
Article:

Blackwell, P.G. orcid.org/0000-0002-3141-4914, Niu, M., Lambert, M.S. et al. (1 more author) (2016) Exact Bayesian inference for animal movement in continuous time. Methods in Ecology and Evolution, 7 (2). pp. 184-195.

https://doi.org/10.1111/2041-210X.12460

Methods in Ecology and Evolution

Methods in Ecology and Evolution 2016, **7,** 184–195 doi: 10.1111/2041-210X.12460

Exact Bayesian inference for animal movement in continuous time

Paul G. Blackwell[1]*, Mu Niu[2], Mark S. Lambert[3] and Scott D. LaPoint[4,5]

[1]School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield S3 7RH, UK; [2]School of Mathematics and Statistics, University Gardens, University of Glasgow, Glasgow G12 8QW, UK; [3]National Wildlife Management Centre, Animal and Plant Health Agency, Sand Hutton, York YO41 1LZ, UK; [4]Department of Migration and Immuno-ecology, Max-Planck-Institute for Ornithology, Am Obstberg 1, 78315 Radolfzell, Germany; and [5]Department of Biology, University of Konstanz, Universitätsstraβe 10, 78464 Konstanz, Germany

---

Summary

1. It is natural to regard most animal movement as a continuous-time process, generally observed at discrete times. Most existing statistical methods for movement data ignore this; the remainder mostly use discrete-time approximations, the statistical properties of which have not been widely studied, or are limited to special cases. We aim to facilitate wider use of continuous-time modelling for realistic problems.

2. We develop novel methodology which allows exact Bayesian statistical analysis for a rich class of movement models with behavioural switching in continuous time, without any need for time discretization error. We represent the times of changes in behaviour as forming a thinned Poisson process, allowing exact simulation and Markov chain Monte Carlo inference. The methodology applies to data that are regular or irregular in time, with or without missing values.

3. We apply these methods to GPS data from two animals, a fisher (*Pekania [Martes] pennanti*) and a wild boar (*Sus scrofa*), using models with both spatial and temporal heterogeneity. We are able to identify and describe differences in movement behaviour across habitats and over time.

4. Our methods allow exact fitting of realistically complex movement models, incorporating environmental information. They also provide an essential point of reference for evaluating other existing and future approximate methods for continuous-time inference.

Key-words: Bayesian statistics, exact simulation, fisher, GPS data, Markov chain Monte Carlo, movement modelling, switching diffusion, wild boar

Introduction

Understanding the movement behaviour of individual animals is an important challenge in ecology, with improvements in tagging technology permitting the collection of data on an increasingly wide range of species. Statistical methodology for such data has received considerable attention in recent years – for example Blackwell (2003); Morales *et al.* (2004); Johnson *et al.* (2008); Langrock *et al.* (2014) – but still lags behind the technology and fieldwork. While it is natural to think of the underlying movement process of an animal as taking place in continuous time (Harris & Blackwell 2013; Fleming *et al.* 2014a), the majority of analysis is based on inherently discrete models, in which the (usually arbitrary) interval between observations is treated as the starting point. This leads to problems with interpretability, handling of irregular or missing observations, etc., discussed at length in Harris & Blackwell (2013) and in a simulation context by Avgar, Deardon & Fryxell (2013). Those analyses that do respect the continuous-time nature of movement generally nevertheless involve approximating the model on the discrete time-scale of the observations, introducing approximation error that is hard to quantify. The few exceptions, where analysis without approximation error has been carried out, mostly relate to specific models with quite limiting assumptions (Dunn & Gipson 1977; Blackwell 2003; Johnson *et al.* 2008), though Fleming *et al.* (2014b) give a rather flexible approach based on the mean and autocorrelation function of the underlying movement process.

In this study, we introduce a new statistical method that allows the fully Bayesian fitting of a rich class of continuous-time models, incorporating behavioural switching and heterogeneity in both space and time, without the need for any approximation error from time discretization. This opens the way for improved understanding of a range of movement data sets, as well as being a starting point for better understanding of the properties of existing methods.

The structure of the remainder of the article is as follows. First, we introduce the class of switching diffusion models that we will use, with some brief examples. The key idea that enables inference for these models is introduced as a simulation technique and then developed as a Markov chain Monte Carlo (MCMC) algorithm. The following sections introduce two

---

*Correspondence author. E-mail: p.blackwell@sheffield.ac.uk

data sets of GPS fixes on different species, fisher *(Pekania [Martes] pennanti)* and wild boar *(Sus scrofa)*, describe continuous-time models that capture key features of their movement, and give results of analyses using our new methods. Finally, we discuss wider implications and further work.

Diffusion models for movement

Continuous-time models for movement are usually taken to be diffusion processes, the simplest of which is Brownian motion, the limit of a random walk as the time interval tends to zero. More general diffusion models for movement could be defined as solutions to stochastic differential equations, as in, for example, Brillinger *et al.* (2002); Preisler, Ager & Wisdom (2013). Here, we follow the general approach of Harris & Blackwell (2013), seen also in the applications discussed below, of building potentially complex models from simple tractable components. Our basic component will be the two-dimensional Ornstein–Uhlenbeck (OU) process, which can be defined by

$$\mathbf{x}(s+t)|\mathbf{x}(s) \sim MVN(\boldsymbol{\mu} + e^{Bt}(\mathbf{x}(s) - \boldsymbol{\mu}), \Lambda - e^{Bt}\Lambda e^{B't}),$$

as in Dunn & Gipson (1977); Blackwell (1997, 2003); Harris & Blackwell (2013); note that the OU process includes as a limiting case Brownian motion with (possibly zero) constant drift $\boldsymbol{\alpha}$,

$$\mathbf{x}(s+t)|\mathbf{x}(s) \sim MVN(\mathbf{x}(s) + t\boldsymbol{\alpha}, t\Sigma).$$

MULTISTATE/SWITCHING DIFFUSION MODELS

For realism in modelling animal movement, we need to combine simple models through a framework of behavioural switching, whereby animals switch between behavioural states with different movement characteristics (Blackwell, 1997, 2003; Morales *et al.* 2004; Harris & Blackwell 2013; Nams 2014). In mathematical terms, we can represent this as a Markov process in continuous time with both a diffusion component – location, in this context – and a discrete one – behaviour – as in Berman (1994).

In the simplest cases, in which the transition rates between behaviours do not depend on location, this reduces to a switching diffusion driven by a continuous-time Markov chain, as in Blackwell (2003). In most applications, however, it is desirable to allow the rate of switching to depend on spatial covariates, through location, and perhaps also on time – for example time of day or year, to model daily or yearly periodicity in behaviour. We can also consider dependence on the time since the previous change in behaviour, that is the semi-Markov case; we omit this for the majority of the paper, to avoid undue notational complexity. Even if the movement process within a particular type of behaviour is simple, this overall structure gives a rich class of models that fit well with ecologists' intuition. Harris & Blackwell (2013) explore this class; after a more precise definition, we outline some specific published examples.

Harris & Blackwell (2013) take the spatial heterogeneity to be discrete; while that suffices for many examples, including those explored in detail here, the methods that we describe generalize immediately to the case of continuous variation in spatial covariates, with switching rates depending in some systematic way on the covariates, and our notation reflects that extra generality.

Separable models

We will restrict attention to what Harris & Blackwell (2013) call 'separable' models. These are defined to be models in which transition rates between behavioural states depend on location, but movement parameters and trajectories do not depend directly on location. So an animal responds to a change in the environment in which it is moving by changing its tendency to start behaving in a particular way – which will then affect its pattern of movement – rather than by instantaneously switching to a new behaviour. As Harris & Blackwell (2013) discuss, this is less of a limitation than it might appear at first.

Mathematical formulation and notation

Let $\mathbf{x}(t)$, $J(t)$ represent the animal's position and behavioural state at time t, and let $\lambda_{ij}(t, \mathbf{x})$ represent the switching rate from behaviour i to behaviour j at time t, given location $\mathbf{x}$, so that

$$Pr(J(t+\delta t) = j|J(t) = i, \mathbf{x}(t) = \mathbf{x}) \approx \lambda_{ij}(t, \mathbf{x})\delta t, i \neq j.$$

The animal starts in some state $J(0) = i, \mathbf{x}(0) = \mathbf{x}_0$, then follows the ith movement model – that is $\mathbf{x}(t)$ is a realization of the ith diffusion process, starting at $\mathbf{x}_0$ – until the time of the first switch in behaviour, at time T_1, when the animal is at $\mathbf{x}(T_1)$. If the behaviour switched to is $J(t) = j$, the next part of the location trajectory is a realization of the jth diffusion process, starting at $\mathbf{x}(T_1)$, and so on.

OUTLINE EXAMPLES

Our two main examples are given after the development of the statistical methodology. Here, we outline a number of other existing applications, to illustrate the class of models covered and hence motivate our methodology. The models of wood mouse movement in Blackwell (1997, 2003) are spatially homogeneous cases of switching OU processes. The analysis by Natvig & Subbey (2011) of vertical movement of cod, based on depth data from recovered data storage tags, uses a mixture of one-dimensional OU processes, with fish switching between points of attraction at different depths. The analysis by Fleming *et al.* (2014a) of the foraging of gazelles, based on semi-variogram methods, does not explicitly use switching models, but their 'OU Motion with Foraging' model is a hybrid of the OU and Integrated OU (IOU; see Discussion) processes that features correlated velocities at short lags and restricted space use at larger scales; see also Fleming *et al.* (2014b). Finally, Harris & Blackwell (2013) discuss how various other published models, not precisely in this class, can be approximately – and sometimes advantageously – represented by switching diffusions.

Simulation

Our aim is to carry out fully Bayesian inference (see e.g. Gelman *et al.* 2013) for the above models, without any approximation due to discretizing time. The key to doing so via Markov chain Monte Carlo is the simulation of trajectories augmented by switching times/locations, appropriately conditioned on the observed data; with that in mind, we introduce the central idea by giving an algorithm for exact simulation of these models.

Note that such simulation is not trivial. If behaviour does not depend on location or time, so switching is 'completely random' and $\lambda_{ij}(t, \mathbf{x}) \equiv \lambda_{ij}$, then it is straightforward to simulate the whole trajectory exactly (Blackwell 2003) by first simulating every switch in behaviour. However, in a more realistic model, $\lambda_{ij}(t, \mathbf{x})$ depends on $\mathbf{x}$, and so we do not know the time of the switch unless we know $\mathbf{x}(t)$, at least up until the switch occurs. An obvious approach is to approximate by a discrete-time representation, but we wish to avoid the poorly understood discretization error involved.

EXACT SIMULATION

The key assumption we make concerns boundedness of the transition rates. This means that switches in behaviour are never forced to occur instantaneously, which fits naturally with the assumption of separability as defined above and, originally, by Harris & Blackwell (2013), that is that instantaneous changes in behaviour do not occur when an animal changes habitat. More precisely, let $\lambda_j(t, \mathbf{x}) = \sum_{i \neq j} \lambda_{ji}(t, \mathbf{x})$ represent the transition rate away from behaviour j at time t, when at location $\mathbf{x}$. We assume that $\lambda_j(t, \mathbf{x})$ is globally bounded above, and define

$$\kappa = \max_{j,t,\mathbf{x}} \{\lambda_j(t, \mathbf{x})\}.$$

(This is an extension of the idea of uniformization, originating with Jensen (1953). See Appendix S1 for further generalizations.) We can then think of the waiting time from any instant until the next switch in behaviour as being bounded below, in a probabilistic sense, by the time that would apply if the rate of switching was always κ. So, starting at $J(0) = j$, $\mathbf{x}(0) = \mathbf{x}$, we can simulate the process forward as follows. Let

$$T \sim \text{Exponential } (\kappa),$$

corresponding to the time of the first event of a process with constant rate κ. This is the first *potential* time at which a change in behaviour might occur. Since the behaviour on $[0,T)$ is then known to be j, we can simulate the location $\mathbf{x}(T)$ forward from $\mathbf{x}$ with movement model j. That in turn gives us the information to determine whether the potential switch at T is an *actual* switch, an event which has probability

$$\frac{\lambda_j(T, \mathbf{x}(T))}{\kappa}.$$

If so, we can sample the new state, with state i having probability proportional to $\lambda_{ji}(T, \mathbf{x}(T))$. Knowing $J(T)$ and $\mathbf{x}(T)$, we can iterate this procedure forwards.

Considered globally, the times of changes of behaviour form a point process (Guttorp & Minin 1995) with a rate dependent on the individual's movement. This point process can then be represented as a Poisson process, with rate κ, which has been 'thinned' – that is each point either retained or deleted, probabilistically (Guttorp & Minin 1995) – in a way that depends on the movement process. This leads to a natural and efficient way to extending the simulation over as long an interval as we desire. Denote the events of this rate κ Poisson process by $T_1, T_2, \ldots$. For each k in turn, we generate location $\mathbf{x}(T_k)$ by forward simulation; then, we decide whether there is a switch at T_k, which happens with probability

$$\frac{\lambda_{J(T_k)}(T_k, \mathbf{x}(T_k))}{\kappa}.$$

If so, we pick the new state, which is j with probability

$$\frac{\lambda_{J(T_k),j}(T_k, \mathbf{x}(T_k))}{\lambda_{J(T_k)}(T_k, \mathbf{x}(T_k))}.$$

In either case, we retain the simulated $\mathbf{x}(T_k)$ as part of the trajectory. Figure 1 illustrates this process. Effectively, we are simulating the process at a 'grid' of time points, but the careful use of random time points, rather than the usual regularly spaced points, enables us to avoid any discretization error.

Inference by Markov chain Monte Carlo

We can build on the simulation idea of the previous section to produce a Markov chain Monte Carlo (MCMC; see e.g. Gelman *et al.* 2013) algorithm for Bayesian inference for these models. Given data $\mathbf{x}(t_0), \ldots, \mathbf{x}(t_N)$, we want to sample from the posterior distributions for the parameters of the movement components and of the switching rates.

Our approach involves augmenting the data with the times of all changes of behavioural state, and associated locations, as in the homogeneous case in Blackwell (2003); to do so, we actually sample times, locations and states for all *potential* changes, that is at all times of a Poisson(κ) process.

The state of our chain is the collection of all times of potential switches throughout the dataset, $\mathcal{T} = \{T_{ck}, k = 1, \ldots, M_c, c = 0, \ldots, N-1\}$, where M_c is the number of potential switches with $t_c < T_{ck} < t_{c+1}$, plus associated locations $\mathbf{x}(T_{ck})$, initial state $J(t_0)$ and changes to states J_{ck} corresponding to some subset of times $\mathcal{R} \subset \mathcal{T}$, and implied states at the times of observations $J(t_1), \ldots J(t_N)$.

TRAJECTORY UPDATES

Our key MCMC step is to sample the trajectory – potential switches, locations and states – over some time interval $t_a, t_b, 1 \leq a < b \leq N$ conditional on the trajectory outside that interval, on the states $J(t_a), J(t_b)$ and on the movement and switching parameters. We use an independence sampler based on our simulation algorithm to propose the new trajectory. We define $\mathcal{T}'_{ab} = \{T'_{ck}, k = 1, \ldots, M'_c, c = a \ldots, b-1\}$ to be a realization of a Poisson process of rate κ on time interval t_a, t_b,

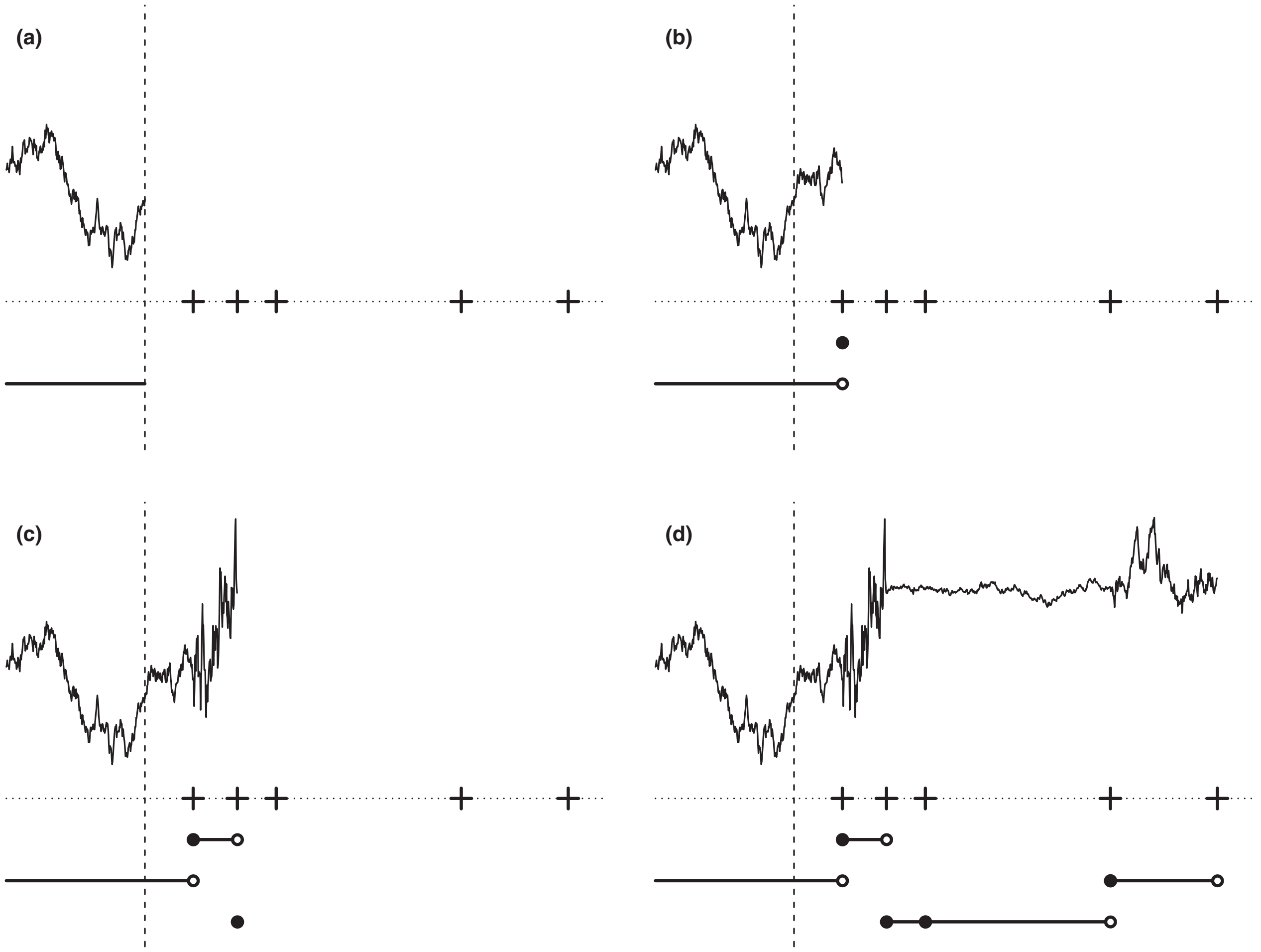

Fig. 1. Diagram of the simulation algorithm. The dotted horizontal line represents the time axis; movement is indicated in 1 dimension by the 'noisy' line above the time axis; behavioural state is indicated by the solid lines below. (a) The process has been simulated up to some initial time indicated by the vertical dashed line. Potential changes of state (+) are simulated as a Poisson process. (b) No change in state can occur until the first potential switch, so the process can be simulated forward in the initial state. The new location determines the probability of an actual switch. In this case, the behaviour changes from its current value (circle) to a new one (solid dot). (c) With the state for the next time interval known, the process can be simulated forward – using the new behaviour – and the next potential switch resolved. (d) This process is iterated to simulate as far as required. Here, the third potential switch shown happens not to be an actual switch.

representing the proposed times of potential switches. Then, we simulate the process forward as before.

Starting with $\mathbf{x}(t_a)$ and $J(t_a)$, we simulate $\mathbf{x}(T'_{a,1})$ from movement model $J(t_a)$, and set the state

$$J'_{a,1} = \begin{cases} j & \text{with probability } \lambda_{J(t_a),j}(T'_{a,1}, \mathbf{x}(T'_{a,1}))/\kappa,\ j \neq J(t_a) \\ J(t_a) & \text{otherwise}. \end{cases}$$

Then, we simulate $\mathbf{x}(T'_{a,k+1})$ from movement model $J'_{a,k}$ and set

$$J'_{a,k+1} = \begin{cases} j & \text{with probability } \lambda_{J'_{a,k},j}(T'_{a,k+1}, \mathbf{x}(T'_{a,k+1}))/\kappa,\ j \neq J'_{a,k} \\ J'_{a,k} & \text{otherwise}, \end{cases} \qquad \text{eqn 1}$$

for $k = 1, \ldots, M'_a - 1$. This defines a new proposed state $J'(t_{a+1})$; $\mathbf{x}(t_{a+1})$ is part of the data, so is fixed. This proposal process is repeated on each subinterval $(t_{c+1} - t_c)$, $c = a, \ldots b - 1$.

Since we are simply simulating forward, we need also to take into account the consistency of our simulated path with the observed data in the interval, $\mathbf{x}(t_{a+1}), \ldots, \mathbf{x}(t_b)$, and with the data and augmentation outside the interval. Therefore, the simulated path is merely a proposal; it may be accepted or rejected, as with a more conventional Metropolis–Hastings update (see e.g. Gelman *et al.* 2013). Given the algorithm as described so far, we require for consistency that the final simulated state at $T'_{b-1,M'_{b-1}}$ matches the existing augmentation $J(t_b)$; if not, rejection is automatic. Otherwise, we need to condition on the locations $\mathbf{x}(t_{a+1}), \ldots, \mathbf{x}(t_b)$. Our proposal comes from the (conditional) prior, so the Hastings ratio is

$$\prod_{c=a}^{b-1} \frac{f(\mathbf{x}(t_{c+1})|\mathbf{x}(T'_{c,M'_c}), J'_{c,M'_c})}{f(\mathbf{x}(t_{c+1})|\mathbf{x}(T_{c,M_c}), J_{c,M_c})} \qquad \text{eqn 2}$$

and we accept or reject the proposed reconstruction accordingly.

Because of the possibility of rejection, it is necessary that the algorithm as a whole starts with a feasible path, even though the proposal does not depend on the current path (except at the fixed endpoints). Initialization of the algorithm therefore involves augmentation with some switching points as well as an initial assignment of states.

PARAMETER UPDATES

Given the augmented trajectory, we know exactly when the animal was following each of its different movement processes, so the inference about the movement parameters is straightforward. Often, we can adequately represent prior knowledge about them via conjugate priors, and then carry out Gibbs sampling (see e.g. Gelman *et al.* 2013); this may be possible even when some parameters are shared between states. See Blackwell (2003) for details. Alternatively, we can use standard random-walk Metropolis–Hastings updates for these parameters.

With the augmentation described, we also know exactly the times of changes in state. However, because the transition rates may depend on location, and location is known only at selected times, we do not have 'full' information about the rates as in the spatially homogeneous case, even with the augmentation. Instead, we have complete information about κ from the process of potential switches, and 'sample' information about the rates in general, relative to κ, from the changes in state (or otherwise) at those potential switches. The likelihood for the rate parameters is thus the product of a likelihood based on $\sum M_c \sim \text{Poisson}((t_N - t_0)\kappa)$ and on single-trial multinomial (or 'categorical') likelihoods based on probabilities like those in eqn 1. We can regard κ as being determined by the $\lambda_{ij}(\cdot)$s, and they can all be updated simultaneously using a simple Metropolis–Hastings step. More simply, we may instead choose to use priors that are bounded above for all $\lambda_{ij}(\cdot)$s, and keep κ fixed at the maximum of those upper bounds. Then the likelihood consists only of multinomial terms, and it may be possible to carry out Gibbs updates on the $\lambda_{ij}(\cdot)$s, depending on how those rates are parameterized.

In principle, the above updates of behaviour, path and parameters are sufficient to give a complete MCMC algorithm. In practice, mixing can be improved substantially by including some 'local' updates to the trajectory (see Appendix S1). Some other details and refinements necessary for the fully worked examples are also described in Appendix S1; a number of other variants or extensions are covered in the closing discussion.

Example: fisher

BACKGROUND AND MODELLING

Our first example uses a subset of data on a male fisher collected by SDL in suburban New York (LaPoint *et al.* 2013a, b). Fishers are medium-sized, terrestrial carnivores typically found within the boreal and northern deciduous forests of much of North America (Powell 1993). A GPS collar attached to the fisher attempted to acquire the animal's location every 10 min; a few observations are missing, leading to longer intervals. A corresponding map of the local environment is also available from the US National Land Cover Database (NLCD2006 Land Cover layer; Fry *et al.* 2011), as a discrete set of known habitat types, defined on a grid with 30m×30m pixels, shown in Fig. 2 along with the observations that we will be using. We selected a 24-h extract from the data, long enough to exhibit a range of behaviours but short enough that the fisher's movement is clearly centred on a single point of attraction, a short-term resting place. Interest thus focusses on the dependence of the dynamics of the movement on the habitat.

A possible type of model for this situation is what might be called 'adaptive' movement, in which there is a single movement process associated with each habitat type. To maintain separability, switching to the associated movement process does not occur instantaneously on entering a given habitat; this would be unrealistic in this example, as in many others, in that the mapped boundaries are clearly not perfectly accurate. Instead, behaviour switches from state i to state j at some finite rate *only* when the animal is in habitat j. In this example, we assume time-homogeneity; writing the n habitat regions as $A_1, \ldots, A_n$, the transition rates are thus

$$\lambda_{ij}(t, \mathbf{x}) = \begin{cases} \alpha_{ij} & \mathbf{x} \in A_j \\ 0 & \text{otherwise}, \end{cases}$$

and for this particular extract from the data, we have n=3, as can be seen in Fig. 2. The behavioural aspects of this model are

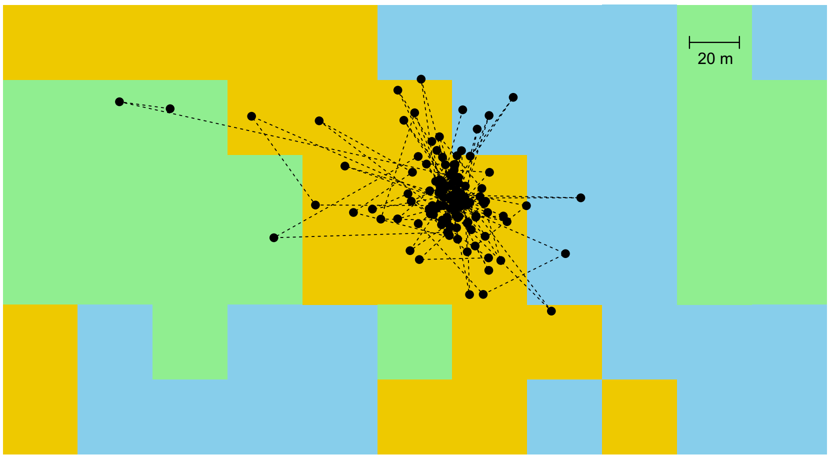

Fig. 2. Fisher locations (points, linked chronologically with dashed lines) and habitat data (coloured 30 m by 30 m grid cells)). The habitats are (1) 'developed open space' (central band, orange); (2) 'deciduous forest' (mostly to the left, green) and (3) 'woody wetlands' (mostly to the right, blue).

similar to those of the 'patch' model of Beyer *et al.* (2013), though they only simulate from discrete-time approximations to their models.

There are several ways in which differences in movement behaviour between the three states could be modelled. To illustrate the idea, bearing in mind the subset of data selected above, we take all three movement processes to be OU processes, with a common point of attraction $\boldsymbol{\mu}$, but different values of the long-term covariance matrix Λ and of the matrix parameter B controlling the dynamics of the process. All three values for both matrices are taken to be multiples of the identity matrix with different scalar multipliers, that is $B_i = b_i I_3, \Lambda_i = v_i I_3, i = 1, 2, 3$, an assumption of circularity which is not essential to the details here. Blackwell (1997) discusses some of the modelling issues in setting the form of B. We consider some variations of this model when fitting to the data.

For longer term modelling of the movement of this fisher, we would need to allow for additional states with different centres of attraction, not unlike the wild boar model below. That in itself does not involve any new difficulties, but we prefer to keep this first example as simple as possible.

GPS error is relatively small compared with other forms of location data, and stationary field tests (LaPoint *et al.* 2013a) suggest that the error here is smaller than the resolution of the habitat data. For the purposes of illustrating the methods, we neglect observation error; this is revisited in the Discussion.

RESULTS

The results here are based on prior distributions that are the same across all states, and on 10 million iterations of the MCMC algorithm after burn-in; for further details of implementation, see Appendix S2.

The posterior distributions for the parameters varying between states, b_i and v_i, are shown in Fig. 3, thinned by a factor of 1000 purely for ease of presentation. There are clear differences in the parameters, and hence the movement, between states and therefore between habitats. States 1 and 3 have similar values for b_i, but very distinct values for v_i; the difference in posterior medians is around 2·6 on the log scale, or a factor of about 14. The parameters for state 2 are less well estimated, because of the small amount of time spent in the corresponding habitat, and there is a clear overlap of posterior distributions with state 3, though the most likely values are quite different for b_i, the speed of 'reversion' to the mean location. Comparison with a typical discrete-time approximation (details omitted) also shows that the method here extracts much more of the limited information available from these observations.

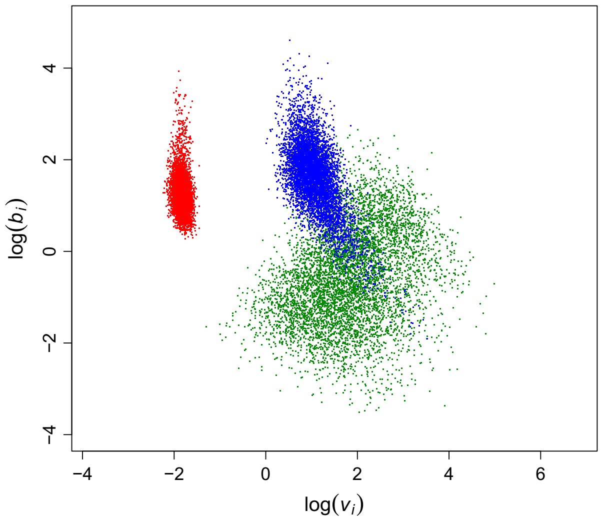

Fig. 3. State-dependent parameters for the fisher model. The three clusters correspond to states 1–3 (red, green, blue, respectively).

Posterior distributions for the other parameters, μ (Fig. S1) and λ_{ij} (Fig. S2), are given in the Supporting Information. The former are much more informative; this is as expected, since all observations are affected by μ, whereas only the (uncertain) transitions give information about the λ_{ij}s.

To illustrate the methodology, Fig. 4 shows two portions (for clarity) of a sampled realization of the trajectory, from the MCMC output. Figure 4(a) covers all three habitats, and states, in an interval of around 20 min; the fisher moves from left to right, with changes in behaviour taking place soon after entering each habitat, as shown by the triangular points. Figure 4(b) covers a 1-h interval which is more typical, with the fisher spending most of the time in habitat (and state) 1, with an excursion into habitat (and state) 3 encompassing a single observation (the rightmost square point). The model with state-dependent b_i and v_i can be compared with a homogeneous model with common values of b and v across all states. Direct use of the MCMC output again allows us to do this straightforwardly, by calculating the Deviance Information Criterion (DIC; Spiegelhalter *et al.* 2002). The more detailed 3-state model is overwhelmingly preferred, with $\Delta DIC \approx 154$. An intermediate model, merging only states 2 and 3, could be investigated in a similar way, for example if it was felt that the distinctions between some habitats might be irrelevant to the animal itself.

This model addresses one criticism of the simple Dunn & Gipson (1977) model, that it can represent only bivariate normal stationary distributions. Dunn & Brisbin (1985) mention a possible remedy based on transforming the underlying space, but the model here gives a more ecologically meaningful solution by exploiting known habitat information.

To explore the robustness of the modelling here, an alternative model was tried, which is simpler but less biologically interpretable, though in fact versions of it are widely used with approximate fitting techniques. Instead of the mixture of three OU processes, we fitted three Brownian motion models, each with a different diffusion parameter and all with zero drift. Details are omitted, but that analysis confirmed differences between the states.

A natural refinement would be to allow the transition rates to depend on the habitat in a more sophisticated way, varying close to boundaries in order to make the adaptation of movement to habitat more symmetric rather than lagging

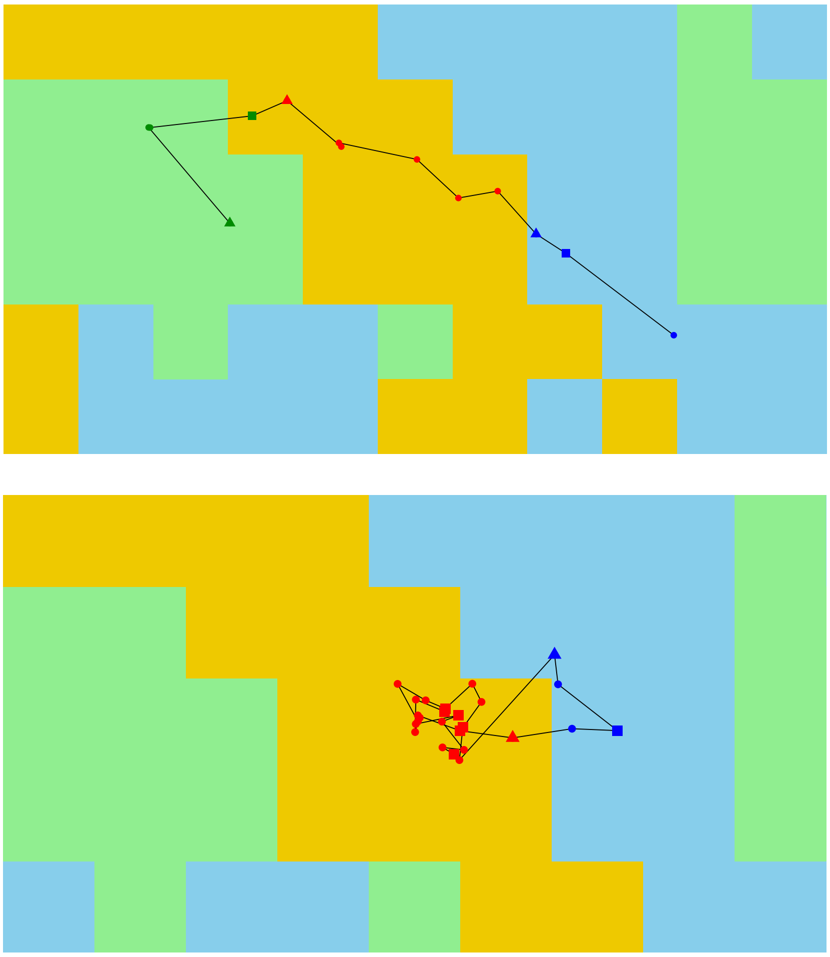

Fig. 4. Two partial reconstructed trajectories for the fisher model. Squares correspond to data points; triangles to switches in behaviour; dots to other potential switches. In part (a) (top), states 1–3 are shown in red, green, blue, respectively; in part (b) (bottom) only states 1 and 3 occur.

behind location. Such a model could still be fitted using the methodology described here.

Example: wild boar

BACKGROUND AND MODELLING

Our second example uses part of an extensive data set on wild boar *Sus scrofa*, collected by the Animal and Plant Health Agency's National Wildlife Management Centre (formerly part of the Food and Environment Research Agency) to investigate the possible side effects of immunocontraceptive vaccines (Quy *et al.* 2014). Animals were tracked using GPS collars, with fixes every 15 or 30 min (night/day). The movement behaviour represented by these data is quite complex; there is a clear, but not completely regular, 24-h cycle of activity, so some inhomogeneity of behaviour over time is necessary in the modelling. Each animal uses a number of distinct nests, returning to the same location sometimes on successive days but also after longer intervals. Here, we use 355 observations, shown in Fig. 5, for one female wild boar over a period of 5 days, short enough to facilitate visualization and interpretation, yet long enough to show the re-use of different nests and other key features of behaviour from the wider data.

An additional complexity is the uneven spacing of data. The intention was to record locations every 15 min while the animals are most active – at night, for this population – and every 30 min at times when they are likely to be less active, but the schedule of fixes is pre-programmed, and does not necessarily correspond to an individual's behaviour on a given day or night. There are also a few missing or slightly irregular observations, leading to a few intervals that are neither 15 nor 30 min.

While there is inevitably some observation error on the locations, it is small on the scale of the movements themselves. An upper bound is provided by the variation in observed locations during the day time, when the animal is less active. This variation is orders of magnitude smaller than occurs during active periods; neglecting the error may affect our assessment of the amount of movement during less active phases, but is unlikely to affect our overall conclusions. Here, we ignore any observation error; see the Discussion for the alternative.

We fit a multistate diffusion model to represent the underlying movement and capture some of these behavioural features; movement in each state is taken to follow an OU process. For present purposes, we treat the number of nests used by an individual as known; we denote this by n. We also divide behaviour into three types: resting at a nest, foraging and heading for a nest. This leads us to a model with $3n$ distinct states: resting in each of the nests, foraging prior to using each of the nests and heading towards each of the nests. For illustration, we assume

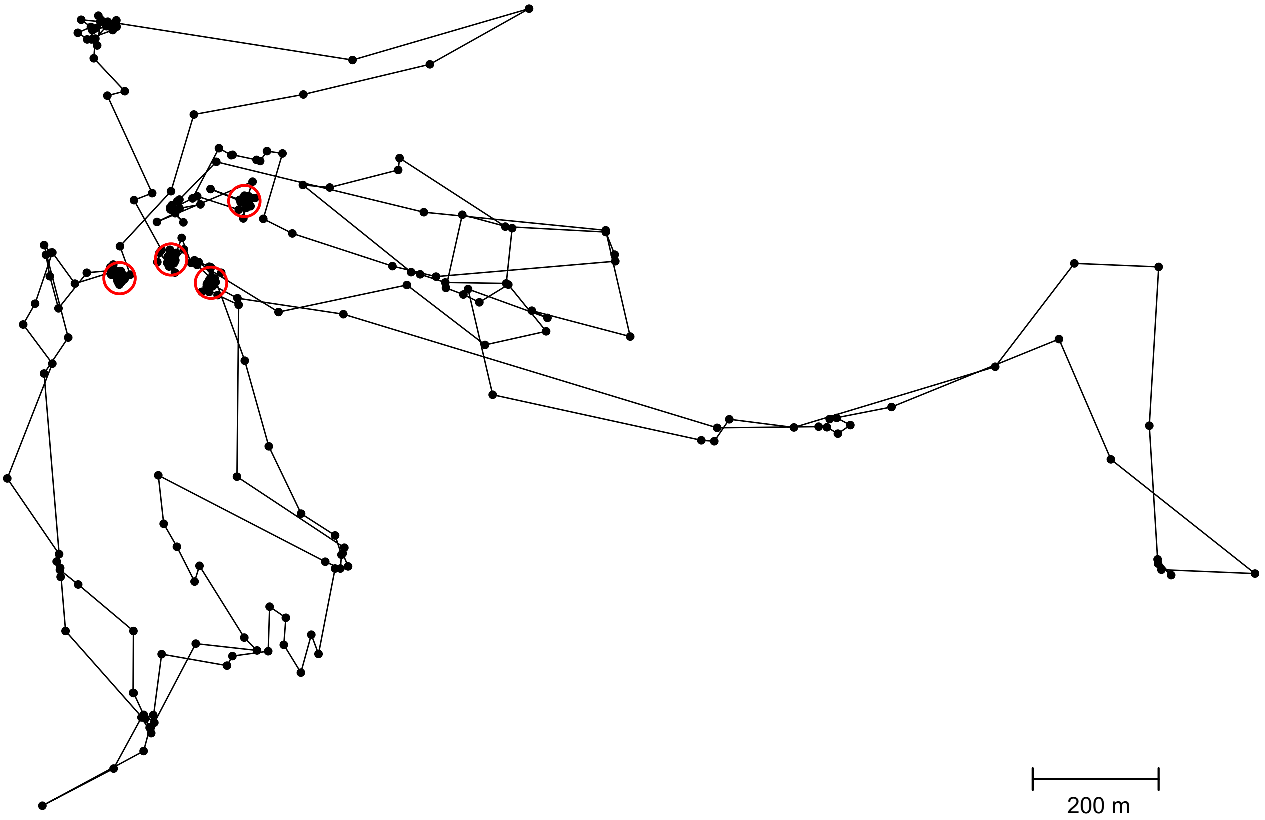

Fig. 5. Data used in wild boar example (black dots, linked in time sequence). The open red circles indicate nest locations; see text for details.

circularity of the OU processes, a common centre for each state relating to a particular nest, and common b and v parameters across all n states of each type (resting, foraging, returning). Thus, the $3n$ movement behaviours are described by a total of $2n + 6$ parameters: $\mu_1^x, \mu_1^y, \ldots, \mu_n^x, \mu_n^y, b_1, b_2, b_3, v_1, v_2, v_3$.

The $3n$ distinct states give a flexible structure to represent a range of specific modelling assumptions, through constraints on how the parameters differ between states. With the parameter constraints actually used, as described above, having n foraging states simply means that the animal's centre of attraction while foraging depends on where it will end that night's activity. Relaxing some constraints could model the idea that which nest is to be used might influence other parameters of the foraging itself. Having three *types* of behaviour seems to be the minimum to fit the range of activity adequately; more detailed modelling of behaviour might require more states, while remaining within this general class of models. A fuller exploration of these models is ongoing, but is beyond the scope of this study.

The form of the transition rates between states is more complex than with the fisher, because of the 'structure' of the states and the need to represent inhomogeneity over time. We give an overview here; details are given in Appendix S2, and examples are shown with the results.

The switch from 'resting' to 'foraging' is assumed to take place at a similar but not identical time each day. We set the rate of 'waking up', $\lambda_i(t), i = 1, \ldots, n$ to be

$$\lambda_i(t) = \frac{\alpha}{1 + \exp(-\alpha(t \bmod 1 - t_\alpha))}$$

for a 'most likely' time of day t_α, and rate parameter α, at time t measured in days; we denote this by $h_\alpha(t)$. This gives a transition rate that, over each day, increases steeply up to the time t_α and then asymptotically approaches α. Each of the n foraging states, $n + 1, \ldots, 2n$, is then assumed to be equally likely.

The transition from foraging to returning to a nest is assumed to have a similar form, denoted by $h_\beta(t)$, but in this case, the state switched *to* is necessarily $j = i + n$. The transition rate is taken to be independent of location, but it would be natural and straightforward to allow it to depend on, for example, distance from the destination nest.

Finally, the transition from returning to resting is taken to occur at a uniform (and presumably high) rate γ once the animal is within some distance ρ of its target nest.

This parameterization of the rates is fairly parsimonious, with just six parameters, $\alpha, t_\alpha, \beta, t_\beta, \gamma, \rho$, in addition to the nest locations already defined. The rates can be collected together as the generator (Guttorp & Minin 1995) of the behaviour process $\Lambda(t, \mathbf{x}) = (\lambda_{ij}(t, \mathbf{x}))$, as shown in Appendix S2.

RESULTS

We used proper but rather uninformative priors on most parameters; the number and locations of the nests were taken as known, since they are very clear from the data. This assumption can be relaxed; see Appendix S2 for details. Similarly, the parameter ρ is fixed on the basis of exploratory analysis. The results here are based on 6 million iterations of the MCMC algorithm after burn-in; further details of implementation are again included in Appendix S2.

The posterior distributions for the movement parameters in the three types of states are shown in Fig. 6, in the form of samples taken from the MCMC output. The results show clear differences between the three types of behaviour. Unsurprisingly, in the resting states, the variance v_i is much smaller, and the mean-reverting tendency b_i much stronger. For the other two types of states, the parameters are more similar, but clearly distinct. Estimation of the product $b_i v_i$ for those two types is also much more precise than estimation of either parameter individually, since that product determines individual step sizes more directly.

This shows that the model in which the boar essentially cycles through these three types of state each day does indeed capture differences in movement behaviour. However, some care is needed in the interpretation of the second and third types of states as 'foraging' and 'returning to nest'; the model only confirms that they have statistically distinct parameters. We need to consider how these types are used in the fitted reconstructions in order to interpret them. Figure 7 shows the assignment of the behavioural states to the observations; those that are essentially uniquely categorized are shown in green, blue and red for resting, foraging and returning to the nest, respectively, while those where there is more than a negligible probability (at least 0.002) of being in either a foraging or a returning state are shown in purple. A very few observations that may be classed as either resting or not are shown in their most likely state, for clarity. The sequence of events is clearer in Fig. 8 in which observations are plotted as x coordinates only, against time, with points coded in the same way. Figures S3 and S4 in the Supporting Information show an individual trajectory, sampled randomly from its posterior distribution and again plotted in the original space and against time, respectively.

There are two key messages from the allocations of states in Figs 7 and 8. First, allocation is usually clear but there can be considerable uncertainty about the transition between the states of types 2 and 3. Secondly, there can be real differences over time in the way in which activity is split between those two types. Bearing in mind the uncertainty on classification, it is clear that during the 1st, 3rd and 5th nights of activity covered by these data, a large proportion of the observations – perhaps nearly all of them – are classified as 'returning' while on the 2nd night, most of the observations are 'foraging'. The overall conclusion is that separation of the non-resting activity into one behaviour with high spread v and low reversion b and another with lower spread and higher reversion is meaningful, but that they should not necessarily be labelled in the way initially suggested. Study of a more extensive data set should help to clarify both the timing and the interpretation.

We can also look directly at the rate functions controlling the transitions between states. Figure 9 shows the rates in transitions per hour as a function of time of day, $h_\alpha(\cdot)$ and $h_\beta(\cdot)$, based on posterior median point estimates of α, t_α, β and t_β. Recall that α and β represent the *maxima* of time-varying rates. The rate of transitions from 'returning' to 'resting', γ, which is constant over time, is also shown for ease of comparison, using its posterior median. It is clear that $h_\alpha(\cdot)$ increases quickly over time, around the time t_α, whereas $h_\beta(\cdot)$ increases more slowly; the rate γ is estimated to be much higher than either. So the daily cycle of behaviour can be summarized as: start 'foraging' at a fairly well-defined time, close to t_α; switch to 'returning' at some time of night that is much more variable; switch to 'resting' very quickly once within range of the appropriate nest. For reference, Fig. S5 shows the posterior densities for the parameters α,β and γ and Fig. S6 shows the posterior densities for t_α and t_β.

Discussion

We have shown that, for a rich class of movement models built from standard modelling components, exact fitting to data

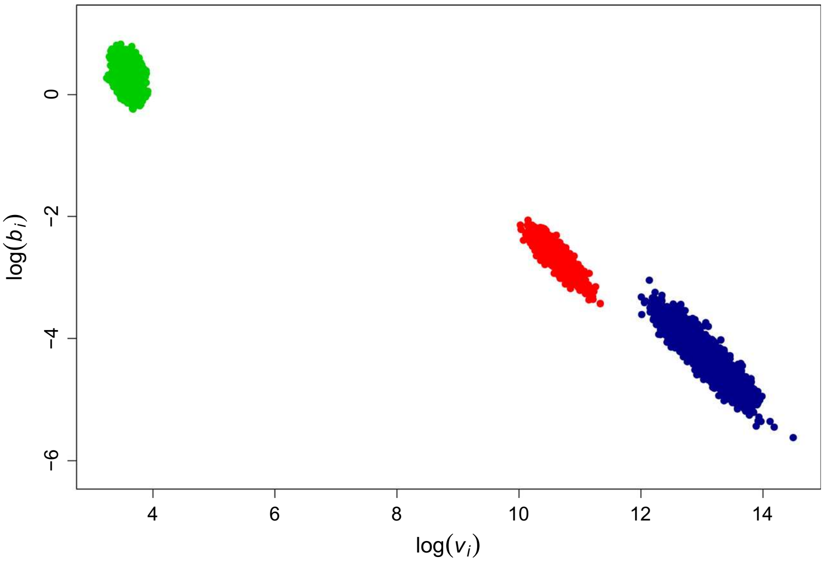

Fig. 6. State-dependent parameters for the wild boar model. The three clusters correspond to resting, returning and foraging states from left to right (green, red and blue, respectively).

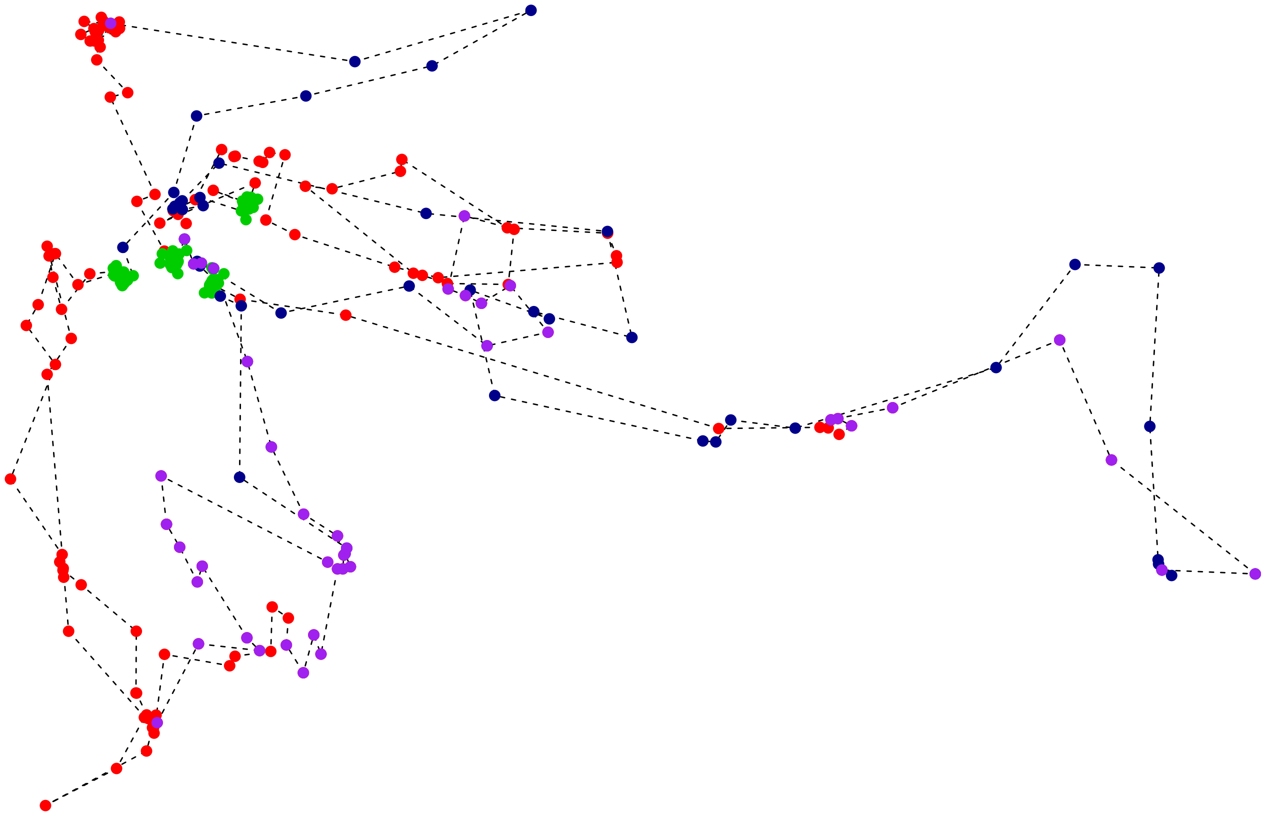

Fig. 7. Classification of observations in the wild boar model plotted in geographic space. Observations are shown here as resting (green), foraging (blue), returning (red) or uncertain (either foraging or returning – purple). See text for details.

using MCMC is possible. The key idea is to relate behavioural switching that may be complex and heterogeneous, both spatially and temporally, to a much simpler homogeneous Poisson process of potential switches. This enables exact inference for realistically complex models and opens up the prospect of more coherent and informative, and better understood, statistical inference for movement data in continuous time. We have illustrated the method for two real examples; both are spatially heterogeneous, and one also incorporates time-heterogeneity in behaviour.

The methodology applies to a wide range of models, and inevitably, for some of them, implementation requires detailed understanding of the modelling, statistical and computational issues. Nevertheless, some cases are rather easier to handle. Models which are time-homogeneous, and where the spatial heterogeneity involves a discrete set of habitats, such as the 'adaptive' model used in the fisher example, can be set up with a minimal level of coding. The implementation details given in that example are likely to be widely applicable, and to give at least reasonable performance, though there will always be scope for improvement through experimentation, as in more conventional MCMC inference. More complex models, like the wild boar example, are likely to require more user input. Work is underway on generic software for this methodology; in the meantime, those interested in applying this approach are encouraged to contact the first author.

In addition to implementation details, some aspects of the method itself could be tailored to particular models. For example, rather than a global upper bound κ on transition rates, it may be feasible to define state-dependent bounds $\kappa(i) = \max_{t,\mathbf{x}}\{\lambda_i(t,\mathbf{x})\}$ or time-dependent bounds $\kappa(t) = \max_{i,\mathbf{x}}\{\lambda_i(t,\mathbf{x})\}$. In general, this will not increase the range of models that can be fitted but potentially gains efficiency, as $\kappa(\cdot)$ can be more closely matched to the actual switching rates in a model, reducing the number of potential switches and hence the computational cost. There is a trade-off, since this loses the convenience of potential switches forming a homogeneous Poisson process when the bound is global; instead, they would form a time-inhomogeneous Poisson process, which would complicate the algorithm and introduce its own computational cost.

We have described and implemented the technique in terms of simulating partial trajectories from the model, as described above. In fact, a wide range of processes can be used for proposing new trajectories, provided allowance is made in the Hastings ratio in eqn 2. We have experimented with proposing locations from a Brownian bridge over the interval being updated, that is from $\mathbf{x}(t_a)$ to $\mathbf{x}(t_b)$ in the notation above, to benefit from the information in $\mathbf{x}(t_b)$. Our results so far suggest that a clear net computational gain may not be easy to obtain, but there are almost certainly cases in which more general proposals are warranted, and our experimentation continues.

In the current modelling, we have ignored the measurement error on locations, but for some applications this will be important. Measurement error can be incorporated by separating the movement and observation processes, and

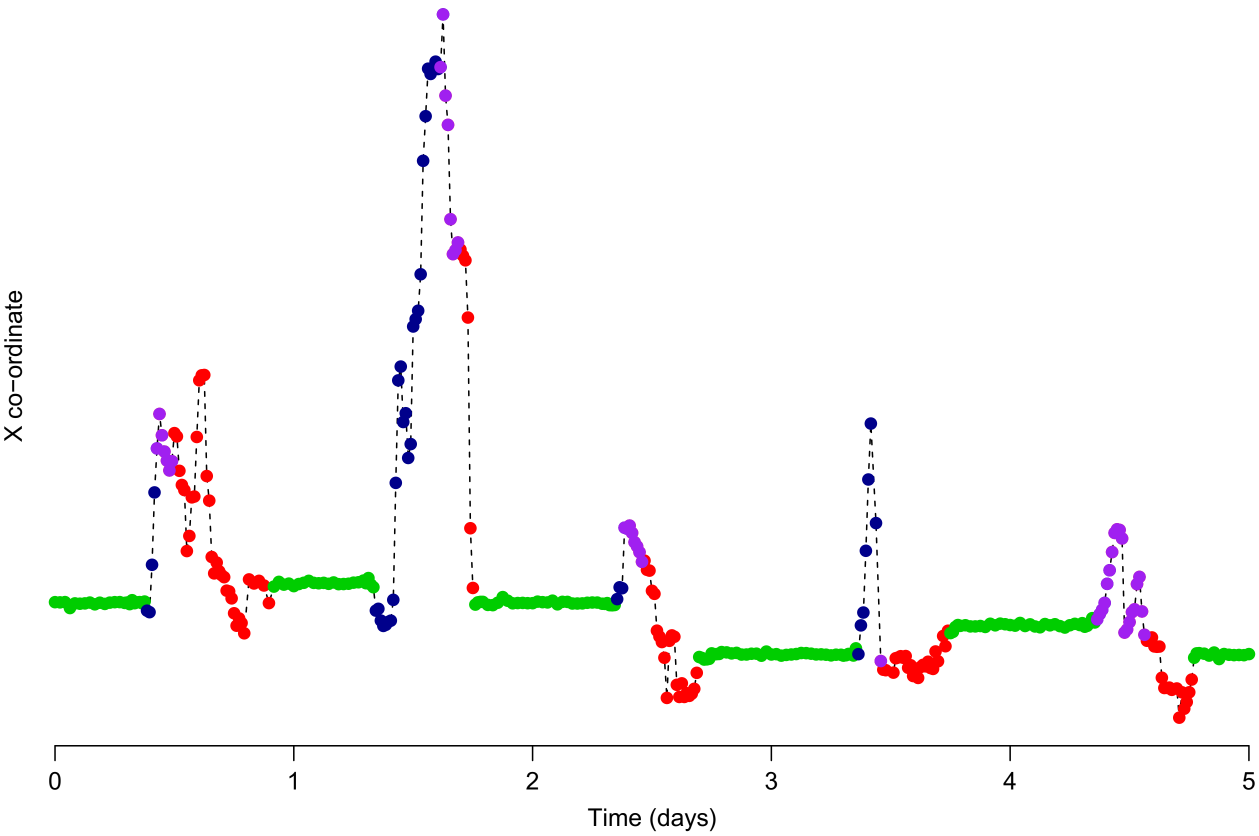

Fig. 8. Classification of observations in the wild boar model plotted against time. Observations are shown here as resting (green), foraging (blue), returning (red) or uncertain (either foraging or returning – purple). See text for details.

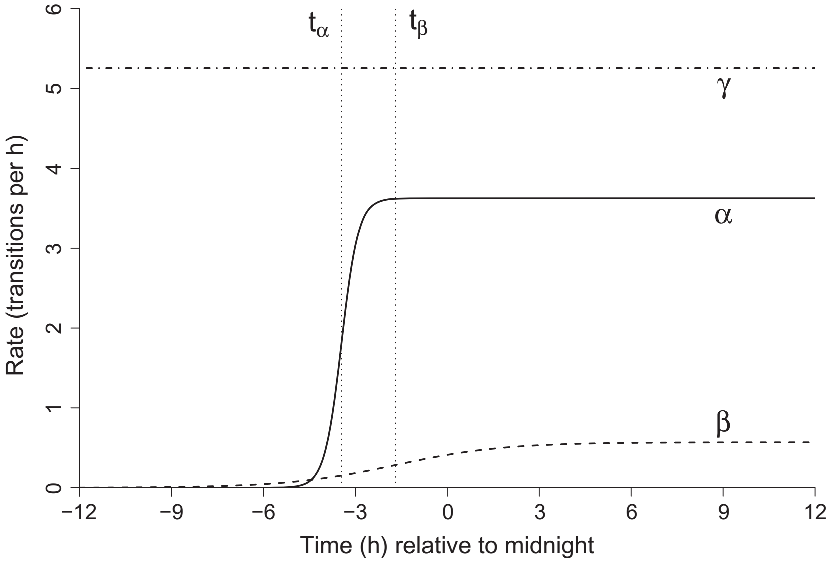

Fig. 9. Posterior median transition rates in the wild boar model in transitions per hour, as a function of time of day. Transition rates from 'resting' to 'foraging': solid line; from 'foraging' to 'returning', dashed line; from 'returning' to 'resting', dots and dashes. Posterior medians for the time-of-day parameters: vertical dotted lines.

including the true locations at the times of observations and of potential switches as state variables that are reconstructed as additional data augmentation within the algorithm. This increases the complexity of implementation and the computational cost, but does not involve any new ideas.

Another extension to the models, as indicated above, is to allow a semi-Markov process for behaviour, that is, to allow the transition rate to depend on the time since the most recent transition; this amounts to allowing the time spent in a behavioural state to have a distribution other than the exponential, and then using the hazard function of that distribution as the transition rate. This would give additional flexibility both in modelling short-term behaviour and in the collection of relevant time-scales at which the model could capture the properties of the data, in the sense of Fleming *et al.* (2014a). Implementing this is straightforward, since in our 'forward simulation' of a proposed path, we will know the time of the previous actual switch, but will add an extra term in the Hastings ratio for path updates.

An important feature of some recent models (Johnson *et al.* 2008; Fleming *et al.* 2014a,b) is the autocorrelation of velocities over time, not represented within the 'building blocks' described here or in Harris & Blackwell (2013) (though of course some autocorrelation of speed is introduced by behavioural switching itself). Autocorrelation of velocity means that the position process itself, $\mathbf{x}(t)$, is no longer Markovian. However, as Johnson *et al.* (2008) note, for their Integrated OU (IOU) model, in which velocity rather than position follows an OU process, the joint process of position and velocity *is* Markovian, and its joint distribution conditional on an earlier value is multivariate normal. Thus, our method can be readily extended to the IOU model with switching, by treating all observations as incomplete – observing only location and not velocity – and incorporating velocities at the times of observations and of potential switches as further augmenting variables.

Finally, we have considered here only a single animal; the approach would apply to multiple simultaneously tracked animals, as in Langrock *et al.* (2014) and the continuous-time models of Niu *et al.* (in press).

Acknowledgements

M. Niu was funded by EPSRC/NERC grant EP/10009171/1 (National Centre for Statistical Ecology). S. D. LaPoint was funded by the Max-Planck-Poland Biodiversity Fund. The work has benefitted from useful discussions with Janine Illian, Roland Kays, Ruth King, Jason Matthiopoulos and Alison Parton. We are very grateful for comments by Justin Calabrese, an anonymous referee and an Associate Editor, which have greatly improved the presentation of this manuscript.

Data accessibility

The full fisher data set is available at LaPoint *et al.* (2013b); the extract used here is included with the code in Appendix S4. The wild boar data set is described in detail in Quy *et al.* (2014); the extract used here is available at Blackwell *et al.* (2015).

References

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.

Berman, S.M. (1994) A bivariate Markov process with diffusion and discrete components. *Stochastic Models*, **10**, 271–308.

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

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

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

Blackwell, P.G., Niu, M., Lambert, M.S. & LaPoint, S.D. (2015) Data from: Exact Bayesian inference for animal movement in continuous time. Dryad Digital Repository. doi:10.5061/dryad.mv02k.

Brillinger, D.R., Preisler, H.K., Ager, A.A., Kie, J.G. & Stewart, B.S. (2002) Employing stochastic differential equations to model wildlife motion. *Bulletin Brazilian Mathematical Society*, **33**, 385–408.

Dunn, J.E. & Brisbin, I.L. (1985) Characterizations of the multivariate Ornstein-Uhlenbeck diffusion process in the context of home range analysis. *Statistical Theory and Data Analysis: Proceedings of the Pacific Area Statistical Conference* (ed. K. Matusita), pp. 181–205. Elsevier, North-Holland.

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

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

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

Fry, J.A., Xian, G., Jin, S., Dewitz, J.A., Homer, C.G., Yang, L., Barnes, C.A., Herold, N.D. & Wickham, J.D. (2011) Completion of the 2006 National Land Cover Database for the conterminous United States. *Photogrammetric Engineering and Remote Sensing*, **77**, 858–864.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. (2013) *Bayesian Data Analysis*. CRC Press, Boca Raton.

Guttorp, P. & Minin, V.N. (1995) *Stochastic Modeling of Scientific Data*. CRC Press, Boca Raton.

Harris, K.J. & Blackwell, P.G. (2013) Flexible continuous-time modelling for heterogeneous animal movement. *Ecological Modelling*, **255**, 29–37.

Jensen, A. (1953) Markoff chains as an aid in the study of Markoff processes. *Scandinavian Actuarial Journal*, **1953**, 87–91.

Johnson, D., London, J., Lea, M. & Durban, J. (2008) Continuous-time correlated random walk model for animal telemetry data. *Ecology*, **89**, 1208–1215.

Langrock, R., Hopcraft, J.G.C., Blackwell, P.G., Goodall, V., King, R., Niu, M., Patterson, T.A., Pedersen, M.W., Skarin, A. & Schick, R.S. (2014) Modelling group dynamic animal movement. *Methods in Ecology and Evolution*, **5**, 190–199.

LaPoint, S., Gallery, P., Wikelski, M. & Kays, R. (2013a) Animal behavior, cost-based corridor models, and real corridors. *Landscape Ecology*, **28**, 1615–1630.

LaPoint, S., Gallery, P., Wikelski, M. & Kays, R. (2013b) Data from: Animal behavior, cost-based corridor models, and real corridors. Movebank Data Repository. doi:10.5441/001/1.2tp2j43g.

Morales, J., 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.

Nams, V.O. (2014) Combining animal movements and behavioural data to detect behavioural states. *Ecology Letters*, **17**, 1228–1237.

Natvig, E. & Subbey, S. (2011) Modelling vertical fish migration using mixed Ornstein-Uhlenbeck processes. In Proceedings Norsk Informatikkonferanse 2011. November 21st to 23rd, Troms, Norway.

Niu, M., Blackwell, P.G. & Skarin, A. (in press) Modelling inter-dependent animal movement in continuous time. *Biometrics*.

Powell, R.A. (1993) *The Fisher: Life History, Ecology, and Behavior*, 2nd edn. University of Minnesota Press, Minneapolis/London.

Preisler, H.K., Ager, A.A. & Wisdom, M.J. (2013) Analyzing animal movement patterns using potential functions. *Ecosphere*, **4**, Article 32.

Quy, R.J., Massei, G., Lambert, M.S., Coats, J., Miller, L.A. & Cowan, D.P. (2014) Effects of a GnRH vaccine on the movement and activity of free-living wild boar (*Sus scrofa*). *Wildlife Research*, **41**, 185–193.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & Van Der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, 583–639.

Received 23 June 2015; accepted 29 July 2015
Handling Editor: Robert B. O'Hara

Supporting Information

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

Appendix S1. Refinements and special cases of the algorithm.

Appendix S2. Modelling and implementation details for the examples.

Appendix S3. Additional plots of posterior samples and distributions.

Appendix S4.R code for the fisher example.