

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

3 Ulrike E. Schlägel^{1,a}, Mark A. Lewis¹.

4 **Short title:** Spatial memory for movement decisions

5 **Word count:** 7000

6 **1.** Centre for Mathematical Biology, Department of Mathematical and Statistical Sciences, University
7 of Alberta, Canada

8 **a.** Corresponding author. Address: Department of Mathematical and Statistical Sciences, CAB 632,
9 University of Alberta, Edmonton, AB, Canada, T6G 1G1. E-mail: ulrike.schlaegel@gmail.ca

10 Summary

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

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

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

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

34 **Keywords:** animal movement, behaviour, information, cognitive map, memory, re-
35 source selection, step-selection function, mechanistic model

36 **Introduction**

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

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

51 Most mechanistic models have concentrated on incorporating relationships between
52 environmental factors and movement behaviour within a static environment (but see Av-
53 gar *et al.*, 2013); however, observations show that animals also take into account dynam-
54 ically changing information and respond with their movements to temporal availability
55 or unavailability of resources (Martin-Ordas *et al.*, 2009). For instance, fruit-eating pri-
56 mates express goal-oriented travel towards those trees in their home range that carry
57 ripe fruit (Asensio *et al.*, 2011), and it has been suggested that monkeys use their daily

58 travels to monitor fruiting histories of trees (Janmaat *et al.*, 2013; Janson & Byrne,
59 2007). On the other hand, many resources, once depleted, need some time before they
60 become available again, providing reason for animals to avoid depleted food patches
61 (Davies & Houston, 1981; Owen-Smith *et al.*, 2010; Bar-Shai *et al.*, 2011). Avoidance
62 behaviour may be a response not only to depletion of resources, such as plant biomass
63 or prey, but also to behavioural depression. Behavioural depression refers to a reduction
64 in prey availability that is caused by behavioural changes of the prey in response to
65 predation (Charnov *et al.*, 1976). For example, prey may show greater alertness or seek
66 shelter. This reduces capture rates, to which predators may respond in turn by changing
67 their hunting areas (Jedrzejewski *et al.*, 2001; Amano & Katayama, 2009). Temporal
68 considerations also become important for movement decisions if territorial defence mech-
69 anisms require animals to visit certain locations regularly, e.g. to scent-mark territory
70 boundaries (Moorcroft & Barnett, 2008; Giuggioli *et al.*, 2011).

71 As the above examples highlight, spatio-temporal information drives movement deci-
72 sions and at the same time movement allows animals to update this information. Exper-
73 imental findings additionally support that animals make decisions based on information
74 that they have obtained through previous experiences. Memory of information about
75 the ‘what, where and when’ of events, obtained through subjective experience, is termed
76 ‘www-memory’ (Martin-Ordas *et al.*, 2009) or ‘episodic-like memory’ (Griffiths *et al.*,
77 1999). It is possible that animals acquire information about current environmental con-
78 ditions through perceptual cues, even over large distances (Tsoar *et al.*, 2011), and that
79 information about the recent travel history is stored in externalized ‘memory’, such as
80 pheromone trails or slime (Deneubourg *et al.*, 1989; Reid *et al.*, 2012). However, it is
81 likely that many animals draw upon internal memory, especially for behaviours that
82 require information about temporal distances (‘how long ago?’) (Griffiths *et al.*, 1999;

83 Martin-Ordas *et al.*, 2009; Janmaat *et al.*, 2013). During recent years, movement models
84 have started to incorporate influences of memorized information on movement decisions
85 (for a review see Fagan *et al.*, 2013). Most of these are simulation models that are used
86 for theoretical considerations only (but see Avgar *et al.*, 2013); however to test our un-
87 derstanding of the feedbacks between movement and information acquisition, we must
88 also interface memory-based models with data (Smouse *et al.*, 2010).

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

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

117 **Methods**

118 For several decades, the basis of many animal movement models have been random walks.
119 In a classical random walk, movement is described as a series of discrete steps that have
120 independent and identical probability distributions. This has been extended to include
121 correlations between steps, biases towards specific locations, and step probabilities that
122 depend on the behavioural state of the individual (Morales *et al.*, 2004; McClintock *et al.*,
123 2012; Breed *et al.*, 2012; Langrock *et al.*, 2013). Random walks and their extensions
124 have been used both to analyze movement behaviour at an individual level (Lagrangian
125 approach; e.g. Smouse *et al.*, 2010) and to derive partial-differential equation models
126 that describe spatio-temporal patterns at a population level or expected space-use of
127 individuals (Eulerian approach; e.g. Codling *et al.*, 2008).

128 We are interested in understanding decision processes that underly movement be-
129 haviour on the scale of individuals. We draw upon a modelling framework that bridges
130 the gap between statistical resource-selection analysis and spatially explicit movement
131 models (Rhodes *et al.*, 2005; Moorcroft & Barnett, 2008; Forester *et al.*, 2009). The

132 framework builds on a random walk and defines movement via step probabilities, which
133 have two components. A resource-independent movement kernel assigns probabilities to
134 steps based on the animal’s general movement tendencies. Given this, a weighting func-
135 tion evaluates the attractiveness of steps according to resource availability and resource
136 preferences. We extend this framework by generalizing the weighting function. In our
137 generalization, the weighting function does not only describe the influence of resources
138 but allows for the inclusion of any information relevant to the animal. Information can
139 pertain to landscape features and resources, as in previous models, but also to memories
140 of past events and timing aspects, which cannot be obtained externally but only through
141 the movement process and the animal’s behaviour itself. We assume that information at
142 a given time is either obtained through direct perception or retrieved from the animal’s
143 cognitive map (i.e. memory) which itself is updated through experience. In our model,
144 the cognitive map is a function that assigns values to locations according to their in-
145 formation content at a given time. Thus, it serves as a mathematical tool without the
146 claim that it truly represents the underlying cognitive mechanism. With the framework
147 of the cognitive map we provide a general method for including an explicit information-
148 acquisition process. The cognitive map itself can take many forms, depending on the
149 species and behaviour of interest. In our candidate models, we demonstrate examples of
150 types of information the cognitive map may contain.

151 **The modelling framework**

152 We consider movement paths of individual animals, and we assume that an individual’s
153 trajectory consists of a series of locations $(\mathbf{x}_1, \dots, \mathbf{x}_N)$ at regular times $T = \{1, \dots, N\}$.
154 Each location has an Easting and a Northing in two-dimensional space, which is dis-
155 cretized into a regular grid of square cells. The resolution of the spatial discretization

156 depends on the available environmental data and should be fine enough compared to
 157 the animal’s movement such that steps generally range over multiple cells. We model
 158 movement as a stochastic process, where the probability of making a step to location \mathbf{x}_t
 159 depends on the location at time $t - 1$ and, if movement is persistent, on the previous
 160 step from \mathbf{x}_{t-2} to \mathbf{x}_{t-1} . We define this step probability as

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

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

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

177 In general, the movement kernel k can be very simple, e.g. constant within the ani-
 178 mal’s maximum movement radius (Rhodes *et al.*, 2005); however, we can also use a more

179 complex kernel that accounts for persistence in movement direction or biases towards
 180 specific locations (Moorcroft & Lewis, 2006). Directions can be measured by either
 181 turning angles (the angles between successive steps) or bearings (the angles of steps with
 182 respect to a fixed direction, e.g. North).

183 We model the weighting function w_t as a resource selection function (Manly *et al.*,
 184 2002; Lele & Keim, 2006). There are several forms available for resource selection func-
 185 tions, and here we present the logistic form,

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

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

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

200 $p(\mathbf{x}_1|\boldsymbol{\theta}) = \frac{1}{|\Omega|}$, and to let

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

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

202 Candidate models

203 We consider four different models that represent biological hypotheses about the types of
204 information that an individual may consider for making movement decisions. In the sim-
205 plest case, the *null model*, we assume that the animal considers no specific information.
206 In the *resource model*, an individual considers static information about the environment,
207 where ‘static’ means that the information content remains constant over the time span
208 of the analysis. Information can be given about any resources pertaining to the animal,
209 e.g. any variables as they are typical in resource-selection analyses. To include dynam-
210 ically changing information, we allow information, and thereby the weighting function,
211 to change through time. If information were only given externally, this would constitute
212 a dynamic version of the resource model. However, our aim is to model a dynamic inter-
213 play of movement decisions and information content. In the *memory model*, we therefore
214 introduce *time since last visit* as new type of information. To account for the possibility
215 that both resources and the dynamic variable time since last visit influence movement
216 decisions simultaneously, we consider a *combination model* as the most complex model.

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

221 **Null model**

222 In the null model, we assume that the information content of all locations is zero. There-
223 fore, the weighting function is homogeneous across the landscape and constant over time,
224 $w_t(\mathbf{x}) = 1$ for all $x \in \Omega, t \in T$. This means that the animal moves only according to the
225 kernel k .

226 **Resource model**

227 In the resource model, information is static and includes all resource variables of interest,
228 $\mathcal{I}_t(\mathbf{x}) = \mathcal{I}(\mathbf{x}) = (r_1(\mathbf{x}), \dots, r_n(\mathbf{x}))$ for every location $x \in \Omega$. It is straightforward to
229 extend this to dynamic resource information to include, e.g. seasonal changes in the
230 landscape or disturbance events.

231 **Memory model**

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

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

240 Because of our spatial discretization, we use $d(\mathbf{y}, \mathbf{z}) = |y_E - z_E| + |y_N - z_N|$ as the distance
 241 between two locations \mathbf{y}, \mathbf{z} with Easting and Northing $\mathbf{y} = (y_E, y_N)$ and $\mathbf{z} = (z_E, z_N)$,
 242 such that all locations within a distance δ of a fixed location \mathbf{z} form a diamond-shaped
 243 area around \mathbf{z} . We assume that $\text{path}(\mathbf{x}_{t-1} \rightarrow \mathbf{x}_t)$ is the straight line between \mathbf{x}_{t-1}
 244 and \mathbf{x}_t . Via eqn 4, an individual counts the number of steps it remains absent from
 245 locations, and therefore $m_t(\mathbf{x})$ is the time since last visit to location \mathbf{x} at time t . A
 246 location is considered visited when the animal comes within a distance $\delta > 0$. Because
 247 m_t is obtained recursively, we have to define appropriate starting values. Here, we use
 248 movement data prior to the trajectory $(\mathbf{x}_1, \dots, \mathbf{x}_N)$ for initialization. If a location \mathbf{x} was
 249 visited during the initialization phase, we calculate the time between the last visit to
 250 this location and the beginning of our actual trajectory and thus reconstruct time since
 251 last visit at time $t = 1$, $m_1(\mathbf{x})$. For all locations not visited during initialization, we
 252 set time since last visit as the length of the initialization phase. The dynamic variable
 253 time since last visit is used in the memory model to inform movement decisions via
 254 $\mathcal{I}_t(\mathbf{x}) = m_{t-1}(\mathbf{x})$. Once \mathbf{x}_t is chosen according to the probability mass function in eqn 1,
 255 m_t is updated via eqn 4. Here, we track time since last visit for the entire spatial domain
 256 Ω . If the selection coefficient with respect to $m_t(\mathbf{x})$ is positive, this leads to any location
 257 eventually becoming highly attractive after long enough absence. If this behaviour is
 258 not desired, one may adjust the definition of the cognitive map or weighting function
 259 appropriately. For example, if prior information about an animal's behaviour is given,
 260 it is possible to track time since last visit only for certain locations of specific interest.

261 **Combination model**

262 In the *combination model*, we allow information types from both the resource and the
 263 memory model to influence movement simultaneously by letting the information vector
 264 be $\mathcal{I}_t(\mathbf{x}) = (r_1(\mathbf{x}), \dots, r_n(\mathbf{x}), m_{t-1}(\mathbf{x}))$. In particular, this models allows for interactive

265 effects of time since last visit and resource variables, e.g. by incorporating multiplicative
 266 terms of the form $\gamma r(\mathbf{x}) m_{t-1}(\mathbf{x})$ into the interaction term $f(\mathcal{I}_t(\mathbf{x}), \gamma)$ in the weighting
 267 function (eqn 2). This is important in situations where return times to locations matter
 268 depending on the resources at the location, e.g. average return times to preferred foraging
 269 areas may differ from those to locations used as shelter.

270 Information-independent kernel

271 We define the movement kernel k based on a step length distribution with density \mathcal{S} and
 272 a distribution for movement directions with density Φ . For step length, we use a Weibull
 273 distribution with scale and shape parameter $\kappa > 0$ and $\lambda > 0$, respectively, because it
 274 has a flexible form and generally shows a good fit with empirical data (Morales *et al.*,
 275 2004). Thus,

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

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

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

280 The wrapped Cauchy distribution is convenient for implementation, and it has been used
 281 commonly to model movement directions (Morales *et al.*, 2004; Codling *et al.*, 2008, note
 282 that eqn 6 is equivalent to their formula with parameter transformation $r = \exp(-\rho)$).

283 One could use alternative distributions, such as the von Mises distribution or wrapped

284 normal distribution (Codling *et al.*, 2008). Assuming that the choices for step length
 285 and movement direction are independent, the kernel becomes the product of \mathcal{S} and Φ ,
 286 describing a correlated random walk,

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

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

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

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

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

300 **Statistical inference**

301 If information \mathcal{I}_t is known, the likelihood function for the collection of parameters $\boldsymbol{\theta} =$
 302 $(\kappa, \lambda, \rho, \alpha, \boldsymbol{\beta}, \boldsymbol{\gamma})$ for the general model is

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

303 In the memory and combination model, \mathcal{I} includes the variable time since last visit
 304 $\mathbf{m} = (m_t, t \in T)$, which represents internal information of the animal that in general
 305 cannot be observed. However, because of the way we define and initialize \mathbf{m} , we are able
 306 to iteratively calculate the time series (m_1, \dots, m_N) based on the movement trajectory.
 307 Therefore, given the data $(\mathbf{x}_1, \dots, \mathbf{x}_N)$, time since last visit becomes a known covariate,
 308 and the likelihood function in eqn 9 is valid for all models.

309 To obtain parameter estimates and their confidence intervals for all models we use
 310 data cloning (Lele *et al.*, 2007). Data cloning uses Markov Chain Monte Carlo (MCMC)
 311 methods, which are usually employed in Bayesian statistical inference. However, data
 312 cloning provides approximations to maximum likelihood estimates (MLE), together with
 313 Wald-type confidence intervals, thus facilitating frequentist inference; see Appendix A.2.

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

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

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

332 This allows to preserve the entire dependency structure of the trajectory and avoids the
 333 need to discard any information. For more information on this, see Appendix A.3.

334 Simulation study

335 To verify the functionality of our method we applied the modelling framework and sta-
 336 tistical inference method to simulated data. Because eqn 1 defines probability mass
 337 functions for movement steps, we can sample from them to iteratively generate individ-
 338 ual movement trajectories according to the four candidate models. These data have the
 339 advantage that we know both a trajectory’s underlying mechanism and the parameter
 340 values that were used to generate the trajectory. By applying our inference procedure
 341 to these data, we investigated whether we were able to identify the true underlying
 342 mechanism of a trajectory and whether we were able to correctly estimate parameter
 343 values.

344 **Simulation of landscapes**

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

353 **Simulation of movement trajectories**

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

357 On each of the five landscape pairs, we simulated a null, resource, memory and
358 combination trajectory, using the same movement parameter values on all landscapes
359 and across all four models, as applicable (Fig. 3). The kernel parameters κ , λ , ρ appear
360 in all models. The resource model has additional parameters α_{res} , β_1 , β_2 , which are
361 the intercept and the selection parameters with respect to the two resources (r_1, r_2)
362 of the weighting function (eqn 2). In this model, we assumed there is no interaction
363 between the two resources. The memory model instead has additional parameters α_{mem}
364 and β_{mem} , which describe the animal's preferences with respect to time since last visit
365 **m**. In the combination model, the weighting function includes all effects, such that it
366 has parameters $\alpha_{\text{com}} = \alpha_{\text{res}} + \alpha_{\text{mem}}$, β_1 , β_2 , and β_{mem} . In this model, we further allowed
367 for interactions between resources and time since last visit by defining the interaction

368 term in the weighting function as $f((r_1, r_2, m_t), \gamma_1, \gamma_2) = \gamma_1 r_1 m_t + \gamma_2 r_2 m_t$, where γ_1
369 and γ_2 are the interaction parameters. We chose the main set of parameter values to
370 represent realistic movement behaviour. To account for scenarios for which parameter
371 values were potentially more difficult to estimate from data, e.g. small values of selection
372 parameters, we generated two supplemental data sets, comprising two additional sets of
373 20 trajectories each generated from alternative sets of parameters; see Appendix B.1 and
374 B.2.

375 For all trajectories, we simulated 2600 time steps, of which we discarded the first
376 1400 steps as initialization. This was particularly important for the memory model, in
377 which we started with a cognitive map having value 0 everywhere. We used the last 400
378 steps from the initialization phase to calculate m_1 . Each final trajectory consisted of
379 1200 time steps, which we considered a length commonly available (e.g. 1200 time steps
380 could represent 50 days of 1-hr data or 100 days of 2-hr data).

381 For an example of how to handle missed observations, we simulated a combination
382 trajectory with 90% fix rate by removing locations from a trajectory, 5% as single loca-
383 tions and 5% as two successive locations. We chose a trajectory from the main data set,
384 which allowed us to compare results for completely and incompletely observed trajectory;
385 see Appendix B.2.2.

386 **Analysis of simulated data**

387 To every simulated trajectory, we fitted all four candidate models (Fig. 3) using data
388 cloning. For details about the data cloning and MCMC procedures, such as number of
389 clones and iterations used, see Appendix A.2. There were two basic types of model fits
390 that we distinguished in our analyses. A model could be fitted to a matching trajectory,
391 i.e. a trajectory that had been simulated using the same model's mechanism (e.g. a

392 resource model fitted to a resource trajectory). Or, a model could be fitted to a non-
393 matching trajectory (e.g. a resource model fitted to a null, memory or combination
394 trajectory). Each model fit generated estimates of the model parameters, together with
395 Wald-type confidence intervals. Here, we used 95% confidence intervals. Using the
396 approximate maximum-likelihood parameter estimates from data cloning, we estimated
397 the maximum value of the model likelihood, and BIC, for each model fit.

398 For all parameter estimates, we examined whether their potential scale reduction
399 factors \hat{R} were close to 1 (Gelman & Rubin, 1992). For an MCMC fit, in which parallel
400 Markov chains are used to generate the posterior distribution of a parameter, the poten-
401 tial scale reduction factor of a parameter indicates whether the chains have mixed well
402 and converged. If this is not the case, the estimate that results from such an MCMC
403 is not meaningful. We considered a potential scale reduction factor $0.9 \leq \hat{R} \leq 1$ to be
404 sufficiently close to 1 (Gelman & Rubin, 1992), and we excluded all parameter estimates
405 that did not meet this condition from our analysis. Whenever such a non-convergent
406 or non-mixing parameter occurred within a model fit, the resultant likelihood and BIC
407 values of the fit were possibly inaccurate. Therefore, if a model fit included one or more
408 parameters with $\hat{R} < 0.9$ or $\hat{R} > 1.1$, we excluded the BIC value from our model-selection
409 analysis.

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

417 For matching model fits, we compared true parameter values that were used to gener-
418 ate a trajectory to the parameter estimates obtained from applying the matching model,
419 and we examined whether 95% confidence intervals of parameters included the true val-
420 ues. This should be achieved 95% of the time if parameters are identifiable and our
421 statistical methodology is functional.

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

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

437 **Simulation results**

438 Here, we present results for data generated with the main set of parameters θ_1 (Fig. 3).
439 Results for supplemental data generated by additional sets of parameters can be found
440 in Appendix B.2.1.

441 Of all 80 model fits (four models fitted to 20 trajectories each), 80% had potential
442 scale reduction factor $0.9 \leq \hat{R} \leq 1.1$ for all model parameters. In the remaining model
443 fits, at least one parameter had $\hat{R} > 1.1$ (Fig. 4). Convergence or mixing problems never
444 occurred when the null model was fitted to a trajectory, even if the trajectory had a
445 more complex underlying mechanism. Large \hat{R} values only occurred if the fitted model
446 contained parameters that were inapplicable to the model that was used to generate the
447 trajectory. This was the case when any of the more complex models was fitted to a null
448 trajectory, when the combination model was fitted to a resource or memory trajectory,
449 or when the memory model was fitted to a resource trajectory and vice versa. In these
450 model fits, the non-convergent parameters were mainly those that did not correspond
451 to true underlying covariate effects. However, when in a model fit problems occurred
452 for multiple parameters, occasionally even applicable parameters failed to converge. In
453 matching model fits, Markov chains always mixed well and converged. For more details
454 on convergence, see Appendix B.3 and B.4.

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

461 Parameter estimates in matching model fits agreed well with true underlying param-
462 eter values. Parameter estimates generally were both close to and balanced around their
463 true values (Fig. 5). 95% confidence intervals (n=115) included the true parameter value
464 91% of the time. If we also considered results from the supplemental data, 94% of all
465 confidence intervals (n=345) included the true value.

466 Our hypothesis test as to whether covariates had an effect agreed with our expect-
467 tations. The combination model fitted to the 20 trajectories lead to 73 estimates of
468 selection parameters, of which 39 corresponded to true underlying effects. Analyzing
469 their confidence intervals, we obtained a false positive rate (Type I error rate) of 0.09
470 and a false negative rate of zero, i.e. Type II errors did not occur. However, if we also
471 considered the supplemental data and thereby increased the amount of resultant param-
472 eter estimates with confidence intervals to a total of 217, we obtained a Type I error rate
473 of 0.04 and again a Type II error rate of 0.09.

474 Discussion

475 In recent years, the link between animal movement and spatial memory has received in-
476 creasing attention (Smouse *et al.*, 2010; Fagan *et al.*, 2013). Studies of animal behaviour
477 and cognition have given useful insights into animals' capacities to remember past experi-
478 ences and use spatial memory. Most results have been obtained through experiments
479 in confined and synthetic settings. However, to better understand how important eco-
480 logical processes such as movement and dispersal are shaped by cognitive processes and
481 memory, we also need to look at animals in their natural environments (Tsoar *et al.*,
482 2011). Understanding the components of individual movement decisions and their inter-
483 actions ultimately will help us to predict how population distribution patterns respond
484 to environmental changes, such as landscape fragmentation and changing climate.

485 We have presented a modelling framework that can be used to detect the influence of
486 memorized information on movement decisions. We recognize that in many situations it
487 is difficult to confirm that animals draw upon memorized information instead of momen-
488 tarily perceived information; however, there is evidence that animals use information
489 that they have obtained during past experiences (Martin-Ordas *et al.*, 2009; Janmaat

490 *et al.*, 2013). As an example of such information, we use time since last visit to locations.
491 In our model, time since last visit is continuously updated during the movement process
492 and at the same time influences movement decisions. We formulate our models in a
493 way that makes them amenable to likelihood-based statistical inference. This allows us
494 to fit our models to data to test whether the timing of events plays a role for move-
495 ment decisions. Fitting the full model (eqn 1), encompassing both general movement
496 tendencies and selective behaviour, to data via the likelihood function (eqn 9) enables
497 simultaneous estimation of parameters of both the general movement kernel and weight-
498 ing function. This distinguishes our method from step selection approaches that use an
499 empirical movement kernel to estimate resource selection parameters in a case-control
500 framework (Fortin *et al.*, 2005; Forester *et al.*, 2009).

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

514 We used simulated landscapes and movement data to verify the functionality of our

515 modelling framework and statistical inference method. Adding the memory process
516 to the modelling framework considerably increased model complexity and the amount
517 of data that had to be processed. We were therefore interested in whether we could
518 correctly detect memory effects in empirical movement patterns and whether parameters
519 that describe the memory process and its interactions with other variables were possible
520 to estimate reliably. To perform inference, we used data cloning, which uses MCMC
521 techniques but facilitates frequentist inference. We used the software package JAGS,
522 which allowed us to define models in an easily understandable language and provides
523 a stable implementation of MCMC sampling. JAGS was able to adapt the sampling
524 process successfully so that parameters of very different magnitude could be reliably
525 estimated. However, this came at the cost of long computation times (ranging 0.5-5
526 days per single chain for different models) and high memory needs (ranging 1-5 GB
527 RAM). Alternatively, we could have used conventional numerical maximization of the
528 likelihood function, which in this case may have been faster but at the same time more
529 limited. Because data cloning is based on MCMC, it is amenable to extensions of our
530 model to include partially observed and hidden processes. We have demonstrated this
531 with our example on missed observations. Any Bayesian method would provide this
532 option and it may be a matter of belief whether frequentist or Bayesian approaches are
533 used. However, data cloning additionally provides tools to detect parameter estimability
534 problems (Lele *et al.*, 2010), which was relevant in our analysis; compare Appendix B.3.
535 At this stage, data cloning via JAGS was computationally intense, and it may be worth
536 to explore alternative options, e.g. a ‘home made’ MCMC sampler in a fast language
537 such as C/C++. Still, with quickly increasing computational capacities and advances in
538 statistical software, we believe that our method has a promising future.

539 Verification of our method was successful. In matching model fits, almost all MCMC

540 runs mixed well and converged. Convergence and mixing problems occurred in non-
541 matching model fits and especially for parameters that were not meaningful to the
542 trajectory (e.g. a resource selection parameter for a memory trajectory). For further
543 application of our method, we have given recommendations how to proceed in cases on
544 non-convergent model fits (Appendix B.4). Model selection via BIC successfully iden-
545 tified trajectories' true underlying mechanisms, and if parameters in a model fit were
546 applicable to the underlying trajectory, we were able to recover true parameter values.
547 Simulated movement trajectories were samples of stochastic processes, and therefore re-
548 alized parameter values were subject to stochasticity. Thus, parameter estimates could
549 not be expected to exactly coincide with the true values. Verifying the functionality
550 of our method was particularly important with respect to the newly introduced mem-
551 ory process. We conclude that if time since last visit is a driver of observed movement
552 trajectories, our framework is able to detect this.

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

560 We have built on the framework of spatially explicit resource-selection models and
561 added the influence of a dynamic memory process on movement decisions by introduc-
562 ing a dynamic cognitive map and linking it with the movement and resource-selection
563 process. The existence of cognitive maps in animals is debated, and there is especially
564 controversy about what form such maps may take, e.g. whether animals use topological

565 cognitive maps for landmark-based navigation or whether animals can create and use
566 geometric cognitive maps that preserve angles and distances between locations (Bennett,
567 1996; Asensio *et al.*, 2011; Collett *et al.*, 2013). This debate also includes the question
568 whether spatial information in the brain is encoded with respect to the position of the
569 viewer, i.e. egocentrically, or independently of the position of the viewer, i.e. allocen-
570 trically (Yeap, 2014). In our models we do not focus on navigational mechanisms but
571 decision making processes, and we use the cognitive map as a useful mathematical tool
572 to model spatial information. Investigation of different navigational mechanisms within
573 a model-selection framework similar to that presented here could be the goal of future
574 research. With our model formulation in terms of a cognitive map, we have provided
575 a general framework for linking movement with information use and acquisition. We
576 emphasize that within this general formulation, a variety of more specific formulations
577 of cognitive maps can be realized, tailored to the situations and behavioural processes
578 of interest.

579 In our candidate models, we have used time since last visit to locations as an example
580 of a form of dynamic information that is mediated by the cognitive map. We have
581 demonstrated how the time since last visit to a location can shape the movement process,
582 either on their own or in interaction with environmental variables. Such behaviours
583 can, for instance, occur when animals patrol their home ranges for defence purposes,
584 when predators counteract behavioural depression, or when animals rely on resources
585 that vary in their availability due to depletion. However, our modelling framework
586 and its elements are flexible and can be extended to include other forms of dynamic
587 information and experiences that animals collect during their movement. For instance,
588 while animals travel they may gather information about seasonally available resources.
589 Observations of Mangabeys show evidence that they remember fruiting statuses of fig

590 trees and use this information to predict the fruiting status of those trees at later times
591 (Janmaat *et al.*, 2013). Prey species can use their movement to collect information about
592 the distribution of predators. Such information can enable prey to reduce costly anti-
593 predatory behaviours and therefore outweigh attack risks connected to the information
594 collection. This has been suggested to explain movement behaviour of caribou towards
595 wolf paths (Latombe *et al.*, 2014).

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

607 **Acknowledgements**

608 We thank Jonathan Potts and Andrew Bateman for helpful discussions and feedback
609 on the manuscript. We are grateful to the editor and anonymous reviewers, whose
610 comments have greatly improved our manuscript. UES was supported by an iCORE
611 Graduate Student Scholarship (iCORE is now part of Alberta Innovates-Technology
612 Futures), a Pacific Institute of Mathematical Sciences International Graduate Training
613 Centre fellowship, and funding from the Department of Mathematical and Statistical

614 Sciences at the University of Alberta. MAL gratefully acknowledges Natural Sciences
615 and Engineering Research Council Discovery and Accelerator grants, a Canada Research
616 Chair and a Killam Research Fellowship.

617 **Data Accessibility**

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

620 **References**

621 Amano, T. & Katayama, N. (2009) Hierarchical movement decisions in predators: effects
622 of foraging experience at more than one spatial and temporal scale. *Ecology*, **90**, 3536–
623 3545.

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

626 Avgar, T., Deardon, R. & Fryxell, J.M. (2013) An empirically parameterized individual
627 based model of animal movement, perception, and memory. *Ecological Modelling*, **251**,
628 158–172.

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

631 Bennett, A. (1996) Do animals have cognitive maps? *Journal of Experimental Biology*,
632 **199**, 219.

- 633 Breed, G.A., Costa, D.P., Jonsen, I.D., Robinson, P.W. & Mills-Flemming, J. (2012)
634 State-space methods for more completely capturing behavioral dynamics from animal
635 tracks. *Ecological Modelling*, **235–236**, 49–58.
- 636 Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference : a*
637 *practical information-theoretic approach*. Springer, New York, 2 edition.
- 638 Charnov, E.L., Orians, G.H. & Hyatt, K. (1976) Ecological implications of resource
639 depression. *American Naturalist*, **110**, 247–259.
- 640 Codling, E.A., Plank, M.J. & Benhamou, S. (2008) Random walk models in biology.
641 *Journal of The Royal Society Interface*, **5**, 813–834.
- 642 Collett, M., Chittka, L. & Collett, T.S. (2013) Spatial Memory in Insect Navigation.
643 *Current Biology*, **23**, R789–R800.
- 644 Dalziel, B.D., Morales, J.M. & Fryxell, J.M. (2008) Fitting Probability Distributions to
645 Animal Movement Trajectories: Using Artificial Neural Networks to Link Distance,
646 Resources, and Memory. *The American Naturalist*, **172**, 248–258.
- 647 Davies, N.B. & Houston, A.I. (1981) Owners and satellites: the economics of territory
648 defence in the pied wagtail, *Motacilla alba*. *The Journal of Animal Ecology*, pp. 157–
649 180.
- 650 Deneubourg, J.L., Goss, S., Franks, N. & Pasteels, J.M. (1989) The blind leading the
651 blind: modeling chemically mediated army ant raid patterns. *Journal of insect behav-*
652 *ior*, **2**, 719–725.
- 653 Fagan, W.F., Lewis, M.A., Auger-Méthé, M., Avgar, T., Benhamou, S., Breed, G.,
654 LaDage, L., Schlägel, U.E., Tang, W., Papastamatiou, Y.P., Forester, J. & Mueller,
655 T. (2013) Spatial memory and animal movement. *Ecology Letters*, **16**, 1316–1329.

- 656 Forester, J.D., Im, H. & Rathouz, P. (2009) Accounting for animal movement in es-
657 timation of resource selection functions: sampling and data analysis. *Ecology*, **90**,
658 3554–3565.
- 659 Forester, J.D., Ives, A., Turner, M., Anderson, D., Fortin, D., Beyer, H., Smith, D. &
660 Boyce, M. (2007) State-space models link elk movement patterns to landscape char-
661 acteristics in Yellowstone National Park. *Ecological Monographs*, **77**, 285–299.
- 662 Fortin, D., Beyer, H., Boyce, M., Smith, D., Duchesne, T. & Mao, J. (2005) Wolves
663 influence elk movements: behavior shapes a trophic cascade in Yellowstone National
664 Park. *Ecology*, **86**, 1320–1330.
- 665 Gelman, A. & Rubin, D.B. (1992) Inference from iterative simulation using multiple
666 sequences. *Statistical science*, pp. 457–472.
- 667 Gillies, C.S., Beyer, H.L. & St Clair, C.C. (2011) Fine-scale movement decisions of
668 tropical forest birds in a fragmented landscape. *Ecological Applications*, **21**, 944–954.
- 669 Gilman, S.E., Urban, M.C., Tewksbury, J., Gilchrist, G.W. & Holt, R.D. (2010) A
670 framework for community interactions under climate change. *Trends in Ecology &*
671 *Evolution*, **25**, 325–331.
- 672 Giuggioli, L., Potts, J.R. & Harris, S. (2011) Animal Interactions and the Emergence of
673 Territoriality. *PLoS Computational Biology*, **7**, e1002008.
- 674 Griffiths, D., Dickinson, A. & Clayton, N. (1999) Episodic memory: what can animals
675 remember about their past? *Trends in Cognitive Sciences*, **3**, 74–80.
- 676 Horne, J.S., Garton, E.O., Krone, S.M. & Lewis, J.S. (2007) Analyzing animal move-
677 ments using brownian bridges. *Ecology*, **88**, 2354–2363.

- 678 Janmaat, K.R., Ban, S.D. & Boesch, C. (2013) Chimpanzees use long-term spatial mem-
679 ory to monitor large fruit trees and remember feeding experiences across seasons. *An-*
680 *imal Behaviour*, **86**, 1183–1205.
- 681 Janson, C.H. & Byrne, R. (2007) What wild primates know about resources: opening
682 up the black box. *Animal Cognition*, **10**, 357–367.
- 683 Jedrzejewski, W., Schmidt, K., Theuerkauf, J., Jedrzejewska, B. & Okarma, H. (2001)
684 Daily movements and territory use by radio-collared wolves (*Canis lupus*) in Bialowieza
685 Primeval Forest in Poland. *Canadian Journal of Zoology*, **79**, 1993–2004.
- 686 Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. & Morales, J.M.
687 (2013) Flexible and practical modeling of animal telemetry data: hidden Markov mod-
688 els and extensions. *Ecology*, **93**, 2336–2342.
- 689 Latombe, G., Fortin, D. & Parrott, L. (2014) Spatio-temporal dynamics in the response
690 of woodland caribou and moose to the passage of grey wolf. *Journal of Animal Ecology*,
691 **83**, 185–198.
- 692 Lele, S.R., Dennis, B. & Lutscher, F. (2007) Data cloning: easy maximum likelihood
693 estimation for complex ecological models using Bayesian Markov chain Monte Carlo
694 methods. *Ecology Letters*, **10**, 551–563.
- 695 Lele, S.R. & Keim, J.L. (2006) Weighted distributions and estimation of resource selec-
696 tion probability functions. *Ecology*, **87**, 3021–3028.
- 697 Lele, S.R., Nadeem, K. & Schmuland, B. (2010) Estimability and likelihood inference for
698 generalized linear mixed models using data cloning. *Journal of the American Statistical*
699 *Association*, **105**, 1617–1625.

- 700 Link, W.A. & Barker, R.J. (2006) Model weights and the foundations of multimodel
701 inference. *Ecology*, **87**, 2626–2635.
- 702 Manly, B.F., McDonald, L.L., Thomas, D.L., McDonald, T.L. & Erickson, W.P. (2002)
703 *Resource selection by animals: statical design and analysis for field studies*. Kluwer
704 Academic Publishers, 2 edition.
- 705 Martin-Ordas, G., Haun, D., Colmenares, F. & Call, J. (2009) Keeping track of time:
706 evidence for episodic-like memory in great apes. *Animal Cognition*, **13**, 331–340.
- 707 McClintock, B.T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B.J. & Morales,
708 J.M. (2012) A general discrete-time modeling framework for animal movement using
709 multistate random walks. *Ecological Monographs*, **82**, 335–349.
- 710 Moorcroft, P.R. & Lewis, M.A. (2006) *Mechanistic Home Range Analysis*. Princeton
711 University Press, Princeton, N.J.
- 712 Moorcroft, P.R. & Barnett, A. (2008) Mechanistic Home Range Models and Resource
713 Selection Analysis: A Reconciliation and Unification. *Ecology*, **89**, 1112–1119.
- 714 Morales, J.M., Haydon, D., Frair, J., Holsinger, K. & Fryxell, J. (2004) Extracting
715 more out of relocation data: building movement models as mixtures of random walks.
716 *Ecology*, **85**, 2436–2445.
- 717 Nathan, R., Getz, W., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse, P.E.
718 (2008) A movement ecology paradigm for unifying organismal movement research.
719 *Proceedings of the National Academy of Sciences of the United States of America*,
720 **105**, 19052–19059.

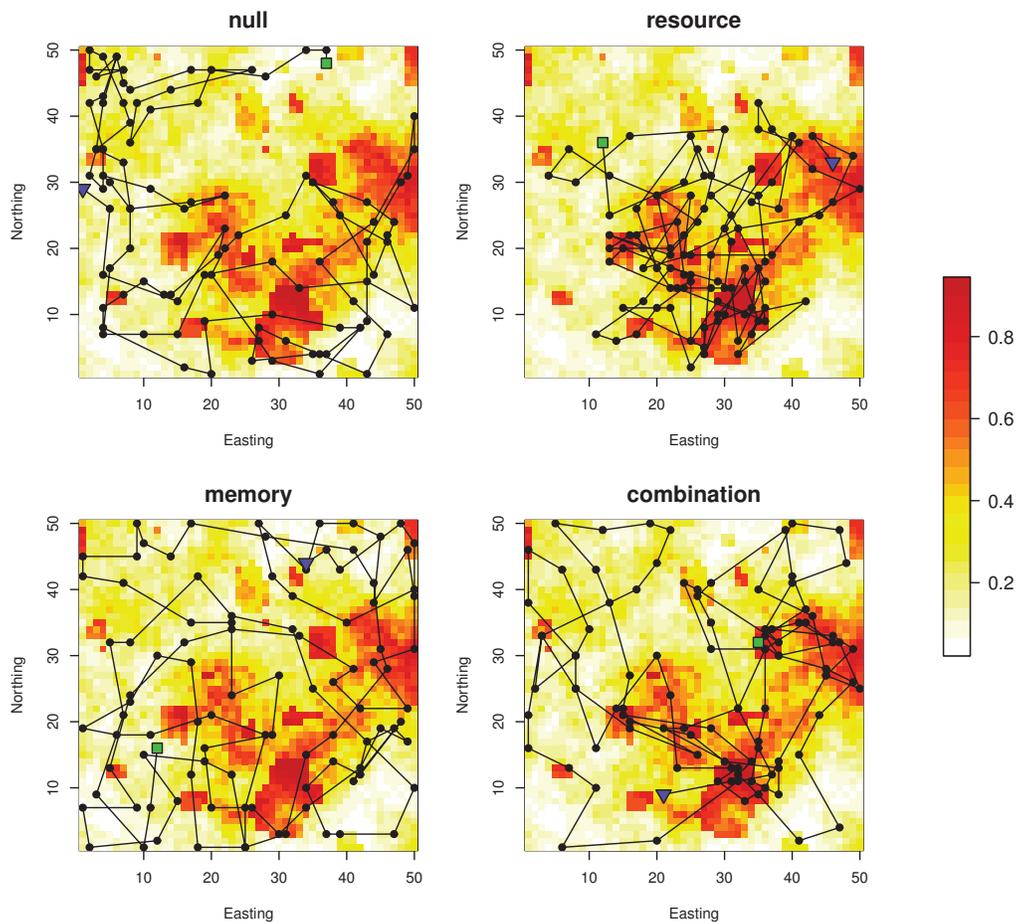
- 721 Owen-Smith, N., Fryxell, J. & Merrill, E. (2010) Foraging theory upscaled: the be-
722 havioural ecology of herbivore movement. *Philosophical Transactions of the Royal*
723 *Society B: Biological Sciences*, **365**, 2267–2278.
- 724 Plummer, M. (2013) *rjags: Bayesian graphical models using MCMC*. R package version
725 3-11. URL <http://CRAN.R-project.org/package=rjags>
- 726 Potts, J.R., Bastille-Rousseau, G., Murray, D.L., Schaefer, J.A. & Lewis, M.A. (2014)
727 Predicting local and non-local effects of resources on animal space use using a mech-
728 anistic step selection model. *Methods in Ecology and Evolution*, **5**, 253–262.
- 729 R Core team (2013) *R: A Language and Environment for Statistical Computing*. R
730 Foundation for Statistical Computing, Vienna. URL <http://www.R-project.org/>
- 731 Reid, C.R., Latty, T., Dussutour, A. & Beekman, M. (2012) Slime mold uses an exter-
732 nalized spatial “memory” to navigate in complex environments. *Proceedings of the*
733 *National Academy of Sciences of the United States of America*, **109**, 17490–17494.
- 734 Rhodes, J.R., McAlpine, C.A., Lunney, D. & Possingham, H.P. (2005) A spatially explicit
735 habitat selection model incorporating home range behavior. *Ecology*, **86**, 1199–1205.
- 736 Smouse, P.E., Focardi, S., Moorcroft, P.R., Kie, J.G., Forester, J.D. & Morales, J.M.
737 (2010) Stochastic modelling of animal movement. *Philosophical Transactions of the*
738 *Royal Society B: Biological Sciences*, **365**, 2201–2211.
- 739 Solymos, P. (2010) dclone: Data Cloning in R. *The R Journal*, **2**, 29–37.
- 740 Squires, J.R., DeCesare, N.J., Olson, L.E., Kolbe, J.A., Hebblewhite, M. & Parks, S.A.
741 (2013) Combining resource selection and movement behavior to predict corridors for
742 Canada lynx at their southern range periphery. *Biological Conservation*, **157**, 187–195.

- 743 Sulikowski, D. & Burke, D. (2011) Movement and memory: different cognitive strategies
744 are used to search for resources with different natural distributions. *Behavioral Ecology*
745 *and Sociobiology*, **65**, 621–631.
- 746 Sutherland, W.J., Freckleton, R.P., Godfray, H.C.J., Beissinger, S.R., Benton, T.,
747 Cameron, D.D., Carmel, Y., Coomes, D.A., Coulson, T. & Emmerson, M.C. (2013)
748 Identification of 100 fundamental ecological questions. *Journal of Ecology*, **101**, 58–67.
- 749 Tolman, E.C. (1948) Cognitive maps in rats and men. *Psychological Review*, **55**, 189–208.
- 750 Tsoar, A., Nathan, R., Bartan, Y., Vyssotski, A., Dell’Omo, G. & Ulanovsky, N. (2011)
751 Large-scale navigational map in a mammal. *Proceedings of the National Academy of*
752 *Sciences*, **108**, E718–E724.
- 753 Yeap, W. (2014) On Egocentric and Allocentric Maps. *Spatial Cognition IX* (eds
754 C. Freksa, B. Nebel, M. Hegarty & T. Barkowsky), pp. 62–75. Springer International
755 Publishing, Cham.

756 **Supporting Information**

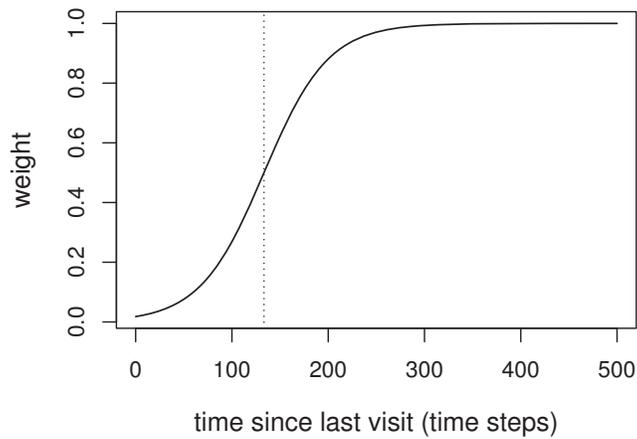
757 Additional Supporting Information may be found in the online version of this article:

- 758 • Appendix A Supplemental methods & Appendix B Supplemental results
- 759 • Appendix C Animated trajectories
- 760 • R and JAGS code, simulated data example



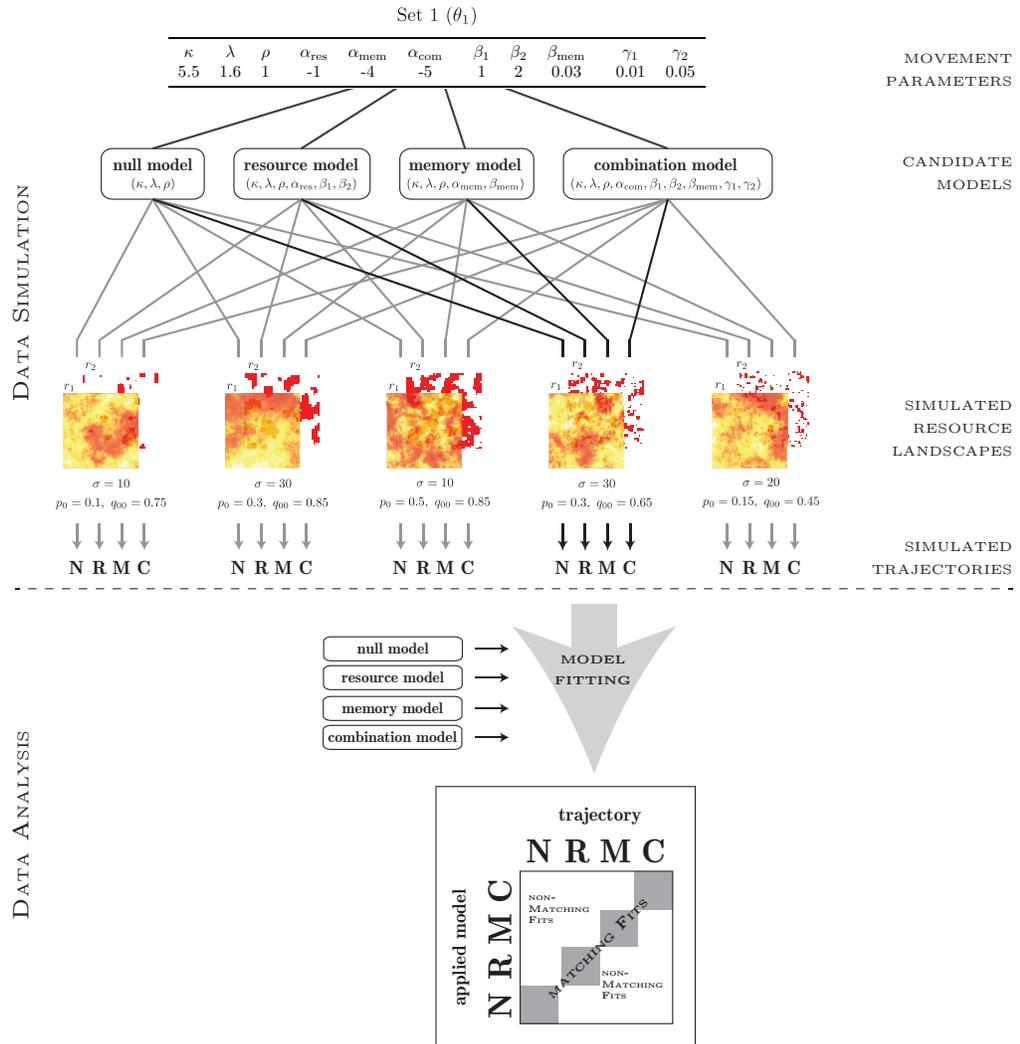
762

763 **Fig. 1.** Example trajectories from the four candidate models, 100 steps long, with
 764 starting location marked by a green box and final location marked by a blue triangle.
 765 All trajectories are plotted on top of an example resource selection function
 766 $w(x; \alpha, \beta) = [1 + \exp(-\alpha - \beta_1 r_1(\mathbf{x}) - \beta_2 r_2(\mathbf{x}))]^{-1}$ generated from two resources r_1
 767 and r_2 . The null model does not consider resource information and therefore the null
 768 trajectory visits locations irrespective of the resource selection function. The memory
 769 model does not consider resource information either, however, the animal avoids
 770 recently visited locations and is attracted to locations with long time since last visit.
 771 Therefore, the memory trajectory efficiently explores the spatial domain in a patrolling
 772 fashion. In contrast, the resource trajectory mainly remains in areas where the resource
 773 selection function has high values. The combination trajectory shows a mixture of
 774 behaviours from the resource and the memory model. The trajectories were generated
 775 using the first landscape pair and main parameter set from the simulation study;
 776 compare Fig. 3 and Appendix A.1.



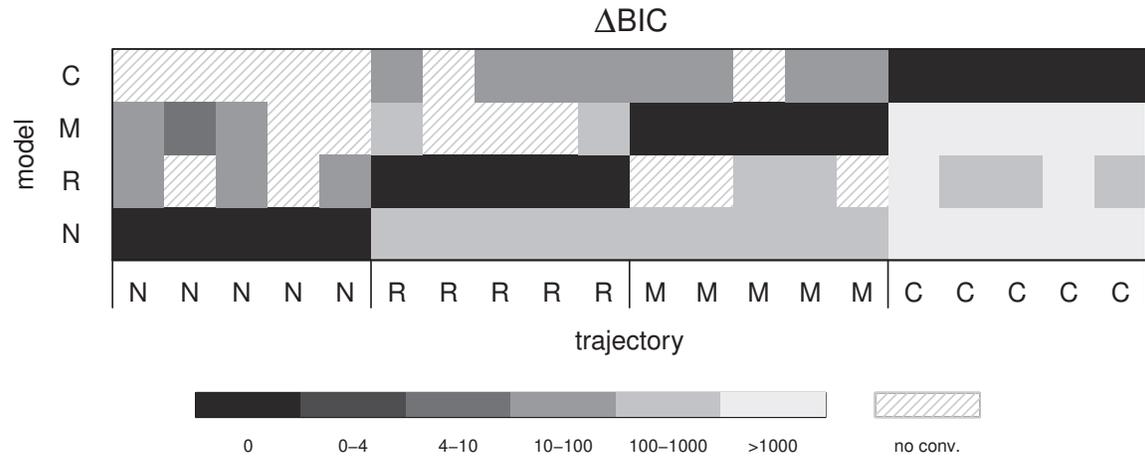
778

779 **Fig. 2.** In the memory model, the weight $w_t(\mathbf{x})$ of a location \mathbf{x} depends on time since
 780 last visit $m_{t-1}(\mathbf{x})$ to that location. Locations that have been visited recently have low
 781 weights and are thus avoided. A weight of 0.5 is attained when $m_{t-1}(\mathbf{x}) = -\frac{\alpha}{\beta}$ (dotted
 783 vertical line).



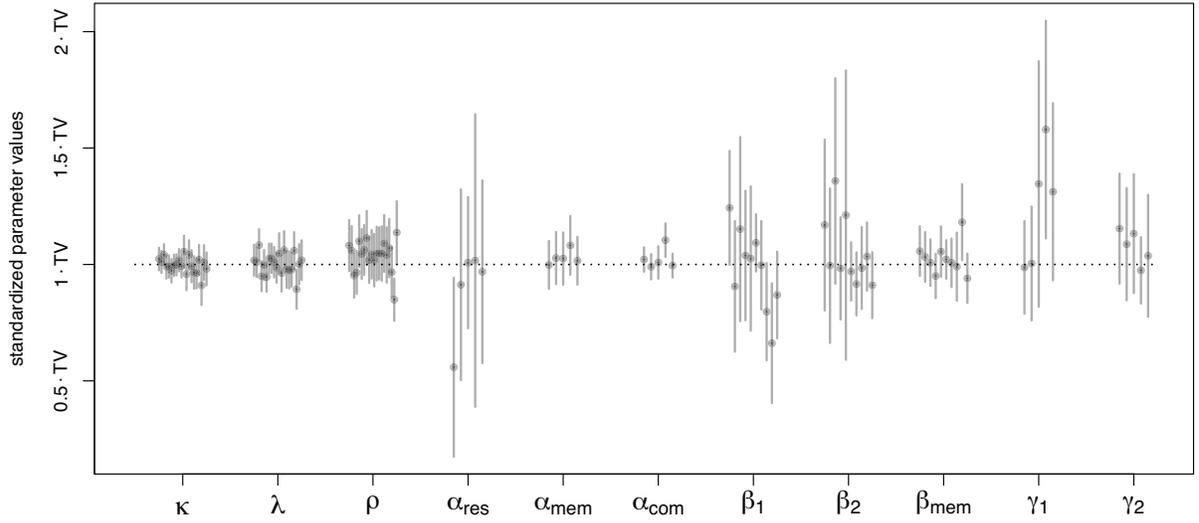
784

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



790

791 **Fig. 4.** Each column shows model selection results for one simulated trajectory when
 792 fitted with the null (N), resource (R), memory (M) and combination(C) model. For
 793 each trajectory, we calculated BIC values for the four fitted models, and the figure
 794 shows differences in BIC with respect to the minimal BIC value, i.e. the model with
 795 minimal BIC has $\Delta BIC = 0$. We excluded model fits with non-convergent MCMC.
 796 Triangles indicate trajectories for which we calculated estimability diagnostics;
 797 Appendix B.3.



799

800 **Fig. 5.** Parameter estimates and their 95% confidence intervals for matching model
 801 fits (each trajectory fitted with the same model that was used to generate the
 802 trajectory). Both parameter estimates and Wald-type confidence intervals are scaled by
 803 the true parameter values (TV): $\kappa = 5.5$, $\lambda = 1.6$, $\rho = 1$, $\alpha_{\text{res}} = -1$, $\alpha_{\text{mem}} = -4$,
 804 $\alpha_{\text{com}} = -5$, $\beta_1 = 1$, $\beta_2 = 2$, $\beta_{\text{mem}} = 0.03$, $\gamma_1 = 0.01$, $\gamma_2 = 0.05$.