

Bayesian Inference for Multistate ‘Step and Turn’ Animal Movement in Continuous Time

A. PARTON and P. G. BLACKWELL

Mechanistic modelling of animal movement is often formulated in discrete time despite problems with scale invariance, such as handling irregularly timed observations. A natural solution is to formulate in continuous time, yet uptake of this has been slow. This lack of implementation is often excused by a difficulty in interpretation. Here we aim to bolster usage by developing a continuous-time model with interpretable parameters, similar to those of popular discrete-time models that use turning angles and step lengths. Movement is defined by a joint bearing and speed process, with parameters dependent on a continuous-time behavioural switching process, creating a flexible class of movement models. Methodology is presented for Markov chain Monte Carlo inference given irregular observations, involving augmenting observed locations with a reconstruction of the underlying movement process. This is applied to well-known GPS data from elk (*Cervus elaphus*), which have previously been modelled in discrete time. We demonstrate the interpretable nature of the continuous-time model, finding clear differences in behaviour over time and insights into short-term behaviour that could not have been obtained in discrete time.

Key Words: Movement modelling; Switching behaviour; Random walk; GPS data; Markov chain Monte Carlo; Elk.

1. INTRODUCTION

The study of individual animal movement is an active area of ecological research, with advances in tracking technologies allowing data collection at increasing precision and frequency. This ability to capture short-term movement has motivated the study of different movement behaviours presented by an animal over time. A number of statistical methodologies have been applied to attempt to tackle questions such as the number of behavioural modes present, when/how often transitions between these occur, and the characteristics of

A. Parton (✉) and P. G. Blackwell, University of Sheffield, Sheffield, UK (E-mail: aparton2@sheffield.ac.uk).

© 2017 The Author(s). This article is an open access publication
Journal of Agricultural, Biological, and Environmental Statistics, Volume 22, Number 3, Pages 373–392
DOI: 10.1007/s13253-017-0286-5

movement they represent. Recent applications include, for example, [Kuhn et al. \(2009\)](#), [McEvoy et al. \(2015\)](#) and [McKellar et al. \(2015\)](#).

Modelling approaches can be classified by their formulation of time: continuous models define movement at any positive, real time, whereas discrete models are defined only on some predetermined ‘grid’ of times. Often, the time scale in a discrete analysis is that given by the sampling scheme of the observations, leading to problems regarding irregular or missing observations ([Patterson et al. in press](#)), along with concerns regarding suitability and interpretability ([Codling and Hill 2005](#); [Rowcliffe et al. 2012](#); [Nams 2013](#); [Harris and Blackwell 2013](#)). This lack of scale invariance places unwarranted importance on the chosen time frame, suggesting no way to combine multiple sources of data or compare analyses. Further, if a discrete-time model is thought of as observations from a continuous-time process, the existence of such a process and the effect of discretisation are not trivial to address. For example, not all discrete-time Markov chains have a continuous-time counterpart. Continuous-time models can therefore be seen as the ‘gold standard’ of movement modelling, avoiding these challenges through being scale invariant and respecting the continuous nature of an animal’s movement.

The continuous-time model of [Johnson et al. \(2008a\)](#) adopts the popular movement assumption of a correlated random walk, modelling velocity via a stochastic differential equation and using a state space framework to incorporate observation error. The ability to incorporate behavioural switching, however, is limited, either being highly restricted [setting velocity to zero for a stationary state at known times based on additional tag information ([Johnson et al. 2008a](#))], or simplifying to a discrete-time behavioural process ([Hanks et al. 2011](#); [McClintock et al. 2014](#)) or movement process ([Breed et al. 2012](#)). Similarly, the correlated and biased movement models of [Kranstauber et al. \(2014\)](#) use discrete-time methods for estimating the behavioural process. [Blackwell et al. \(2015\)](#) overcome these limitations by modelling location and allowing for a rich class of behavioural processes dependent on both environmental covariates and time via continuous-time Markov chains. A set of models able to incorporate a range of movement assumptions including the home range movement of [Blackwell et al. \(2015\)](#) are given in [Fleming et al. \(2014\)](#), basing inference on the semivariance function of the underlying movement. This approach offers a flexible range of models, but the user is unable to associate behaviours directly with environmental information or identify the behavioural state of the animal at a specific point in time. The functional model of [Buderman et al. \(2016\)](#) fits splines to infer movement in continuous time, offering much versatility. However, as the estimable quantities of this approach are parameters of splines, rather than mechanistic parameters such as a ‘mean speed’, the interpretation of these quantities is unclear. A recent generalisation using basis functions by [Hooten and Johnson \(in press\)](#) is a promising development, able to incorporate a wide range of movement and observation error. An alternative approach to those above is given by [Hanks et al. \(2015\)](#) in which movement is defined in discrete space, using a Markov chain to model location switches. The inference method they propose, however, requires imputing continuous-time movement paths via some other movement model [examples include [Johnson et al. \(2008a\)](#) and [Buderman et al. \(2016\)](#)], therefore inheriting such a model’s associated assumptions and limitations.

The uptake of continuous-time approaches has been somewhat limited, owing in part to the difficulty for the practitioner to interpret the estimated instantaneous movement and behavioural parameters (McClintock et al. 2014). In contrast, a class of discrete-time movement models based on ‘step lengths’ and ‘turning angles’ (Kareiva and Shigesada 1983; Morales et al. 2004) attract widespread use (McClintock et al. 2012). The behaviour of the animal is assumed to follow a Markov chain, with movement evolving according to behaviour-specific parameters. Within a behaviour, movement is defined by the straight line ‘step length’ between two consecutive locations and the ‘turning angle’ between three consecutive locations, following parametric distributions such as the Weibull and the wrapped Cauchy, respectively (Morales et al. 2004; McClintock et al. 2014). Popular variants on this include state space models to incorporate observation error (Patterson et al. 2010; Jonsen et al. 2013), hidden Markov models for efficiency (Langrock et al. 2012) and change point analysis rather than Markov chains to identify behavioural switches (Gurarie et al. 2009; Nams 2014).

Parton et al. (2017) introduce a continuous-time movement model based on similar quantities to those of the popular discrete-time ‘step and turn’ models. This provides familiar descriptive parameters for estimation, whilst respecting the inherent continuous-time characteristic of movement, having the ability to handle missing and irregular observations with ease. The inference method involves simulating realisations of the underlying movement trajectory at a finer time scale than that observed, furthering our goal of providing easily understood movement analysis through the ability to visualise and relate estimated parameters to the movement they describe. This method is demonstrated on noisy observations of a reindeer (*Rangifer tarandus*), taken at mostly 2 min intervals. In Fig. 2 of Parton et al. (2017), the examples of reconstructed movement paths highlight that the characteristics of movement inferred from the observations are markedly different from a simple linear interpolation of such observations. Without accounting for observation error, as in many discrete-time methods, linearly interpolating between observations would lead to a small number of large ($\pm\pi$) turning angles. To account for these, inference would describe movement that is tortuous (correlated random walk with low correlation). However, if observation error is accounted for, Parton et al. (2017) show that the information provided by all the observations suggests movement that is persistent (correlated random walk with high correlation).

Describing only single-state movement limits Parton et al. (2017) to applications with short-term sampling periods. Our aim here is to introduce a statistical, multistate movement model in continuous time able to provide intuitive and easily interpretable estimated parameters for the non-statistical user. Multistate switching movement is introduced by extending Parton et al. (2017) to include a continuous-time Markov chain behavioural process. Section 2 introduces our proposed model, and an approach for fully Bayesian inference given observed telemetry data is outlined in Sect. 3. The interpretability of this method is demonstrated in Sect. 4 on well-known GPS data from a single elk (*Cervus elaphus*).

2. MULTISTATE MOVEMENT BASED ON STEPS AND TURNS

2.1. SINGLE-STATE MOVEMENT MODEL

The basic component for movement follows that of [Parton et al. \(2017\)](#), in which the animal has both a bearing $\theta(t)$ and a speed $\psi(t)$ at time $t \geq 0$. The bearing process describes the direction the animal is facing, assumed to evolve according to Brownian motion with volatility σ_θ^2 so that

$$d\theta(t) = \sigma_\theta dW(t),$$

where $W(t)$ is the Wiener process ([Guttorp 1995](#)). This reflects the common assumption of persistence, where the animal will most likely travel in the same direction over a short period of time. Over a finite period of time, the change in direction of facing will be a wrapped Gaussian with mean zero and a variance which is a linear function of time.

The direction an animal is facing at any time is constrained to $[-\pi, \pi]$; however, here $\theta(t)$ is not constrained in this way and can take any real value. For example, given times $0 \leq t < s$, let $\theta(t) = 0$ and $\theta(s) = 2\pi$. Although the animal was facing the same direction at both times, there is information about the behaviour of the process between these points, as the animal has turned an entire ‘loop’ over this time frame (with the distribution of this constrained process being a Brownian bridge)

A one-dimensional Ornstein–Uhlenbeck process ([Iacus 2008](#)) is assumed to govern the speed with which the animal is travelling, with parameters $\{\mu, \beta, \sigma_\psi^2\}$ so that

$$d\psi(t) = \beta(\mu - \psi(t))dt + \sigma_\psi dW(t).$$

Hence, the animal’s speed is stochastic but correlated, with long-term average μ and variance $\sigma_\psi^2/2\beta$.

Alternate modelling assumptions to those presented may be desired dependent upon application. A more direct comparison with discrete-time correlated random walk models would be to model speed as Brownian motion so that distances travelled over disjoint time periods are independent. Similarly, directed/biased movement could be achieved by altering the Brownian motion on the bearing process, or assuming some Ornstein–Uhlenbeck process.

The joint process given by the bearing and speed of the animal completely defines the location process $\mathbf{Z} = \{X, Y\}$, given by

$$dX(t) = \psi(t) \cos(\theta(t)), \quad dY(t) = \psi(t) \sin(\theta(t)).$$

2.2. MULTISTATE SWITCHING MODEL

To reflect the changing behaviours of an animal over time, a switching model is employed, with different movement characteristics for each state ([Blackwell 1997](#); [Morales et al. 2004](#); [McClintock et al. 2012](#); [Blackwell et al. 2015](#)). The behavioural process is taken to be a continuous-time Markov chain with switching rates λ and probabilities q ([Guttorp 1995](#)).

The animal will follow behavioural state i for a length of time exponentially distributed with rate λ_i , before switching to state j with probability $q_{i,j}$. Within a behaviour there is a corresponding set of parameters describing the movement, as in Sect. 2.1. With this extension in place the marginal joint process of bearing and speed is not Markovian; however, the joint process of behaviour, bearing and speed is. The movement of the animal is therefore parametrised by the set $\Phi = \{\Phi_B, \Phi_M\}$, with $\Phi_B = \{\lambda_i, q_{i,j}\}$ and $\Phi_M = \{\sigma_{\theta,i}^2, \mu_i, \beta_i, \sigma_{\psi,i}^2\}$ for $i \neq j \in \{1, \dots, n\}$, where n is the number of behavioural states.

2.3. SIMULATING MULTISTATE MOVEMENT

Realisations of movement given parameters Φ can be easily simulated, with an example of such in Fig. 1. The behavioural process is simulated according to a continuous-time Markov chain with generator matrix defined by Φ_B . Given a current behaviour $B(t) = s$, this involves drawing the time until the next behavioural switch from an exponential distribution with rate λ_s and then choosing the new behaviour $j \neq s$ with probability $q_{s,j}$.

Given a realisation of the behavioural process, movement is simulated at an approximate time scale δt , which can be arbitrarily fine. If the behaviour at time t is $B(t) = s$, then the bearing and speed are given as

$$\theta(t + \delta t) \mid \theta(t), s \sim N(\theta(t), \sigma_{\theta,s}^2 \delta t), \tag{1}$$

$$\psi(t + \delta t) \mid \psi(t), s \sim N\left(\mu_s + \exp\{-\beta_s \delta t\}(\psi(t) - \mu_s), \frac{\sigma_{\psi,s}^2}{2\beta_s} (1 - \exp\{-2\beta_s \delta t\})\right). \tag{2}$$

Given this approximation, the familiar notion of a ‘step’ is recovered by $v(t) = \psi(t)\delta t$.

Given the joint processes $\{\theta, v\}$, the Euler–Maruyama approximation of location in two-dimensional space is given by the cumulative sums

$$X(t_i) = X(t_0) + \sum_{j=1}^{i-1} v(t_j) \cos(\theta(t_j)), \quad Y(t_i) = Y(t_0) + \sum_{j=1}^{i-1} v(t_j) \sin(\theta(t_j)). \tag{3}$$

3. THE MARKOV CHAIN MONTE CARLO ALGORITHM

Observations \mathbf{Z} of an animal’s two-dimensional location are taken at a finite, but irregular, series of times t . The likelihood of these observations given parameters Φ is intractable due to the complicated relationship between the locations and parameters when the bearing and speed processes are unobserved. This is further complicated by the unobserved behavioural process, where there is the possibility of multiple switches between observations. The following describes the Markov chain Monte Carlo algorithm used to carry out inference given observations.

Following Blackwell (2003) a data augmentation approach is taken, simplifying the relationship between observations and parameters by augmenting the data with the times

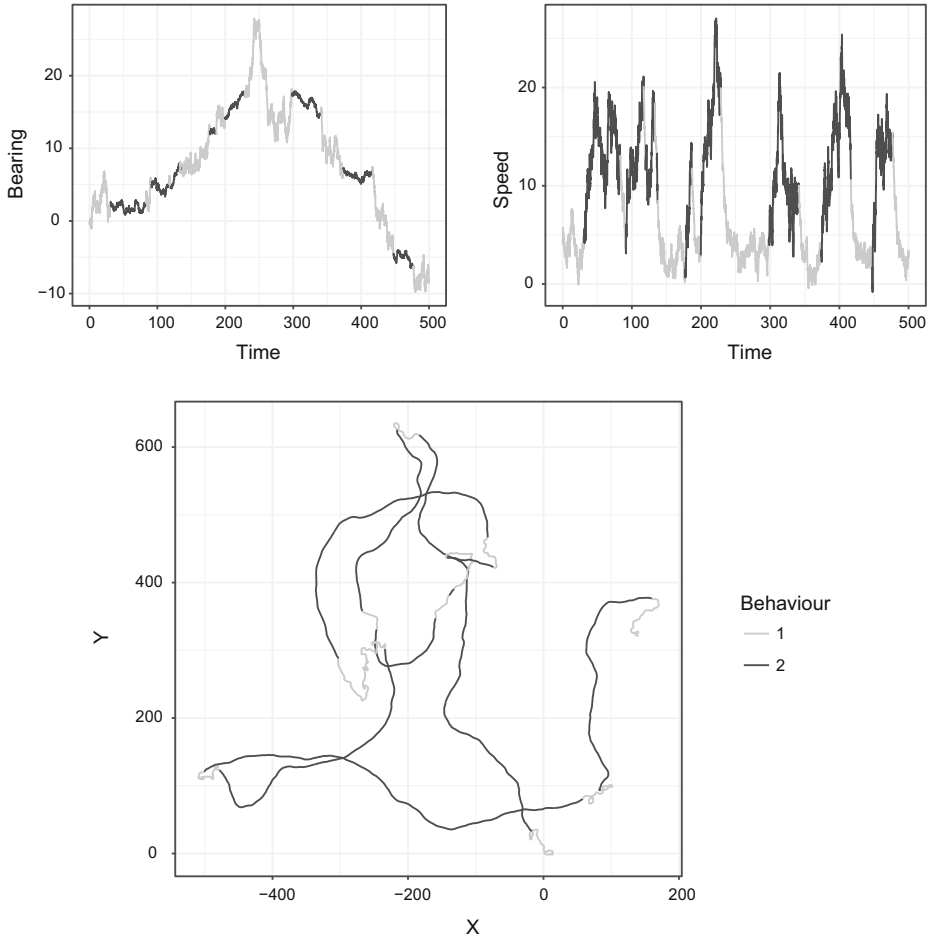


Figure 1. An example of a simulated movement path with two behavioural states. The simulated bearing and speed processes are shown, coloured by the simulated behavioural process, along with the resulting two-dimensional locations.

of all behavioural switches. Here, augmentation also includes an approximation to the underlying bearing and speed processes on some (arbitrarily fine) time scale. The hybrid Markov chain Monte Carlo algorithm used splits the quantities of interest into three groups to update separately, in each case conditional on all other quantities. In cases where the full conditional distribution can be directly sampled from, Gibbs sampling is employed, and in all other scenarios the Metropolis–Hastings sampler is used (see, for example, [Gelman et al. \(2013\)](#) for general sampling methods). The groups to be separately sampled from are the behavioural parameters (Φ_B), the movement parameters (Φ_M), and the unobserved refined path consisting of behavioural switches, bearings and speeds ($\mathbf{B}, \boldsymbol{\theta}, \mathbf{v}$).

Sections 3.1 and 3.2 describe the sampling schemes used for the behavioural and movement parameters, respectively. In both cases the sampling is standard, employing Gibbs sampling and a random walk Metropolis–Hastings algorithm. Section 3.3 describes the

Metropolis–Hastings algorithm used for the reconstruction of the unobserved refined path, in which a novel method of simulation is used to create the independent proposals within this sampling scheme.

3.1. SAMPLING THE BEHAVIOURAL PROCESS PARAMETERS

The behavioural process parameters are sampled conditional on the complete observation of the behavioural process. Conjugate distributions for the switching rates (λ) and probabilities (q) of a continuous-time Markov chain are gamma and Dirichlet, respectively. Assuming such conjugate priors allows direct sampling from the posterior conditional as a Gibbs steps (Blackwell 2003). Further details are given in Section A.1.

3.2. SAMPLING THE MOVEMENT PROCESS PARAMETERS

The movement process parameters are sampled conditional on the complete observation of the refined path (both behaviour and movement) and the behavioural parameters. The movement parameters are updated simultaneously using a random walk Metropolis–Hastings step, with independent proposals for each parameter. Since all movement parameters are constrained to be positive, independent univariate Gaussians truncated below at zero are used as proposal distributions to generate the step in the random walk.

In a simultaneous update of the movement parameters, the likelihood of the refined movement path is calculated for the current and proposed parameters and combined with the appropriate prior probability. The standard Metropolis–Hastings acceptance ratio is used to decide on the acceptance of the proposal. Further details are given in Section A.2.

3.3. RECONSTRUCTING THE UNOBSERVED REFINED PATH

The key step for inference is to sample the unobserved ‘refined path’—given by the behavioural process, and the bearing and speed processes at a refined time scale—conditional on the parameters. As the dimension of the full movement path will be large (the example of Sect. 4 leads to a path with around 2300 locations at the chosen refined time scale), reconstruction is carried out on random short sections. The aim is to simulate the refined path between two observation times a and b , conditional on the fixed path outside of these times and a set of parameters. This can easily be extended to span multiple observed locations. A diagram of this scenario is given in Fig. 2, with two circular points showing the fixed observations that the path will be simulated between.

The quantities to simulate are those in black in Fig. 2 consisting of the behavioural process \mathbf{B} between times a and b , the bearings $\{\theta_1, \dots, \theta_{n-1}\}$ and the steps $\{v_1, \dots, v_{n-1}\}$. The fixed values that are to be conditioned upon are displayed in grey in Fig. 2 consisting of the locations $\{\mathbf{Z}(a), \mathbf{Z}(b)\}$, the behaviours $\{B(a), B(b)\}$, the bearings $\{\theta_0, \theta_n\}$ and the steps $\{v_0, v_n\}$. As the bearing and step processes are given by a discrete-time approximation, the fixed points are the values of the respective process at the refined point immediately before and after the path section of interest, as shown in Fig. 2.

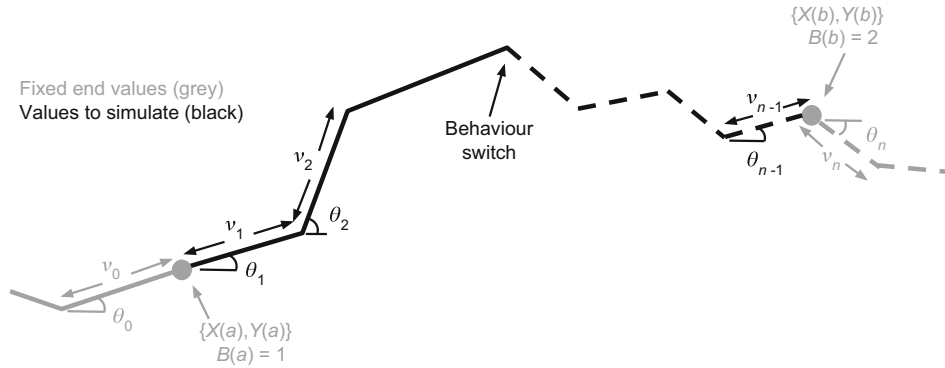


Figure 2. Diagram of a section of the refined path, with fixed endpoint locations at the times a and b . The behavioural process, B (represented as two states with *solid* and *dashed* lines here), is simulated with fixed endpoints $\{B(a), B(b)\}$. The bearing and step processes, $\{\theta_1, \dots, \theta_{n-1}, v_1, \dots, v_{n-1}\}$, are simulated, given fixed endpoints $\{\theta_0, \theta_n, v_0, v_n\}$.

Simulating the quantities of interest conditional on all fixed values is not possible due to the nonlinearity of the location process (see Eq. 3), and so a proposal path section is simulated from a simpler distribution that is then accepted or rejected using a Metropolis–Hastings ratio. An independence sampler is employed using a novel simulation method to propose a new path section, described below. Further details on the acceptance condition is given in Section A.3.

3.3.1. Simulating a Refined Path Proposal

A behavioural proposal B^* is simulated between the times a and b , given fixed values $\{B(a), B(b)\}$ and parameters Φ_B , by a rejection method. A continuous-time Markov chain with parameters Φ_B starting at $B(a)$ at time a and ending at time b is simulated (see Sect. 2.3). If the final state is not equal to $B(b)$, then the proposal is instantly rejected. Otherwise, the path proposal continues (still with the possibility of rejection in the Metropolis–Hastings step). Less naive approaches to this simulation could be implemented [see, for example, [Hobolth and Stone \(2009\)](#), [Rao and Teh \(2013\)](#) and [Whitaker et al. \(2016\)](#)]; however, this naive method performed well in our examples.

Given the behavioural simulation, the set of refined times $\{t_1 = a, \dots, t_{n-1}\}$ is created. This must be a sequence of times between a and b that includes behavioural switch times, and is chosen to approximately be on some time scale δt , the choice of which is discussed in Sect. 5. This forms the times to simulate the bearings and speed over, as in Fig. 2.

The bearing proposal θ^* over the times $\{t_1, \dots, t_{n-1}\}$ is simulated conditional on the fixed bearings $\{\theta_0, \theta_n\}$ at the times $\{t_0, t_n = b\}$, the behaviours B^* and the parameters Φ . The distribution of this process is a Brownian bridge with time-varying volatility parameter, dependent on behaviour. The times $\{t_1, \dots, t_{n-1}, t_n\}$ are transformed, weighted by the turn volatility at each respective time, to give a process with constant volatility. The Brownian bridge is then simulated on the transformed times $\{t'_1, \dots, t'_{n-1}\}$, given the values $\{\theta_0, \theta_n\}$ at the end times $\{t_0, t'_n\}$ (see [Iacus \(2008\)](#) for Brownian bridge simulation).

Simulating the step proposal To propose the steps \mathbf{v}^* over the times $\{t_1, \dots, t_{n-1}\}$, the joint distribution of \mathbf{v} and $\mathbf{Z}(b)$, given by

$$\begin{pmatrix} \mathbf{v} \\ \mathbf{Z}(b) \end{pmatrix} \mid \Phi, \mathbf{B}^*, \boldsymbol{\theta}^*, \mathcal{F} \sim N \left(\begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}, \begin{pmatrix} \Sigma_1 & \Sigma_{1,2} \\ \Sigma_{1,2}^T & \Sigma_2 \end{pmatrix} \right), \tag{4}$$

where $\mathcal{F} = \{\mathbf{Z}(a), B(a), B(b), \theta_0, \theta_n, v_0, v_n\}$, is first constructed. The marginal distribution of \mathbf{v} (dimension $n - 1$) given a known behavioural process and fixed end steps is $N(\mathbf{m}_1, \Sigma_1)$ (discussed further below). The location $\mathbf{Z}(b)$ is given by $\mathbf{Z}(a) + A\mathbf{v}$, where

$$A = \begin{pmatrix} \cos(\theta_1^*) & \dots & \cos(\theta_{n-1}^*) \\ \sin(\theta_1^*) & \dots & \sin(\theta_{n-1}^*) \end{pmatrix}.$$

The marginal distribution of $\mathbf{Z}(b)$ (dimension 2) is $N(\mathbf{m}_2, \Sigma_2)$, and $\Sigma_{1,2}$ is the $(n - 1) \times 2$ covariance between the steps \mathbf{v} and the location $\mathbf{Z}(b)$. Given $\mathbf{m}_1, \Sigma_1, A$, values for $\mathbf{m}_2, \Sigma_2, \Sigma_{1,2}$ can be easily calculated due to $\mathbf{Z}(b)$ being a linear combination of the normally distributed \mathbf{v} .

The form of \mathbf{m}_1, Σ_1 arises from the speed process (from which \mathbf{v} is derived) being an Ornstein–Uhlenbeck bridge with inhomogeneous parameters, calculated by the following method. The fixed values v_0, v_n are transformed to give speeds $\psi_0 = v_0/\delta t_0$ and $\psi_n = v_n/\delta t_n$. The joint distribution $\psi_1, \dots, \psi_n \mid \psi_0, \mathbf{B}^*$ is created by iteratively applying

$$\psi_i \mid \psi_{i-1}, B(t_i) \sim N(\mu, \sigma^2), \tag{5}$$

where μ, σ^2 are given by Eq. 2. This joint distribution is then partitioned into $\psi_1, \dots, \psi_{n-1}$ and ψ_n in order to condition upon the known value for ψ_n using standard conditioning of a multivariate normal (Eaton 2007) to give the joint distribution $\psi_1, \dots, \psi_{n-1} \mid \psi_0, \psi_n, \mathbf{B}^*$. This distribution can be transformed back to steps v_1, \dots, v_{n-1} to give \mathbf{m}_1, Σ_1 through a transformation by multiplying the speeds $\psi_1 \dots, \psi_{n-1}$ by the times $\delta t_1, \dots, \delta t_{n-1}$.

The step proposal \mathbf{v}^* is simulated by further conditioning \mathbf{v} in Eq. 4 on the known $\mathbf{Z}(b)$ by standard conditioning of a normal distribution (Eaton 2007), given by

$$\mathbf{v} \mid \Phi, \mathbf{B}^*, \boldsymbol{\theta}^*, \mathcal{F}, \mathbf{Z}(b) \sim N \left(\mathbf{m}_1 + \Sigma_{1,2} \Sigma_2^{-1} (\mathbf{Z}(b) - \mathbf{m}_2), \Sigma_1 - \Sigma_{1,2} \Sigma_2^{-1} \Sigma_{1,2}^T \right).$$

The steps are being conditioned upon a linear constraint (the fixed $\mathbf{Z}(b)$), leading to a singular distribution. Simulation of such follows the ‘conditioning by Kriging’ procedure in Rue and Held (2005), by first simulating from the unconditioned $\mathbf{x} \sim N(\mathbf{m}_1, \Sigma_1)$ and adjusting for the constraint by

$$\mathbf{v}^* = \mathbf{x} - \Sigma_{1,2} \Sigma_2^{-1} (A\mathbf{x} - \mathbf{Z}(b)).$$

This path proposal method does not take into account the fixed location at the end of the section when simulating the behaviours and bearings. Therefore, a Metropolis–Hastings step (ratio details in Section A.3) assesses whether this proposal is accepted.

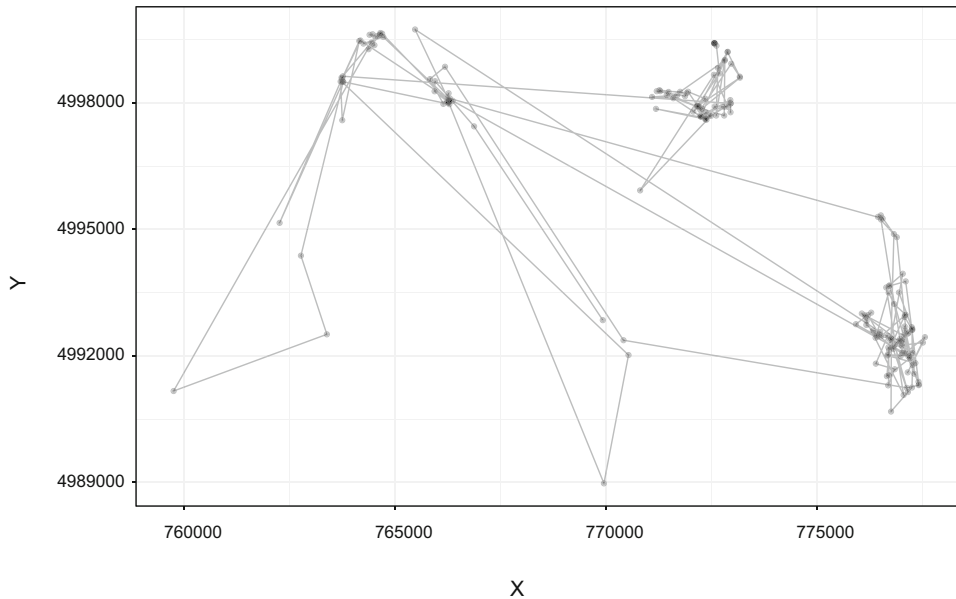


Figure 3. Observed daily observations of elk-115 (points linked chronologically with lines). Note that observed points are displayed here with transparency to highlight the times where multiple observations were captured in the same/similar location.

4. TWO-STATE SWITCHING MOVEMENT IN ELK

A set of 194 daily GPS observations from the elk (*C. elaphus*) tagged as ‘elk-115’ are used in this example (see https://bitbucket.org/a_parton/elk_example). These observations were introduced and modelled as part of a larger set consisting of four elk in the discrete-time ‘step and turn’ model of Morales et al. (2004), and more recently modelled in the vignette of the R package `moveHMM` (Michelot et al. 2016) applying the hidden Markov model of Langrock et al. (2012). Observations are shown in Fig. 3, appearing to display two distinct movement modes: slow, volatile movement where observations are over-plotted, and fast, directed movement.

Morales et al. (2004) fit a number of models to the larger dataset containing the observations from elk-115, with the model most similar to ours being the ‘double switch’ model. Fixed switching probabilities between the two states were modelled, governing a mixture of correlated random walks. In the vignette of `moveHMM` the larger dataset is used to demonstrate a two-state hidden Markov model with switching dependent on environment. For comparison with the methods here, the reproduction of analysis shown in Fig. 6 does not include this environmental information and so is the same underlying movement model as the ‘double switch’ in Morales et al. (2004). In both these discrete-time applications, ‘travelling’ and ‘foraging’ states were identified as having mean daily turning angles of close to zero and π , respectively. The implications of turn distributions not centred at zero are discussed in Sect. 5.

In this example, the model of Sect. 2 with two behaviours is applied to the elk-115 observations. The original analysis in [Morales et al. \(2004\)](#) described observations as being mostly daily, but with some taken at 22- and 26-h intervals. In order to handle this irregularity, they divided the observed straight line step lengths by the sampling time frame to approximate daily steps. A method transforming the observed turning angles to some daily approximation is unclear, and so these remained as the observed values in their analysis. The open-access version of the elk data does not include the times of the observations, and rounding of the [Morales et al. \(2004\)](#) ‘daily step lengths’ meant that the original observation times could not be ascertained. The analysis performed here therefore followed that in the vignette of `moveHMM`, using the observed locations, but assuming that these were all at 24-h intervals. The continuous-time formulation of our model, however, would easily allow for these irregularly timed observations (and missing observations, if applicable) to be handled if exact observation times were known.

Applying our presented methodology to multiple animals in the same way as `moveHMM`, by pooling information across individuals and estimating a set of population parameters, could be implemented by a simple extension to the current R code, but is not attempted here for simplicity. Following [Morales et al. \(2004\)](#) and the vignette of `moveHMM`, observation error is assumed to be negligible here (though see Sect. 5). Interest thus involves inference on the eight movement parameters, consisting of a bearing volatility and three speed parameters for each state. Using daily observations leaves large portions of the elk’s movement unobserved, and so it is expected that the reconstructed movement paths, and thus parameters, for this example will be very uncertain. Rather than a full ecological analysis, this example is therefore included as a proof of concept for the presented methods and to highlight some of the possible dangers when analysing daily observations in discrete time. Readers are directed to [Parton et al. \(2017\)](#) for an example of single-state movement on a dataset with a sampling scheme of 2 min to compare the uncertainty of movement reconstructions.

4.1. PRIOR AND INITIAL INFORMATION

A prior distribution specifying an upper bound on the ratio of the speed parameters to avoid the presence of negative speeds in both states was applied. To define state 2 as ‘travelling’, a Gaussian prior with mean 0.05 and standard deviation of 0.1 was placed on the turn volatility. All remaining movement parameters had flat priors. The same prior was on both switching rates, being a gamma distribution with rate 4 and shape 0.1. This was chosen to limit the rate of behavioural switching, strongly discouraging switching occurring at a shorter time frame than 4 h, with 90% prior credible interval for residency time of (6.7×10^{13}) h. This prior is fairly vague when comparing with the posterior credible intervals (see below).

An initial movement path was created at a time scale of 2 h by taking an interpolating cubic spline between observations. The choice of a 2-h time scale gives around 11 unknown locations for reconstruction between each pair of observations, thought to provide an acceptable trade-off between computational cost and approximation to continuous time (see Sect. 5 for further discussion of δt). The corresponding initial behavioural configuration was set by

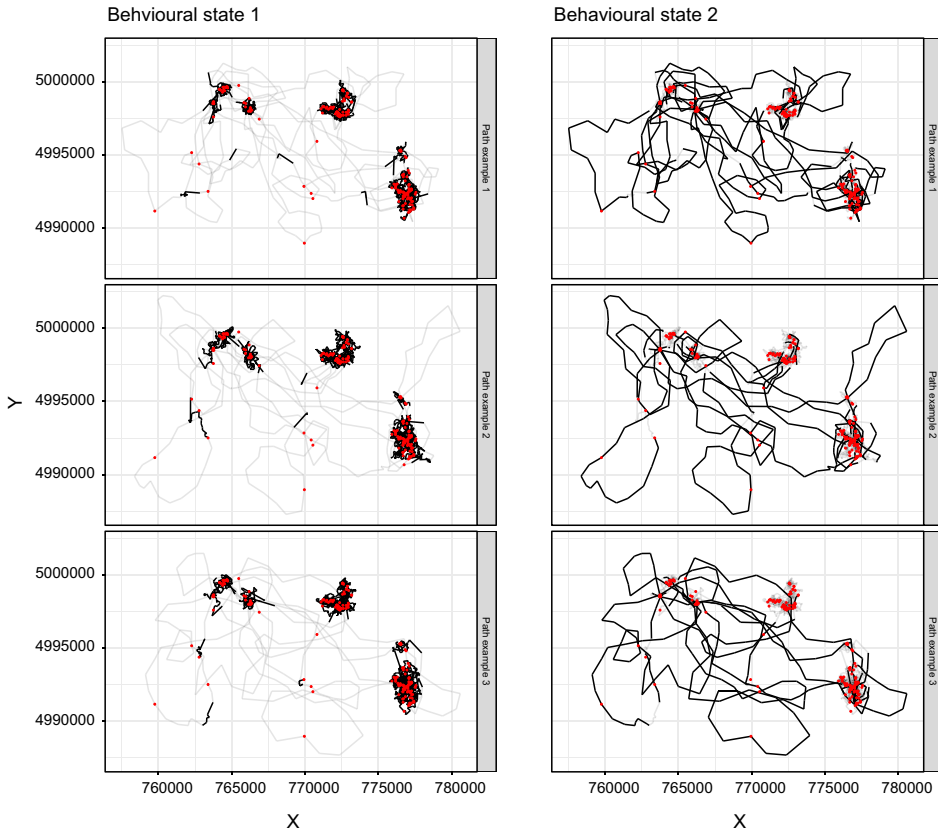


Figure 4. Three examples of reconstructed refined movement paths for elk-115. For each example, the observed locations are shown as *red points* and the reconstructed refined path is displayed as linearly interpolated lines. The *left and right panels* both show the full reconstructed refined path (in *grey and black*), but differ by the behavioural state highlighted: the *left panel* highlights in *black* the parts of the path labelled as behavioural state 1 and the *right panel* highlights in *black* the parts of the path labelled as state 2. This separation of behavioural segments clearly highlights the difference in movement characteristics resulting from the parameters associated with the two behavioural states (Color figure online).

identifying any points on this path with speed above 100 m/h. Initial parameters were set as estimates from this initial path configuration.

The algorithm in Sect. 3 was applied for 48×10^5 iterations, with each iteration consisting of a single parameter update and 100 refined path updates on random sections of path with lengths ranging 4–24 points (i.e., 8–48 h). Samples were thinned by a factor of 1000 and the first quarter were treated as a ‘burn-in’ period, leaving 3600 stored samples of parameters and reconstructed refined paths. Long subpath lengths are desirable as the proportion of path being updated is high. However, this incurs computational cost and has low acceptance due to high dimensionality. A mixture of short subpath lengths (easily accepted) helps with mixing, following on from such a discussion in Blackwell et al. (2015). The choice here was based on acceptance rates in pilot runs: lengths higher than 24 had too low acceptance to be feasible, and lengths of 4 allowed these short section updates that helped with mixing.

4.2. RESULTS

Figure 4 shows three examples (separated vertically) of the reconstructed refined movement path. Red points show the observations, and the combination of grey and black lines shows the three example path reconstructions. Each reconstruction is shown in two panels: the left panel highlights in black the segments of the refined path categorised as behavioural state 1, and the right panel highlights in black the segments of the path labelled as state 2. This highlights the difference in movement types between the two identified states, appearing in many ways similar in interpretation to those of [Morales et al. \(2004\)](#) and the vignette of `moveHMM`, having a slow ‘foraging’ state and fast ‘travelling’ state. These reconstructions aid in the interpretation of the movement parameters and give insight into the space use of the animal between observation times.

Samples from the posterior distributions for the movement parameters, split by state, are shown in Fig. 5, showing the clear differences between the two states. Posterior summary statistics of the parameters are given in Table 1. Behavioural state 1 has high σ_θ^2 and low μ , defining volatile, slow movement categorised here as ‘foraging’. The level of σ_θ^2 for state 1 (median given by 5.61 rad/h) is high enough to produce turns that are uniform over the sampling scheme of the observations. The median for long-term travelling speed for state 1 is given by 77.3 m/h. State 1 has a higher β and lower σ_ψ^2 than state 2, describing speeds that are less correlated in the short term (the mean expression of the speed process in Eq. 2 is dominated by the first term involving the ‘mean speed’ parameter rather than the second term involving the ‘current speed’) and have lower variation in the long term. The movement parameters for state 1 have a low effective sample size and do not pass standard convergence diagnostics. This is due to the turn volatility being so high as to produce uniform turns, and so this parameter is ‘drifting’.

Behavioural state 2, the ‘travelling’ state, has low σ_θ^2 and high μ , reflecting fast, straight movement. The median long-term travelling speed for state 2 is 638 m/h, with speeds that are highly correlated in the short term (through a low β) but with high variation in the long term (through a high σ_ψ^2). The movement parameters for state 2 pass standard convergence diagnostics (Heidelberger and Welch) with effective sample size of over 75.

Samples from the posterior distributions for the two rates of switching defining the behavioural process are shown in the left panel of Fig. 6. Posterior summary statistics for the switching rates are given in Table 1, with the 90% credible intervals leading to a mean residence time in state 1 being between 4 and 11 days and in state 2 between 10 and 36 h. The behavioural parameters pass standard convergence diagnostics, with effective sample size of over 125. The right panel of Fig. 6 displays the probability of being in behavioural state 2 throughout the course of the sampling period. Additionally, the corresponding state probabilities estimated by fitting a hidden Markov model as in the vignette of `moveHMM` (but using the larger dataset of tracks from four elk) are shown below. The two models can be seen to identify the same areas of the movement path as being in the ‘travelling’ state; however, the residence times in this state differ between the two models, with the hidden Markov model classifying three long stays in state 2 in the middle of the observation period.

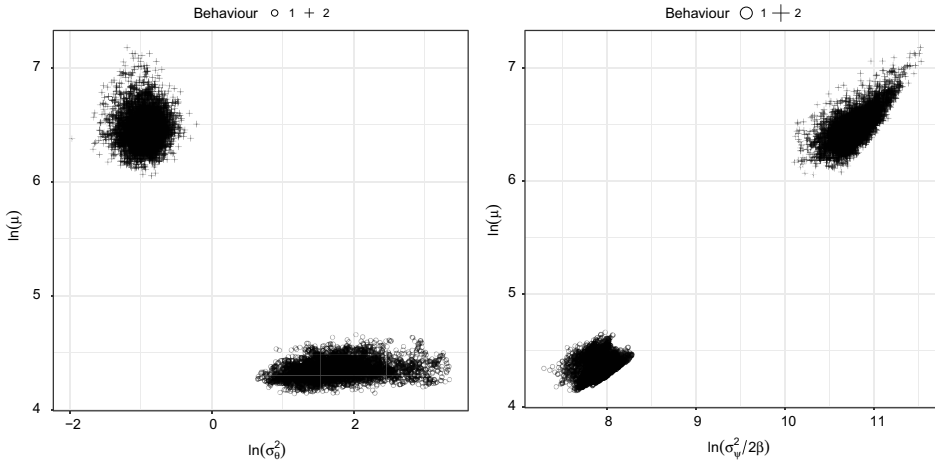


Figure 5. Sampled state-dependent movement parameters (on log scale) for the example using observations of elk-115. *Left plot* the joint sample space between the turn volatility (σ_θ^2) and the mean speed (μ). *Right plot* the joint sample space between the mean speed and the long-term speed variance ($\sigma_\psi^2/2\beta$).

Table 1. Posterior summary statistics (5, 50, 95% quantiles) for the sampled movement and behavioural parameters, split by state, in the elk-115 example.

	Parameter	5%	50%	95%
Behaviour 1 ('foraging')	λ_1 (switching rate)	0.00391	0.00651	0.0105
	σ_θ^2 (turn volatility)	2.87	5.61	16.4
	μ (long-term speed mean)	68.8	77.3	90.2
	β	0.627	1.45	1.94
	σ_ψ^2	2900	7920	11,300
	$\sigma_\psi^2/2\beta$ (long-term speed variance)	2160	2820	3390
Behaviour 2 ('travelling')	λ_2 (switching rate)	0.0275	0.0520	0.0959
	σ_θ^2 (turn volatility)	0.274	0.389	0.521
	μ (long-term speed mean)	519	638	855
	β	0.170	0.245	0.340
	σ_ψ^2	16,000	23,600	29,700
	$\sigma_\psi^2/2\beta$ (long-term speed variance)	34,300	47,600	66,400

5. DISCUSSION

We have provided a methodology for Bayesian inference for continuous-time, multistate movement. The behavioural process leads to a flexible range of movement patterns, whilst the continuous-time formulation allows missing and irregular observations to be handled with ease. Movement within a behaviour has some similarities with the velocity-based continuous-time model of Johnson et al. (2008a) but is more intuitive, enabling a separation of speed and direction that matches empirical observations well. Parameter interpretation is simpler when separated in this way, describing aspects of movement such as a mean travelling speed and a volatility to the direction of movement. Although continuous-time models based

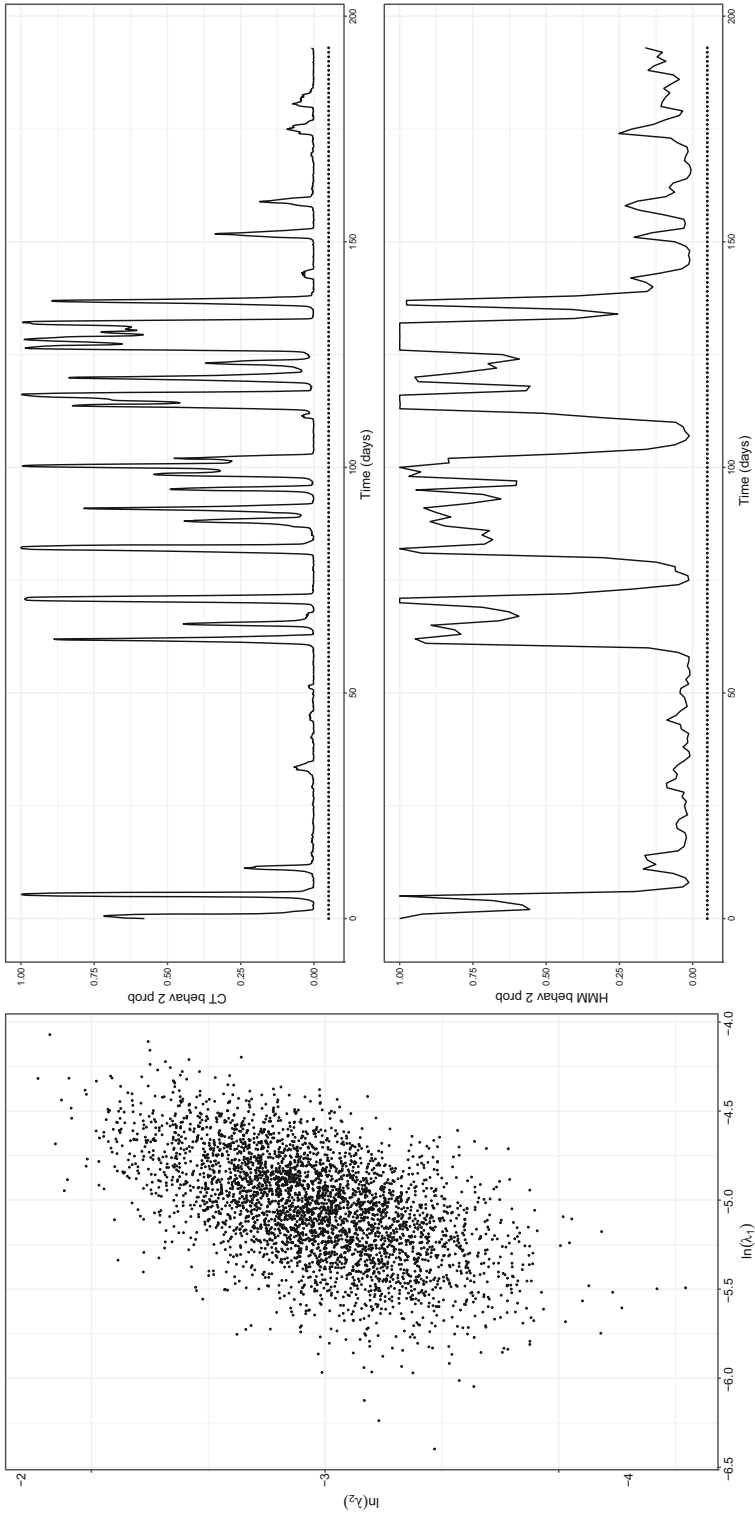


Figure 6. *Left plot* sampled behavioural parameters (on log scale) for elk-115; λ_1 is the switching rate out of the ‘foraging’ state and λ_2 is the switching rate out of the ‘travelling’ state. *Upper right plot* probability of residing in behaviour 2 (‘travelling’) over time. *Lower right plot* probability of residing in behaviour 2 using the R package moveHMM (Michelot et al. 2016). In both plots on the right, points are included to highlight the times/frequency of observations.

on (x, y) locations (Johnson et al. 2008a; Blackwell et al. 2015) could be applied, with post-processing to determine the distribution of speed and bearing, the covariance structure of such distributions, and hence the implicit shapes of the paths, will not be the same as that presented here. Ecological justification for such a covariance structure may be difficult or lacking, whereas our model is directly defined by these quantities and therefore initially motivated by ecological ideas.

For a given state and time interval, the distribution of the change in direction given by our model will always be a wrapped Gaussian centred at zero. A von Mises distribution (often used in discrete models; McClintock et al. 2012) centred at zero is very similar to this, but a von Mises (or other circular) distribution centred at $\pm\pi$ is not. In fact, no natural continuous-time process for change in direction would lead to such a distribution when observed at regular intervals. Such a distribution would require the expected rate of change of bearing to be nonzero, leading to paths that consistently form loops. Whilst this may be appropriate occasionally (Boakes et al. 2011) we do not feel it is realistic in our example or in most published applications. It seems more likely that such a distribution emerges only as an artefact of some other process, e.g. ignored measurement error (Hurford 2009) or attraction to a particular location. The classification of a foraging state with a mean turning angle of $\pm\pi$ in many discrete-time applications is therefore questionable. The ecological interpretation of a ‘foraging’ state would be better modelled as having a uniform turning angle, such as $\sigma_\theta^2 \rightarrow \infty$ in our model.

Modelling in continuous time allows us to consider movement/behaviour between observation times, something not possible in discrete time. The estimated residency rate of the travelling state in the elk example suggests that there are parts of the movement path where short sojourns of fast movement occur. In fact, 72% of the sampled values from the posterior distribution of λ_2 lead to a mean residence time of less than the 24-h sampling scheme. In Fig. 4, it can be seen in a number of places that the reconstruction involves a switch into and back out of state 1 between two consecutive observations. The exact time when these short (between observation) switches in behaviour occur varies over the sampled reconstructions, but their presence has high probability. There is therefore information in the observed locations indicating a behavioural sojourn has occurred, but the precise time of its occurrence is very uncertain. Being able to extract such qualitative information on short-term behavioural switches from observations, albeit with uncertainty, gives extra insight into the movement that is not possible when switches can only occur at the observation time scale.

Although the approach for inference here is an approximation to the underlying continuous-time model, advantages remain over discrete time: behavioural switching can occur continuously in contrast to strictly at observation times and the parameters of the model are scalable (representing parameters of a continuous-time model) rather than ‘per observation time’. Reducing the refined time scale will provide a ‘better’ approximation to the underlying model, but does come with a computational cost. Simulation experiments on the effect of varying δt (details omitted here for brevity) show that great improvements to parameter estimation can be made against using only observations by augmenting as little as four locations between observation pairs. Improving the approximation with further refinement was found to increase accuracy of parameter estimation further, but incurred additional computation time.

The methods described here assume that observation error is negligible. Extending this to observation error is easily implemented, included in the single behavioural method of Parton et al. (2017). This simple model assumed normally distributed errors, independent in space and time. There is therefore a single additional parameter describing the observation error (a mean error of zero is assumed). An extension to the inference method described here allows for such a parameter to be sampled as a Gibbs step, and the path reconstruction method can be extended to include error around observed locations. Extending further to allow for errors to be correlated in time could also be implemented without difficulty.

The augmentation approach furthers our aim for comprehensible inference. The ability to view examples of path reconstructions, such as in Fig. 4, aids in understanding the movement type associated with a given combination of parameters. Sampling a large number of reconstructions displays the uncertainty in the times at which behavioural switches occur and can easily be used to estimate the space/resource use of the animal at the local scale. With the resolution of environmental covariates increasing, this information can be correctly combined with local scale movement rather than assuming that only the covariate values corresponding to directly observed locations are important. For discussion of the wider issues of linking movement and resource use, see, for example, Johnson et al. (2008b).

We have assumed here that transition rates between behaviours are constant. It would be desirable to allow these to depend on spatial covariates (Morales et al. 2004) or on location itself. Depending on the duration of study, it may also be useful to allow varying rates with time, perhaps periodically to reflect daily or annual cycles. Both these extensions could be addressed, without any additional approximation, using the framework in Blackwell et al. (2015), applied there to movement models directly based on location (rather than velocity or steps and turns) with heterogeneity in both space and time. More generally, we could capture some more of the complexity of behaviour by including an additional ‘resting’ state, likely to occur at particular times of the day, with low or zero speed and perhaps a high volatility to represent the ‘forgetting’ of bearing whilst resting. We do not explore that approach further here, preferring to illustrate the key ideas as simply as possible.

ACKNOWLEDGEMENTS

The authors thank Théo Michelot and anonymous referees for their helpful and insightful comments that have improved this text.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

A TECHNICAL DETAILS OF THE INFERENCE ALGORITHM

A.1 CONDITIONAL DISTRIBUTION FOR BEHAVIOURAL PARAMETERS

The full conditional distribution, $\mathcal{L}(\Phi_B ; \mathbf{B}, \boldsymbol{\theta}, \mathbf{v}, \mathbf{z}, \Phi_M)$, simplified as $\mathcal{L}(\Phi_B ; \mathbf{B})$, is the posterior for a fully observed continuous-time Markov chain. Sufficient statistics for such a process are given by a_i , the total time spent in state i , and $b_{i,j}$, the number of transitions from state i to state j . Given independent, exchangeable prior distributions of $\lambda_i \sim \text{Gamma}(c_i, d_i)$ and $q_{i,1}, \dots, q_{i,n} \sim \text{Dirichlet}(\mathbf{f}_i)$, the posterior distribution is

$$\lambda_i | \mathbf{B} \sim \text{Gamma} \left(c_i + \sum_{j=1}^n b_{i,j}, d_i + a_i \right),$$

$$q_{i,1}, \dots, q_{i,n} | \mathbf{B} \sim \text{Dirichlet}(\mathbf{f}_i + \mathbf{b}_i), \quad \text{where } \mathbf{b}_i = \{b_{i,1}, \dots, b_{i,n}\}.$$

A.2 CONDITIONAL DISTRIBUTION FOR MOVEMENT PARAMETERS

The full conditional distribution, $\mathcal{L}(\Phi_M ; \Phi_B, \mathbf{B}, \boldsymbol{\theta}, \mathbf{v}, \mathbf{Z})$, simplified to $\mathcal{L}(\Phi_M ; \Phi_B, \mathbf{B}, \boldsymbol{\theta}, \mathbf{v})$ when there is no observation error present, is given as

$$\mathcal{L}(\Phi_M ; \Phi_B, \mathbf{B}, \boldsymbol{\theta}, \mathbf{v}) \propto \mathcal{L}(\Phi_M ; \Phi_B, \mathbf{B}) \mathcal{L}(\boldsymbol{\theta}, \mathbf{v} ; \Phi, \mathbf{B}),$$

up to a constant. Above, $\mathcal{L}(\Phi_M ; \Phi_B, \mathbf{B}) = \mathcal{L}(\Phi_M)$ is the density of the prior of the movement parameters and

$$\mathcal{L}(\boldsymbol{\theta}, \mathbf{v} ; \Phi, \mathbf{B}) = \pi_\theta(\theta_1) \pi_v(v_1 | \Phi) \prod_{i=2}^M \pi_\theta(\theta_i | \theta_{i-1}, \Phi, \mathbf{B}) \pi_v(v_i | v_{i-1}, \Phi, \mathbf{B}),$$

where $\theta_1 \sim \text{U}(-\pi, \pi)$,

$$\pi_v(v_1 | \Phi_M, \Phi_B) = \sum_{i=1}^n \pi_v(v_1 | \Phi_M, s_0 = i) \mathcal{L}(s_0 = i ; \Phi_B),$$

using the equilibrium distribution of the Ornstein–Uhlenbeck process for the initial speed likelihood. The conditional likelihoods in the product above are given by Eqs. 1.

A.3 CONDITIONAL DISTRIBUTION FOR THE UNOBSERVED REFINED PATH

The full conditional distribution of a section of the refined path, up to a constant, needed for the Metropolis–Hastings step can be written as

$$\begin{aligned} & \mathcal{L} \{ \mathbf{B}^*, \boldsymbol{\theta}^*, \mathbf{v}^* ; \Phi, \mathcal{F}, \mathbf{Z}(b) \} \\ &= \mathcal{L} \{ \mathbf{B}^*, \boldsymbol{\theta}^* ; \Phi, \mathcal{F}, \mathbf{Z}(b) \} \mathcal{L} \{ \mathbf{v}^* ; \mathbf{B}^*, \boldsymbol{\theta}^*, \Phi, \mathcal{F}, \mathbf{Z}(b) \} \\ &\propto \mathcal{L} \{ \mathbf{B}^*, \boldsymbol{\theta}^* ; \Phi, \mathcal{F} \} \mathcal{L} \{ \mathbf{Z}(b) ; \mathbf{B}^*, \boldsymbol{\theta}^*, \Phi, \mathcal{F} \} \mathcal{L} \{ \mathbf{v}^* ; \mathbf{B}^*, \boldsymbol{\theta}^*, \Phi, \mathcal{F}, \mathbf{Z}(b) \} \end{aligned}$$

where $\mathcal{F} = \{B(a), B(b), \theta_0, \theta_n, v_0, v_n, \mathbf{Z}(a)\}$. The simulation method employed to create a proposal for a refined path section, described in Sect. 3.3.1, results in a proposal distribution proportional to $\mathcal{L}\{\mathbf{B}^*, \boldsymbol{\theta}^*; \Phi, \mathcal{F}\} \mathcal{L}\{\mathbf{v}^*; \mathbf{B}^*, \boldsymbol{\theta}^*, \Phi, \mathcal{F}, \mathbf{Z}(b)\}$, and so the Metropolis–Hastings acceptance ratio is based only on the marginal distribution $\mathcal{L}\{\mathbf{Z}(b); \mathbf{B}^*, \boldsymbol{\theta}^*, \Phi, \mathcal{F}\}$ given in Eq. 4.

REFERENCES

- P.G. Blackwell. Random diffusion models for animal movement. *Ecological Modelling*, 100:87–102, 1997.
- P.G. Blackwell. Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika*, 90(3):613–27, 2003.
- P.G. Blackwell, M. Niu, M.S. Lambert, and S.D. LaPoint. Exact Bayesian inference for animal movement in continuous time. *Methods in Ecology and Evolution*, 2015. 10.1111/2041-210X.12460.
- D.E. Boakes, E.A. Codling, G.J. Thorn, and M. Steinke. Analysis and modelling of swimming behaviour in *Oxyrrhis marina*. *Journal of Plankton Research*, 33(4):641–9, 2011.
- G.A. Breed, D.P. Costa, I.D. Jonsen, P.W. Robinson, and J. Mills-Flemming. State-space methods for more completely capturing behavioral dynamics from animal tracks. *Ecological Modelling*, 225–236:49–58, 2012.
- F.E. Buderman, M.B. Hooten, J.S. Ivan, and T.M. Shenk. A functional model for characterizing long-distance movement behaviour. *Methods in Ecology and Evolution*, 7:264–73, 2016.
- E. Codling and N.A. Hill. Sampling rate effects on measurements of correlated and biased random walks. *Journal of Theoretical Biology*, 233:573–88, 2005.
- M.L. Eaton. *Multivariate statistics: a vector space approach*. Lecture notes monograph series. Institute of Mathematical Statistics, Beachwood, OH, 2 edition, 2007.
- C.H. Fleming, J.M. Calabrese, T. Mueller, K.A. Olson, P. Leimgruber, and W.F. Fagan. From fine-scale foraging to home ranges: a semivariance approach to identifying movement modes across spatiotemporal scales. *The American Naturalist*, 183(5):E154–67, 2014.
- A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. *Bayesian Data Analysis*. Texts in statistical science series. CRC Press, Boca Raton, FL, 3 edition, 2013.
- E. Gurarie, R.D. Andrews, and K.L. Laidre. A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, 12:395–408, 2009.
- P. Guttorp. *Stochastic modeling of scientific data*. Stochastic modeling series. Chapman and Hall/CRC, Boca Raton, FL, 1995.
- E.M. Hanks, M.B. Hooten, D.S. Johnson, and J.T. Sterling. Velocity-based movement modelling for individual and population level inference. *PLoS ONE*, 6(8):e22795, 2011.
- E.M. Hanks, M.B. Hooten, and M.W. Allredge. Continuous-time discrete-space models for animal movement. *The Annals of Applied Statistics*, 9(1):145–65, 2015.
- K.J. Harris and P.G. Blackwell. Flexible continuous-time modelling for heterogeneous animal movement. *Ecological Modelling*, 255:29–37, 2013.
- A. Hobolth and E.A. Stone. Simulation from endpoint-conditioned, continuous-time Markov chains on a finite state space, with applications to molecular evolution. *Annals of Applied Statistics*, 3(3), 2009.
- M.B. Hooten and D.S. Johnson. Basis function models for animal movement. *Journal of the American Statistical Association*, in press. [arXiv:1601.05408](https://arxiv.org/abs/1601.05408) [stat.ME].
- A. Hurford. GPS measurement error gives rise to spurious 180 degrees turning angles and strong directional biases in animal movement data. *PLOS ONE*, 4(5), 2009.
- S.M. Iacus. *Simulation and inference for stochastic differential equations*. Springer series in statistics. Springer, New York, NY, 2008.
- D.S. Johnson, J.M. London, M. Lea, and J.W. Durban. Continuous-time correlated random walk model for animal telemetry data. *Ecology*, 89(5):1208–15, 2008a.

- D.S. Johnson, D.L. Thomas, J.M. Ver Hoef, and A. Christ. A general framework for the analysis of animal resource selection from telemetry data. *Biometrics*, 64:968–76, 2008b.
- I.D. Jonsen, M. Basson, M.V. Bravington, T.A. Patterson, M.W. Pederson, R. Thomson, U.H. Thygesen, and S.J. Wotherspoon. State space models for bio-loggers. *Deep-Sea Research II*, 88–9:34–46, 2013.
- P.M. Kareiva and N. Shigesada. Analyzing insect movement as a correlated random walk. *Oecologia*, 56(2):234–8, 1983.
- B. Kranstauber, K. Safi, and F. Bartumeus. Bivariate Gaussian bridges: directional factorization of diffusion on brownian bridge models. *Movement Ecology*, 2(5):1–10, 2014.
- C.E. Kuhn, D.S. Johnson, R.R. Ream, and T.S. Gelatt. Advances in the tracking of marine species: using GPS locations to evaluate satellite track data and a continuous-time movement model. *Marine Ecology Progress Series*, 393:97–109, 2009.
- R. Langrock, R. King, J. Matthiopoulos, L. Thomas, D. Fortin, and J.M. Morales. Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, 93(11):2336–42, 2012.
- B.T. McClintock, R. King, L. Thomas, J. Matthiopoulos, B.J. McConnell, and J.M. Morales. A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs*, 82(3):335–49, 2012.
- B.T. McClintock, D.S. Johnson, M.B. Hooten, J.M.V. Hoef, and J.M. Morales. When to be discrete: the importance of time formulation in understanding animal movement. *Movement Ecology*, 2(21), 2014.
- J.F. McEvoy, D.A. Roshier, R.F.H. Ribot, and A.T.D. Bennett. Proximate cues to phases of movement in a highly dispersive waterfowl, *Anas Supercilliosa*. *Movement Ecology*, 3(21), 2015.
- A.E. McKellar, R. Langrock, J.R. Walters, and D.C. Kesler. Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird. *Behavioral Ecology*, 26(1):148–57, 2015.
- T. Michelot, R. Langrock, and T.A. Patterson. moveHMM: an R package for the statistical modelling of animal movement data using hidden markov models. *Methods in Ecology and Evolution*, 7(11):1308–15, 2016.
- J.M. Morales, D.T. Haydon, J. Frair, K.E. Holsinger, and J.M. Fryxell. Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, 85(9):2436–45, 2004.
- V.O. Nams. Sampling animal movement paths causes turn autocorrelation. *Acta Biotheoretica*, 61(2):269–84, 2013.
- V.O. Nams. Combining animal movement and behavioural data to detect behavioural states. *Ecological Letters*, 17(10):1228–37, 2014.
- A. Parton, P.G. Blackwell, and A. Skarin. Bayesian inference for continuous time animal movement based on steps and turns. In R. Argiento, E. Lanzaron, I. Antoniano Villalobos, and A. Mattei, editors, *Bayesian Statistics in Action*, volume 194 of *Springer Proceedings in Mathematics & Statistics*. Springer International Publishing, 2017. 10.1007/978-3-319-54084-9. [arXiv:1608.05583](https://arxiv.org/abs/1608.05583).
- T.A. Patterson, B.J. McConnel, M.A. Fedak, M.V. Bravington, and M.A. Hindell. Using GPS data to evaluate the accuracy of state-space methods for correction of Argos satellite telemetry error. *Ecology*, 91(1):279–85, 2010.
- T.A. Patterson, A. Parton, R. Langrock, P.G. Blackwell, L. Thomas, and R. King. Statistical modelling of animal movement: A myopic review and a discussion of good practice. *Advances in Statistical Analysis*, in press. [arXiv:1603.07511](https://arxiv.org/abs/1603.07511) [stat.AP].
- V. Rao and Y.W. Teh. Fast MCMC sampling for Markov jump processes and extensions. *Journal of Machine Learning Research*, 4:3295–320, 2013.
- J.M. Rowcliffe, C. Carbone, R. Kays, B. Kranstauber, and P.A. Jansen. Bias in estimating animal travel distance: the effect of sampling frequency. *Methods in Ecology and Evolution*, 2012. 10.1111/j.2041-210X.2012.00197.x.
- H. Rue and L. Held. *Gaussian Markov random fields. Theory and applications*. Monographs on Statistics and Applied Probability 104. Chapman and Hall/CRC, Boca Raton, FL, 2005.
- G.A. Whitaker, A. Golightly, R.J. Boys, and C. Sherlock. Improved bridge constructs for stochastic differential equations. *Statistics and Computing*, (4):885–900, 2016.