Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2014, **5,** 1236–1246 doi: 10.1111/2041-210X.12284

Detecting effects of spatial memory and dynamic information on animal movement decisions

Ulrike E. Schlägel* **and Mark A. Lewis**

Centre for Mathematical Biology, Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, Canada, T6G 2G1

Summary

1. Animals of many species demonstrate movement behaviour in which decisions are based on a variety of information. Effects of resources have been studied widely, often under the assumption that the environment is constant over the course of the study. Much less understood is the role of dynamic information about continuously changing resource availability and past experiences. Such information can be acquired during movement bouts and used for future decisions via memory.

2. We present a new class of animal movement models, which incorporates a dynamic interplay of movement and information gain processes. Information is contained in a dynamic cognitive map. As an example, we consider time since last visit to locations and how this interacts with environmental information to shape movement patterns. Our models can be fitted to empirical movement trajectories and are therefore amenable to statistical inference (parameter estimation and model selection).

3. We tested the functionality of our method using simulated data. Parameter estimates were in accordance with true values used in the simulations, and model selection via Bayesian information criterion (BIC) was able to identify true underlying mechanisms of simulated trajectories. Thus, if time since last visit to locations influences movement decisions, our method is able to detect this mechanism.

4. The use of dynamic information such as the one demonstrated in our example models likely requires cognitive abilities such as spatial memory. Therefore, our method can be used to reveal evidence of spatial memory in empirical movement data. Understanding the components of individual movement decisions and their interactions ultimately helps us to predict how population distribution patterns respond to environmental changes, such as landscape fragmentation and changing climate.

Key-words: animal movement, behaviour, cognitive map, information, mechanistic model, memory, resource selection, step selection function

Introduction

Animal movement serves important needs such as food acquisition, escape from predators and travel to reproduction sites. Consequently, many species have evolved capacities to move efficiently and purposefully by considering varying sources of information for their movement decisions (Janson & Byrne 2007; Sulikowski & Burke 2011). Explaining the mechanisms that underly such informed movement behaviour will allow us to better understand animal space-use patterns and their responses to environmental changes (Dalziel, Morales & Fryxell 2008; Nathan *et al.* 2008; Sutherland *et al.* 2013).

Most animals live in heterogenous environments, and the link between movement and environment has received much attention. Using classical resource-selection analyses (Manly *et al.* 2002), a wide range of studies have demonstrated that animals selectively use the biotic and abiotic features that are available to them (Fortin *et al.* 2005; Gillies, Beyer & St Clair 2011; Squires *et al.* 2013). Analyses of movement characteristics have shown that animals express different movement behaviours, for example encampment or travel, in different habitats (Morales *et al.* 2004; Forester *et al.* 2007).

Most mechanistic models have concentrated on incorporating relationships between environmental factors and movement behaviour within a static environment (but see Avgar, Deardon & Fryxell 2013); however, observations show that animals also take into account dynamically changing information and respond with their movements to temporal availability or unavailability of resources (Martin-Ordas *et al.* 2009). For instance, fruit-eating primates express goal-oriented travel towards those trees in their home range that carry ripe fruit (Asensio *et al.* 2011), and it has been suggested that monkeys use their daily travels to monitor fruiting histories of trees (Janson & Byrne 2007; Janmaat, Ban & Boesch 2013). On the other hand, many resources, once depleted, need some time before they become available again, providing reason for animals to avoid depleted food patches (Davies & Houston 1981; Owen-Smith, Fryxell & Merrill 2010; Bar-Shai, Keasar & Shmida 2011). Avoidance behaviour may be a response not only to depletion of resources, such as plant biomass or prey, but also to behavioural depression. Behavioural depression refers to a

---

*Correspondence author. E-mail: ulrike.schlaegel@gmail.com

reduction in prey availability that is caused by behavioural changes of the prey in response to predation (Charnov, Orians & Hyatt 1976). For example, prey may show greater alertness or seek shelter. This reduces capture rates, to which predators may respond in turn by changing their hunting areas (Jedrzejewski *et al.* 2001; Amano & Katayama 2009). Temporal considerations also become important for movement decisions if territorial defence mechanisms require animals to visit certain locations regularly, for example to scent-mark territory boundaries (Moorcroft & Barnett 2008; Giuggioli, Potts & Harris 2011).

As the above examples highlight, spatio-temporal information drives movement decisions and at the same time movement allows animals to update this information. Experimental findings additionally support that animals make decisions based on information that they have obtained through previous experiences. Memory of information about the 'what, where and when' of events, obtained through subjective experience, is termed 'www-memory' (Martin-Ordas *et al.* 2009) or 'episodic-like memory' (Griffiths, Dickinson & Clayton 1999). It is possible that animals acquire information about current environmental conditions through perceptual cues, even over large distances (Tsoar *et al.* 2011), and that information about the recent travel history is stored in externalized 'memory', such as pheromone trails or slime (Deneubourg *et al.* 1989; Reid *et al.* 2012). However, it is likely that many animals draw upon internal memory, especially for behaviours that require information about temporal distances ('how long ago?') (Griffiths, Dickinson & Clayton 1999; Martin-Ordas *et al.* 2009; Janmaat, Ban & Boesch 2013). During recent years, movement models have started to incorporate influences of memorized information on movement decisions (for a review, see Fagan *et al.* 2013). Most of these are simulation models that are used for theoretical considerations only (but see Avgar, Deardon & Fryxell 2013); however, to test our understanding of the feedbacks between movement and information acquisition, we must also interface memory-based models with data (Smouse *et al.* 2010).

Here, we present a new model for animal movement that is amenable to likelihood-based inference, and in which we mechanistically incorporate the interplay of movement decisions, environmental information and dynamically changing temporal information. Our model is similar in its form to recent spatially explicit resource-selection models (e.g. Rhodes *et al.* 2005; Forester, Im & Rathouz 2009), in which movement steps are assigned probabilities based on general movement tendencies and resource preferences. In previous models, resource information enters as a static covariate, providing knowledge about features of the landscape, such as land cover type or topographical features. In our model, we add dynamic information obtained through experiences made during movement. To realize the interplay of movement and information acquisition in our model, we draw on the concept of a cognitive map (Tolman 1948; Asensio *et al.* 2011). We use this concept here as a helpful mathematical construct that provides a map-like representation of the animal's environment containing all relevant information. For an example of a dynamic information gain process, we introduce information about the time since last visit to locations. Time since last visit is useful information that can play a role, for example, in the process of patrolling in canids or food acquisition across species if food availability varies (Davies & Houston 1981). With the inclusion of this information acquisition process, we present a practical model that incorporates both dynamic information and spatial memory.

We place our model into a model selection framework that allows us to identify which types of information most likely shape the movement decision process. We first outline the general formulation of our model and how memory effects can be integrated. Subsequently, we present the details of several candidate models that correspond to different underlying mechanisms of animal movement behaviour. Next, we show how the models can be fitted to empirical movement trajectories to perform statistical inference. Finally, using simulated data, we test the functionality of our framework and assess whether our method can correctly detect effects of static resource information and dynamically changing temporal information and whether we can estimate model parameters reliably.

Materials and methods

For several decades, the basis of many animal movement models have been random walks. In a classical random walk, movement is described as a series of discrete steps that have independent and identical probability distributions. This has been extended to include correlations between steps, biases towards specific locations, and step probabilities that depend on the behavioural state of the individual (Morales *et al.* 2004; Breed *et al.* 2012; McClintock *et al.* 2012; Langrock *et al.* 2013). Random walks and their extensions have been used both to analyse movement behaviour at an individual level (Lagrangian approach; e.g. Smouse *et al.* 2010) and to derive partial differential equation models that describe spatio-temporal patterns at a population level or expected space use of individuals (Eulerian approach; e.g. Codling, Plank & Benhamou 2008).

We are interested in understanding decision processes that underly movement behaviour on the scale of individuals. We draw upon a modelling framework that bridges the gap between statistical resource-selection analysis and spatially explicit movement models (Rhodes *et al.* 2005; Moorcroft & Barnett 2008; Forester, Im & Rathouz 2009). The framework builds on a random walk and defines movement via step probabilities, which have two components. A resource-independent movement kernel assigns probabilities to steps based on the animal's general movement tendencies. Given this, a weighting function evaluates the attractiveness of steps according to resource availability and resource preferences. We extend this framework by generalizing the weighting function. In our generalization, the weighting function does not only describe the influence of resources but allows for the inclusion of any information relevant to the animal. Information can pertain to landscape features and resources, as in previous models, but also to memories of past events and timing aspects, which cannot be obtained externally but only through the movement process and the animal's behaviour itself. We assume that information at a given time is either obtained through direct perception or retrieved from the animal's cognitive map (i.e. memory), which itself is updated through experience. In our model, the cognitive map is a function that assigns values to locations according to their information content at a given time. Thus, it serves as a mathematical tool without the claim that it truly represents the underlying cognitive mechanism. With the framework of the cognitive map, we provide a general method for including an explicit information acquisition process. The cognitive map itself can take many forms, depending on the species and behaviour of interest. In our candidate models, we demonstrate examples of types of information the cognitive map may contain.

THE MODELLING FRAMEWORK

We consider movement paths of individual animals, and we assume that an individual's trajectory consists of a series of locations $(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)$ at regular times $T = \{1,\ldots,N\}$. Each location has an Easting and a Northing in two-dimensional space, which is discretized into a regular grid of square cells. The resolution of the spatial discretization depends on the available environmental data and should be fine enough compared to the animal's movement such that steps generally range over multiple cells.

We model movement as a stochastic process, where the probability of making a step to location $\boldsymbol{x}_t$ depends on the location at time $t-1$ and, if movement is persistent, on the previous step from $\boldsymbol{x}_{t-2}$ to $\boldsymbol{x}_{t-1}$. We define this step probability as

$$p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{x}_{t-2},\boldsymbol{\theta}) = \frac{k(\boldsymbol{x}_t;\boldsymbol{x}_{t-1},\boldsymbol{x}_{t-2},\boldsymbol{\theta}_1)w_t(\boldsymbol{x}_t;\boldsymbol{\theta}_2)}{\sum_{\boldsymbol{y}\in\Omega}k(\boldsymbol{y};\boldsymbol{x}_{t-1},\boldsymbol{x}_{t-2},\boldsymbol{\theta}_1)w_t(\boldsymbol{y};\boldsymbol{\theta}_2)}, \qquad \text{eqn 1}$$

where k is an information-independent movement kernel, w_t is an information-based weighting function, and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1,\boldsymbol{\theta}_2)$ is a collection of model parameters. The sum in the denominator ensures that p is an appropriately normalized probability mass function over space. The spatial domain Ω is the area within which the animal can choose to travel during the time relevant to the study.

Using the conceptual framework of Nathan *et al.* (2008), we can interpret the kernel k as describing the animal's motion capacity and w_t as formulating the influence of external factors, to which we add memorized information. Both k and w_t can be affected by the animal's internal goal. For instance, if a herbivore is foraging, it is likely that it moves slowly, changes its movement direction frequently and generally stays in an environment with suitable foraging material. It may additionally prefer to forage in an area with low predation risk. Such behaviour could be implemented by a kernel that assigns higher probabilities to locations in the animal's close vicinity with the same values in all directions and a weighting function that has highest values in preferred foraging habitat. The weighting function could also include information about previously experienced presence of predators (Latombe, Fortin & Parrott 2014).

In general, the movement kernel k can be very simple, for example constant within the animal's maximum movement radius (Rhodes *et al.* 2005); however, we can also use a more complex kernel that accounts for persistence in movement direction or biases towards specific locations (Moorcroft & Lewis 2006). Directions can be measured by either turning angles (the angles between successive steps) or bearings (the angles of steps with respect to a fixed direction, e.g. North).

We model the weighting function w_t as a resource-selection function (Manly *et al.* 2002; Lele & Keim 2006). There are several forms available for resource-selection functions, and here we present the logistic form,

$$w_t(\boldsymbol{x};\alpha,\boldsymbol{\beta},\boldsymbol{\gamma}) = \left[1 + \exp\left(-\alpha - \mathcal{I}_t(\boldsymbol{x})\cdot\boldsymbol{\beta} - f(\mathcal{I}_t(\boldsymbol{x}),\boldsymbol{\gamma})\right)\right]^{-1}, \qquad \text{eqn 2}$$

where $\cdot$ denotes the dot product of two vectors. The vector $\mathcal{I}_t(\boldsymbol{x}) \in \mathbb{R}^n$ is the cognitive map content at location $\boldsymbol{x}$ at time t containing the values of all information variables of location $\boldsymbol{x}$ at time t, and $\boldsymbol{\beta} \in \mathbb{R}^n$ is a parameter vector describing the animal's preference for a location of type $\mathcal{I}_t(\boldsymbol{x})$. The intercept $\alpha \in \mathbb{R}$ determines the baseline weight of a location when all information variables are zero. The function f and parameter vector $\boldsymbol{\gamma}$ account for possible interactions between different information variables. Locations with preferred features have high weights, thereby increasing the chance that an animal will visit those. The logistic form of the weighting function restricts weights to be between zero and one, and therefore the weighting function can in fact be viewed as a resource-selection probability function (Lele & Keim 2006).

Because of the dependence structure of the step probabilities in eqn 1, they are only valid for times $t \geq 3$. Here, we chose to define an initial probability for the first two locations, $p(\boldsymbol{x}_1, \boldsymbol{x}_2|\boldsymbol{\theta}) = p(\boldsymbol{x}_2|\boldsymbol{x}_1,\boldsymbol{\theta})p(\boldsymbol{x}_1|\boldsymbol{\theta})$. A simple option is to assume that every location in the spatial domain has the same probability to be the first location, $p(\boldsymbol{x}_1|\boldsymbol{\theta}) = \frac{1}{|\Omega|}$, and to let

$$p(\boldsymbol{x}_2|\boldsymbol{x}_1,\boldsymbol{\theta}) = \frac{\tilde{k}(\boldsymbol{x}_2;\boldsymbol{x}_1,\kappa,\lambda)w_t(\boldsymbol{x}_2;\alpha,\boldsymbol{\beta},\boldsymbol{\gamma})}{\sum_{\boldsymbol{y}\in\Omega}\tilde{k}(\boldsymbol{y};\boldsymbol{x}_1,\kappa,\lambda)w_t(\boldsymbol{y};\alpha,\boldsymbol{\beta},\boldsymbol{\gamma})}, \qquad \text{eqn 3}$$

where $\tilde{k}$ is possibly a simplified form of k in case that k describes persistent movement.

CANDIDATE MODELS

We consider four different models that represent biological hypotheses about the types of information that an individual may consider for making movement decisions. In the simplest case, the *null model*, we assume that the animal considers no specific information. In the *resource model*, an individual considers static information about the environment, where 'static' means that the information content remains constant over the time span of the analysis. Information can be given about any resources pertaining to the animal, for example any variables as they are typical in resource-selection analyses. To include dynamically changing information, we allow information, and thereby the weighting function, to change through time. If information was only given externally, this would constitute a dynamic version of the resource model. However, our aim is to model a dynamic interplay of movement decisions and information content. In the *memory model*, we therefore introduce *time since last visit* as new type of information. To account for the possibility that both resources and the dynamic variable time since last visit influence movement decisions simultaneously, we consider a *combination model* as the most complex model.

We implement the different models by varying the information variable $\mathcal{I}_t$ in the weighting function (eqn 2) while using the same movement kernel for all models. For example, trajectories demonstrating the different movement patterns resulting from the four candidate models, see Fig. 1 and animations in Appendix S3.

Null model

In the null model, we assume that the information content of all locations is zero. Therefore, the weighting function is homogeneous across the landscape and constant over time, $w_t(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \Omega$, $t \in T$. This means that the animal moves only according to the kernel k.

Resource model

In the resource model, information is static and includes all resource variables of interest, $\mathcal{I}_t(\boldsymbol{x}) = \mathcal{I}(\boldsymbol{x}) = (r_1(\boldsymbol{x}), , r_n(\boldsymbol{x}))$ for every location $\boldsymbol{x} \in \Omega$. It is straightforward to extend this to dynamic resource

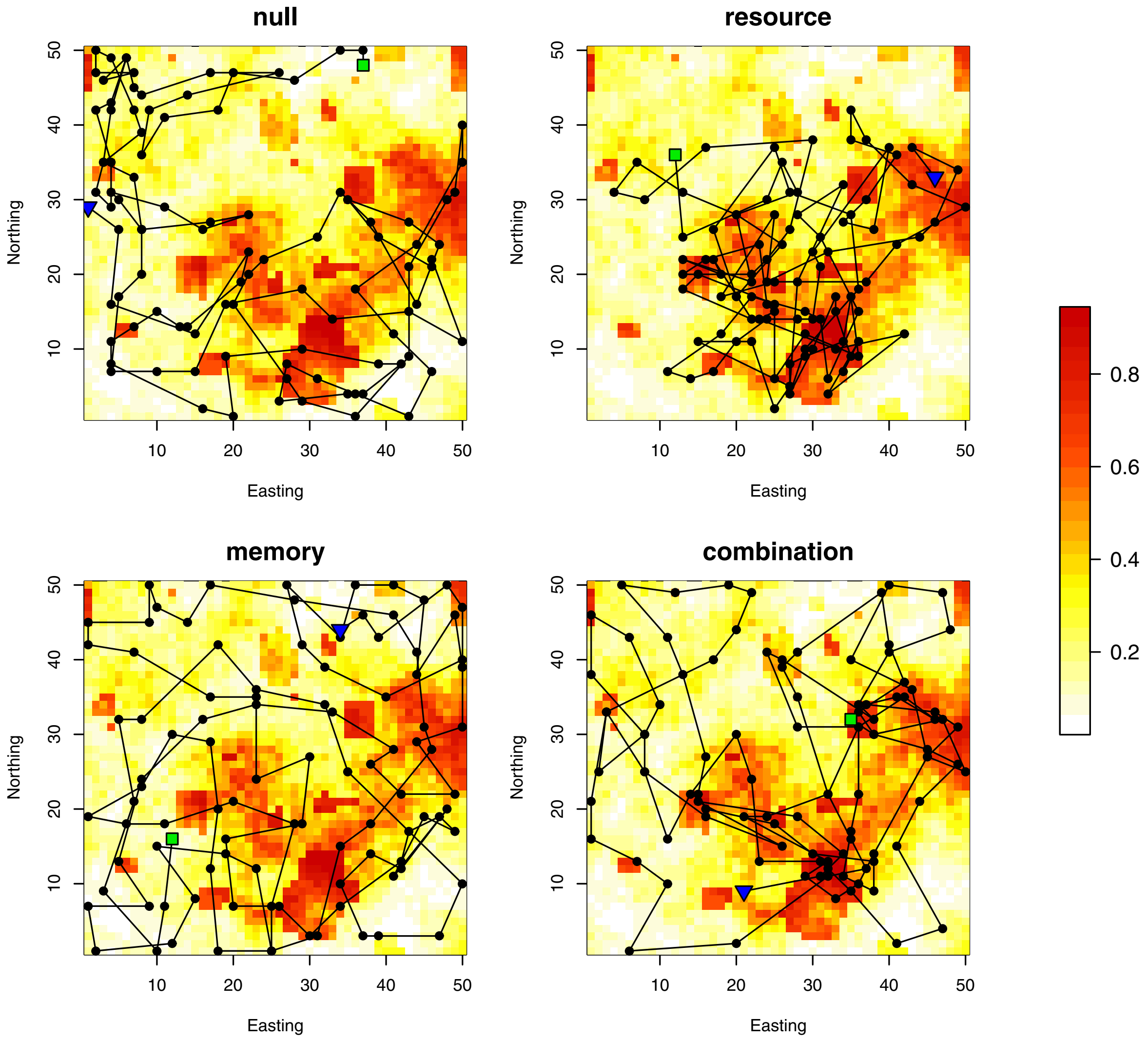

Fig. 1. Example trajectories from the four candidate models, 100 steps long, with starting location marked by a green box and final location marked by a blue triangle. All trajectories are plotted on top of an example resource-selection function $w(\boldsymbol{x};\alpha,\boldsymbol{\beta}) = \left[1+\exp\left(-\alpha-\beta_1 r_1(\boldsymbol{x})-\beta_2 r_2(\boldsymbol{x})\right)\right]^{-1}$ generated from two resources r_1 and r_2. The null model does not consider resource information, and therefore, the null trajectory visits locations irrespective of the resource-selection function. The memory model does not consider resource information either, however, the animal avoids recently visited locations and is attracted to locations with long time since last visit. Therefore, the memory trajectory efficiently explores the spatial domain in a patrolling fashion. In contrast, the resource trajectory mainly remains in areas where the resource selection function has high values. The combination trajectory shows a mixture of behaviours from the resource and the memory model. The trajectories were generated using the first landscape pair and main parameter set from the simulation study; compare Fig. 3 and Appendix S1.

information to include, for example, seasonal changes in the landscape or disturbance events.

Memory model

In the memory model, we assume that while the animal moves through the environment, it monitors the time since last visit from locations and uses this information for movement decisions. For instance, recently visited areas may be avoided for a period of time, whereas locations with long absence may be attractive. In our model, we include this feature by defining the cognitive map as $m_t : \Omega \to \mathbb{N}$, which at any time assigns values to all locations in the spatial domain based on the map values at the previous time and the most recent movement step. If the animal moves from location $\boldsymbol{x}_{t-1}$ to $\boldsymbol{x}_t$ between times $t-1$ and t, we define for any location y in the spatial domain

$$m_t(\boldsymbol{y}) = \begin{cases} 0, & \text{if } d(\boldsymbol{y},\boldsymbol{z}) \leq \delta \text{ for any } \boldsymbol{z} \in \text{path}(\boldsymbol{x}_{t-1} \to \boldsymbol{x}_t) \\ m_{t-1}(\boldsymbol{y}) + 1, & \text{otherwise.} \end{cases} \qquad \text{eqn 4}$$

Because of our spatial discretization, we use $d(\boldsymbol{y},\boldsymbol{z}) = |y_E - z_E| + |y_N - z_N|$ as the distance between two locations $\boldsymbol{y}$, $\boldsymbol{z}$ with Easting and Northing $\boldsymbol{y} = (y_E, y_N)$ and $\boldsymbol{z} = (z_E, z_N)$, such that all locations within a distance δ of a fixed location $\boldsymbol{z}$ form a diamond-shaped area around $\boldsymbol{z}$. We assume that path($\boldsymbol{x}_{t-1} \to \boldsymbol{x}_t$) is the straight line between $\boldsymbol{x}_{t-1}$ and $\boldsymbol{x}_t$. Via eqn 4, an individual counts the number of steps it remains absent from locations, and therefore, $m_t(\boldsymbol{x})$ is the time since last visit to location

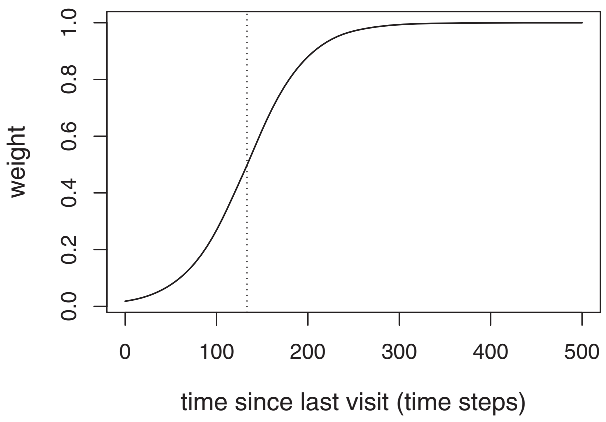

Fig. 2. In the memory model, the weight $w_t(\boldsymbol{x})$ of a location $\boldsymbol{x}$ depends on time since last visit $m_{t-1}(\boldsymbol{x})$ to that location. Locations that have been visited recently have low weights and are thus avoided. A weight of 0·5 is attained when $m_{t-1} = -\frac{\alpha}{\beta}$ (dotted vertical line).

$\boldsymbol{x}$ at time t. A location is considered visited when the animal comes within a distance $\delta > 0$. Because m_t is obtained recursively, we have to define appropriate starting values. Here, we use movement data prior to the trajectory $(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)$ for initialization. If a location $\boldsymbol{x}$ was visited during the initialization phase, we calculate the time between the last visit to this location and the beginning of our actual trajectory and thus reconstruct time since last visit at time $t = 1$, $m_1(\boldsymbol{x})$. For all locations not visited during initialization, we set time since last visit as the length of the initialization phase. The dynamic variable time since last visit is used in the memory model to inform movement decisions via $\mathcal{I}_t(\boldsymbol{x}) = m_{t-1}(\boldsymbol{x})$; compare also Fig. 2. Once $\boldsymbol{x}_t$ is chosen according to the probability mass function in eqn 1, m_t is updated via eqn 4. Here, we track time since last visit for the entire spatial domain Ω. If the selection coefficient with respect to $m_t(\boldsymbol{x})$ is positive, this leads to any location eventually becoming highly attractive after long enough absence. If this behaviour is not desired, one may adjust the definition of the cognitive map or weighting function appropriately. For example, if prior information about an animal's behaviour is given, it is possible to track time since last visit only for certain locations of specific interest.

Combination model

In the *combination model*, we allow information types from the resource and memory model to influence movement simultaneously by letting the information vector be $\mathcal{I}_t(\boldsymbol{x}) = (r_1(\boldsymbol{x}),\ldots,r_n(\boldsymbol{x}),m_{t-1}(\boldsymbol{x}))$. In particular, this models allows for interactive effects of time since last visit and resource variables, for example, by incorporating multiplicative terms of the form $\boldsymbol{\gamma}\boldsymbol{r}(\boldsymbol{x})m_{t-1}(\boldsymbol{x})$ into the interaction term $f(\mathcal{I}_t(\boldsymbol{x}),\boldsymbol{\gamma})$ in the weighting function (eqn 2). This is important in situations where return times to locations matter depending on the resources at the location, for example average return times to preferred foraging areas may differ from those to locations used as shelter.

Information-independent kernel

We define the movement kernel k based on a step length distribution with density $\mathcal{S}$ and a distribution for movement directions with density Φ. For step length, we use a Weibull distribution with scale and shape parameter $\kappa > 0$ and $\lambda > 0$, respectively, because it has a flexible form and generally shows a good fit with empirical data (Morales *et al.* 2004). Thus,

$$\mathcal{S}(\|\boldsymbol{x}_t-\boldsymbol{x}_{t-1}\|;\kappa,\lambda) = \frac{\kappa}{\lambda}\left(\frac{\|\boldsymbol{x}_t-\boldsymbol{x}_{t-1}\|}{\lambda}\right)^{\kappa-1}\exp\left(-\left(\frac{\|\boldsymbol{x}_t-\boldsymbol{x}_{t-1}\|}{\lambda}\right)^{\kappa}\right). \qquad \text{eqn 5}$$

To measure movement directions, we use bearings, and we denote the bearing of the step from $\boldsymbol{x}_{t-1}$ to $\boldsymbol{x}_t$ by $\phi(\boldsymbol{x}_{t-1},\boldsymbol{x}_t) \in [-\pi,\pi)$. We include directional persistence by choosing a wrapped Cauchy distribution for bearings with scale parameter $\rho > 0$ and mode at the previous step's bearing $\phi(\boldsymbol{x}_{t-2},\boldsymbol{x}_{t-1})$,

$$\Phi(\varphi(\boldsymbol{x}_{t-1},\boldsymbol{x}_t);\varphi(\boldsymbol{x}_{t-2},\boldsymbol{x}_{t-1}),\rho) = \frac{1}{2\pi}\frac{\sinh\rho}{\cosh\rho-\cos(\varphi(\boldsymbol{x}_{t-1},\boldsymbol{x}_t)-\varphi(\boldsymbol{x}_{t-2},\boldsymbol{x}_{t-1}))}. \qquad \text{eqn 6}$$

The wrapped Cauchy distribution is convenient for implementation, and it has been used commonly to model movement directions (Morales *et al.* 2004; Codling, Plank & Benhamou 2008, note that eqn 6 is equivalent to their formula with parameter transformation $r = \exp(-\rho)$). One could use alternative distributions, such as the von Mises distribution or wrapped normal distribution (Codling, Plank & Benhamou 2008). Assuming that the choices for step length and movement direction are independent, the kernel becomes the product of $\mathcal{S}$ and Φ, describing a correlated random walk,

$$k(\boldsymbol{x}_t;\boldsymbol{x}_{t-1},\boldsymbol{x}_{t-2},\kappa,\lambda,\rho) = \mathcal{S}(\|\boldsymbol{x}_t-\boldsymbol{x}_{t-1}\|;\kappa,\lambda)\Phi(\varphi(\boldsymbol{x}_{t-1},\boldsymbol{x}_t);\varphi(\boldsymbol{x}_{t-2},\boldsymbol{x}_{t-1}),\rho). \qquad \text{eqn 7}$$

Because the kernel formulates persistent movement and takes into account the bearing of the previous step, we define a simplified kernel for $t = 2$ as

$$\tilde{k}(\boldsymbol{x}_2;\boldsymbol{x}_1,\kappa,\lambda) = \frac{1}{2\pi}\frac{\kappa}{\lambda}\left(\frac{\|\boldsymbol{x}_t-\boldsymbol{x}_{t-1}\|}{\lambda}\right)^{\kappa-1}\exp\left(-\left(\frac{\|\boldsymbol{x}_t-\boldsymbol{x}_{t-1}\|}{\lambda}\right)^{\kappa}\right). \qquad \text{eqn 8}$$

This means that we assume a uniform distribution for the first bearing.

Note that this definition of the movement kernel from step length and bearing distributions does not mean that we obtain the kernel from empirical step lengths and bearings in advance and then use this observed kernel to estimate the weighting function parameters in a case–control study, as has been previously suggested for resource-selection analysis (Fortin *et al.* 2005; Forester, Im & Rathouz 2009). Because movement and resource selection are not independent processes, a decoupled treatment of the processes can lead to biased estimates. We circumvent this problem, and we use the formulation in terms of step length and bearing only to define the functional form of the information-independent movement kernel. During model fitting (see next section), we estimate all model parameters simultaneously from the data.

STATISTICAL INFERENCE

If information $\mathcal{I}_t$ is known, the likelihood function for the collection of parameters $\boldsymbol{\theta} = (\kappa,\lambda,\rho,\alpha,\boldsymbol{\beta},\boldsymbol{\gamma})$ for the general model is

$$L(\boldsymbol{\theta}) = p(\boldsymbol{x}_1,\boldsymbol{\theta})p(\boldsymbol{x}_2|\boldsymbol{x}_1\boldsymbol{\theta}) \times \prod_{t=3}^{N} p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{x}_{t-2},\boldsymbol{\theta}). \qquad \text{eqn 9}$$

In the memory and combination model, $\mathcal{I}$ includes the variable time since last visit $\boldsymbol{m} = (m_t,\ t \in T)$, which represents internal information of the animal that in general cannot be observed. However, because of the way we define and initialize $\boldsymbol{m}$, we are able to iteratively calculate the time series ($m_1,\ldots,m_N$) based on the movement trajectory. Therefore, given the data ($\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N$), time since last visit becomes a known covariate, and the likelihood function in eqn 9 is valid for all models.

To obtain parameter estimates and their confidence intervals for all models, we use data cloning (Lele, Dennis & Lutscher 2007). Data cloning uses Markov Chain Monte Carlo (MCMC) methods, which are usually employed in Bayesian statistical inference. However, data cloning provides approximations to maximum-likelihood estimates (MLE), together with Wald-type confidence intervals, thus facilitating frequentist inference; see Appendix S2.

We use the approximate MLEs for the model parameters in eqn 9 to calculate the corresponding approximate maximum-likelihood values. From these, we obtain the Bayesian information criterion (BIC) for each of the four models (Burnham & Anderson 2002). Alternatively, we could have used Akaike information criterion (AIC); however for large data sets, AIC tends to favour overly complex models (Link & Barker 2006). For each trajectory, we select the model with smallest BIC as the one that explains the decision mechanism of the trajectory best. We use the BIC of this best model as a reference to calculate BIC differences for all alternative models ($\Delta\text{BIC} = \text{BIC}_{\text{alternative}} - \text{BIC}_{\text{best}}$).

A common problem in statistical inference are missed observations. Missed locations in an otherwise regular movement trajectory occur, for instance when GPS devices fail to acquire satellite signal due to closed canopy or otherwise limited available sky. In an autocorrelated trajectory, with each missed location we additionally lose associated information. Calculations of step lengths and bearings require two successive locations. In models with persistent movement, we require not only the current but also the previous bearing for step probabilities. Therefore, in a correlated random walk, one missed location can effectively lead to a gap of two full steps. In MCMC-based data cloning, we can treat missed locations explicitly as unknown variables and account for this in the likelihood function,

$$L(\boldsymbol{\theta}) = \int p(\boldsymbol{x}_{\text{avail}}, \boldsymbol{x}_{\text{miss}})d\boldsymbol{x}_{\text{miss}}. \qquad \text{eqn 10}$$

This allows to preserve the entire dependency structure of the trajectory and avoids the need to discard any information. For more information on this, see Appendix S1.

SIMULATION STUDY

To verify the functionality of our method, we applied the modelling framework and statistical inference method to simulated data. Because eqn 1 defines probability mass functions for movement steps, we can sample from them to iteratively generate individual movement trajectories according to the four candidate models. These data have the advantage that we know both a trajectory's underlying mechanism and the parameter values that were used to generate the trajectory. By applying our inference procedure to these data, we investigated whether we were able to identify the true underlying mechanism of a trajectory and whether we were able to correctly estimate parameter values.

Simulation of landscapes

Because movement decisions in the resource and combination model are based on environmental information, we first simulated landscapes of covariate values. We consider two resources (r_1,r_2), one having a continuous range of values, for example a biomass measure or elevation, and the other representing presence or absence of a feature, for example a preferred food source, via a binary variable that takes either value 1 or 0. To include biological realism, we accounted for spatial correlations in the covariate values. We simulated five pairs of landscapes with varying spatial structures. For more information, see Appendix S1.

Simulation of movement trajectories

We generated movement trajectories using the four candidate models presented above. When we used the null model, we called the resultant trajectory a *null trajectory*, and we named trajectories analogously for the other models.

On each of the five landscape pairs, we simulated a null, resource, memory and combination trajectory, using the same movement parameter values on all landscapes and across all four models, as applicable (Fig. 3). The kernel parameters κ, λ, ρ appear in all models. The resource model has additional parameters α_{res}, β_1, β_2, which are the intercept and the selection parameters with respect to the two resources (r_1,r_2) of the weighting function (eqn 2). In this model, we assumed there is no interaction between the two resources. The memory model instead has additional parameters α_{men} and β_{mem}, which describe the animal's preferences with respect to time since last visit $\boldsymbol{m}$. In the combination model, the weighting function includes all effects, such that it has parameters $\alpha_{\text{com}} = \alpha_{\text{res}} + \alpha_{\text{mem}}$, β_1, β_2 and β_{mem}. In this model, we further allowed for interactions between resources and time since last visit by defining the interaction term in the weighting function as $f\big((r_1, r_2, m_t), \gamma_1, \gamma_2\big) = \gamma_1 r_1 m_t + \gamma_2 r_2 m_t$, where γ_1 and γ_2 are the interaction parameters. We chose the main set of parameter values to represent realistic movement behaviour. To account for scenarios for which parameter values were potentially more difficult to estimate from data, for example small values of selection parameters, we generated two supplemental data sets, comprising two additional sets of 20 trajectories each generated from alternative sets of parameters; see Appendix S2.

For all trajectories, we simulated 2600 time steps, of which we discarded the first 1400 steps as initialization. This was particularly important for the memory model, in which we started with a cognitive map having value 0 everywhere. We used the last 400 steps from the initialization phase to calculate m_1. Each final trajectory consisted of 1200 time steps, which we considered a length commonly available (e.g. 1200 time steps could represent 50 days of 1-h data or 100 days of 2-h data).

For an example of how to handle missed observations, we simulated a combination trajectory with 90% fix rate by removing locations from a trajectory, 5% as single locations and 5% as two successive locations. We chose a trajectory from the main data set, which allowed us to compare results for completely and incompletely observed trajectory; see Appendix S2.

Analysis of simulated data

To every simulated trajectory, we fitted all four candidate models (Fig. 3) using data cloning. For details about the data cloning and MCMC procedures, such as number of clones and iterations used, see Appendix S1. There were two basic types of model fits that we distinguished in our analyses. A model could be fitted to a matching trajectory, that is a trajectory that had been simulated using the same model's mechanism (e.g. a resource model fitted to a resource trajectory). Or, a model could be fitted to a non-matching trajectory (e.g. a resource model fitted to a null, memory or combination trajectory). Each model fit generated estimates of the model parameters, together with Wald-type confidence intervals. Here, we used 95% confidence intervals. Using the

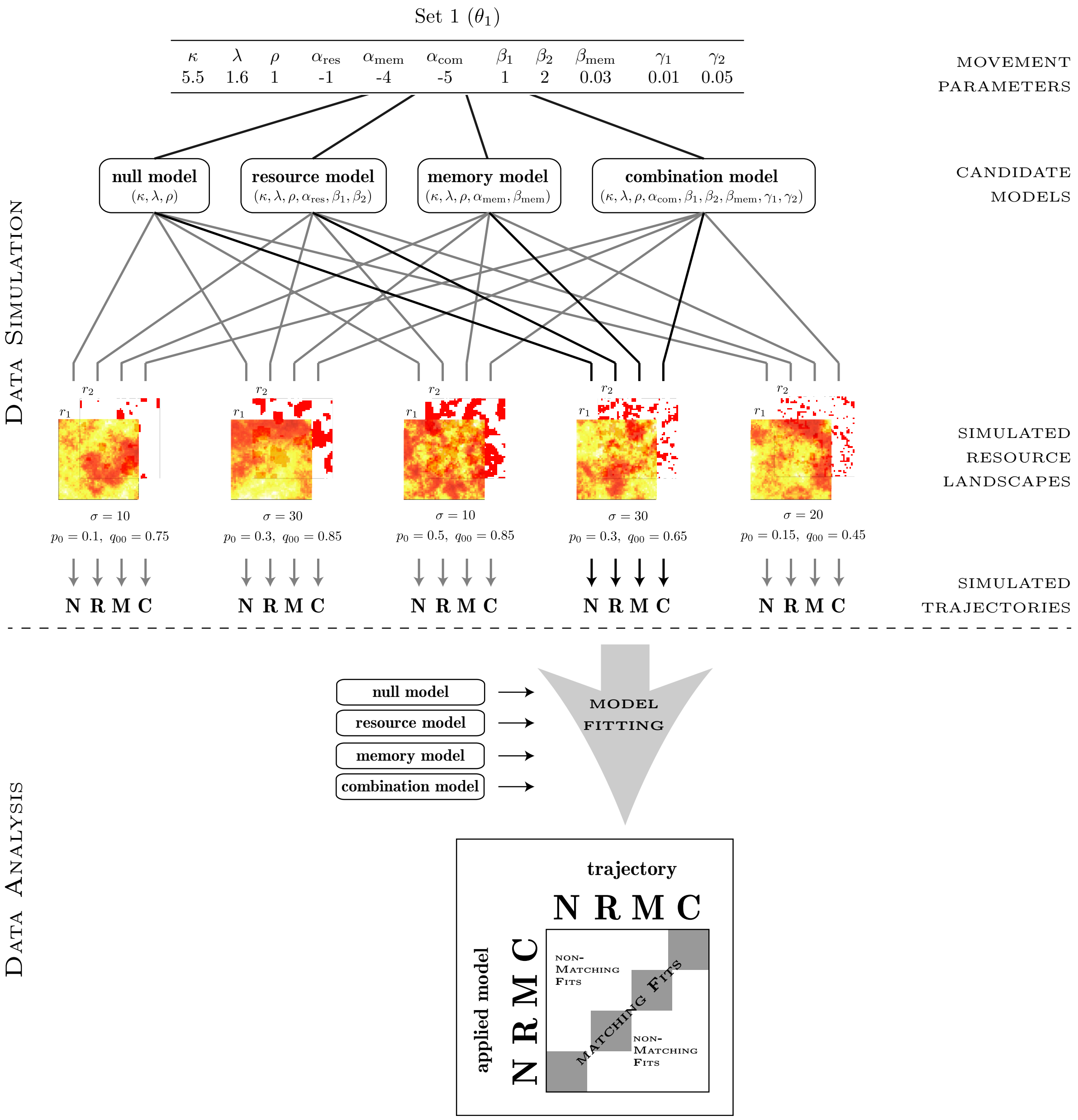

Fig. 3. Overview of data simulation and model fitting. For a set of parameter values, we generated trajectories using all four candidate models. Using each model, we simulated trajectories on five different landscapes resulting in 20 trajectories. Each trajectory was then fitted with all four models, leading to a total of 80 model fits.

approximate maximum-likelihood parameter estimates from data cloning, we estimated the maximum value of the model likelihood, and BIC, for each model fit.

For all parameter estimates, we examined whether their potential scale reduction factors $\hat{R}$ were close to 1 (Gelman & Rubin 1992). For an MCMC fit, in which parallel Markov chains are used to generate the posterior distribution of a parameter, the potential scale reduction factor of a parameter indicates whether the chains have mixed well and converged. If this is not the case, the estimate that results from such an MCMC is not meaningful. We considered a potential scale reduction factor $0 \cdot 9 \leq \hat{R} \leq 1$ to be sufficiently close to 1 (Gelman & Rubin 1992), and we excluded all parameter estimates that did not meet this condition from our analysis. Whenever such a non-convergent or non-mixing parameter occurred within a model fit, the resultant likelihood and BIC values of the fit were possibly inaccurate. Therefore, if a model fit included one or more parameters with $\hat{R} < 0 \cdot 9$ or $\hat{R} > 1 \cdot 1$, we excluded the BIC value from our model selection analysis.

For each trajectory, we compared whether the best model according to model selection via BIC coincided with the true underlying model of the trajectory. Under the assumption that our framework is functional, we expected the model that matched a trajectory's underlying mechanism to have minimal BIC. Because both the resource model and the memory model are nested within the combination model, we further expected the combination model, when applied to a resource or memory trajectory, to perform better than the simple alternative (e.g. a memory model applied to a resource trajectory).

For matching model fits, we compared true parameter values that were used to generate a trajectory to the parameter estimates obtained from applying the matching model, and we examined whether 95%

confidence intervals of parameters included the true values. This should be achieved 95% of the time if parameters are identifiable and our statistical methodology is functional.

In resource-selection analysis, it is usual to use hypothesis testing to determine whether a covariate has an effect or not. We performed an equivalent analysis and examined confidence intervals of the selection parameters β_1, β_2, β_{mem}, γ_1, γ_2 in those model fits, in which the combination model was fitted to a trajectory. The combination model includes all possible covariates, but not all covariates were simulated to have an effect in all trajectories, for example a resource trajectory includes effects of the resource variables but not time since last visit. Confidence intervals that corresponded to true underlying effects should exclude zero and vice versa. However, by definition, an α-level hypothesis test results in a Type I error of α, which we expected to observe approximately in this analysis. Additionally, we expected a Type II error to occur, where a confidence interval included zero although the corresponding covariate had an effect. We compared the outcome of this method with the results from model selection via BIC.

We performed all simulations of movement trajectories and statistical analyses in R (R Core team, 2013), using additionally package 'dclone' (Solymos 2010). To generate MCMC samples, we used JAGS via the R package 'rjags' (Plummer 2013).

RESULTS

Here, we present results for data generated with the main set of parameters $\boldsymbol{\theta}_1$ (Fig. 3). Results for supplemental data generated by additional sets of parameters can be found in Appendix S2.

Of all 80 model fits (four models fitted to 20 trajectories each), 80% had potential scale reduction factor $0 \cdot 9 \leq \hat{R} \leq 1 \cdot 1$ for all model parameters. In the remaining model fits, at least one parameter had $\hat{R} > 1 \cdot 1$ (Fig. 4). Convergence or mixing problems never occurred when the null model was fitted to a trajectory, even if the trajectory had a more complex underlying mechanism. Large $\hat{R}$ values only occurred if the fitted model contained parameters that were inapplicable to the model that was used to generate the trajectory. This was the case when any of the more complex models was fitted to a null trajectory, when the combination model was fitted to a resource or memory trajectory, or when the memory model was fitted to a resource trajectory and vice versa. In these model fits, the non-convergent parameters were mainly those that did not correspond to true underlying covariate effects. However, when in a model fit problems occurred for multiple parameters, occasionally even applicable parameters failed to converge. In matching model fits, Markov chains always mixed well and converged. For more details on convergence, see Appendix S2.

Our model selection framework was able to correctly identify the true underlying model for all trajectories (Fig. 4). When a trajectory had underlying resource or memory mechanism, the next best model was always the combination model with ΔBIC being a magnitude smaller than for the alternatives. This pattern was only disturbed if the combination model experienced convergence problems and was therefore excluded from further analysis.

Parameter estimates in matching model fits agreed well with true underlying parameter values. Parameter estimates generally were both close to and balanced around their true values (Fig. 5). The 95% confidence intervals (n = 115) included the true parameter value 91% of the time. If we also considered results from the supplemental data, 94% of all confidence intervals (n = 345) included the true value.

Our hypothesis test as to whether covariates had an effect agreed with our expectations. The combination model fitted to the 20 trajectories lead to 73 estimates of selection parameters, of which 39 corresponded to true underlying effects. Analysing their confidence intervals, we obtained a false-positive rate (Type I error rate) of 0·09 and a false-negative rate of zero, that is Type II errors did not occur. However, if we also considered the supplemental data and thereby increased the amount of resultant parameter estimates with confidence intervals to a total of 217, we obtained a Type I error rate of 0·04 and again a Type II error rate of 0·09.

Discussion

In recent years, the link between animal movement and spatial memory has received increasing attention (Smouse *et al.* 2010; Fagan *et al.* 2013). Studies of animal behaviour and cognition have given useful insights into animals' capacities to remember past experiences and use spatial memory. Most results have been obtained through experiments in confined and synthetic settings. However, to better understand how important ecological processes such as movement and dispersal are shaped by cognitive processes and memory, we also need to look at animals in their natural environments (Tsoar *et al.* 2011). Understanding the components of individual movement decisions and their interactions ultimately will help us to predict how population distribution patterns respond to environmental changes, such as landscape fragmentation and changing climate.

We have presented a modelling framework that can be used to detect the influence of memorized information on move-

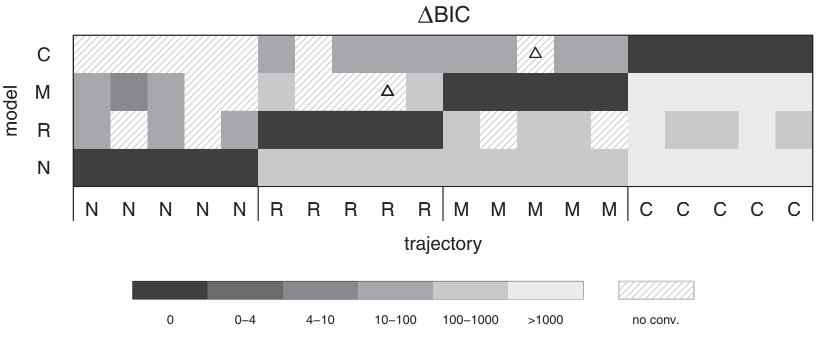

Fig. 4. Each column shows model selection results for one simulated trajectory when fitted with the null (N), resource (R), memory (M) and combination (C) model. For each trajectory, we calculated BIC values for the four fitted models, and the figure shows differences in BIC with respect to the minimal BIC value, that is the model with minimal BIC has ΔBIC = 0. We excluded model fits with non-convergent MCMC. Triangles indicate trajectories for which we calculated estimability diagnostics; Appendix S2.

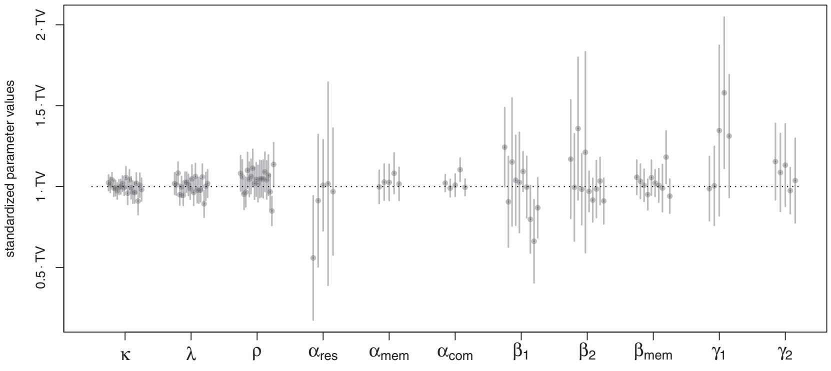

Fig. 5. Parameter estimates and their 95% confidence intervals for matching model fits (each trajectory fitted with the same model that was used to generate the trajectory). Both parameter estimates and Wald-type confidence intervals are scaled by the true parameter values (TV): $\kappa = 5{\cdot}5$, $\lambda = 1{\cdot}6$, $\rho = 1$, $\alpha_{\mathrm{res}} = -1$, $\alpha_{\mathrm{mem}} = -4$, $\alpha_{\mathrm{com}} = -5$, $\beta_1 = 1$, $\beta_2 = 2$, $\beta_{\mathrm{mem}} = 0{\cdot}03$, $\gamma_1 = 0{\cdot}01$, $\gamma_2 = 0{\cdot}05$.

ment decisions. We recognize that in many situations, it is difficult to confirm that animals draw upon memorized information instead of momentarily perceived information; however, there is evidence that animals use information that they have obtained during past experiences (Martin-Ordas *et al.* 2009; Janmaat, Ban & Boesch 2013). As an example of such information, we use time since last visit to locations. In our model, time since last visit is continuously updated during the movement process and at the same time influences movement decisions. We formulate our models in a way that makes them amenable to likelihood-based statistical inference. This allows us to fit our models to data to test whether the timing of events plays a role for movement decisions. Fitting the full model (eqn 1), encompassing both general movement tendencies and selective behaviour, to data via the likelihood function (eqn 9) enables simultaneous estimation of parameters of both the general movement kernel and weighting function. This distinguishes our method from step selection approaches that use an empirical movement kernel to estimate resource-selection parameters in a case–control framework (Fortin *et al.* 2005; Forester, Im & Rathouz 2009).

In our definition of the weighting function (eqn 1), we followed the classical formulation of resource-selection functions and evaluated a movement step based on the information at the endpoint of the step. In the memory model, this means that an animal may cross recently visited locations on its path although these have low weights. Depending on the behaviour of the study species, it may be appropriate to change this so that cognitive map values along the entire path are considered, thus following the idea of step selection functions (Fortin *et al.* 2005; Potts *et al.* 2014). In our framework, it is straightforward to define the weighting function as a function not only of $\boldsymbol{x}_t$ but also $\boldsymbol{x}_{t-1}$ and to include any information related to the step from $\boldsymbol{x}_{t-1}$ to $\boldsymbol{x}_t$. Endpoints are observed locations and therefore have certainly be used. To include information about entire steps, we must make an assumption about which locations were visited between observed locations. In the memory model, we assume this is a straight line; however, one may use more sophisticated methods similar to Brownian Bridges (Horne *et al.* 2007).

We used simulated landscapes and movement data to verify the functionality of our modelling framework and statistical inference method. Adding the memory process to the modelling framework considerably increased model complexity and the amount of data that had to be processed. We were therefore interested in whether we could correctly detect memory effects in empirical movement patterns and whether parameters that describe the memory process and its interactions with other variables were possible to estimate reliably. To perform inference, we used data cloning, which uses MCMC techniques but facilitates frequentist inference. We used the software package JAGS, which allowed us to define models in an easily understandable language and provides a stable implementation of MCMC sampling. JAGS was able to adapt the sampling process successfully so that parameters of very different magnitude could be reliably estimated. However, this came at the cost of long computation times (ranging 0·5–5 days per single chain for different models) and high memory needs (ranging 1–5 GB RAM). Alternatively, we could have used conventional numerical maximization of the likelihood function, which in this case may have been faster but at the same time more limited. Because data cloning is based on MCMC, it is amenable to extensions of our model to include partially observed and hidden processes. We have demonstrated this with our example on missed observations. Any Bayesian method would provide this option and it may be a matter of belief whether frequentist of Bayesian approaches are used. However, data cloning additionally provides tools to detect parameter estimability problems (Lele, Nadeem & Schmuland 2010), which was relevant in our analysis; compare Appendix S2. At this stage, data cloning via JAGS was computationally intense, and it may be worth to explore alternative options, for example a 'home-made' MCMC sampler in a fast language such as C/C++. Still, with quickly increasing computational capacities and advances in statistical software, we believe that our method has a promising future.

Verification of our method was successful. In matching model fits, almost all MCMC runs mixed well and converged. Convergence and mixing problems occurred in non-matching model fits and especially for parameters that were not meaningful to the trajectory (e.g. a resource-selection parameter for a memory trajectory). For further application of our method, we have given recommendations how to proceed in cases on non-convergent model fits (Appendix S2.). Model selection via BIC successfully identified trajectories' true underlying mechanisms, and if parameters in a model fit were applicable to the underlying trajectory, we were able to recover true parameter values. Simulated movement trajectories were samples of stochastic processes, and therefore realized parameter values were subjected to stochasticity. Thus, parameter estimates could not be expected to exactly coincide with the true values. Verifying the functionality of our method was particularly important with respect to the newly introduced memory process. We conclude that if time since last visit is a driver of observed movement trajectories, our framework is able to detect this.

When we compared results from model selection to outcomes of hypothesis tests, we found that model selection was better able to distinguish true underlying mechanisms of trajectories. By definition, hypothesis tests allow for a Type I error, the size of which is influenced by the level of the test. However, decreasing the Type I error simultaneously decreases the power to correctly detect effects of covariates and increases the Type II error. The model selection framework is not based on this concept, and it proved to be more accurate in our analysis.

We have built on the framework of spatially explicit resource-selection models and added the influence of a dynamic memory process on movement decisions by introducing a dynamic cognitive map and linking it with the movement and resource-selection process. The existence of cognitive maps in animals is debated, and there is especially controversy about what form such maps may take, for example whether animals use topological cognitive maps for landmark-based navigation or whether animals can create and use geometric cognitive maps that preserve angles and distances between locations (Bennett 1996; Asensio *et al.* 2011; Collett, Chittka & Collett 2013). This debate also includes the question whether spatial information in the brain is encoded with respect to the position of the viewer, that is egocentrically, or independently of the position of the viewer, that is allocentrically (Yeap 2014). In our models, we do not focus on navigational mechanisms but decision-making processes, and we use the cognitive map as a useful mathematical tool to model spatial information. Investigation of different navigational mechanisms within a model selection framework similar to that presented here could be the goal of future research. With our model formulation in terms of a cognitive map, we have provided a general framework for linking movement with information use and acquisition. We emphasize that within this general formulation, a variety of more specific formulations of cognitive maps can be realized, tailored to the situations and behavioural processes of interest.

In our candidate models, we have used time since last visit to locations as an example of a form of dynamic information that is mediated by the cognitive map. We have demonstrated how the time since last visit to a location can shape the movement process, either on their own or in interaction with environmental variables. Such behaviours can, for instance, occur when animals patrol their home ranges for defence purposes, when predators counteract behavioural depression, or when animals rely on resources that vary in their availability due to depletion. However, our modelling framework and its elements are flexible and can be extended to include other forms of dynamic information and experiences that animals collect during their movement. For instance, while animals travel, they may gather information about seasonally available resources. Observations of primates show evidence that they remember fruiting statuses of individual trees and use this information to predict the fruiting status of those trees at later times (Janmaat, Ban & Boesch 2013). Prey species can use their movement to collect information about the distribution of predators. Such information can enable prey to reduce costly antipredatory behaviours and therefore outweigh attack risks connected to the information collection. This has been suggested to explain movement behaviour of caribou towards wolf paths (Latombe, Fortin & Parrott 2014).

Although our models describe movement behaviour of individual animals, the ideas we have presented can also apply to other systems. A specific feature of our models is the interaction between a movement process and an information, or memory, process. A similar dynamic interplay can arise on a larger scale when a species disperses and expands its range. While moving into a new environment, the dispersing species might alter the environment and its species composition, which in turn could affect the dispersing species (Gilman *et al.* 2010). Such processes could be analysed with the same mathematical ideas and modelling tools as we have presented here. Thus, we have presented a powerful modelling approach to identify spatial memory and dynamic information as drivers of movement decisions, and our framework and its elements promise a wide range of applications within movement ecology.

Acknowledgements

We thank Jonathan Potts and Andrew Bateman for helpful discussions and feedback on the manuscript. We are grateful to the editor and anonymous reviewers, whose comments have greatly improved our manuscript. UES was supported by an iCORE Graduate Student Scholarship (iCORE is now part of Alberta Innovates-Technology Futures), a Pacific Institute of Mathematical Sciences International Graduate Training Centre fellowship, and funding from the Department of Mathematical and Statistical Sciences at the University of Alberta. MAL gratefully acknowledges Natural Sciences and Engineering Research Council Discovery and Accelerator grants, a Canada Research Chair and a Killam Research Fellowship.

Data accessibility

R scripts to perform statistical analyses, simulated data (examples) and JAGS model definitions can be found in the online supporting information.

References

Amano, T. & Katayama, N. (2009) Hierarchical movement decisions in predators: effects of foraging experience at more than one spatial and temporal scale. *Ecology*, **90**, 3536–3545.

Asensio, N., Brockelman, W.Y., Malaivijitnond, S. & Reichard, U.H. (2011) Gibbon travel paths are goal oriented. *Animal Cognition*, **14**, 395–405.

Avgar, T., Deardon, R. & Fryxell, J.M. (2013) An empirically parameterized individual based model of animal movement, perception, and memory. *Ecological Modelling*, **251**, 158–172.

Bar-Shai, N., Keasar, T. & Shmida, A. (2011) The use of numerical information by bees in foraging tasks. *Behavioral Ecology*, **22**, 317.

Bennett, A. (1996) Do animals have cognitive maps? *Journal of Experimental Biology*, **199**, 219.

Breed, G.A., Costa, D.P., Jonsen, I.D., Robinson, P.W. & Mills-Flemming, J. (2012) State-space methods for more completely capturing behavioral dynamics from animal tracks. *Ecological Modelling*, **235–236**, 49–58.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference : A Practical Information-theoretic Approach*, 2 edn. Springer, New York.

Charnov, E.L., Orians, G.H. & Hyatt, K. (1976) Ecological implications of resource depression. *American Naturalist*, **110**, 247–259.

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

Collett, M., Chittka, L. & Collett, T.S. (2013) Spatial memory in insect navigation. *Current Biology*, **23**, R789–R800.

Dalziel, B.D., Morales, J.M. & Fryxell, J.M. (2008) Fitting probability distributions to animal movement trajectories: using artificial neural networks to link distance, resources, and memory. *The American Naturalist*, **172**, 248–258.

Davies, N.B. & Houston, A.I. (1981) Owners and satellites: the economics of territory defence in the pied wagtail, Motacilla alba. *The Journal of Animal Ecology*, **50**, 157–180.

Deneubourg, J.L., Goss, S., Franks, N. & Pasteels, J.M. (1989) The blind leading the blind: modeling chemically mediated army ant raid patterns. *Journal of insect behavior*, **2**, 719–725.

Fagan, W.F., Lewis, M.A., Auger-Méthé, M., Avgar, T., Benhamou, S., Breed, G., LaDage, L., Schlägel, U.E., Tang, W., Papastamatiou, Y.P., Forester, J. & Mueller, T. (2013) Spatial memory and animal movement. *Ecology Letters*, **16**, 1316–1329.

Forester, J.D., Im, H. & Rathouz, P. (2009) Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, **90**, 3554–3565.

Forester, J.D., Ives, A., Turner, M., Anderson, D., Fortin, D., Beyer, H., Smith, D. & Boyce, M. (2007) State-space models link elk movement patterns to landscape characteristics in Yellowstone National Park. *Ecological Monographs*, **77**, 285–299.

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

Gelman, A. & Rubin, D.B. (1992) Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**, 457–472.

Gillies, C.S., Beyer, H.L. & St Clair, C.C. (2011) Fine-scale movement decisions of tropical forest birds in a fragmented landscape. *Ecological Applications*, **21**, 944–954.

Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010) A framework for community interactions under climate change. *Trends in Ecology & Evolution*, **25**, 325–331.

Giuggioli, L., Potts, J.R. & Harris, S. (2011) Animal interactions and the emergence of territoriality. *PLoS Computational Biology*, **7**, e1002008.

Griffiths, D., Dickinson, A. & Clayton, N. (1999) Episodic memory: what can animals remember about their past? *Trends in Cognitive Sciences*, **3**, 74–80.

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

Janmaat, K.R., Ban, S.D. & Boesch, C. (2013) Chimpanzees use long-term spatial memory to monitor large fruit trees and remember feeding experiences across seasons. *Animal Behaviour*, **86**, 1183–1205.

Janson, C.H. & Byrne, R. (2007) What wild primates know about resources: opening up the black box. *Animal Cognition*, **10**, 357–367.

Jedrzejewski, W., Schmidt, K., Theuerkauf, J., Jedrzejewska, B. & Okarma, H. (2001) Daily movements and territory use by radio-collared wolves (*Canis lupus*) in Bialowieza Primeval Forest in Poland. *Canadian Journal of Zoology*, **79**, 1993–2004.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. & Morales, J.M. (2013) Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, **93**, 2336–2342.

Latombe, G., Fortin, D. & Parrott, L. (2014) Spatio-temporal dynamics in the response of woodland caribou and moose to the passage of grey wolf. *Journal of Animal Ecology*, **83**, 185–198.

Lele, S.R., Dennis, B. & Lutscher, F. (2007) Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. *Ecology Letters*, **10**, 551–563.

Lele, S.R. & Keim, J.L. (2006) Weighted distributions and estimation of resource selection probability functions. *Ecology*, **87**, 3021–3028.

Lele, S.R., Nadeem, K. & Schmuland, B. (2010) Estimability and likelihood inference for generalized linear mixed models using data cloning. *Journal of the American Statistical Association*, **105**, 1617–1625.

Link, W.A. & Barker, R.J. (2006) Model weights and the foundations of multimodel inference. *Ecology*, **87**, 2626–2635.

Manly, B.F., McDonald, L.L., Thomas, D.L., McDonald, T.L. & Erickson, W.P.(2002) *Resource Selection by Animals: Statical Design and Analysis for Field Studies*, 2 edn. Kluwer Academic Publishers, Dordrecht.

Martin-Ordas, G., Haun, D., Colmenares, F. & Call, J. (2009) Keeping track of time: evidence for episodic-like memory in great apes. *Animal Cognition*, **13**, 331–340.

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

Moorcroft, P.R. & Lewis, M.A. (2006) *Mechanistic Home Range Analysis.* Princeton University Press, Princeton, NJ.

Moorcroft, P.R. & Barnett, A. (2008) Mechanistic home range models and resource selection analysis: a reconciliation and unification. *Ecology*, **89**, 1112–1119.

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

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

Owen-Smith, N., Fryxell, J. & Merrill, E. (2010) Foraging theory upscaled: the behavioural ecology of herbivore movement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2267–2278.

Plummer, M. (2013) *rjags: Bayesian Graphical Models Using MCMC*. R package version 3-11. URL http://CRAN.R-project.org/package=rjags

Potts, J.R., Bastille-Rousseau, G., Murray, D.L., Schaefer, J.A. & Lewis, M.A. (2014) Predicting local and non-local effects of resources on animal space use using a mechanistic step selection model. *Methods in Ecology and Evolution*, **5**, 253–262.

R Core team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. URL http://www.R-project.org/

Reid, C.R., Latty, T., Dussutour, A. & Beekman, M. (2012) Slime mold uses an externalized spatial "memory" to navigate in complex environments. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 17490–17494.

Rhodes, J.R., McAlpine, C.A., Lunney, D. & Possingham, H.P. (2005) A spatially explicit habitat selection model incorporating home range behavior. *Ecology*, **86**, 1199–1205.

Smouse, P.E., Focardi, S., Moorcroft, P.R., Kie, J.G., Forester, J.D. & Morales, J.M. (2010) Stochastic modelling of animal movement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 2201–2211.

Solymos, P. (2010) dclone: Data cloning in R. *The R Journal*, **2**, 29–37.

Squires, J.R., DeCesare, N.J., Olson, L.E., Kolbe, J.A., Hebblewhite, M. & Parks, S.A. (2013) Combining resource selection and movement behavior to predict corridors for Canada lynx at their southern range periphery. *Biological Conservation*, **157**, 187–195.

Sulikowski, D. & Burke, D. (2011) Movement and memory: different cognitive strategies are used to search for resources with different natural distributions. *Behavioral Ecology and Sociobiology*, **65**, 621–631.

Sutherland, W.J., Freckleton, R.P., Godfray, H.C.J., Beissinger, S.R., Benton, T., Cameron, D.D., Carmel, Y., Coomes, D.A., Coulson, T. & Emmerson, M.C. (2013) Identification of 100 fundamental ecological questions. *Journal of Ecology*, **101**, 58–67.

Tolman, E.C. (1948) Cognitive maps in rats and men. *Psychological Review*, **55**, 189–208.

Tsoar, A., Nathan, R., Bartan, Y., Vyssotski, A., Dell'Omo, G. & Ulanovsky, N. (2011) Large-scale navigational map in a mammal. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, E718–E724.

Yeap, W. (2014) On egocentric and allocentric maps. *Spatial Cognition IX* (eds C. Freksa, B. Nebel, M. Hegarty & T. Barkowsky), pp. 62–75. Springer International Publishing, Cham.

Received 25 April 2014; accepted 17 September 2014
Handling Editor: Luca Börger

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Supplemental methods.

Appendix S2. Supplemental results.

Appendix S3. Animated trajectories.

Data S1. R and JAGS code, simulated data example.