segments into functional classes, is often key to understanding the dynamics of processes. Based
352 on bivariate time-series of metrics such as location coordinates (northing and easting) or speed and turning
353 angles, segclust2d has the potential to facilitate discovery in the field of movement ecology. If necessary,
354 this approach can apply similarly to three and more dimensions, so as to consider ancillary variables such as
355 activity, as well as other metrics such as distance to a nest, proxies of habitat quality, or any other variable

that may be relevant when studying animal movements. Finally, as it can deal with two or more variables of any nature, our approach should be useful not only in movement ecology but also in many other fields.

Acknowledgement

We thank Clément Calenge for interesting discussions about segmentation issues. The study was partially funded by the grant ANR-16-CE02-0001-01 of the French 'Agence Nationale de la Recherche', and the Zone Atelier program of the CNRS. Computer simulations were programed in Pascal and ran thanks to the FreePascal compiler (www.freepascal.org).

Data accessibility

The code of the method is already publicly available as an R package (https://cran.r-project.org/package=segclust2d). There are only a few data used as examples, and they will be made publicly available as well.

Supporting Information 1: Complements about segclust2d

Expectation-Maximization algorithm and smoothing procedure

The Expectation-Maximization (EM) algorithm used to estimate the distribution parameters in the segmentation–clustering model is known to be sensitive to initialisation, so that it may converge to local maxima of the likelihood. This behaviour has some consequences on the parameters estimates but also makes the choice of the number of segments or states complicated. The classical initialisation solution consists in running the algorithm numerous times and just choose the point with the highest value of the log-likelihood, but this strategy is too computationally demanding. To minimize the risk of reaching local maxima within an unacceptable computation time, we propose the following initialisation strategy: (1) perform a pure segmentation of the signal and (2) use a hierarchical cluster algorithm, based on the log-likelihood ratio distance to assign segments to states.

Even with smart initialisation points, however, the EM algorithm may still converge to local maxima. This situation appears when looking, for a given number of states M, at the log-likelihood as a function of the number of segments K: whereas it is expected to be somewhat regular, this function can be quite noisy. To solve this issue, we propose to 'smooth' the function based on the parameter estimates of the distribution (π_m, μ_m, Σ_m) obtained for all the 'reliable' solutions, which correspond to the points that lie on the convex hull of log-likelihood curve, as smart initial points for any 'non-reliable' solution, i.e. for which an initialisation problem can be suspected. The improvement in terms of regularity of the log-likelihood curve obtained thanks to this procedure is illustrated in figure 6.

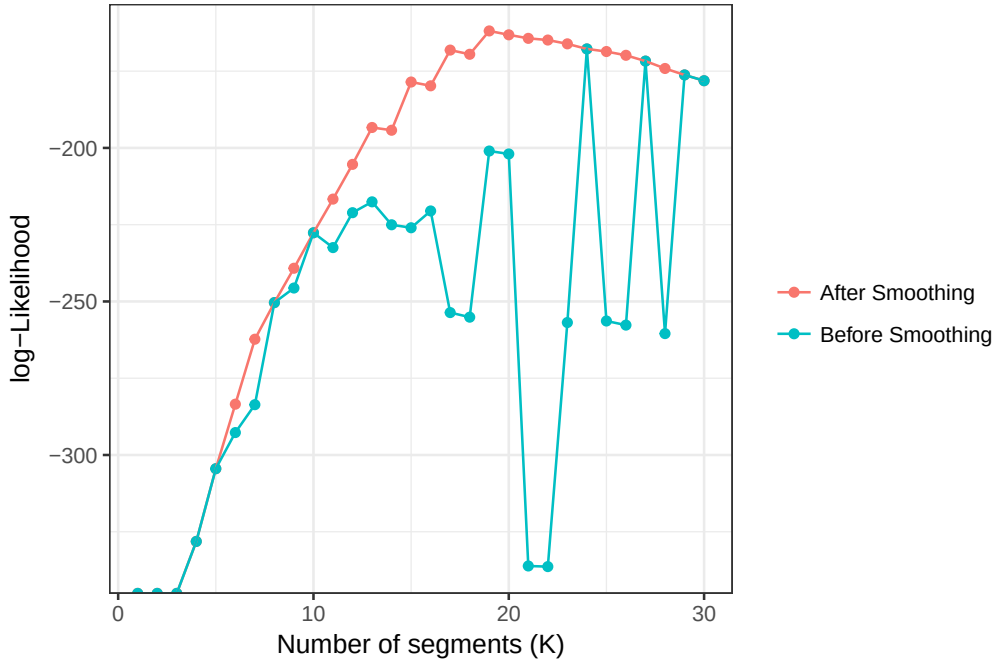


Figure 6: Maximum Likelihood estimates as a function of the number of segments before (in blue) and after (in red) smoothing. For instance the 'reliable' solution obtained with $K = 24$ segments was used to provide starting points for the EM algorithm for $K = 23$ and for $K = 25$, and this smoothing procedure gradually spreads over adjacent points.

Model selection

Choice of the number of segments K in the pure segmentation model

We used the adaptive model selection strategy proposed in Lavielle (2005) consisting in choosing the value of K that maximizes the following penalized log-likelihood : $\mathcal{L}_K - C\,K$ where $\mathcal{L}_K$ is the log-likelihood of the optimal segmentation in K segments and C is a unknown positive constant to be calibrated. The heuristic proposed by Lavielle (2005) consists in detecting the value of K for which the log-likelihood ceases to increase significantly. More specifically, consider the normalised log-likelihood defined as

$$\tilde{\mathcal{L}}_K = (K_{max} - 1)\frac{\mathcal{L}_{K_{max}} - \mathcal{L}_K}{\mathcal{L}_{K_{max}} - \mathcal{L}_1} + 1$$

Then, K is chosen as the value such that $\tilde{\mathcal{L}}_K$ displays the largest slope change. Namely, we take :

$$\hat{K} = argmin_K \left\{ (\tilde{\mathcal{L}}_K \tilde{\mathcal{L}}_{K+1}) - (\tilde{\mathcal{L}}_{K+1} \tilde{\mathcal{L}}_{K+2}) > S \right\}$$

where the value of threshold S is set to a predefined value (we used $S = 0.7$ as proposed in Lavielle, 2005).

As the selection relies on a predefined threshold, it is worth checking where the point corresponding to the selected number of segment lies on a plot of the log-likelihood curve (fig. S3.2). The optimal K value obtained in this way should correspond to a noticeable slope change.

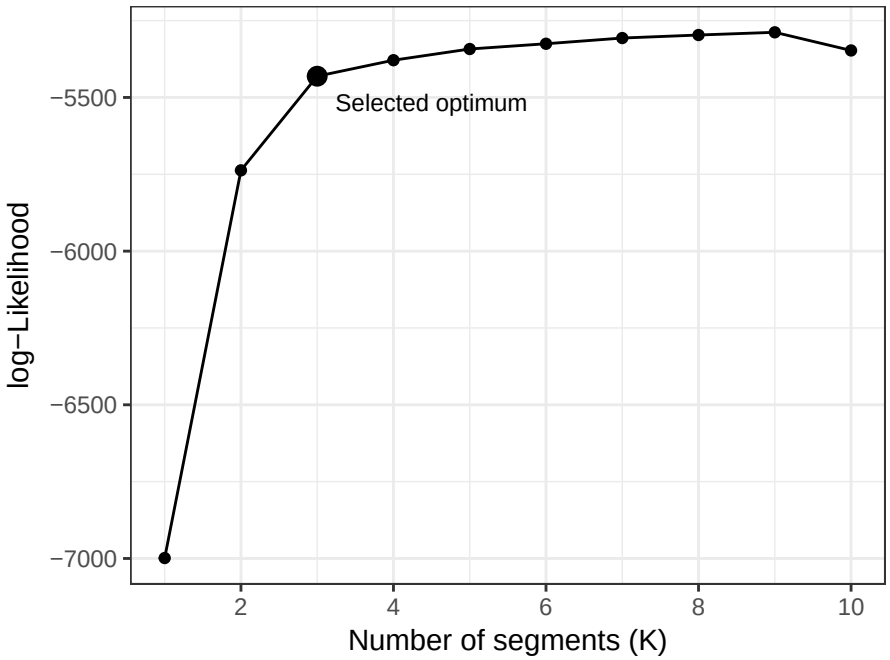


Figure 7: log-likelihood of a segmentation as a function of the number of segment. The optimum selected by the criterion from Lavielle (2005) should be located at a break in the increase of the curve.

Choice of the number of segments K and states M in the segmentation–clustering model

Selection of the best segmentation-clustering model (i.e. of the best couple of K and M values) is a hard task as no method has been yet proposed for this purpose. A log-likelihood is expected to increase with the number of parameters. However, as explained in Picard et al. (2007), if the log-likelihood increases with the number of clusters M, it does not always increases with the number of segments K. Indeed a phenomenon of self-penalization occurs at the 'true' number of segments when the detection of breakpoints is easy, stressing to choose K simply as the value that maximizes the log-likelihood. However when the detection

of breakpoints is more difficult, choosing the maximum value would tend to overestimate K. Picard et al. (2007) suggested to add a penalty. A Bayesian Information Criterion (BIC)-based penalty appeared to be sufficient in this case, although it does not work in pure segmentation (Picard et al., 2005). As BIC is the most popular criterion to choose the optimal number of clusters in a mixture model (Frühwirth-Schnatter, 2006), we propose to use the maximum value of the following BIC-based penalised likelihood $\mathcal{B}_{K,M}$ for the selection of both K and M parameters:

$$\mathcal{B}_{K,M} = \mathcal{L}_{K,M} - \frac{5 \times M - 1}{2} log(2n) - \frac{K-1}{2} log(2n),$$

where $\mathcal{L}_{K,M}$ stands for the log-likelihood for the optimal segmentation-clustering with K segments classified into M states. The penalization terms in the BIC criterion is half the number of parameters times the logarithm of the size of the dataset. For our model the number of parameters to be estimated is $2M$ means $+ 2M$ variances $+ M - 1$ proportions for the states, and $K - 1$ breakpoints for the segments, and the size of the dataset for n bivariate values is $2n$.

Although this procedure appears to work well for choosing the optimal number of segments, it has been observed to be less reliable for choosing the optimal number of states, which tends to be overestimated. We therefore advise users to set an a priori number of states M, based on biological knowledge. We also advise to look at the plot of the BIC-penalized log-likelihood, as in figure 4b of main text, to check that the solution obtained makes sense.

Supporting Information 2: Interpolating entrance and exit points of a circle

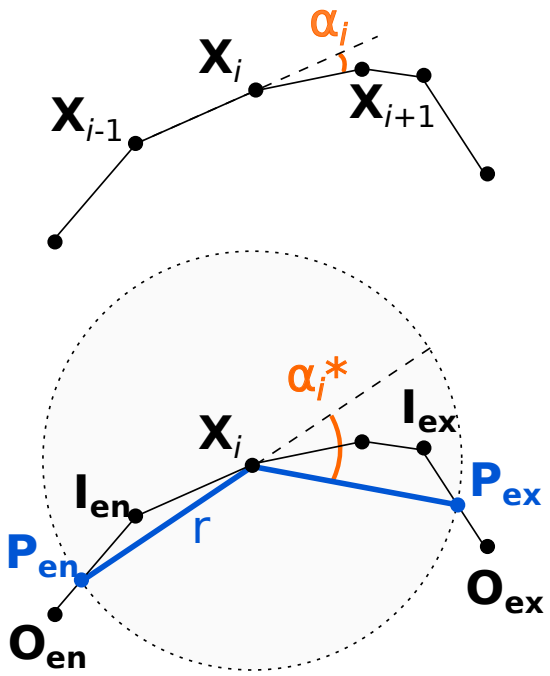


Figure S2.1: Interpolating entrance and exit points of a circle.

Given a series of locations $\boldsymbol{X}_i = (x_i, y_i)$ recorded at constant time intervals Δt. Whereas turning angle at constant time intervals α_i (top) corresponds to the change in direction between vectors $\boldsymbol{X}_{i-1} \rightarrow \boldsymbol{X}_i$ (with length L_i) and $\boldsymbol{X}_i \rightarrow \boldsymbol{X}_{i+1}$ (with length L_{i+1}) and therefore act as a proxy for angular speeds $\left(\frac{\alpha_i}{\Delta t}\right)$, turning angle at constant length interval α_i^* (bottom) corresponds to the change in direction between vectors $\boldsymbol{P_{en}} \rightarrow \boldsymbol{X}_i$ and $\boldsymbol{X}_i \rightarrow \boldsymbol{P_{ex}}$, where $\boldsymbol{P_{en}}$ and $\boldsymbol{P_{ex}}$ are the last entrance and first exit locations, respectively, of a virtual circle with radius r and centred on current location $\boldsymbol{X}_i$. Let $\boldsymbol{I} = (x_{in}, y_{in})$ and $\boldsymbol{O} = (x_{out}, y_{out})$ be the last inside and first outside recorded locations, respectively, of the first passage at the circle perimeter, either backwards ($\boldsymbol{I} = \boldsymbol{I_{en}}$, $\boldsymbol{O} = \boldsymbol{O_{en}}$, $\boldsymbol{I_{en}} = \boldsymbol{X}_i$ if $L_i > r$) to determine $\boldsymbol{P_{en}}$, or forwards ($\boldsymbol{I} = \boldsymbol{I_{ex}}$, $\boldsymbol{O} = \boldsymbol{O_{ex}}$, $\boldsymbol{I_{ex}} = \boldsymbol{X}_i$ if $L_{i+1} > r$) to determine $\boldsymbol{P_{ex}}$. The location $\boldsymbol{P}$ (either $\boldsymbol{P_{en}}$ or $\boldsymbol{P_{ex}}$) corresponds to the point where the vector $\boldsymbol{I} \rightarrow \boldsymbol{O}$ intersects the circle perimeter. The length of this vector is $d_{IO} = (d_x^2 + d_y^2)^{0.5}$, with $d_x = x_{out} - x_{in}$, $d_y = y_{out} - y_{in}$, and its orientation is θ, with $\cos(\theta) = \frac{d_x}{d_{IO}}$ and $\sin(\theta) = \frac{d_y}{d_{IO}}$. In a new orthonormal frame of reference (U, V) originating at $\boldsymbol{I}$ and with U axis running through $\boldsymbol{O}$, the coordinates of current location $\boldsymbol{X}_i$ become $u_i = (x_i - x_{in})\cos(\theta) + (y_i - y_i n)\sin(\theta)$ and $v_i == (y_i - y_{in})\cos(\theta) + (x_i - x_i n)\sin(\theta)$. By applying Pythagoras' theorem, one gets $r^2 = (d_{IP} - u_i)^2 + v_i^2$, where d_{IP} corresponds to the distance between $\boldsymbol{I}$ and $\boldsymbol{P}$, with $d_{IP} > u_i$. Entrance or exit location can therefore be linearly interpolated as $\boldsymbol{P} = \boldsymbol{I} + (\boldsymbol{O} - \boldsymbol{I})\frac{d_{IP}}{d_{IO}}$ (i.e. $x_P = x_{in} + \cos(\theta) d_{IP}$ and $y_P = y_{in} + \sin(\theta) d_{IP}$), with $d_{IP} = u_i + (r^2 - v_i^2)^0.5$.

Supporting Information 3: Efficiency of the segmentation-clustering method

1 Comparison with HMM for low ($\zeta = 0.2$) and high ($\zeta = 0.4$) noise levels

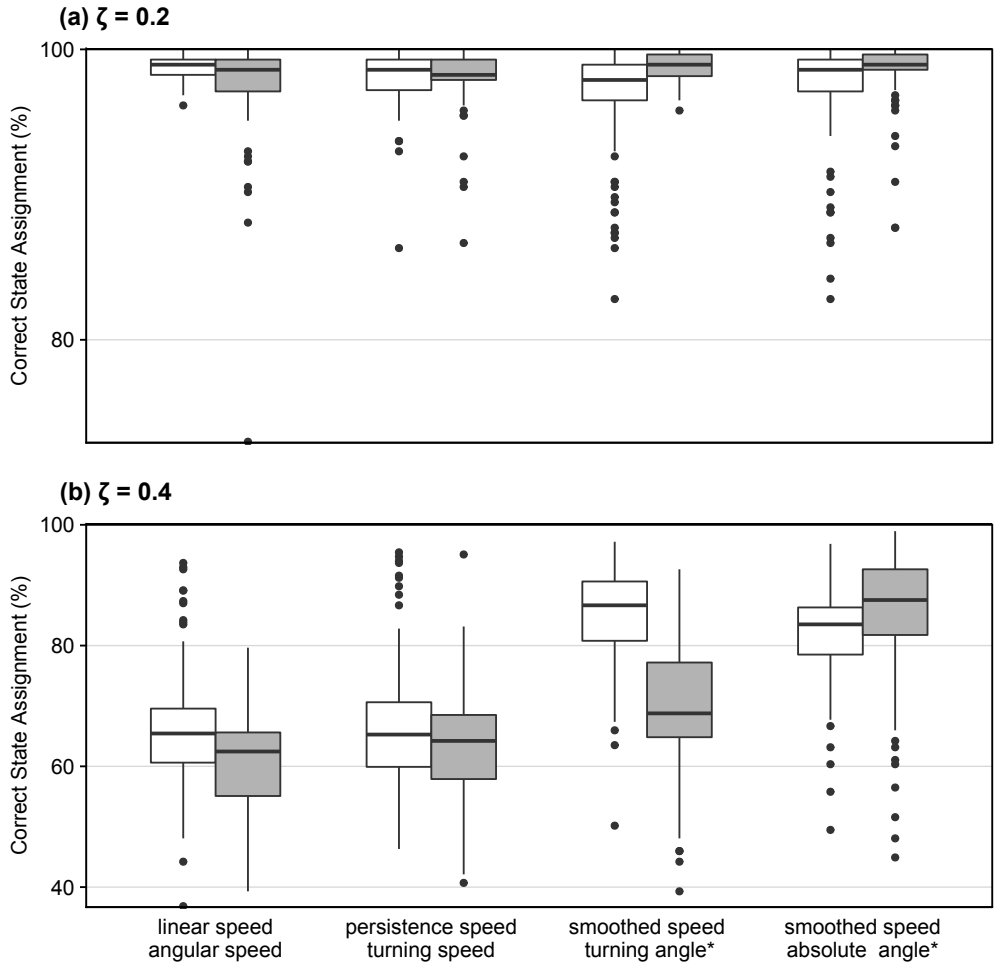


Figure S3.1: The boxplots show the proportion of correct state assignments, obtained for various bivariate signals when the true number of states is known (M = 3) with noise level $\zeta = 0.2u$ (a) or $\zeta = 0.4u$ (b), as estimated from 100 replicates. The star (*) indicates turning angles computed with a constant step length, in terms of arithmetic (α_i^*) or absolute (—α_i^*—) values. The white boxplots show the results obtained with HMM-based R package momentuHMM (McClintock and Michelot, 2018), with informative initial state-dependent probability distribution parameters set to the true values of the various metrics in the different states (using the following distributions: Gaussian for persistence and turning speeds, wrapped Cauchy for angular speed and turning angle α_i^*, Weibull for linear speed, smooth speed and —turning angle*—). The grey boxplots shows the results obtained using the segclust2d/segmentation-clustering procedure with $L_{min} = 10$.

2 Estimation of the number of states

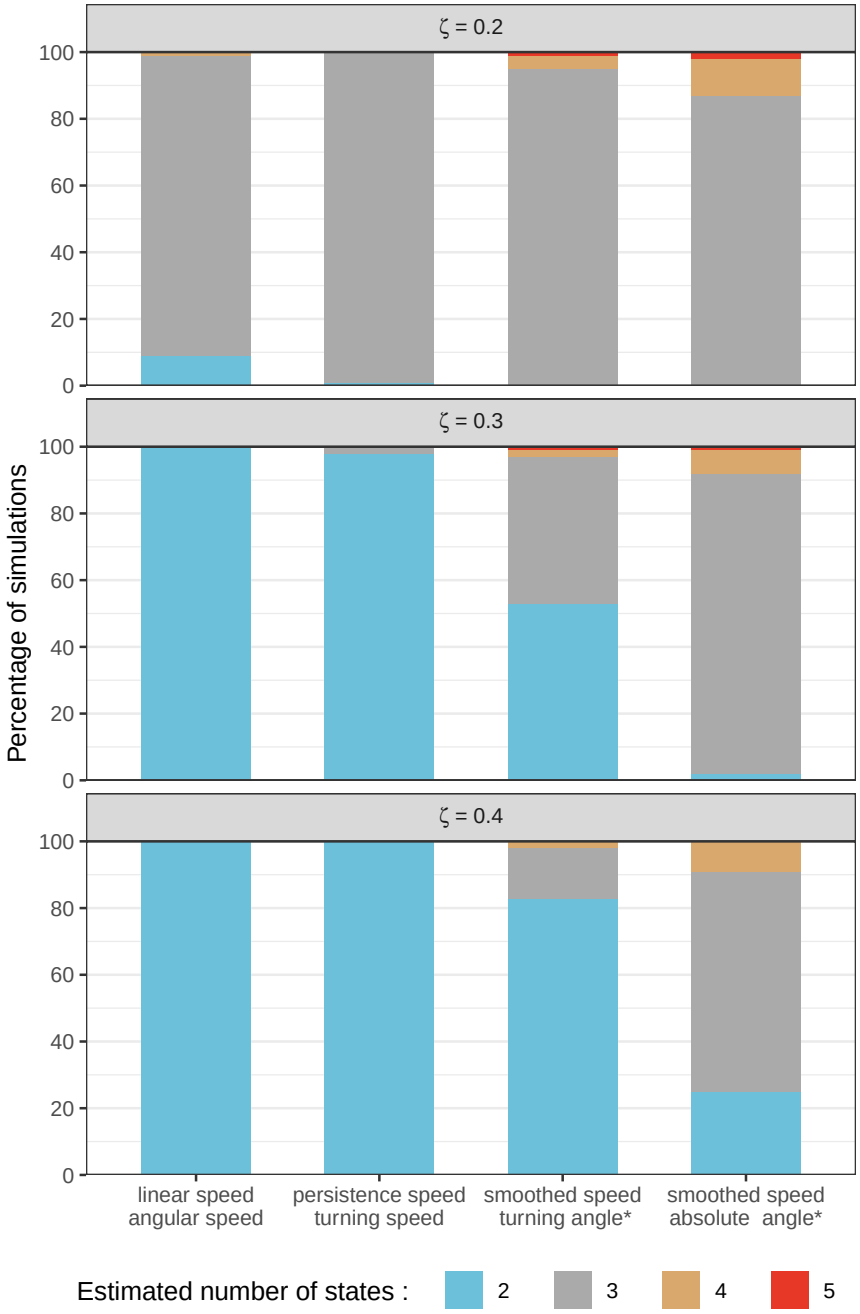


Figure S3.2: The various bars show the proportions of simulations resulting in a predicted number of states (i.e. behavioural modes) equal to 2, 3, 4, or 5, for the three noise levels considered ($\zeta = 0.2u$, $\zeta = 0.3u$ and $\zeta = 0.4u$) and the four types of couples of metrics considered. The true number of states is 3. The star (*) indicates turning angles computed with a constant step length, in terms of arithmetic (α_i^*) or absolute ($—\alpha_i^*—$) values. The couple of metrics leading to best segmentation when the true number of states is known – absolute turning angle computed with a constant step length and smoothed speed – also leads to the best estimation of the number of states, but this latter estimation is not fully satisfactory, and should be worse with actual data because of possible mixing of movement behaviours.

References

Barraquand, F. and S. Benhamou, 2008. Animal movements in heterogeneous landscapes: identifying profitable places and homogeneous movement bouts. Ecology 89:3336–48. URL http://www.ncbi.nlm.nih.gov/pubmed/19137941.

Bellman, R., 1954. The theory of dynamic programming. Bulletin of the American Mathematical Society 60:503–515. URL https://www.ams.org/home/page/.

Benhamou, S., 1989. An olfactory orientation model for mammals' movements in their home ranges. Journal of Theoretical Biology 139:379–388. URL http://www.sciencedirect.com/science/article/pii/S0022519389802164.

———, 2014. Of scales and stationarity in animal movements. Ecology Letters 17:261–272. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.12225.

Benhamou, S. and P. Bovet, 1989. How animals use their environment: a new look at kinesis. Animal Behaviour 38:375–383.

Beyer, H. L., J. M. Morales, D. Murray, and M.-J. Fortin, 2013. The effectiveness of bayesian state-space models for estimating behavioural states from movement paths. Methods in Ecology and Evolution 4:433–441.

Breed, G. A., E. A. Golson, and M. T. Tinker, 2017. Predicting animal home-range structure and transitions using a multistate Ornstein-Uhlenbeck biased random walk. Ecology 98:32–47. URL http://onlinelibrary.wiley.com/doi/10.1002/ecy.1615/abstract.

Bunnefeld, N., L. Börger, B. van Moorter, C. M. Rolandsen, H. Dettki, E. J. Solberg, and G. Ericsson, 2011. A model-driven approach to quantify migration patterns: individual, regional and yearly differences. Journal of Animal Ecology 80:466–476. URL http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2656.2010.01776.x/abstract.

Cagnacci, F., S. Focardi, A. Ghisla, B. van Moorter, E. H. Merrill, E. Gurarie, M. Heurich, A. Mysterud, J. Linnell, M. Panzacchi, R. May, T. Nygøard, C. Rolandsen, and M. Hebblewhite, 2016. How many routes lead to migration? Comparison of methods to assess and characterize migratory movements. Journal of Animal Ecology 85:54–68. URL http://onlinelibrary.wiley.com/doi/10.1111/1365-2656.12449/abstract.

Couriot, O., A. M. Hewison, S. Saïd, F. Cagnacci, S. Chamaillé-Jammes, J. D. Linnell, A. Mysterud, W. Peters, F. Urbano, M. Heurich, et al., 2018. Truly sedentary? the multi-range tactic as a response to resource heterogeneity and unpredictability in a large herbivore. Oecologia 187:47–60.

Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 39:1–38. URL https://www.jstor.org/stable/2984875.

Dias, M. P., J. P. Granadeiro, and J. M. Palmeirim, 2009. Searching behaviour of foraging waders: does feeding success influence their walking? Animal Behaviour 77:1203–1209. URL http://www.sciencedirect.com/science/article/pii/S0003347209000657.

Frühwirth-Schnatter, S., 2006. Finite mixture and markov switching models. Springer Science & Business Media.

Gloaguen, P., S. Mahévas, E. Rivot, M. Woillez, J. Guitton, Y. Vermard, and M. P. Etienne, 2015. An autoregressive model to describe fishing vessel movement and activity. Environmetrics 26:17–28. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/env.2319.

Gurarie, E., R. D. Andrews, K. L. Laidre, and E. Letters, 2009. A novel method for identifying behavioural changes in animal movement data. Ecology Letters 12:395–408.

Gurarie, E., F. Cagnacci, W. Peters, C. Fleming, J. M. Calabrese, T. Müller, and W. F. Fagan, 2017. A framework for modeling range shifts and migrations: Asking whether, whither, when, and will it return. The Journal of Animal Ecology .

Karlin, S. and H. M. Taylor, 1975. A first course in stochastic processes.

Langrock, R., R. King, J. Matthiopoulos, L. Thomas, D. Fortin, and J. M. Morales, 2012. Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. Ecology 93:2336–2342. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/11-2241.1.

Lavielle, M., 2005. Using penalized contrasts for the change-point problem. Signal Processing 85:1501–1510. URL http://www.sciencedirect.com/science/article/pii/S0165168405000381.

Levin, S. A., 1992. The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. Ecology 73:1943–1967. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1941447.

McClintock, B. T., R. King, L. Thomas, J. Matthiopoulos, B. J. McConnell, and J. M. Morales, 2012. A general discrete-time modeling framework for animal movement using multistate random walks. Ecological Monographs 82:335–349. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/11-0326.1.

McClintock, B. T. and T. Michelot, 2018. momentuHMM: R package for generalized hidden Markov models of animal movement. Methods in Ecology and Evolution 9:1518–1530. URL https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12995.

Michelot, T., R. Langrock, and T. A. Patterson, 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. URL https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12578.

Monsarrat, S., S. Benhamou, F. Sarrazin, C. Bessa-Gomes, W. Bouten, and O. Duriez, 2013. How predictability of feeding patches affects home range and foraging habitat selection in avian social scavengers? PLOS ONE 8:e53077. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0053077.

Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell, 2004. Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology 85:2436–2445. URL https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/03-0269.

Mueller, T. and W. F. Fagan, 2008. Search and navigation in dynamic environments - from individual behaviors to population distributions. Oikos 117:654–664. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0030-1299.2008.16291.x.

Naidoo, R., P. D. Preez, G. Stuart-Hill, M. Jago, and M. Wegmann, 2012. Home on the range: Factors explaining partial migration of African buffalo in a tropical environment. PLOS ONE 7:e36527. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0036527.

Nams, V. O., 2014. Combining animal movements and behavioural data to detect behavioural states. Ecology Letters 17:1228–1237. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.12328.

Picard, F., S. Robin, M. Lavielle, C. Vaisse, and J.-J. Daudin, 2005. A statistical approach for array CGH data analysis. BMC Bioinformatics 6:27. URL https://doi.org/10.1186/1471-2105-6-27.

Picard, F., S. Robin, E. Lebarbier, and J.-J. Daudin, 2007. A Segmentation/Clustering Model for the Analysis of Array CGH Data. Biometrics 63:758–766. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1541-0420.2006.00729.x.

Pohle, J., R. Langrock, F. M. van Beest, and N. M. Schmidt, 2017. Selecting the number of states in hidden markov models: Pragmatic solutions illustrated using animal movement. Journal of Agricultural, Biological and Environmental Statistics 22:270–293. URL https://doi.org/10.1007/s13253-017-0283-8.

Pyke, G. H., 1978. Are animals efficient harvesters? Animal Behaviour 26:241–250. URL http://www.sciencedirect.com/science/article/pii/0003347278900246.

Rigaill, G., 2015. A pruned dynamic programming algorithm to recover the best segmentations with 1 to k_{max} change-points. Journal de la Société Française de Statistique 156:180–205.

Schwarz, G. et al., 1978. Estimating the dimension of a model. The annals of statistics 6:461–464.

Smith, J. N. M., 1974. The food searching behaviour of two european thrushes. II: The adaptiveness of the search patterns. Behaviour 49:1–60. URL http://booksandjournals.brillonline.com/content/journals/10.1163/156853974x00390.