

How to use random walks for modeling the movement of wild animals

Geoffroy Berthelot^{1,2,3,✉}, Sonia Saïd⁴, and Vincent Bansaye¹

¹Centre de mathématiques appliquées (CMAP), Ecole Polytechnique, Route de Saclay, 91128 Palaiseau, France

²REsearch LABoratory for Interdisciplinary Studies (RELAIS), Paris, France

³Institut national du sport, de l'expertise et de la performance (INSEP), Avenue Du Tremblay, Paris, France

⁴Office Français de la Biodiversité, Direction Recherche et Appui Scientifique, Unité Ongulés Sauvages-Unité Flore et Végétation, Montfort, 01330 Birieux, France

Animal movement has been identified as a key feature in understanding animal behavior, distribution and habitat use and foraging strategies among others. At the same time, technological improvements now allow for generating large datasets of high sampled GPS data over a long period of time. However, such datasets often remain unused or used only in part due to the lack of practical models that can directly infer the desired features from raw GPS locations and the complexity of existing approaches. Some of them being disputed for their lack of rational or biological justifications in their design. We propose a simple model of individual movement with explicit parameters based on essential features of animal behavior. The main thrust was to stick to empirical observations, rather than using black-box models that could possibly perform very well while providing little insight from an ecological perspective. We used a simple model, based on a two-dimensional biased and correlated random walk with three forces related to advection, attraction and immobility of the animal. These forces can be directly estimated using individual raw GPS data. The performance of the model is assessed through 5 statistics that describe the spatial features of animal movement. We demonstrate the approach by using GPS data of 5 roe deer with high frequency sampling. We show that combining the three forces significantly improves the model performance. We also found that the model's parameters are not affected by the sampling rate of the GPS, suggesting that our model could also be used with low frequency sampling GPS devices. Additionally, the practical design of the model was verified for detecting spatial feature abnormalities (such as voids) and for providing estimates of density and abundance of wild animals. Our results show that a simple and practicable random walk template can account for the spatial complexity of wild animals. Integrating even more additional features of animal movement, such as individuals' interactions or environmental repellents, could help to better understand the spatial behavior of wild animals.

spatial ecology | movement modeling | animal movement | random walk | GPS data

Correspondence: geoffroy.berthelot@insep.fr

Introduction

Animals live in an environment that is patchy and hierarchical, and the manner in which individuals search for spatially dispersed resources is crucial to their success in exploiting them (1). At the same time, the tracking of animals using the modern global positioning system (GPS) now allows for the collection of important datasets on animal locations (2). They are often used for the analysis of the home range behavior i.e., restrict their movements to self-limited portions

of space far smaller than expected from their sole locomotion capacities (3) and, more generally, to better understand the spatial and temporal behavior of animals (4, 5). New, smaller and reliable devices allow for gathering large datasets (e.g. locations or activity data for instance) at a finer temporal and spatial scale and offer a greater opportunity to investigate animal movement at the individual scale. However, datasets often remain only partially used due to both the lack of practical models that can directly infer the desired features from raw GPS locations and the complexity of existing approaches. Meanwhile, ecologists in particular are called to develop new capabilities to deal with these large datasets (6, 7).

The modeling of animal movement includes a wide range of methodologies: biased and/or correlated random walks (BCR), the disputed Lévy Flight/walk (8–12), Stochastic Differential Equation (SDE) (13–16) including diffusion models based on the two-dimensional Ornstein-Uhlenbeck process (17–21) and other more exotic algorithms using *ad-hoc* rules to mimic movement features such as memory (22, 23). Lévy Flight has convenient patterns but ecological motivations are scarce (11). SDE -the continuous analog of BCR- or the Brownian bridge and Movement Model (24) may be used to interpolate the trajectory between two observations. SDE includes a drift (directional) and one or several random diffusion processes (16, 25). BCRs are convenient tools to model animal movement as the discrete time is well adapted to regular GPS data (25) and the parameters of the BCR can be directly interpreted in terms of the behavior of the animal. They correspond to the attraction of some locations, the inertia and memory feature of the movement, time dependence of the movement, local interactions with other individuals, etc. Some key features of animal movement have already been identified by previous studies, including diffusion (or randomness) which corresponds to an isotropic random motion, where the individual has the same probability to go in all directions; Attraction (directional bias) where the movement of the animal is anisotropic and is confined in an area or domain, according to (26) and other studies (3), while the attraction may depend on the distance from the isobarycenter of locations (27); Inertia where the movement of the animal is also shaped by foraging tasks where the animal alternates exploration periods -the path has high tortuosity- with straightforward movements (28). These three features can be implemented as parameters of a BCR.

This study aims at modeling animal movement of sedentary

Table 1. Data summary. For each animal, the total number of observations n is given along with the period of collection (date and time), the sampling rate \bar{T} (*i.e.* the average time between 2 observations) (in min.) and corresponding standard deviation, total distance (in kilometers), total recording time (in days) and average speed s (in 10^{-2} m.s $^{-1}$).

Animal	n	period of collection	$\bar{T} \pm \text{s.d. (min.)}$	tot. distance (km.)	time (d.)	$s (10^{-2} \text{ m.s}^{-1})$
Deer 1	29520	09/01/2010 00:00 - 11/08/2010 23:50	10.49 \pm 3.18	945.67	214.99	5.09
Deer 2	27324	10/12/2009 00:00 - 24/06/2010 23:50	10.38 \pm 3.69	1030.97	196.99	6.06
Deer 3	23301	16/03/2010 00:00 - 07/09/2010 23:41	10.88 \pm 8.67	876.33	175.99	5.76
Deer 4	24735	22/01/2010 00:00 - 21/07/2010 23:51	10.54 \pm 5.16	898.38	180.99	5.74
Deer 5	21451	16/01/2010 00:01 - 24/06/2010 07:31	10.69 \pm 10.11	785.93	159.31	5.71

individuals over short periods (29), in a homogeneous landscape using GPS data sets and a BCR. As a first approach, we consider one individual of a given species with no interaction and simulate its movement in continuous space and discrete time in 2 dimensions using a BCR with the aforementioned parameters -diffusion, attraction, inertia- and one additional term: immobility. This last parameter takes into account the absence of movement between a pair of locations (*i.e.* distance is 0). This can be accredited to technological limitations with the satellite telemetry due to a weak GPS signal strength (*i.e.* due to natural elements: such as when the animal was standing underneath a rock or due to dense clouds, dust particles, mountains or flying objects, such as airplanes). However, this can also be part of the behavior of the animals, during specific times: sleep cycles for instance. The introduced model is general, simple and informative as the three parameters are directly inferred from the GPS data set. The model can be sophisticated by including more complicated environmental aspects of individual movement, such as spatial memory (30, 31), reinforcement and site fidelity (32), environmental predictability (33) including landscape effect (34), interacting individuals and prey-predator dynamics. We also introduce five statistics that describe the spatial features of animal movement, with a particular interest in census (*i.e.* population estimation) using transects. These statistics allow for estimating the ‘performance’ of the BCR model, or in other words its ability to mimic the spatial characteristics of an animal’s movement. Finally, we investigate how our model can address ecological questions including census and spatial issues, using GPS data sets.

Material and Methods

Data. The locations of 5 GPS-collared red deer (*Cervus elaphus*) were gathered at La Petite Pierre National Hunting and Wildlife Reserve (NHWR), in north-eastern of France (48.8321 (Lat.) / 7.3514 (Lon.)). The reserve is an unfenced 2670 ha forest area characteristics by deciduous trees (mostly *Fagus sylvatica*) in the western part and by coniferous species (mostly *Pinus sylvestris* and *Abies alba*) in the eastern part in nature reserve surrounded by crops and pastures. It is located at a low elevation area of the Vosges mountain range, which rises up to 400 m a. s. l. The climate is continental with cool summers and mild winters (mean January and July temperatures of 1.4 and 19.6°C, respectively, data from Phalsbourg weather station, Météo France, from 2004 to 2017). Three ungulate species are present and mainly managed through hunting in the NHWR: wild boar, red deer and roe deer. The

present study focuses on female red deer for test model. A detailed overview of the landscape and surroundings is given in (35). The GPS data had regular observation frequencies with high frequency sampling (Table 1). In the following text, we note $X_i = [X_i^1, X_i^2]$ the successive locations of the individual with $X_i \in \mathbb{R}^2$ and $i = 1, 2, \dots, n$. We use t_i ($t_1 = 0$) as the time elapsed between two successive locations X_{i-1} and X_i and

$$\bar{T} = \frac{1}{n} \sum_{i=1}^n t_i$$

as the average sampling time. The trajectory of the animal, or ‘path’, was interpolated using linear interpolation between each pair of recorded observations (Figure 1 and detailed in Supplementary file 1 (eq. 18) and associated Graph 2). It approximates the animal travels in straight lines at constant velocity between each pair of locations (36). The home/attractor X_F of one individual was estimated as the isobarycenter of all recorded locations:

$$X_F = \bar{X} = \left[\bar{X}^1, \bar{X}^2 \right] = \left[\frac{1}{n} \sum_{i=1}^n X_i^1, \frac{1}{n} \sum_{i=1}^n X_i^2 \right]$$

Modeling framework. The movement of an animal was modeled by a two-dimensional BCR in discrete time and continuous space which included 3 parameters coupled with isotropic diffusion:

- *Diffusion*: We considered a random direction with uniform spatial distribution in a 2D plane,
- *Attraction* (p_F): A natural way to include this feature is to increase the probability to go to a fixed attractive point named attractor (37). This yielded a drift or advection parameter in the direction of X_F ,
- *Inertia* (p_I): This parameter increased the probability to move forward (*i.e.* to perform one step in the direction of the previous step),
- *Immobility* (p_s): We included this as a specific parameter and the movement was stopped for one step.

Thus, the movement of the animal was modeled by a chain X_n , characterized by a transition matrix:

$$\begin{bmatrix} 1 & 1 + p_I & 1 \\ 1 & p_s & 1 \\ 1 + p_F & 1 & 1 \end{bmatrix} \quad (1)$$

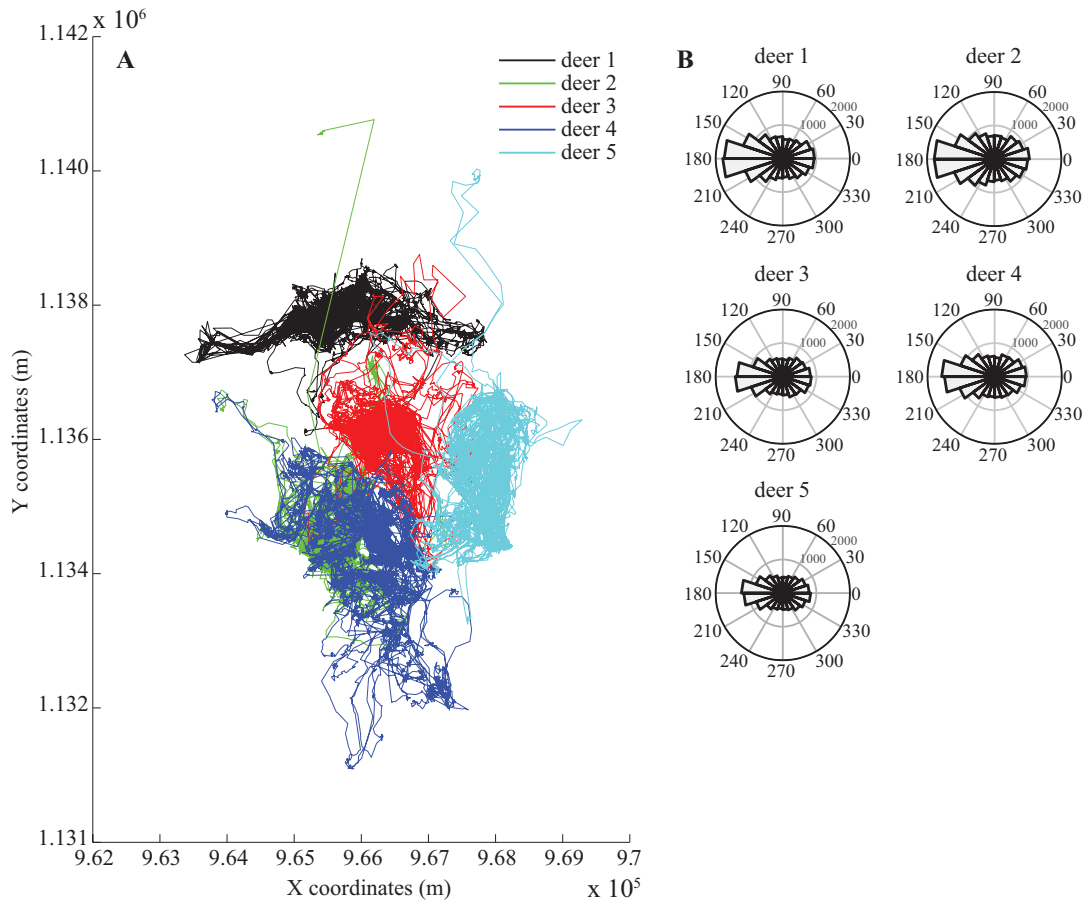


Fig. 1. Individual paths of the five red deer. Individual paths of the five red deer. The individual paths are plotted for the five red deer (left panel, **A**) along with the distribution of the relative turning angles (degrees) in polar plots (right panel, **B**). An angular value of 0 consists in a straight motion from the previous location, while a relative turning angle of 180° corresponds to a turn back.

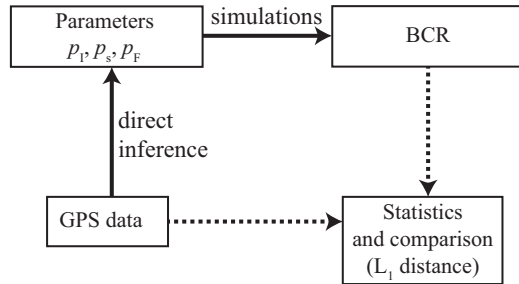


Fig. 2. Framework used for testing the BCR model performance, for one animal. Black lines detail the two operations processed from the GPS dataset. The 3 parameters are estimated from the GPS data and -using these parameters- 1000 simulations of the BCR model are computed. No particular operations are associated with the dotted black lines, but they show how the BCR and the GPS dataset are evaluated and compared using the statistics. We use the same framework to investigate the performance of the BCR with only one or two parameters.

renormalized by $8 + p_I + p_s + p_F$ to obtain the probability distribution of the next position compared to the current one. In the presented case (eq. 1), the immediate past movement of the animal is coming from the down part of the matrix (*i.e.* upward vertical direction: \uparrow) and the attractor is estimated to be located southwest from the actual position of the individual.

We analyzed the distribution of the distances covered between each pair of locations (Supplementary file 2) and used

the log-normal law to model the distance covered by the individual between each time step. Thus, if $p_F = p_I = p_s = 0$, the BCR resumes to a typical two-dimensional random walk with a log-normal step size distribution $\ln \mathcal{N}(\mu, \sigma^2)$. The three parameters were accordingly tuned to the corresponding dataset using a straightforward estimation procedure (Supplementary file 1). We then simulated 1000 BCR and used 5 statistics that describe the spatial features of an individual's movement to assess the BCR performance (framework detailed in Figure 2). We additionally perform a sensitivity analysis by testing the performance of the BCR but using only one or two parameters instead of the three parameters.

Statistics for describing animal movement. The 5 statistics were designed to assess the model reliability on spatial features including: (*i*) the distribution of relative turning angles which provides information about the local movement of the animal, (*ii*) the home range which provides information about the spatial density of observations and (*iii*) observation counts using still and mobile transects, providing information on absolute observation abundance (38). We did not keep track of locations during transect counts, thus ignoring spatial information, as it was already collected in (*ii*). Additional details are provided in Supplementary file 3.

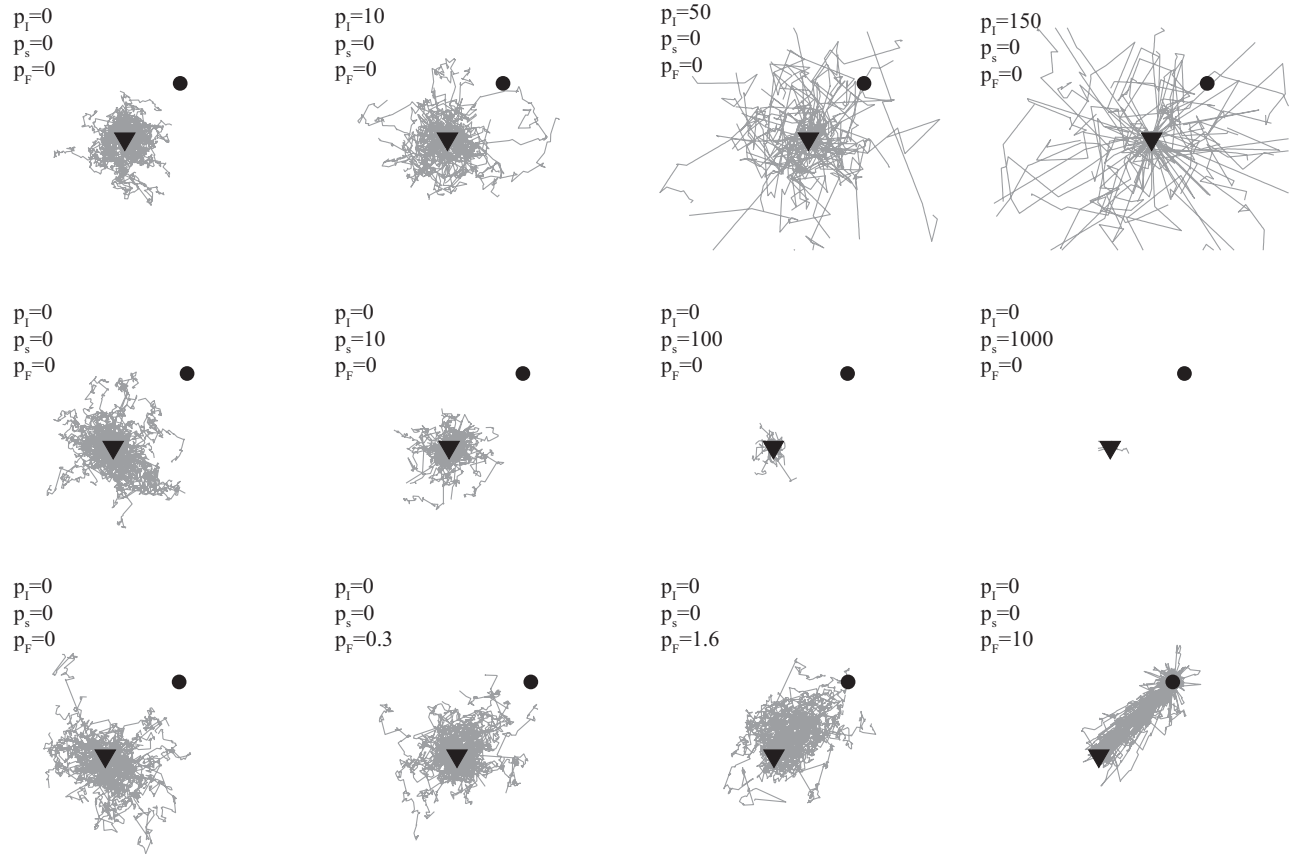


Fig. 3. Simulated animal motions over arbitrary parameter values. Fifty motions of length $n_s = 100$ steps are simulated and originate from a common centroid (downward-pointing triangle) with increased levels of inertia (p_I), immobility (p_s) and attractor (p_F). Both the location of the attractor (black dot) and the log-normal parameters controlling the step size distribution are fixed ($\mu = 3$, $\sigma^2 = 1$).

Distribution of turning angles. For each individual, the distribution of counter-clockwise relative turning angles ($\widehat{X_{i-1}X_iX_{i+1}}$) was gathered, provided $d(X_{i-1}, X_i) > d_{\min}$ and $d(X_i, X_{i+1}) > d_{\min}$ with d_{\min} the immobility threshold distance between two locations. We set $d_{\min} < 10\text{m}$, which corresponds to the magnitude of the error typically found in GPS locations (39). This means that we only kept the angles from observations that were separated by an Euclidean distance greater than d_{\min} .

Home range. We used an adaptive kernel density estimator (matlab package kde2d - kernel density estimation version 1.3.0.0) as an estimator of the utilization distribution (40) to represent the home range of the animal. The approach of Z.I. Botev provided an estimate of observation density using a bivariate (Gaussian) kernel with diagonal bandwidth matrix (41). The density was estimated over a grid of 210×210 nodes and we computed the home range area (in m^2) for various values: 100, 99, 95, 90, 80, \dots , 20, 10% of the estimated density. Similarly to the distribution of turning angles, we compared each value of the data's home range against the simulated one.

Dilation. Dilation is generally used to account for the spatial attributes of an object such as to measure an area around the path or the volume of a brownian motion (see Wiener sausage (42) and Gromov-Hausdorff distance). In our approach, we

use dilation of both simulated and GPS paths for two reasons: to have a real -and comparable- number that accounts for how a trajectory has explored space and because it is natural tool from a census point of view (the dilated path corresponds to the area where the animal can be detected). Each simulated or real path was plotted in binary format in a window and dilated with a disk shape. The window size was set to a huge value in order to encapsulate the dilated path while preventing boundary effects, *i.e.* the convex envelope of the dilated area did not collide with any window border. We then estimated the surface covered by the dilated path for 100 different sizes of the disk, from disk size 1 to disk size 100. We compared each value of the data's estimated surface against the simulated one.

Immobile transects. We used still transects that counted the number of times the animal was seen in their line of sight. We arbitrary set the line-of-sight value at 200m. The number of sightings of each transect was gathered and ordered in decreasing order, thus breaking the spatial dependence. We then compared the bins of the resulting histogram in the data and in the simulated path.

Mobile transects. First, the movement of the animal was linearly interpolated from the GPS data, meaning that between two recorded locations the individual followed a linear path. The speed of the animal between two locations was accord-

ingly reconstructed using the recorded times t_i between each location. Second, we used mobile transects as the ecological sampling method, where each transect ‘count’ the intersection between its path and the animal’s one. The mobile transects followed a predefined path at a given constant speed as time increased. The area of vision of each transect was defined as a circle of a given radius. Each time the path of an individual collided with an area of vision, the count of the corresponding transect increased by 1. Two types of movements were used: linear and clockwise rotational transects. The initial locations of both types of transects are X_1 and X_F . Both the animal and mobile transects started to move at the same time. At each of the two locations X_1 , X_F , 8 linear transects moved in the 8 cardinal directions, totalizing 16 transects. For the linear transects, every 10000 time steps, we set 2×8 new transects starting at the same locations and following the same directions. Clockwise rotational transects were rotated around X_1 and X_F using a 500m radius. When we reached t_n , we gathered the total count (*i.e.* the count of all transects). For the two types of transects, we gathered the total count for 6 different lines of sight: 50, 100, 200, 400, 500, 1000m. and 4 speeds: $s/4$, $s/2$, s , $2.s$ with s the average speed of the animal. We then aggregated the overall count in each of the two types of transects, and compared the results from the data and the simulated path (Supplementary file 1 and 3).

Scale invariance. We also studied how the scale affected the BCR parameters. The movement path formed by the GPS observations X_i was subsampled (decimated) for each individual. We only kept every k^{th} observation starting with the first one and $k \in [1, 10]$. For $k = 1$ the path corresponded to the original one. The time spent between each successive observation was also accordingly reconstructed in order to keep track of \bar{T} in subsampled movement paths. The time between two locations X_i and X_{i+k} was reconstructed as: $t'_j = \sum_{i=j}^{i+(k-1)} t_i$ with $j \in [1, 1+k, 1+2k, \dots, n-(k-1)]$. We then compared the resulting parameters p_I , p_F and p_s as the subsampling parameter k increased.

Deterministic aspects of the statistics. Whereas the BCR is a stochastic process, the deterministic aspects of the 5 statistics were tested with an increasing number of steps n_s . The statistic associated with each realization of the model (a simulated path) is a random variable. If the distribution of these random variables has low concentration (high variance) then it is not a convenient statistic as it cannot be used as a reference for assessing the model’s performance, even when averaging over multiples realizations. On the opposite, if the statistic is deterministic it can provide a reliable tool to assess the model’s performance. This was numerically tested over a range of increasing n_s values with $n_s = 10^4, 2 \times 10^4, \dots, 4 \times 10^5$. For each of those step values, a set of 100 BCR was simulated with parameters p_I , p_F and p_s estimated from the first deer (see Table 2) and we studied the variance of the statistics.

Table 2. Estimated parameters. For each deer, the estimated parameters p_I , p_F , p_s and the two parameters μ , σ that control the step size distribution are given.

Animal	p_I	p_F	p_s	μ	σ
Deer 1	0.01	0.01	2.01	2.94	1.01
Deer 2	0.06	0.13	1.44	3.15	0.97
Deer 3	0.12	0.05	1.70	3.07	1.04
Deer 4	0.10	0.06	1.52	3.10	0.98
Deer 5	0.22	0.24	1.66	3.03	1.06

Application: detecting anomalous voids. The proposed model could be used to infer environmental and behavior information from the dataset. For example, it was possible to detect anomalous voids (or holes) in the spatial territory of the individual using Monte-Carlo simulations of the model. Anomalous means that the observed void is not related to the randomness of the movement, but rather related to a geographical artifact. The three parameters p_I , p_F , p_s and parameters of the log-normal distribution for step-size were accordingly estimated from the data of each individual, similarly to previous experiments (Figure 2 and Table 2). A simple heuristic was used to find voids in empirical and simulated paths for each individual: we computed the alpha shape of all locations using a fixed alpha radius of 60m. This allowed for determining the surface covered by all locations while preserving the voids. We then collected the area of each void provided they had an area of at least 100m². We focused on voids near the center of the alpha shape in order to avoid artificial voids, generated by the weak density of locations at the boundaries. We ran 10,000 iterations of the model for each animal and estimated the probability p_\emptyset of finding voids of different sizes in the simulated paths. This probability was then compared to voids found in the GPS datasets and available environmental information was used to determine whether any geographical element(s) could explain the unexpected voids.

Results

The parameters estimated for each individual are given in Table 2 for the model. We showed that the parameters μ and σ^2 were close for all individuals, and estimates of the three parameters for individuals 3 and 4 were similar. The values of p_I and p_F showed that inertia and attractor play a greater role in the movement of deer 5 ($p_I = 0.22$, $p_F = 0.24$), compared to the other individuals. The immobility p_s was stable across the individuals while p_I and p_F varied together (Table 2). The latter is a mechanistic effect, as they act as opposite forces. We also detailed additional configurations of the BCR where we used only some of the parameters (Supplementary files 8 and 9).

Evaluation of the BCR model. The distribution of error of each of the model’s 5 configurations in the 5 statistics is provided in Figure 4 for deer 5 (see Supplementary file 4 for the complete results). The mean error and standard deviation, median error and interquartile range are also provided in Supplementary file 9, for all deer and statistics. We showed that combining the parameters plays an important role in model-

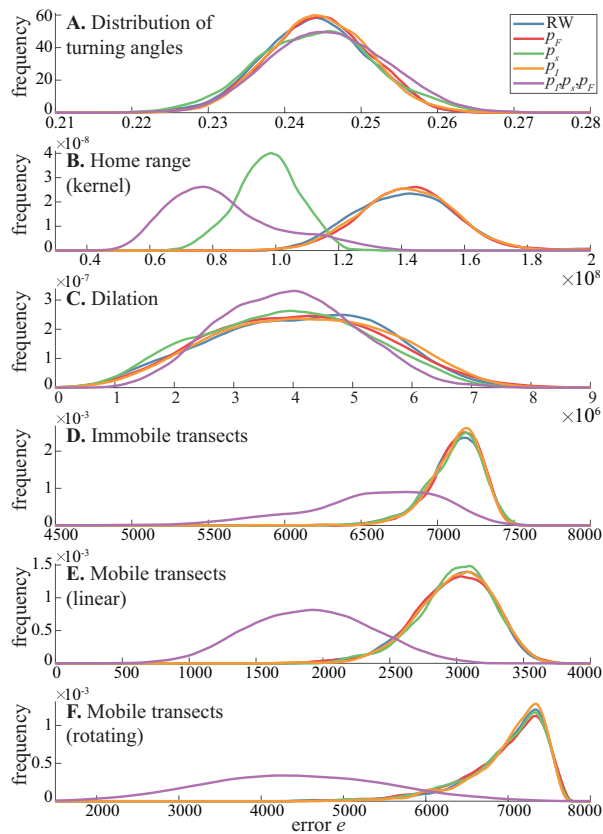


Fig. 4. Density of error e of all 5 configurations tested in each statistic for deer 5. Densities are fitted by the Epanechnikov kernel function. A complete comparison of results for all red deer is provided in Supplementary file 1.

ing deer behavior. Configurations with only one parameter did not perform well on average while further investigations showed that combining p_I , p_F and p_s allows for a better description of movement, especially regarding the census statistics for both linear and rotational transects and home range estimates (Supplementary files 4 and 9). Four of the 5 statistics became more and more deterministic and concentrated around their mean value (Supplementary file 5). On the contrary, the variance of the dilation seemed to increase for the first n_s steps. The trend observed in the linear mobile transects also seemed to increase, however the trend was not clear and presented small fluctuations.

Scale invariance. The resampling of movement paths showed that p_s decreased as the subsampling rate increased in all five deer. The three other parameters remain roughly constant (Supplementary file 6).

Application: detecting spatial voids. The resulting alpha shapes and detected voids (holes) are presented for each deer in Supplementary file 7. The probability p_\emptyset of observing such voids is computed and showed in Figure 5. Many voids whose area fell in the interval $[0, 1.5 \times 10^4]$ were related to boundary conditions, where the alpha shape produced artificial voids due to less dense areas. However, the Monte Carlo simulations show that 3 voids, located inside the alpha shape (void 1 (deer 1) and voids 1 and 2 (deer 4)), should not appear. In other words, these voids are possibly not related

to movement randomness but to other spatial features, with good probability. The distributions of errors for each configuration varied in each deer.

Discussion

In this work we aim at providing valuable and reliable ecological information regarding the components of animal movement. We introduce a simple and tractable model to deal with animal movement, based on a two-dimensional BCR in discrete time and continuous space that allows for combining the ecological forces in a simple way. The parameters of the BCR are directly estimated using the GPS data recorded in a large herbivore and its performance is assessed in 5 spatial and ecologically-related statistics. Four of them differ from the typical signals or parameters calculated based on empirical relocations (43) and address the home range size, census issues and animal behavior. The framework is presented in Figure 2 and explains how synthetic (or simulated) paths parameterized from empirical observations are ultimately compared to the empirical paths throughout the statistics. It describes how any field data (30), whatever the sampling rate and type, should be analyzed whenever possible. However, we emphasize that the attributes of the data (such as the sampling rate for instance) should be consistent amongst different individuals in order to allow for inter-individual comparison (see (44) and subsequent discussion). It is also important to have a sufficient number of locations ('sample size') as precision in parameter estimation scales with the sample size, meaning that the more observations, the higher the precision. We focus on three essential forces that may allow for an efficient description of animal motion over large periods of time: inertia p_I , immobility p_s and attractor p_F . Other perturbations are encapsulated into a random noise. The BCR is then designed to embed the main ecological features driving the movement of most terrestrial animals: exploratory or foraging behavior (inertia), patch exploitation or sleeping periods (immobility), attractor (reference and working-memory). The results display that by using those parameters, we get a much better description of animal movement compared to an unbiased random-walk with a log-normal step size distribution.

The model reproduces the distribution of relative turn angles observed in a large herbivore, provided the parameters are tuned accordingly. This distribution of turning angles is similar in all five individuals and resembles a π -oriented oval (Figure 1). This oval pattern was already noted in many other species such as the elk (45), the brushtail possum ('possum': *Trichosurus vulpecula*) (44), the caribou (46), the Mediterranean mouflon (*Ovis gmelini musimon Ovis*) (47) and even in Cinereous Antshrike (*Thamnomanes caesius*) flocks (48). GPS errors are known to possibly generate such directional bias where a stationary animal is most likely to be measured as turning 180° or moving towards (49). We constructed turning angles using locations separated by at least d_{\min} and this pattern remains and carry additional analyses using the adehabitatHR package in R (50) to i compute the residence time as the time spent into a 100m. circular area

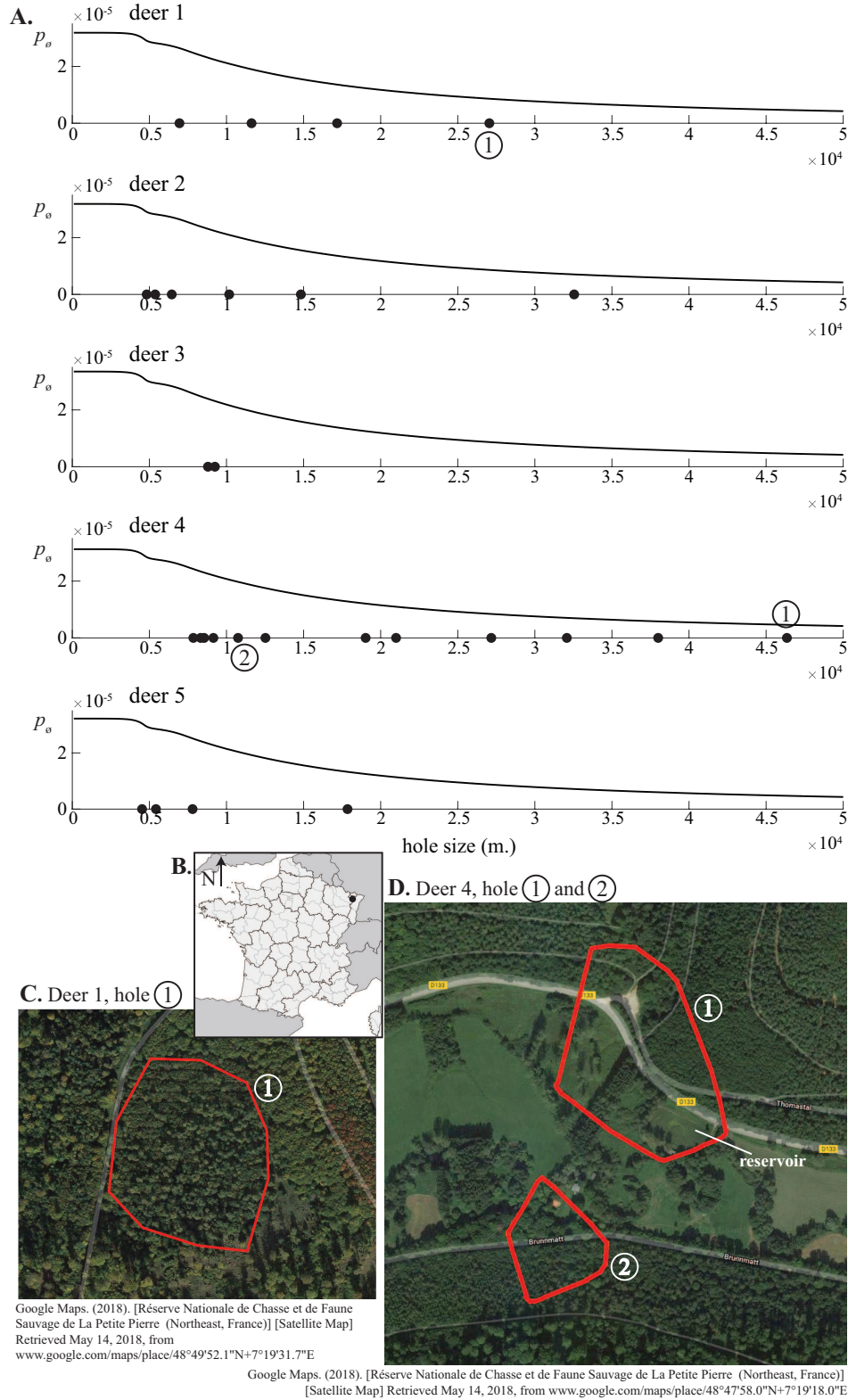


Fig. 5. Using the framework and model to identify spatial voids in movement paths. In panel **A** the probability p_0 of finding voids of different sizes in the simulated paths is given for each deer (black lines). All voids $> 100\text{m}^2$ detected in the empirical GPS location are given (black dots in x-axis). Selected voids (circled numbers) correspond to voids that both are near the center of the alpha shape and have a low p_0 (see alpha shape figures in Supplementary file 1). **B** Geographic location of the study area (black dot). The three voids detected in deer 1 and 4 are detailed in panel **C** and **D** along with the environmental features. Image in **B** was created by TomKr and is distributed under GNU Free Documentation License. It corresponds to the map of France with regions and departments in equirectangular projection and was realized with free IGN data base GeoFla (www.ign.fr).

for each localization and to *ii* study the relative turn angles associated with high ($> 100\text{m.}$) or small ($< 100\text{m.}$) distances covered during a movement. We find that frequent turn backs are associated with high residence time. We also find that turn angles are more evenly distributed in the small distances and that small turn angles are more associated with small distances values. These analyses suggest that the animals have an exploratory / foraging behavior in a local patch. Exploratory or foraging behavior is often modeled using stochastic processes satisfying the Markov property, such as the bivariate Ornstein-Uhlenbeck diffusion process which produces similar distribution patterns of relative turn angles (51, 52), in line with our BCR approach.

Comparison of model configurations reveals that both the two-dimensional random walk and the configurations with only one parameter do not perform well in most cases. On the contrary, combining all three parameters p_I (inertia), p_s (immobility) and p_F (attractor) provides better results in the vast majority of cases (Figure 4 and Supplementary files 4 and 9). This confirms that animal movement is a complex process, driven by several forces instead of a single and dominant one.

The inertia, describing the short-term memory effect, is the first force introduced in this approach. Whether the use of land space by the animal is dependent on short-term or long-term memory is a debated topic. It gave rise to a series of studies that emphasized the importance of memory in animal movement from a biological or modeling perspective (30, 31, 53–57). These studies also underlined that inferring memory effects directly from relocations is not a trivial task. Those relocations instead depend on a mixture of effects, including landscape and territorial constraints, resource patches and possibly long-term memory. Using a single memory feature p_I might be a too simple approximation for efficiently capturing the memory effect. In our approach it is possible to alter p_I in order to include several previous steps instead of just one.

Immobility combines several features of animal movement including animal at rest, in vigilant state, and GPS noise. Multiple factors are known to affect GPS noise, including topographic exposure, canopy cover, vegetation height and the slow movement of the ionosphere. The latter changes by a few centimeters during 30sec intervals (2), possibly introducing up to 20 fold this bias in each of the recorded GPS observations. However, this is small regarding the average step size of non-immobile movements, ranging from 41m (red deer 1) to 46m (red deer 3 and 5). Thus, we assume that the measured step lengths and turning angles reflect the reality. Immobile (*i.e.* $\leq d_{\min}$) observations represent a large proportion in our total datasets: 25.0% (red deer 1), 17.2% (red deer 2), 22.0% (red deer 3), 19.0% (red deer 4) and 23.6% (red deer 5), associated with specific behaviors such as on-site foraging, eating, resting, etc. The estimates of p_s in all five animals are greater than inertia or attractor (Table 2), underlining the importance of considering immobility when analyzing the movement of red deer. This is in line with previous experimental studies that showed the high frequency of

feeding, resting cycles in red deer and labile diet (58).

Site fidelity is the recurrent visit of an animal to a previously occupied location. This is a well-known and wide-spread behavior in the animal kingdom (59). The animal favors locations that are ecologically valuable and related to a foraging or explorative behavior. In our approach, we rather and simply depict site fidelity using one single attractor p_F . The fact it improves the performance of the model when combined with inertia and immobilism confirms that site fidelity (or a simplified estimation of it) should be taken into account when modeling deer movement.

Other analyzes were carried out to ensure the robustness and consistency of the BCR model, including the impact of the GPS sampling rate on the estimated parameters. Several authors pointed out that the temporal resolution of the discretization is of importance: it should be relevant to the considered behavioral mechanisms (5, 44, 60, 61). Schlägel and Lewis focused on the quantification of movement models' robustness under subsampled movement paths (61). They found that increased subsampling leads to a strong deviation of the central parameter in resource selection models (61, 62). They also underlined that the parameter estimates vary with sampling rate when movement models are fitted to data. Postlethwaite and Dennis highlighted the difficulty of comparing model results amongst tracking-datasets that vary substantially in temporal grain (44)). We use data with a relatively high sampling rate (roughly 10m.) and a period of study that is appropriate to the analysis of animal movement at the year scale (Table 1). More importantly, the five animals have the same sampling rate (Table 1). We changed the sampling rate of the movement path to ensure that the parameters related to directional movement are scale invariant (Supplementary file 6). We found that they remain almost constant with increased subsampling, thus strengthening their essential role in animal movement (Supplementary file 6). On the other hand, distance-related parameters such as p_s , μ , σ^2 are highly sensitive to the subsampling rate, but it is a mechanistic effect of the subsampling procedure: as we progressively prune the i^{th} values, the distance between each GPS location increases. The distribution of turning angles or the home range estimates can vary with the temporal scale of study and sampling rate of datasets (44, 63). Schlägel and Lewis further underlined that important quantities derived from empirical data (*e.g.* travel distance or sinuosity) can differ based on the temporal resolution of the data (61, 62). This could have introduced bias in the aforementioned statistics and in parameter inference. However, both the sampling rate and temporal scale of study are similar in all our datasets (Table 1), thus allowing for unbiased inference, proper evaluation of the model's performance and inter-individual comparison. We also investigated the deterministic feature of the 5 statistics and found that the variance decreases or does not change as the number of simulated steps increases in most of the statistics. Thus, 4 of the 5 statistics are robust and add limited randomness to the results when the number of steps increases. The long-term trend is not clear in the mobile transects case as we investigated the variance over 4×10^5 steps

and we may only observe a transient increase or stagnation. This statistic is expected to be similar to the one of immobile transects but the speed of convergence to the null variance may be very slow and it may take a much larger number of steps. The variance of the estimated areas in the dilation statistic increases with n_s because we dilated the simulated paths in a huge window, encapsulating the whole path including a very large portion of empty space around it. This was done to prevent boundary effects when assessing the area of dilated paths, *i.e.* to make sure that dilated paths do not hit any of the window bounds. Otherwise this would produce biased, underestimated areas. However, using a smaller window or, again, a much larger number of steps would result in a null-variance.

Another leading rationale of this work is to investigate the model's ability to address ecological challenges such as estimating the abundance of a given species (census) or detecting anomalous spatial features. We use both mobile and immobile simulated transects to illustrate how the model could be used in the first problem. The probability of counting the same animal multiple times can then be estimated using Monte Carlo simulations. The second issue is, for instance, to detect spatial voids in empirical movement paths. The location of animals may present empty spatial voids (or holes) of various sizes. This may be related to environmental conditions such as urban areas, water, cliff, or other ecological reasons (such as interactions with other individuals *e.g.* repulsive marks) and other factors. Using numerical simulations of the model, we are able to detect anomalous voids in the dataset, that are not related to randomness but to human activity. The void 1 (deer 1) reveals that some environmental changes took place between the recording time of the GPS location in 2010 and the satellite image in 2018. After cross-checking with additional information from the OFB, we learned that the identified area was a forest enclosure. This explains why the deer was not able to reach this area. Both voids 1 and 2 of deer 4 also are related to human activities: forest roads, buildings and one artificial reservoir.

Conclusions

This work introduces a tractable model, based on a two-dimensional BCR for describing animal movement in discrete time and continuous space. The model allows for a direct and explicit estimation of the three parameters that provide the optimal design regarding the studied statistics. Moreover, it allows for deriving reliable (*i.e.* independent from the GPS sampling rate) and quantitative information about the components of animal movement. Results show that combining the parameters is a key component in modeling movement, while allowing for an accurate description of the turning angles, home range size and census issues. The model also has practical applications for addressing ecological issues such as census or spatial anomalies. While we only focus on 5 animals to demonstrate the approach, the model is general and can be applied to any other species. We considered only one attractor per animal in the proposed approach and both the existence and influence of multiple attractors are

yet to be investigated. Additional behavior features such as the spatial reinforcement, memory of n previous steps, activity rhythms (such as the circadian cycle), distance from the attractor, landscape/habitat effect (34), interactions with other animals and topological issues are currently being investigated and will be included in a future work. Another point of interest is the development of a continuous version of the proposed model, where the direction of the step is drawn from a specific distribution, whose parameters are yet to be empirically characterized.

ACKNOWLEDGEMENTS

We are very grateful to the French Biodiversity Office (Office Français de la Biodiversité) for the data. We also thank Sylvain Billard, Jean-Michel Gaillard and Clément Calenge for the discussions and Catherine Carter for her review of the manuscript. This work was supported by the Chair 'Modélisation Mathématique et Biodiversité' of Veolia Environnement-Ecole Polytechnique-Museum National d'Histoire Naturelle-Fondation X and by the ANR ABIM 26(ANR-16-CE40-0001) and ANR Mov-It (ANR-16-CE02-0010).

Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All applicable institutional and/or national guidelines for the care and use of animals were followed. The research program is hosted by the French National Hunting and Wildlife Agency (Office National de la Chasse et de la Faune Sauvage). This institution has granted all consents necessary for the fieldwork. Game captures were conducted in accordance with European and French laws. The experiment was designed to minimize animal stress and handling time and to ensure animal welfare, as defined in guidelines for the ethical use of animals in research. A specific accreditation was also delivered for capturing animals for scientific and wildlife management purposes. Animal captures and experimental procedures were in line with the French Environmental Code (Art. R421-15 to 421-31 and R422-92 to 422-94-1).

Bibliography

1. Graham H Pyke. Optimal foraging theory: a critical review. *Annual review of ecology and systematics*, 15(1):523–575, 1984.
2. S.M. Tomkiewicz, M.R. Fuller, J.G. Kie, and K.K. Bates. Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2163–2176, 2010. doi: 10.1098/rstb.2010.0090.
3. L. Börger, B.D. Dalziel, and J.M. Fryxell. Are there general mechanisms of animal home range behaviour? a review and prospects for future research. *Ecology Letters*, 11(6):637–650, 2008. ISSN 1461-0248. doi: 10.1111/j.1461-0248.2008.01182.x.
4. M. Hebblewhite and D.T. Haydon. Distinguishing technology from biology: a critical review of the use of gps telemetry data in ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2303–2312, 2010. doi: 10.1098/rstb.2010.0087.
5. B.T. McClintock, D.S. Johnson, M.B. Hooten, J.M. Ver Hoef, and J.M. Morales. When to be discrete: the importance of time formulation in understanding animal movement. *Movement ecology*, 2(1):21, 2014. doi: 10.1186/s40462-014-0021-6.
6. S.E. Hampton, C.A. Strasser, J.J. Tewksbury, W.K. Gram, A.E. Budden, A.L. Batcheller, C.S. Duke, and J.H. Porter. Big data and the future of ecology. *Frontiers in Ecology and the Environment*, 11(3):156–162, 2013. doi: 10.1890/120103.
7. U. Demšar, K. Buchin, F. Cagnacci, K. Safi, B. Speckmann, N. Van de Weghe, D. Weiskopf, and R. Weibel. Analysis and visualisation of movement: an interdisciplinary review. *Movement ecology*, 3(1):5, 2015. doi: 10.1186/s40462-015-0032-y.
8. G.M. Viswanathan, S.V. Buldyrev, S. Havlin, M.G.E. da Luzk, E.P. Raposok, and H.E. Stanley. Optimizing the success of random searches. *Nature*, 401:911–914, 1999. doi: 10.1038/44831.

9. S. Benhamou. How many animals really do the lévy walk? *Ecology*, 88:1962–1969, 2007. doi: 10.1890/06-1769.1.
10. S. Benhamou and J. Collet. Ultimate failure of the lévy foraging hypothesis: two-scale searching strategies outperform scale-free ones even when prey are scarce and cryptic. *Journal of theoretical biology*, 2015. doi: 10.1016/j.jtbi.2015.09.034.
11. G.H. Pyke. Understanding movements of organisms: it's time to abandon the lévy foraging hypothesis. *Methods in Ecology and Evolution*, 6(1):1–16, 2015. doi: 10.1111/2041-210X.12298.
12. Hervé Fritz, Sonia Said, and Henri Weimerskirch. Scale-dependent hierarchical adjustments of movement patterns in a long-range foraging seabird. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1520):1143–1148, 2003.
13. D.R. Brillinger, K.P. Haiganoush, A.A. Ager, J.G. Kie, and B.S. Stewart. Employing stochastic differential equations to model wildlife motion. *Bull Braz Math Soc*, 33:385–408, 2002. doi: 10.1007/s005740200021.
14. D.R. Brillinger. Simulating constrained animal motion using stochastic differential equations. *Lecture Notes-Monograph Series*, 41: pp. 35–48, 2003. ISSN 07492170.
15. H.K. Preisler, A.A. Ager, B.K. Johnson, and J.G. Kie. Modeling animal movements using stochastic differential equations. *Environmetrics*, 15:643–657, 2004. doi: 10.1002/env.636.
16. P.E. Smouse, S. Focardi, P.R. Moorcroft, J.G. Kie, J.D. Forester, and J.M. Morales. Stochastic modelling of animal movement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2201–2211, 2010. doi: 10.1098/rstb.2010.0078.
17. S. Benhamou. Spatial memory and searching efficiency. *Animal Behaviour*, 47:1423–1433, 1994. doi: 10.1006/anbe.1994.1189.
18. S. Benhamou. Detecting an orientation component in animal paths when the preferred direction is individual-dependent. *Ecology*, 87:518–528, 2006. doi: 10.1890/05-0495.
19. E.A. Codling, M.J. Plank, and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5:813–834, 2008. doi: 10.1098/rsif.2008.0014.
20. O. Ovaskainen. Habitat-specific movement parameters estimated using mark-recapture data and a diffusion model. *Ecology*, 85:242–257, 2004. doi: 10.1890/02-0706.
21. P.G. Blackwell, M. Niu, M.S. Lambert, and S.D. LaPoint. Exact bayesian inference for animal movement in continuous time. *Methods in Ecology and Evolution*, 7(2):184–195, 2016. ISSN 2041-210X. doi: 10.1111/2041-210X.12460.
22. A.O. Gaustestad and I. Mysterud. Intrinsic scaling complexity in animal dispersion and abundance. *The American Naturalist*, 165(1):44–55, 2005. doi: 10.1086/426673.
23. A.O. Gaustestad. Memory matters: influence from a cognitive map on animal space use. *Journal of Theoretical Biology*, 287:26–36, 2011. doi: 10.1016/j.jtbi.2011.07.010.
24. J.S. Horne, E.O. Garton, S.M. Krone, and J.S. Lewis. Analyzing animal movements using brownian bridges. *Ecology*, 88:2354–2363, 2007. doi: 10.1890/06-0957.1.
25. S. Benhamou. Dynamic approach to space and habitat use based on biased random bridges. *PLoS ONE*, 6:e14592, 2011. doi: 10.1371/journal.pone.0014592.
26. W.H. Burt. Territoriality and home range concepts as applied to mammals. *J. Mammal.*, 24: 346–352, 1943. doi: 10.2307/1374834.
27. P.R. Moorcroft and M.A. Lewis. *Mechanistic Home Range Analysis*. Princeton University Press, 2006. doi: 10.2193/2007-147.
28. J.M. Morales and S. Ellner. Scaling up animal movements in heterogeneous landscapes: the importance of behavior. *The Ecological Society of America*, 83:2240–2247, 2002. doi: 10.1890/0012-9658(2002)083[2240:SUAAMIH]2.0.CO;2.
29. J. Martin, V. Tolon, N. Morellet, H. Santin-Janin, A. Licoppe, C. Fischer, J. Bombois, P. Patthey, E. Pesenti, D. Chenesseau, et al. Common drivers of seasonal movements on the migration-residency behavior continuum in a large herbivore. *Scientific reports*, 8(1):7631, 2018. doi: 10.1038/s41598-018-25777-y.
30. M.E. Nelson. Experimental evidence of spatial memory and home range affinity in white-tailed deer (*odocoileus virginianus*). *The Canadian Field-Naturalist*, 129(1):1–7, 2015. doi: 10.22621/cfn.v129i1.1661.
31. W.F. Fagan, M.A. Lewis, M. Auger-Méthé, T. Avgar, S. Benhamou, G. Breed, L. LaDage, U.E. Schlägel, W. Tang, Y.P. Papastamatiou, et al. Spatial memory and animal movement. *Ecology letters*, 16(10):1316–1329, 2013. doi: 10.1111/ele.12165.
32. D. Boyer and C. Solis-Salas. Random walks with preferential relocations to places visited in the past and their application to biology. *arXiv preprint*, page arXiv:1403.6069, 2014. doi: 10.1103/PhysRevLett.112.240601.
33. Louise Riotte-Lambert and Jason Matthiopoulos. Environmental predictability as a cause and consequence of animal movement. *Trends in ecology & evolution*, 2019.
34. George Wittemyer, Joseph M Northrup, and Guillaume Bastille-Rousseau. Behavioural valuation of landscapes using movement data. *Philosophical Transactions of the Royal Society B*, 374(1781):20180046, 2019.
35. Gaudry William, Gaillard Jean-Michel, Saïd Sonia, Bonenfant Christophe, Mysterud Atle, Morellet Nicolas, Pellerin Maryline, and Calenge Clément. Same habitat types but different use: evidence of context-dependent habitat selection in roe deer across populations. *Scientific reports*, 8(1):1–13, 2018.
36. M. Lonergan, M. Fedak, and B. McConnell. The effects of interpolation error and location quality on animal track reconstruction. *Marine Mammal Science*, 25(2):275–282, 2009. doi: 10.1111/j.1748-7692.2008.00264.x.
37. Bram Van Moorter, Darcy Visscher, Simon Benhamou, Luca Börger, Mark S Boyce, and Jean-Michel Gaillard. Memory keeps you at home: a mechanistic model for home range emergence. *Oikos*, 118(5):641–652, 2009.
38. S. Mandujano and S. Gallina. Comparison of deer censusing methods in tropical dry forest. *Wildlife society bulletin*, pages 180–186, 1995.
39. K.E. Ironside, D.J. Mattson, T.R. Arundel, and J.R. Hansen. Is gps telemetry location error screening beneficial? *Wildlife Biology*, pages wlb-00229, 2017. doi: 10.2981/wlb.00229.
40. Nathanael I Lichti and Robert K Swihart. Estimating utilization distributions with kernel versus local convex hull methods. *The Journal of Wildlife Management*, 75(2):413–422, 2011.
41. Z.I. Botev, J.F. Grotowski, and D.P. Kroese. Kernel density estimation via diffusion. *Annals of Statistics*, 38:2916–2957, 2010. doi: 10.1214/10-AOS799.
42. D-Y Yang, Yu A Makhnovskii, S-Y Sheu, and SH Lin. Simulation of the wiener sausage. *Physical Review E*, 62(3):3116, 2000.
43. H. Edelhoff, J. Signer, and N. Balkenhol. Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns. *Movement ecology*, 4(1):21, 2016. doi: 10.1186/s40462-016-0086-5.
44. C.M. Postlethwaite and T.E. Dennis. Effects of temporal resolution on an inferential model of animal movement. *PLoS ONE*, 8(5):e57640, 05 2013. doi: 10.1371/journal.pone.0057640.
45. D. Fortin, H.L. Beyer, M.S. Boyce, D.W. Smith, T. Duchesne, and J.S. Mao. Wolves influence elk movements: behavior shapes a trophic cascade in yellowstone national park. *Ecology*, 86(5):1320–1330, 2005. doi: 10.1890/04-0953.
46. C.M. Bergman, J.A. Schaefer, and S.N. Luttich. Caribou movement as a correlated random walk. *Oecologia*, 123(3):364–374, 2000. doi: 10.1007/s004420051023.
47. P. Marchand, M. Gareil, G. Bourgoïn, D. Dubray, D. Maillard, and A. Loison. Coupling scale-specific habitat selection and activity reveals sex-specific food/cover trade-offs in a large herbivore. *Animal behaviour*, 102:169–187, 2015. doi: 10.1016/j.anbehav.2015.01.011.
48. J.R. Potts, K. Mokross, P.C. Stouffer, and M.A. Lewis. Step selection techniques uncover the environmental predictors of space use patterns in flocks of amazonian birds. *Ecology and evolution*, 4(24):4578–4588, 2014. doi: 10.1002/ece3.1306.
49. Amy Hurford. Gps measurement error gives rise to spurious 180 turning angles and strong directional biases in animal movement data. *PLoS One*, 4(5), 2009.
50. C. Calenge. The package "adehabitat" for the r software: a tool for the analysis of space and habitat use by animals. *Ecological modelling*, 197(3):516–519, 2006. doi: /10.1016/j.ecolmodel.2006.03.017.
51. J.E. Dunn and P.S. Gipson. Analysis of radio telemetry data in studies of home range. *Biometrics*, pages 85–101, 1977. doi: 10.2307/2529305.
52. P.G. Blackwell. Random diffusion models for animal movement. *Ecological Modelling*, 100(1):87–102, 1997. doi: /10.1016/S0304-3800(97)00153-1.
53. W.H. Piper. Making habitat selection more 'familiar': a review. *Behavioral Ecology and Sociobiology*, 65(7):1329–1351, 2011. doi: 10.1007/s00265-011-1195-1.
54. L.D. Howery, D.W. Bailey, and E.A. Laca. Impact of spatial memory on habitat use. *Grazing behavior of livestock and wildlife*, 70:91–100, 1999.
55. A.O. Gaustestad, L.E. Loe, and A. Mysterud. Inferring spatial memory and spatiotemporal scaling from gps data: comparing red deer cervus elaphus movements with simulation models. *Journal of Animal Ecology*, 2013. doi: 10.1111/1365-2656.12027.
56. T. Avgar, J.A. Baker, G.S. Brown, J.S. Hagens, A.M. Kittle, E.E. Mallon, M.T. McGreer, A. Mosser, S.G. Newmaster, B.R. Patterson, et al. Space-use behaviour of woodland caribou based on a cognitive movement model. *Journal of Animal Ecology*, 2015. doi: 10.1111/1365-2656.12357.
57. A.O. Gaustestad and A. Mysterud. The lévy flight foraging hypothesis: forgetting about memory may lead to false verification of brownian motion. *Movement ecology*, 1:1–9, 2013. doi: 10.1186/2051-3933-1-9.
58. David Storms, Philippe Aubry, Jean-Luc Hamann, Sonia Saïd, Hervé Fritz, Christine Saint-Andrieux, and François Klein. Seasonal variation in diet composition and similarity of sympatric red deer cervus elaphus and roe deer capreolus capreolus. *Wildlife Biology*, 14(2): 237–250, 2008.
59. P.J. Greenwood. Mating systems, philopatry and dispersal in birds and mammals. *Animal behaviour*, 28(4):1140–1162, 1980. doi: 10.1016/S0003-3472(80)80103-5.
60. P.G. Ryan, S.L. Petersen, G. Peters, and D. Grémillet. Gps tracking a marine predator: the effects of precision, resolution and sampling rate on foraging tracks of african penguins. *Marine Biology*, 145(2):215–223, 2004. doi: 10.1007/s00227-004-1328-4.
61. U.E. Schlägel and M.A. Lewis. Robustness of movement models: can models bridge the gap between temporal scales of data sets and behavioural processes? *Journal of mathematical biology*, 73(6-7):1691–1726, 2016. doi: 10.1007/s00285-016-1005-5.
62. U.E. Schlägel and M.A. Lewis. A framework for analyzing the robustness of movement models to variable step discretization. *Journal of mathematical biology*, 73(4):815–845, 2016. doi: 10.1007/s00285-016-0969-5.
63. K.J. Mills, B.R. Patterson, D.L. Murray, and J. Bowman. Effects of variable sampling frequencies on gps transmitter efficiency and estimated wolf home range size and movement distance. *Wildlife Society Bulletin*, 34(5):1463–1469, 2006. doi: 10.2193/0091-7648(2006)34[1463:EOVSFO]2.0.CO;2.