# Statistical Modelling of Collective Animal Movement

## with an application to reindeer movement in northern Sweden.


**Fay Frost**

School of Mathematics and Statistics
Faculty of Science
University of Sheffield

This dissertation is submitted for the degree of
*Doctor of Philosophy*

March 2021

# Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and in the following paragraph.

The work of Chapter 3 was a collaborative effort of myself, Mu Niu, Paul Blackwell, Jordan Milner and Anna Skarin and published in Niu et al. (2020). Their contributions are described in the introduction of that chapter. Section 4.4 is also presented in Niu et al. (2020). Additionally, Paul Blackwell provided a supervisory role for all of the work within this thesis.

Fay Frost
March 2021

# Acknowledgements

Firstly, I would like to thank my supervisor Paul Blackwell for all the support during my PhD. His patience and endless calm has been a pleasure to work with and I value the fact that he never complained when asked to repeat himself multiple times. I would also like to extend my appreciation to my co-supervisor Anna Skarin, who has been a friendly support throughout, and the reindeer herders of Sweden who kindly granted me the unique experience of visiting the reindeer herding district.

I would like to thank all my friends at my second home at St. Barnabas Rd for all the living room downtime and boardgames; thanks to Alex, Dave, Francis, Harry, Josh, Sarah, Nessie and especially Beastie, the house mascot and the constant thread during the pandemic. Equally, I would like to thank my own housemates; thanks to Ellie, especially for her great food, Mim who has been a huge support through times of need and Matt who has been an all-round inspiring life-coach.

I would like to give a special thank you to Charlotte Ward, who has been a close and wonderful friend since my first arrival at Sheffield University as an undergraduate and whose baked goods have got me through my PhD.

I would like to give my acknowledgements to my Dad, John, my Mum, Louise and my sister, Libby who have always been encouraging and given me belief and confidence in what I do. Finally, and by no means least, I would like to express my immense gratitude to my partner Emma for all her love, support and care.

# Abstract

The ways in which animals move are a complex phenomena, from small scale interactions to larger migratory movement. Internal and external stimuli govern a variety of behavioural patterns whose processes are vital for species survival. Analysing these movement and behavioural processes can have significant applications for conservation and management. Although there are many statistical tools readily available for investigating animal movement, they are largely directed towards individual-level cases and do not consider the group movement present in collective species such as ungulates.

This thesis aims to redress the shortcomings of statistical literature by providing a modelling framework for collective animal movement in continuous time. Our modelling approach builds upon general themes of group movement originally put forward by Langrock et al. (2014), where each individual in the group is at times attracted to an unobserved leading point. However, the behaviour of each individual can switch between 'following the group' and 'moving independently', modelled as an Ornstein Uhlenbeck process and Brownian motion respectively. The movement of the leading point is also modelled as an Ornstein-Uhlenbeck process or, if we forgo the leader's drift term, as Brownian motion. An inhomogeneous Kalman filter Markov chain Monte Carlo algorithm is used to estimate the diffusion and switching parameters and the behavioural states of each individual at a given time point.

We assess the model's performance in a variety of simulated settings before providing a real world application using the location data of semi-domesticated reindeer (*rangifer tarandus*). We extend this methodology by allowing switching to depend explicitly on covariate information. We define a general auxiliary model for the inclusion of covariate data which accounts for a wide range of environmental heterogeneity. We give a simulated illustration where the animals switch behaviour sinusoidally depending on the time of day. Then, we revisit the reindeer application by including covariate data on insect harassment, which is thought to influence reindeer movement.

# Table of contents

# List of figures

# List of tables

# Chapter 1

# Introduction

The ways in which animals move are a complex phenomena, from small scale interactions with the environment to larger migratory movement. Internal and external stimuli govern behavioural patterns in animals which lead to individuals exhibiting a variety of movement processes. These processes are vital for species survival. Historically, our understanding of animal movement was by reason of observational studies which are clearly limited to non-elusive focal species. As animal tracking technology has advanced we have been gifted with a wealth of data from myriad species. However, the statistical tools required to analyse such data have not evolved at the same rate. That said, there has been a drive to develop tools that underpin movement mechanisms and give us better insight into the intricacy of the natural world, helping us to overcome ecological challenges such as mitigating species loss.

Knowing where animals go and why can contribute to a detailed understanding of animal ecology (Cagnacci et al., 2010). The increasing wealth of knowledge surrounding animals' resource use, movement patterns and behavioural drivers can have significant applications for conservation, food and disease spread (Kays et al., 2015). Existing research in animal movement has shed light on home ranges (Merkle et al., 2017; Moorcroft et al., 1999), territory formation (Potts and Lewis, 2014), social interactions within collective migration (Torney et al., 2018), sites of interest (Munden et al., 2019), resource selection (Wang et al., 2019) and group cohesion (Bode et al., 2011) as well as a multitude of other topics of interest.

At present we live in constant change where the natural world suffers the consequences of rapid alterations in which species face extinction (Thomas et al., 2004). The climate is changing drastically, especially around the polar regions (Pörtner et al., 2019) and both the direct and indirect effects can force unnecessary demands on species (Pape and Löffler, 2012; Tyler et al., 2007; Vistnes et al., 2004). For example, direct consequences of climate

change can be seen in Arctic ice sheet melt leading to higher sea levels and threats to coastal habitats (Pörtner et al., 2019) and warmer winters see more rain-on-snow events where vital forage is compacted beneath hardened ice (Forbes et al., 2016). On the other hand, the increase in the abundance of warm weather parasitic insects such as mosquitoes is an indirect consequence of climate change (Gunn, 2016). Additionally, in an effort to avoid anthropogenic activities such as infrastructure development and agriculture, animals may relocate and select alternative resources that may not be as well suited (Hall et al., 2006; Polfus et al., 2011; Rowland et al., 2000; Skarin and Åhman, 2014; Vistnes et al., 2008).

This has motivated considerations of animal movement modelling in response to environmental cues. Enriching our understanding of movement by quantifying behavioural drivers can help to put provisions in place to mitigate biodiversity loss and promote the well-being and conservation of species.

## 1.1 Animal Movement: Data and Modelling

When an interest in animal ecology first emerged, understanding animal behaviour was a challenging task. Often, it involved laborious hours of direct field observations, clearly limited by observer bias and visibility of the species. Later, technologies which recorded locations of animals were used. These involved recapturing the individual to obtain the data and only very low resolution observations could be made. However, a rapid improvement in technology has led to an increase in accuracy, duration and taxa studied (Hooten et al., 2018). High resolution location observations with multiple fixes per day, hour and even minutes can now be made remotely in an unobtrusive fashion, allowing animals to behave more naturally. This has provided the ecological community with exciting revelations such as the circumpolar migrations of the albatross (Croxall et al., 2005) or the complex and collective decision processes of baboons (Strandburg-Peshkin et al., 2015); discoveries that would not have been possible using traditional methods. Thus, our understanding of animal movement is growing in complexity and more interesting biological and ecological questions can be answered with improved precision.

Modern animal movement data is a collection of an animal's locations throughout time. This data is usually collected in the form of an electronic tag. However, these devices have changed significantly since their early use in the 1960's (Cagnacci et al., 2010; Hebblewhite and Haydon, 2010). In the past, very high frequency (VHF) radio transmitters were popular tags. Early instances of their use were in studies to track Yellowstone grizzly bear winter

den sites (Craighead and Craighead, 1972) and incidentally the respiration of a mallard duck whilst in a flying motion (Lord et al., 1962). However, VHF telemetry devices are variable in accuracy with precision of 200-600m (Thomas et al., 2011; Zimmerman and Powell, 1995). Now, more commonly used telemetry devices for the collection of animal locations are the Global Positioning System (GPS) and the Argos system, both of which have been heavily implemented since the early 1990's. There has been much discussion on the accuracy of these systems but it is generally agreed upon that the location error of GPS devices are usually smaller than that of the Argos system (Frair et al., 2010; Patterson et al., 2010), with GPS devices providing a location accuracy of <30m (Tomkiewicz et al., 2010). Throughout this thesis we will be concerned only with GPS movement data. However, it is worth noting that there are many other types of movement data available such as capture-recapture and ringing which are not discussed in this thesis.

GPS data is usually recorded in two dimensions for land-based animals. This is naturally given as the latitude and longitudinal coordinates. For marine or aerial animals GPS data is typically given in the two-dimensional horizontal plane with an additional recording of depth or altitude. Location data can be used simultaneously with other bio-logging devices such as accelerometers and heart-rate monitors which provide physiological information e.g. energetic expenditure (Cooke et al., 2004). In recent developments, oceanographers have used bio-logging devices attached to marine mammals to record environmental information such as temperature, salinity and currents to create detailed profiles of the Southern Indian ocean (Roquet et al., 2014).

In recent years GPS tracking device technology has significantly improved. The batteries which once powered them are now lighter and in some cases can be recharged on the move with solar power. That said, it is now possible to ethically tag animals for longer with the possibility of increasingly higher resolutions; sampling frequencies can now be as little as every minute for multiple years. The sampling frequency often affects the intensity of data management and ease of computational analysis and with tracking data now entering the realm of 'big data' (Kays et al., 2015) it is important that the sampling frequency is carefully chosen by what is of interest. Taking location observations once a day will be of little use to questions about small scale behavioural patterns, but may well help in studies of migration.

In addition, a decrease in the price of telemetry devices has made multi-individual tagging more economically feasible. This has opened up new avenues for going beyond the conventional studies of individual-level movement and exploring how individuals move in relation to each other. This is particularly useful for investigating collective animal movement or grouping dynamics of a species, which has generally been overlooked in favour of individual-

level studies. However, GPS tags are still significantly more expensive than VHF devices meaning that researchers have to make trade-offs between the accuracy of using GPS units and the sample size in which they can afford to tag (Tomkiewicz et al., 2010).

Despite the rapid and continuing development of animal tracking equipment, the statistical tools needed to analyse such data have not advanced as quickly (Breed et al., 2011; Jonsen et al., 2003). Descriptions of individual-level movement are commonplace in the literature but there has been slower progress for collective movement, which is essential for social species whose movement is dependent on others. Inferences of group-level behaviour necessitate the collection of location data from multiple animals which has only recently become technologically feasible. Thus, there is a new drive for getting up to speed with statistical methods for analysing group-level movement. Some studies have approached this gap by providing metrics for collective movement such as statistics that quantify proximity and co-ordination but these are usually limited to dyadic interactions (Joo et al., 2018; Long et al., 2014). Whilst these practices can be useful, the results are unlikely to be consistent depending on which animals are tagged within a group. As with individual-level studies, we require an explicit and coherent model of movement but with additional properties to account for dependency between individuals. Although the affordability of electronic tags has enabled extensive deployment to simultaneously track multiple individuals, some species collectives are too large to practically or affordably tag every individual. Thus, a useful and favourable model of movement is one whose inference is invariant on who or how many individuals are tagged.

Langrock et al. (2014) provide an explicit model of group movement where each animal is linked indirectly via a *'leading'* point of attraction (see Section 3.1 for a more detailed description). This offers a good foundation for modelling group movement however, their model and inference method are limited to discrete time. This is not an uncommon practice since location data from an electronic tag are usually obtained at discrete time points and many models use this directly for convenience. However, many problems can arise with this technique. Firstly, regardless of resolution, location data from electronic tags come with their own errors and irregularities such as missing and delayed observations which can hinder fitting models (Niu et al., 2016) and inferences on movement and behavioural patterns (Frair et al., 2010). The inference of movement parameters in discrete-time models are also limited to a fixed 'grid' of times usually determined by the sampling frequency of the observations. The scale-dependent nature of these models means that it can be difficult to use the same model to compare datasets with different observation sampling schemes or to analyse data with missing/delayed observations. One avenue to combat these problems is to develop

continuous-time models which naturally fit with the way in which animals move (Harris and Blackwell, 2013). However, this area has been under-developed largely due to difficulty in parameter interpretation and computational demand.

Niu et al. (2016) alleviates some of these issues by presenting novel methodology for modelling group animal movement in continuous time using diffusion processes (see Section 2.5.1). The underlying theme echoes that of Langrock et al. (2014) insofar as each individual of the group is attracted to a (generally unobserved) leading point, and the movement of the leading point exhibits its own movement process with attraction to an unknown attractor. However, this methodology restricts movement to a single underlying process. In reality, animals display a range of behaviours and movement patterns which ought to be accounted for. Thus, a model of group movement with the possibility of behavioural switching would relieve this naivety.

## 1.2 Thesis Aims

This thesis aims to redress the shortcomings of group movement literature by providing a statistical framework for collective animal movement in continuous time with behavioural switching. The material extends previous work of Niu et al. (2016), by acknowledging the diversity of animal movement we allow individuals to switch between 'following the group' and 'independent movement'. In addition, we give flexibility to the model by allowing the leader to have no attraction point. This is particularly useful if there is no tractable point of attraction on the time scale of the data, for example if the duration is short relative to the movement towards a particular location; this work is presented in Niu et al. (2020).

We assess whether the model can distinguish between different behavioural modes under contrasting environmental settings. Specifically, we use the real location data of semi-domesticated reindeer (*rangifer tarandus*) during two different time periods in summer, early and peak, where we expect low or high parasitic harassment respectively.

We formalise this methodology by allowing behavioural switching to depend on covariate information. We define a general auxiliary model for the inclusion of covariate data to account for a wide range of environmental heterogeneity. We use this covariate model to revisit reindeer grouping dynamics in response to insect harassment. We use weather data as a proxy for insect presence and investigate a variety of harassment models as a covariate for movement. We give a second illustration using a simulation experiment where the animals switch behavioural sinusoidally depending on the time of day.

## 1.3 Thesis Outline

**Chapter 2: Statistical Methodology and Literature Review**

This chapter initially presents the statistical methodology required for reviewing existing literature of animal movement models. Specifically, we discuss state space models (SSM), Kalman filters and Markov chain Monte Carlo (MCMC). Then we review current and common approaches of modelling animal movement in both discrete and continuous time. We give detailed descriptions of popular tools used in animal movement models such as random walks, Hidden Markov models, Brownian motion and Ornstein-Uhlenbeck diffusion processes. This chapter is restricted to individual-level movement models and we reserve discussion of group-level models for Chapter 3.

**Chapter 3: Modelling Collective Movement**

We begin by discussing the importance of collective movement in the animal kingdom and highlight that cohesion and coordination in some species is necessary for their survival. We acknowledge that the statistical tools used to analyse such phenomena are largely understudied which motivates the need for research in modelling collective animal movement. We review existing literature of modelling group movement with special focus on the Ornstein Uhlenbeck approach presented in Niu et al. (2016). We highlight the possibility of a leading point having no attractor and demonstrate a non-stationary version of the model in this case. Section 3.4 presents an extension to this model which allows for behavioural switching between two modes; following the group and moving independently. Section 3.5 provides inference for the model including details of the trajectories simulation, state space formulation and application of the inhomogeneous Kalman filter. The chapter is finalised with several simulation experiments using data-sets with varying underlying processes.

**Chapter 4: Case Study: Collective Reindeer Movement**

This chapter gives an application of the movement models presented in Chapter 3 using the real location data of reindeer (*rangifer tarandus*). We begin by understanding the importance of reindeer populations and what conservation and management implications modelling their movement may have. We provide insight into reindeer habitat, behavioural traits and the challenges they face as a result of the direct and indirect consequences of anthropogenic activities and climatic change. Section 4.2 gives a description of the reindeer movement data collected by my co-supervisor, Anna Skarin. Section 4.3 applies the original non-switching model of Niu et al. (2016). We then compare this to an application of the switching model presented in Section 4.4 to the same location data.

**Chapter 5: Reindeer Grouping Strategies for the Relief of Insect Harassment**

This chapter presents a case study of the switching model by highlighting reindeer grouping strategies for the relief from insect harassment. We begin by introducing the concept of insect harassment and stress the detrimental effects parasitic pressure can have on reindeer. Section 5.1 reviews indices of insect harassment, primarily using weather variables as a proxy for insect presence. In Section 5.3 and 5.4 we apply the switching model presented in Section 3.4 to two datasets of reindeer locations assumed to be during contrasting levels of insect harassment. We investigate whether the model is able to distinguish between different behavioural patterns and give results consistent with the knowledge of reindeer movement. We show that the model captures contrasting movement and behavioural processes during those two time periods.

**Chapter 6: A Model for Group Movement with Switching and Covariate Data**

Motivated by the results of Chapter 5, we acknowledge that animal movement is often influenced by internal or external stimuli. As a result, in this chapter we develop a framework for explicitly including covariate information by allowing switches in behaviour to be dependent on ancillary data. Section 6.2 provides an illustration of this covariate model using simulated data whose switching rates depend sinusoidally on the time of day. We conclude the chapter by revisiting the reindeer application given in Chapter 5 where we formally include covariate data on insect harassment into the model. We compare this to the previous application using the non-covariate model where the switching rates were not dependent on insect harassment.

**Chapter 7: Discussion and Further Work**

This chapter provides a summary of the thesis and gives a discussion of the conducted research. I conclude with a description of potential avenues for future work which include general modifications to the modelling approach and developments specific to reindeer applications.

# Chapter 2

# Statistical Methodology and Literature Review

In this chapter we present statistical methodology which will reoccur throughout this thesis and be useful for reviewing the existing literature of animal movement models. We begin by introducing state-space models (SSMs); these are a popular and useful framework now used extensively in ecology. We review the Kalman filter algorithm which allows for exact inference on a specific group of SSMs. We also discuss other methods for fitting models to ecological data using frequentist approaches such as maximum likelihood estimation (MLE) and Bayesian methods such as Markov chain Monte Carlo. Then we discuss current and common approaches to modelling animal movement, separating our overview into discrete and continuous time models.

## 2.1 State-space Models (SSM)

State-space models (SSMs) are a popular modelling framework for time-series data; they can alleviate the stumbling blocks of analysing inherently complex movement patterns coupled with measurement errors associated with location data. This in part is because an SSMs are structured in a way that accounts for two levels of variability (Auger-Méthé et al., 2016); stochasticity in the underlying process is separated from that of the measurement error (Jonsen et al., 2005, 2003; Patterson et al., 2008). For example, with location data, variation in the animal movement is modelled separately to the variation in measurement error from an electronic tag i.e. the difference between the observation of the tag and the animal's true

location. In addition, they are often used to estimate the dynamics of a phenomenon that cannot be observed directly; this may be discrete such as an animal's behavioural state or continuous for example, the true location. As a result, SSMs have been a favoured approach in movement modelling for over a decade (Breed et al., 2012; Patterson et al., 2008).

The main framework of an SSM is a hierarchy of two models; the process model and the observation model. The process model, which in some literature may be referred to as the transition or state equation, is given by

$$\boldsymbol{y_t} = g(\boldsymbol{y}_{t-1}, \eta_t). \tag{2.1}$$

This describes the process of the unobservable (or hidden) states with process stochasticity given by $\eta_t$. Note that the equation has the Markovian property where the state of the system at time $t$ depends only on the immediate previous state. The observation or measurement model given by

$$\boldsymbol{z_t} = h(\boldsymbol{y}_t, \varepsilon_t), \tag{2.2}$$

gives the relationship between what is observed at time $t$, $z_t$, and the corresponding true hidden state, $y_t$, with observation error parameter $\varepsilon_t$. Thus, the observations are independent once we account for their dependence on the state. See Figure 2.1.

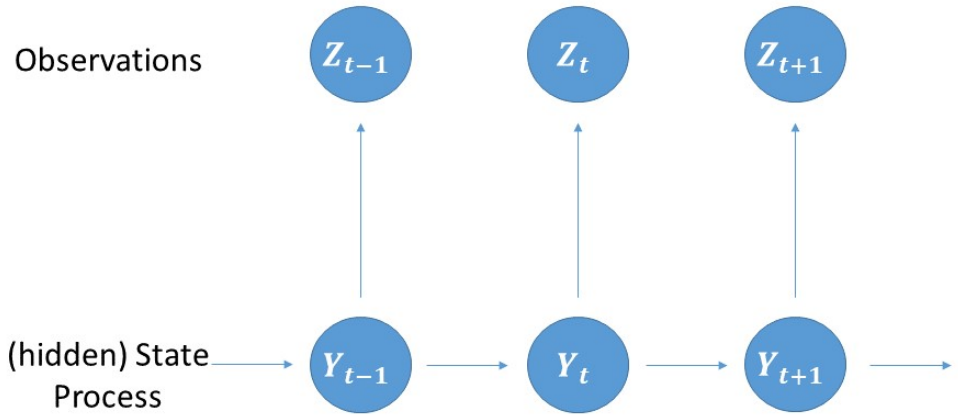


Fig. 2.1 The general structure of state-space models. The $Z_t$ represent observations and the $Y_t$ represent the true but unobserved states. Horizontal arrows depict the process model (Equation 2.1). Vertical arrows depict the observation model (Equation 2.2).

A state-space model in its simplest form is linear and Gaussian; it is said to be linear if the state and observation equations are linear and Gaussian if the transition density of the state process and the observation noise are normal. However, SSMs are flexible in that they can model linear and nonlinear processes. A simple linear Gaussian SSM is sometimes referred to as a linear dynamical system.

State-space models have many benefits. Firstly, they account for spatio-temporal autocorrelation inherent in movement data. Historically autocorrelation was often treated as a nuisance and in an effort to deal with it resulted in sub-sampling data to reduce autocorrelation between locations (Patterson et al., 2010); however, SSMs avoid throwing away possibly useful information by accounting for autocorrelation within their modelling framework. Secondly, although movement data is almost always obtained at discrete intervals, models for animal movement can be conceptualised in either discrete or continuous time. SSMs are able to account for either representation of movement (McClintock et al., 2014). They can also be useful for dealing with irregularly observed data (Jonsen et al., 2005).

## 2.2 The Kalman Filter

Filtering is an umbrella term for tools which smooth noisy data by updating knowledge of a system based on observations (Durbin and Koopman, 2001). The Kalman filter is a specific filtering algorithm named after Rudoph Kalman (Kalman, 1960), which allows for exact inference in a linear dynamical system; that is, where the state of the system at time $t+1$ evolves linearly from some prior state at time $t$. It is particularly useful if the true state of the system cannot be directly observed and instead is estimated from some measurement and system dynamics.

At each iteration of the algorithm the Kalman filter makes a prediction of the state at the next time point using the models of the system. This prediction is then updated by augmentation with observed data. The augmentation is via a weighted average of the prediction and observation known as the Kalman gain, with larger weight given to more confident measurements. The intuition behind this is that a better estimate can be made by the combination of knowledge from the prediction and measurements.

The Kalman filter has been applied in a multitude of engineering and computer science concepts, with an early but highly celebrated application being its use in the Apollo navigation system. More recent times have seen an increased use in ecological settings to estimate trajectories from location data, incorporate location measurement error, estimate movement

parameters and compute likelihoods (Breed et al., 2012; McClintock et al., 2012; Patterson et al., 2010). The Kalman filter's computational efficiency allows for speedy maximum-likelihood estimation of system parameters (Fleming et al., 2017; Patterson et al., 2010) where irregularly spaced data may be dealt with via continuous-time process models (Johnson et al., 2008).

The details of the Kalman filter are as follows. Given the state $y_t$, the system process follows a linear dynamical system given by

$$y_t = F_t y_{t-1} + B_t u_t + w_t, \tag{2.3}$$

where,

- $y_t$ is the vector containing the states of the system which are of interest,
- $F_t$ is the state transition matrix of the model,
- $B_t$ is the control input which is applied to the control vector $u_t$ and
- $w_t$ is a noise vector assumed to be drawn from a multivariate normal $w_t \sim \mathcal{N}(0, Q_t)$.

Also, measurements $z_t$, of the system at time $t$ are represented according to

$$z_t = H_t y_t + v_t \tag{2.4}$$

where,

- $H_t$ maps the state vector into the measurement domain and
- $v_t$ is a vector containing the measurement noise assumed to be the multivariate normal $v_t \sim \mathcal{N}(0, R_t)$.

The Kalman filter then alternates iteratively between two steps, prediction and updating. The prediciton equations are given by

- **Prediction Step:**

$$m_{t|t-1} = F_t m_{t-1|t-1} + B_t u_t,$$
$$P_{t|t-1} = F_t P_{t-1|t-1} F_t^T + Q_t.$$

The measurement updating equations are given by

- **Updating Step:**

$$\begin{aligned} m_{t|t} &= m_{t|t-1} + K_t(z_t - H_t m_{t|t-1}), \\ P_{t|t} &= P_{t|t-1} - K_t H_t P_{t|t-1}, \\ K_t &= P_{t|t-1} H_t^T \left(H_t P_{t|t-1} H_t^T + R_t\right)^{-1} \quad \textbf{(Kalman Gain)}. \end{aligned}$$

The Kalman filter iterations output an estimate of the state $y_{t_i}$, denoted by $m_{t_i}$ and covariance matrix $P_{t_i}$ for each time $t_i$. These estimates are conditional on all previous observations.

In the case of modelling animal movement, $y_t$ may represent the true (unobservable) location of an animal with movement model given by $F_t$ and zero control input. Whereas, $z_t$ may represent an measurement of that location, given by an electronic tag, with $H$ as the identity matrix and $v_t$ denoting any measurement error.

We can use this to calculate, as a by-product, the marginal likelihood of our observations given the parameters. This is especially useful as it enables us to calculate the likelihood using only what we observe. The log-likelihood for our set of observations, $\mathbf{z} = \{z_0, \ldots, z_T\}$ is given as

$$\begin{aligned} -\sum_{i=0}^{T} \frac{1}{2} \big\{ & n \log 2\pi + \log |H_i P_{t_i|t_{i-1}} H_i^{\mathrm{T}}| \\ & + (Z_{t_i} - H_i m_{t_i|t_{i-1}})^{\mathrm{T}} (H_i P_{t_i|t_{i-1}} H_i^{\mathrm{T}})^{-1} (Z_{t_i} - H_i m_{t_i|t_{i-1}}) \big\}, \end{aligned} \tag{2.5}$$

where $n$ is the dimension of the measurement vector. The equations may be simplified if we assume no process or observational noise. For more information on the derivation of these equations consult Särkkä (2013).

## 2.3 Markov chain Monte Carlo

Markov chain Monte Carlo (MCMC) methods are a class of algorithms for sampling from a probability distribution. They are particularly useful in Bayesian statistics when evaluating the posterior distribution of model parameters $\theta$ given observed data $D$,

$$P(\theta|D) = \frac{P(\theta)P(D|\theta)}{\int P(\theta)P(D|\theta)d\theta}, \tag{2.6}$$

is analytically infeasible. In light of intractable evaluations, MCMC methods can approximate a posterior distribution by drawing samples from the distribution in question. To explain how this works we must first introduce some theory of Markov chains.

A Markov chain is a stochastic process $X_t$ which has a '*lack-of-memory*' property. In other words at each time $t \geq 0$ the next state in the chain $X_{t+1}$ is only dependent on its current state, $X_t$, and no further in the past i.e.

$$P(X_{t+1}|X_t, X_{t-1}, \ldots X_0) = P(X_{t+1}|X_t). \tag{2.7}$$

Under certain technical conditions, this process will eventually converge to a stationary distribution independent of $t$ and its initial value $X_0$. As the name suggests, Markov chains are fundamental to the theory of MCMC methods. The general idea is to construct a Markov chain whose stationary distribution is the desired posterior distribution $\pi(.)$. A popular algorithm for constructing a chain with these properties is the Metropolis-Hastings (MH) algorithm (Hastings, 1970; Metropolis et al., 1953).

The MH algorithm has two parts; a proposal step and an acceptance step. At each iteration of the algorithm a random sample $X'$ is made, with dependence only on the previous sample $X_t$, from the proposal density $q(X'|X_t)$ . The new sample is then either accepted or rejected probabilistically. The algorithm is as follows:

Initialise the algorithm by selecting a starting point, $X_0$. Then at each iteration $t$, generate a candidate point $X'$ from a proposal distribution $q(.|X_t)$. This candidate point is then accepted with the probability

$$\alpha(X', X_t) = \min\left\{1, \frac{\pi(X')q(X_t|X')}{\pi(X_t)q(X'|X_t)}\right\}. \tag{2.8}$$

If acceptance is successful then the point is retained and we set $X_{t+1}$ to $X'$. Otherwise, the point is rejected and we set $X_{t+1}$ to $X_t$.

Regardless of what form the proposal distribution $q(.)$ takes, the stationary distribution of the chain will still be $\pi$ (see Gilks et al. (1996) for details). This has a very useful consequence insofar as we can choose a symmetric proposal distribution where $q(x|y) = q(y|x)$ so that the Hastings ratio simplifies to

$$\alpha(X_{t+1}, X_t) = \min\left\{1, \frac{\pi(X_{t+1})}{\pi(X_t)}\right\}. \tag{2.9}$$

Although, the convergence rate to the stationary distribution and mixing will depend on the relationship between the proposal distribution $q(.)$ and the posterior $\pi(.)$. Proposal distributions with smaller steps will generally have a higher acceptance rate in the MH algorithm so it is advised to tune proposals such that acceptance rates lie between $20-30\%$ (**?**). Throughout this thesis I make extensive use of this simplifying property by choosing symmetric Gaussian proposal distributions.

## 2.4 Models for Animal Movement

Traditionally, ecologists have focussed attention toward large scale examinations such as species population fluctuations, biodiversity and general patterns of space use (Buckland et al., 2004; Kays et al., 2015). Early low resolution tracking technologies enabled studies of spatial utilisation, home-range and territories using tools such as minimum convex polygons (Holgate, 1971; Macdonald et al., 1980; Moorcroft et al., 1999; Odum and Kuenzler, 1955). Some studies provided methods for quantifying the interaction of individuals termed either static or dynamic; static interaction is defined as the joint-space use between individuals, usually measured by spatial overlap of home ranges whereas dynamic interactions are concerned with correlation in movement (Doncaster, 1990) or how the movement of individuals are related, for example through attraction or avoidance behaviour (Macdonald et al., 1980). Whilst these methods have their uses, they essentially provide summaries of location data rather than giving a description of the animal's trajectory or movement.

Modern tracking technologies are able to collect data at higher resolutions, thus easily capturing short-term movement. This has motivated a more in-depth study of the different movement processes that animals may exhibit over time e.g. (Parton et al., 2016). Recent modelling techniques and inference methods can provide intricate details of animal movement and specific interactions with others or their environment.

There have been many approaches to modelling the movement of animals. These models can be classified by their formulation of time (Parton et al., 2016) namely discrete or continuous. Continuous-time models define movement at any real and positive valued instances whereas, discrete-time models are limited to a fixed 'grid' of times usually determined by the sampling frequency of the observations. Whilst the former seems natural in its assumption of animal

movement, parameter estimation and interpretation is often difficult; as a result discrete-time models are the predominant approach (McClintock et al., 2012)).

The remainder of this chapter aims to give an overview of some common techniques used to model animal movement and discuss their advantages and disadvantages where necessary. We separate our review into models using a discrete-time framework and those using a continuous one. Popular discrete-time methods we discuss are step and turn models and hidden Markov models whereas, for continuous-time models we mainly consider diffusion processes. In both cases we restrict our review to models of individual-level movement and reserve descriptions of group movement for Chapter 3.

### 2.4.1 Discrete-time Models for Movement

In its simplest form, a model for animal movement can be framed as a random walk in space. A random walk is a stochastic Markov process formed of a series of steps. Many discrete-time models of animal movement are based on the random walk and its extensions (Turchin, 1999). A simple example of this is a random walk in 1-dimension with equal probability of moving left or right. Biologically we might think of this as an animal moving along an edge where at each time increment it has moved forwards or backwards. A random walk process can be extended to any number of dimensions. The theory of random walks have been extended to include biased random walks (BRW), where there is some bias towards a certain direction; in ecological terms this could be bias towards a nest site or foraging patch; correlated random walks (CRW), where there exists a persistence in steps i.e. some tendency to continue in the same direction, which are a particularly useful model for animal movement patterns such as migration or exploring behaviour; finally, biased correlated random walks (BCRW) which are a combination of both.

Siniff and Jessen (1969) and Cody (1971) provide early examples of random walk models for movement. Both innovatively describe an animal's trajectory as a bivariate time-series of step lengths and turning angles. Step lengths are the distance between two consecutive locations whilst turning angles are the directional change between three successive observations. Rather than having fixed parameters for the step lengths and turning angles, Siniff and Jessen (1969) explore a range of possible probability distributions which have since been popularised for example, distributions for step lengths need to be positive and continuous e.g. Weibull and Gamma (McClintock et al., 2013; Mckellar et al., 2015) and for turning angles some circular distribution such as Wrapped Cauchy or Von Mises (McClintock et al., 2012; Roever et al., 2014; Turchin, 1999).

Whilst this methodology has laid the foundations for many publications in the area, its simplicity assumes that an animal moves with just one movement process. In reality, animals live in a complex world where they are constantly subjected to both internal and external stimuli such as hunger, fatigue, predators and competition of resources which affect their behaviour and consequently the way they move. Random walk techniques became highly popularised after the introduction of multi-state models where at any instant an animal is following one of a finite number of movement processes which reflect heterogeneous behaviours (Morales and Ellner, 2002; Morales et al., 2004). In Morales et al. (2004), the methodology is illustrated with location data of elk and movement processes defined by mixture of CRWs to reflect behaviours such exploring and foraging. The exploratory behaviour can typically be described with large step length and persistence in direction. In contrast, foraging behaviour which they term '*encamped*' may have smaller step lengths and frequent change in direction. Others have termed similar processes as '*travelling*' and '*foraging*' for example, Breed et al. (2009)'s study on grey seals. As a caveat, it is worth remembering that we have to be careful when interpreting the states as these are just names given to a statistical processes and the actual behaviour of the animal may not reflect the state's name.

In a similar vein, a bivariate process of bearings (the direction of the animal) and step lengths have been used in McClintock et al. (2012). The authors adopted the discrete-time, continuous-space modelling approach of Morales et al. (2004) but increased flexibility to allow for biased movement towards one or many centres of attraction as well as exploratory movement and correlated random walks. This gave rise to a class of models which could be combined to reflect complex behavioural states.

### 2.4.2 Hidden Markov Models for Animal Movement

Hidden Markov models (HMM) are time series models comprised of two components, an observed series and an unobserved, hidden, state sequence. The state sequence in these models are generated by a first-order Markov chain, which leads to temporal autocorrelation in both the behavioural states and in the observed movement patterns (Langrock et al., 2012). HMMs can be thought of as a special case of the SSM as the have they exhibit the same dependence structure but whose number of states is finite and often observations are regularly spaced. A HMM with $m$ states is referred to as an $m$-state HMM. They have become an extremely popular modelling approach in ecology due to their simplicity, parameter interpretation, ease of fitting and computational tractability. They have been used

to categorise behavioural modes in many taxa including marine, terrestrial and avian species (Beest et al., 2016; Farhadinia et al., 2020; Langrock et al., 2012).

Multistate random walks such as those proposed in Morales et al. (2004) may be easily framed as a HMM since the movement process of an animal is dependent on an assumed underlying behavioural state (e.g. foraging or resting). Here, each state of the hidden Markov process is associated with a distinct random walk behaviour and the state-dependent distributions i.e. the observation processes are categorised by a bivariate distribution of step lengths and turning angles.

Since HMMs focus on discrete, regularly spaced observations the fitting is a rather simple process via a frequentist likelihood based approach. If the process is homogeneous i.e. the transitions probabilities are invariant over time then the likelihood is easily obtained using a computationally efficient recursive scheme called the forward algorithm (Langrock et al., 2012). The likelihood is given by

$$\mathscr{L} = \delta^{(1)} P(z_1) \Gamma P(z_2) * \cdots * \Gamma P(z_{t-1}) \Gamma P(z_t) 1^{'} \tag{2.10}$$

where $z_t$ are the observations (for example comprising of step lengths and turning angles). $\delta^{(1)}$ is the initial distribution of the Markov chain; $P$ is a diagonal matrix of conditional density functions of $z_t$ given the state at time $t$; $\Gamma$ is the matrix of all possible state transition probabilities; finally, $1^{'}$ is a column vector of ones.

Whilst hidden Markov models offer a flexible and computationally efficient modelling approach with easy fitting to real movement data, they are limited by the requirement that the data is evenly spaced with no missing values. This is not often possible with movement data. Inference of HMMs are restricted to the sampling scheme of the observations which leads to difficulties in comparisons between different data sets. Some work has been done to allow for multiple time scales within HMMs. Leos-Barajas et al. (2017) use a nested structure where the data depends on a fine scale HMM and this in turn depends on a coarse scale HMM. This technique is capable of capturing both small and large scale movements such as foraging and migration behaviours. However, it still suffers from the shortcomings of operating within a discrete-time framework.

HMMs have become even more popular due to R packages made accessible to practitioners such as moveHMM (Michelot et al., 2016) and momentumHMM (McClintock and Michelot, 2018). What's more, estimates of the underlying state at each time point can be obtained via the Viterbi algorithm (Zucchini et al., 2016), although I will not go in to the details of this

here. For more a more detailed description of HMMs and other R packages see Zucchini et al. (2016).

## 2.5 Continuous-time Models for Movement

So far I have discussed a wealth of movement models which are growing in sophistication. For simplicity these models operate in a discrete time framework. In some cases this is reasonable methodology such as Morales et al. (2004) where observations are taken daily and within-day behaviours are not important for the purposes of the study. However, discrete-time models such as the CRWs presented in Morales et al. (2004) confound the sampling and movement process meaning that they are sensitive to the sampling schedule and can produce conflicting results when the same data is sampled using different sampling schemes (Calabrese et al., 2016). Thus, there is no guarantee that they have a meaningful interpretation on other time scales making it difficult to compare results between different datasets (Harris and Blackwell, 2013). In addition, often there are unwanted irregularities associated with location data such as missing observations and uneven sampling times. These require some form of interpolation or discretisation approximation with poorly understood error.

Continuous-time models alleviate these problems often by separating the movement and sampling process which in turn avoids any approximations. That said, progress in developing continuous-time models for animal movement has been slower than for their discrete-time counterpart. These developments have typically been avoided due to difficulty in parameter interpretation or computation but are necessary to providing statistically robust and biologically pertinent models. Moreover, real movement data emerging from animals, cells, pollen or particles are typified by inherent autocorrelation which discrete-time models fail to capture. Continuous-time stochastic processes such as diffusion processes are a popular modelling approach in continuous-time for handling the inherently large autocorrelation associated with movement data and the use of such approaches to infer relationships between animal movement and habitat based on telemetry data is increasing (Hooten et al., 2018).

In this section we discuss a common approach to modelling movement in continuous time, diffusion processes. We begin by describing diffusion processes in general terms before discussing a simple diffusion process, Brownian motion. We then visit Ornstein Uhlenbeck diffusion processes and their extensions, which become the main foundations of the research in this thesis.

### 2.5.1 Diffusion Processes

A diffusion process is the solution to a stochastic differential equations (SDE). A general SDE has the form

$$dX_t = \mu(x_t,t)dt + \sigma(X_t,t)dW_t, \tag{2.11}$$

where $X_t$ is a stochastic process and $W_t$ is a Wiener process. A Wiener process is a continuous time stochastic process with the Markov property i.e. the process has independent increments. Written in statistical notation the Wiener process is defined to be

$$W_{t+\delta} - W_t \sim N(0,\delta). \tag{2.12}$$

### 2.5.2 Brownian Motion Diffusion Process

The most simplistic diffusion process, which lays the foundations of many other models, is a result of setting $\mu(X_t,t) = 0$ and letting the function $\sigma(X_t,t)$ be the constant $\sigma$. This process is known as a Brownian motion (BM) process (sometimes referred to as simply as a Wiener process). Thus, the BM process is a solution of the SDE

$$dX_t = \sigma dW_t. \tag{2.13}$$

The Brownian motion process $X_t$ is Normally distributed with

$$X_{t+\delta}|X_t = x_t \sim N(x_t, \sigma^2\delta). \tag{2.14}$$

In this 1-parameter diffusion equation, $\sigma$ can be thought of as controlling the speed of the process. From the equation above we can see that Brownian motion is a Markovian process with independent Gaussian increments. The Brownian motion process is the continuous-time analogue or limit of a random walk.

### 2.5.3 Ornstein-Uhlenbeck Diffusion Processes

The Ornstein-Uhlenbeck diffusion process (Uhlenbeck and Ornstein, 1930) is a popular continuous-time movement model in animal ecology. The model is simplistic in its nature

but offers more flexibility than Brownian motion alone. In essence, it describes a biased random walk with some drift or tendency towards a long term mean. As a result the process is often referred to as 'mean-reverting'. The process has convenient properties such as being stationary, Gaussian and Markovian. In one dimension the OU diffusion process is the solution to the stochastic differential equation

$$dX_t = -\beta(X_t - \mu)dt + \sigma dW_t. \tag{2.15}$$

The process has three parameters namely $\beta, \sigma$ and $\mu$. The term $\beta$ is usually referred to as the drift term and controls to speed at which the process reverts to the mean $\mu$ whilst $\sigma$ scales the Wiener process. The Ornstein Uhlenbeck process has a closed form normal distribution with conditional distribution given by

$$X_{t+\delta}|X_t = x_t \sim N\left(\mu(1-e^{-\beta\delta}) + x_t e^{-\beta\delta}, \frac{\sigma^2}{2\beta}(1-e^{-2\beta\delta})\right) \tag{2.16}$$

This may be extended to any number of dimensions. Here, we will describe the process in $d$-dimensional space. Let $\boldsymbol{X}(t)$ be the locations in $d$ dimensions of an individual at time $t$. The conditional distribution of its location at time $s+t$ given its location at time $t$ has the multivariate normal distribution

$$\boldsymbol{X}_{s+t}|\boldsymbol{X}_t \sim N\left(\boldsymbol{\mu} + e^{Bs}(\boldsymbol{X}_t - \boldsymbol{\mu}), \Lambda - e^{Bs}\Lambda e^{B's}\right), \tag{2.17}$$

where $\boldsymbol{\mu}$ is a $d$-dimensional vector and $\Lambda$, $B$ are $d \times d$ matrices. The vector $\boldsymbol{\mu}$ gives the long-term mean of the process or centre of attraction whilst $B$ controls the rate of attraction towards $\boldsymbol{\mu}$ with covariance $\Lambda$.

### 2.5.4 OU Developments

Rather than using the standard OU process to model the locations of an animal, there exists a number of developments. Two main adaptations specific to animal movement are the Integrated Ornstein Uhlenbeck process (IOU) (Johnson et al., 2008; Michelot and Blackwell, 2019) and the Ornstein Uhlenbeck Foraging (OUF) (Fleming et al., 2014).

The IOU introduced by Johnson et al. (2008) allows the velocity of an animal to be defined by an OU process rather than the locations. Specifically, the locations at time, $X(t)$, are given

as

$$X(t) = X(0) + \int_0^t V(u)du, \tag{2.18}$$

where,

$$dX(t) = V(t)dt, \tag{2.19}$$

and $V(t)$ is given by an OU process such that

$$V_{t+\delta}|V_t = v_t \sim N\left(\mu_v(1-e^{-\beta_v\delta}) + v_t e^{-\beta_v\delta}, \frac{\sigma_v^2}{2\beta_v}(1-e^{-2\beta_v\delta})\right). \tag{2.20}$$

Similar to Equation 2.16, $\mu_v$ can be thought of as the mean velocity or drift term while $\beta_v$ and $\sigma_v$ control the autocorrelation and variability respectively in the velocity. This suggests that the variance of the velocity at time $t+\delta$ gets larger with $\delta$.

The IOU is non-stationary, similar to Brownian motion but with autocorrelated velocities at short sampling by frequencies, thus it is particularly useful for modelling movement with duration short enough to not display any home-range.

The Ornstein Uhlenbeck Foraging (OUF) (Fleming et al., 2014) is similar to the IOU except that it has an additional term to Equation 2.19 which make the process stationary and enables its use in modelling home-range. Explicitly, locations of an animal are given by

$$dX(t) = V(t)dt - \beta_x(X_t - \mu_x)dt, \tag{2.21}$$

where $V(t)$ is the velocity as given in Equation 2.20, $\mu_x$ is the mean location and $\beta_x$ controls the scale of attraction toward $\mu_x$. In both Equation 2.20 and 2.21 it is common for $\mu_v$, the mean velocity parameter, to be taken as 0. In fact, in the latter it is essential to ensure the model is stationary and reflects foraging behaviour.

### 2.5.5 Examples

Early examples of using diffusion processes as a model for animal movement are presented in Dunn and Gipson (1977) where movement is assumed to follow an OU diffusion process. The authors give biological incentive for its predictive nature suggesting that forecasting

where an animal may be at a future time point given an observation could be useful for locating and administering medication.

The nature of the attraction term in the OU process allows modelling of resource-driven movement by setting the centre of the OU attraction to a preferential location such as a foraging or nest site (Wang et al., 2019). Conversely, attraction away from particular areas i.e. repulsions, can be used to analyse the effects of human-animal interactions such as avoidance behaviour of human activity zones (Preisler et al., 2013). Brillinger et al. (2002) gives an overview of the uses of SDEs as an animal movement model with application to elk and elephant seals. In the latter case, the species dives below ice sheets thus moving in a 3-dimensional space. Here, the authors give a SDE for diffusion on a sphere. Analysis for the elk data suggest a multi-modal behavioural process where animals follow a circadian rhythm. Despite this, the model uses only a single movement process. A multi-modal model such as a continuous-time version of the Morales et al. (2004) switching model may have been more appropriate here. In Section 2.5.6 we discuss diffusion models for animal movement which allow for behavioural switching.

### 2.5.6 Diffusion Models for Movement with Behavioural Switching

Parton et al. (2016) gives a continuous-time analogue of the step and turn approach presented in Morales et al. (2004) in an effort to provide interpretable parameters and 'bolster' the usage of continuous-time models. Rather than using step lengths and turning angles, they model the bearing and speed of the animal using a Wiener process and an OU process respectively. They extend this work in Parton and Blackwell (2017) to allow for behavioural switching between 'foraging' and 'travelling' using a similar notion to the 'double switch' method of Morales et al. (2004). The modelling approach and corresponding Bayesian inference was demonstrated using the same elk dataset as in Morales et al. (2004). Although they made similar conclusions from the data, they also found that residence times in a particular state were often unsurprisingly less than the 24hour sampling frequency confirming that their approach offers extra insight to coarsely sampled data. Here, the authors use constant switching rates but acknowledge that spatial or temporal heterogeneity would be more biologically relevant.

Blackwell (1997) present a class of models for animal movement in continuous time. Movement is represented as a continuous-time Markov process with a diffusion component (location) and a discrete component (behaviour) as in Berman (1994). The models have the flexibility to incorporate certain movement processes governed by behavioural or psycho-

logical states. Here, it is assumed that there are a finite number of behavioural states, each coupled with an Ornstein Uhlenbeck movement process as given in Dunn and Gipson (1977). Each OU diffusion process has a fully-specified parametric form whose parameters reflect the behaviour in question. At any given time an individual exhibits one behavioural state but the model allows for behavioural switching whose process is given by a continuous time Markov chain. Presenting a model in this way enables capturing many biologically realistic features.

Following this, Blackwell (2003) gives fully Bayesian inference methodology for the switching models presented in Blackwell (1997). The method for inference involves separately sampling from the full conditional distributions of the switching process, the behaviour parameters and the diffusion process parameters, in sequence, using a mixture of Gibbs and random walk Metropolis Hastings MCMC. They assume that the location and behavioural state of the animal is observed and regard the switching process as 'missing data'.

Both models presented in Blackwell (1997) and Blackwell (2003) rely on the limiting assumption of spatial homogeneity i.e. that the switching process is independent of the animals' location. In reality, animals will be likely to move differently depending on their environment. In an extension, Harris and Blackwell (2013) detail a modelling framework for animal movement in which the movement processes depend on location in terms of a discrete set of regions and also an underlying behavioural state. Much like Blackwell (1997), the study gives a flexible class of models with illustration of a range of movement behaviours where the switching process is a continuous-time Markov chain; the difference here is that there are separate generator matrices for each individual region.

Blackwell et al. (2016) give methodology for fully Bayesian inference of the models given in Harris and Blackwell (2013) that include behavioural switching and spatial heterogeneity as well as extending this to temporally varying movement. The switching process is given as a thinned Poisson process which allows for exact simulation and avoids the usual time discretisation error. The authors focus on what Harris and Blackwell (2013) term 'separable models' where the transition rates between behaviours depend on the location of the animal but the movement parameters and trajectories do not. i.e. when an animal moves to a different environment it will change its tendency to behave in a particular way which will indirectly change its movement process. This is dissimilar to the model for spatially heterogeneous behaviour presented in Ovaskainen (2004) which involves an edge-mediated behaviour i.e. movement at the boundary of distinct habitat will be bias towards a preferred habitat. A principle component of the methodology in Harris and Blackwell (2013) is to simulate trajectories augmented with switching times and locations conditioned on the observed data

and use MCMC methods to accept or reject those proposed samples. Section 2.6 reviews the algorithm for exact simulation of trajectories as discussed in Blackwell et al. (2016).

## 2.6 Exact Simulation of Trajectories

The main assumption of Blackwell et al. (2016) and Harris and Blackwell (2013) is the boundedness of switching rates. As a consequence, an animal is never forced to instantly change behaviour upon entering a new environment. This is consistent with what Harris and Blackwell (2013) term 'separability'. Let

$$\lambda_j(t,\boldsymbol{x}) = \sum_{i \neq j} \lambda_{ji}(t,\boldsymbol{x}),$$

represent the switching rate out of behaviour $j$ at time $t$ and location $\boldsymbol{x}$. Then let

$$\kappa = \max_{j,t,x} \left(\lambda_j(t,\boldsymbol{x})\right).$$

We then consider the waiting time from any point to the next behavioural switch to be bounded below by what would be the time if the rate was always $\kappa$. In other words we can formulate a series of potential switches in time, $\{T_1, T_2, \ldots, T_k\}$, forming a Poisson process of rate $\kappa$. We then decide whether the potential switch $T_i$ is an actual switch given by the probability

$$\frac{\lambda_j(T_i,\boldsymbol{x}(T_i))}{\kappa}.$$

As an example, imagine we initialise our simulation at location $\boldsymbol{x}$ in behaviour $j$ i.e. $J(0) = j$ and $\boldsymbol{x}(0) = \boldsymbol{x}$. We can then simulate the first potential switching time $T_1$ and correspondingly the location $\boldsymbol{x}(T_1)$ by a forward simulation from the movement parameters associated with behavioural state $J(0)$. This potential switch is then deemed an actual switch with probability

$$\frac{\lambda_{j(T_1)}(T_1,\boldsymbol{x}(T_1))}{\kappa}.$$

If it is an actual switch, a new behavioural state $i$ is sampled with probability proportional to $\lambda_{ji}(T,\boldsymbol{x}(T))$. Now, with the updated location and behavioural state we can simulate the

process forwards for as long an interval as we like. The collection of actual switches can be thought of as a Poisson process with rate $\kappa$ that has been thinned.

More generally, consider the times of the potential switches (realisations of the Poisson process with rate $\kappa$) to be given as $T_1, T_2, T_3, \ldots$. Then for each $T_k$ simulate the location $\boldsymbol{x}(T_k)$ via the movement process defined by the behavioural state $J(T_{k-1})$ and decide whether the potential switch given at $T_k$ is an actual switch with probability

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

If we deem this to be an actual switch, a new state $j$ is picked with probability

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

Note that we have described the case where behavioural switching depends on the location of an individual however, the principle remains for simpler modelling approaches where we do not account for this heterogeneity.

### 2.6.1 Inference for Exact Simulation Method

In this section we discuss the inference method given in Blackwell et al. (2016) for the exact simulation presented in Section 2.6. The algorithm for Bayesian inference of the above model alternates between updating the trajectory over some time interval, that is the potential switches, locations and behavioural states, and updating the movement parameters and switching process parameters. The inference method is formulated as a MCMC algorithm which is used to sample from the posterior distributions of the movement parameters and switching rates. The key to the inference is the augmentation of observed data with the simulated potential switches and locations.

### 2.6.2 Trajectory Updates

To update the trajectory over some time interval $[t_a, t_b]$, where $1 \leq a < b \leq N$, conditional on the states immediately either side of the interval $J(t_a)$ and $J(t_b)$, the movement parameters and the switching parameters, define

$$T'_{a,b} = \{T_{c,k}, k = 1, \ldots, M'_c, c = a, \ldots, b-1\},$$

as the collection of potential switching times between $t_a$ and $t_b$, where $M'_c$ is the number of potential switches. In other words, $T'_{a,b}$ is a realisation of a Poisson process with rate $\kappa$.

Starting with the location and behaviour at the beginning of the interval $\boldsymbol{x}(t_a)$ and $J(t_a)$, simulate forward a new location $\boldsymbol{x}(T'_{a,1})$ using the movement model defined by the behavioural state $J(t_a)$. Then pick a new state $j$ with the following criteria

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

The process is then simulated forward to get the location $\boldsymbol{x}(T'_{a,k+1})$ using the movement model defined by the behavioural state $J(T'_{a,k})$ and letting the new state

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

for $k = 1, \ldots, M'_a - 1$. This process is repeated for each subinterval between observations until the entire interval $[t_a, t_b]$ has been sampled.

This trajectory sample may then be accepted or rejected using a Metropolis-Hastings update. However, if the final sampled behavioural state is not consistent with the pre-existing behaviour $J(t_b)$ the rejection is automatic. Conditioning on the observed data the Metropolis-Hastings ratio is given as

$$\prod_{c=a}^{b-1} \frac{f(\boldsymbol{x}(t_{c+1}) | \boldsymbol{x}(T'_{c,M_c}), J'_{c,M_c})}{f(\boldsymbol{x}(t_{c+1}) | \boldsymbol{x}(T_{c,M_c}), J_{c,M_c})}.$$

### 2.6.3 Parameter Updates

Given a trajectory of an individual, the times of switches in states are known. Thus, we have a precise description of when the animal was following each of its different movement processes. With this, parameter inference can be made straightforwardly using standard random-walk Metropolis-Hastings updates.

## 2.7 Conclusions

Research in animal movement is growing rapidly. In recent years, ecologists have adopted useful techniques from other fields such as computer science and engineering, and the widespread use of statistical tools for example state-space models and MCMC methods have enabled more sophisticated and rigorous analysis of animal movement data.

Recent advancements in computational power have eased fitting increasingly complex movement models to ecological data by providing methods for estimating likelihoods that would otherwise be intractable. Simulation methods for parameter estimation such as MCMC are able to approximate high dimensional integrals that cannot be solved analytically.

There are myriad approaches to modelling the movement of animals and so far no single method is considered as the 'gold standard'. Conceptualising movement in either a discrete or continuous time framework both have their merits. Popular discrete time models such as step and turn random walks allow for ease of fitting and interpretation of parameters but usually comes at a cost of statistical robustness, whereas movement models in continuous-time can suffer from slow fitting and difficulty in parameter interpretation. We have acknowledge the underdevelopment of modelling animal movement in continuous time; however, despite computational challenges it is often a worthwhile trade off for the inherent statistical robustness and the modelling of movement in a more natural way.

We have discussed only models of individual-level movement. For most taxa this is an unnatural assumption. Many species live in collectives whose movement decisions are influenced by their neighbours, possibly of a different species. In the Chapter 3 we will discuss methodology for modelling collective animal movement, although this is not nearly as developed as its individual-level counterpart.

For the remainder of thesis our focus will be toward developing continuous-time models of collective animal movement using Ornstein-Uhlenbeck diffusion processes.

# Chapter 3

# Modelling Collective Movement

As mentioned, many of the world's taxa do not movement independently of other individuals. For most, conspecific or heterospecific interactions influence decision making, behavioural choices and movement (Couzin et al., 2005; Delgado et al., 2018; Merkle et al., 2016; Schlägel et al., 2019). Many species of birds, fish, insects and ungulates demonstrate highly cohesive and coordinated movements whose social interactions are vital for survival (Buhl et al., 2006; Croft et al., 2015; Herbert-Read, 2016; Westley et al., 2018). In ecology, when localised interactions amongst individuals lead to large scale patterns in movement and behaviour we term this as collective movement (Dalziel et al., 2016; Mueller et al., 2011).

In the past, simulation models have provided useful insights into the movement and decision making of animal groups (Aoki, 1982; Huth and Wissel, 1992). By assuming underlying laws of interaction, these predictive models help us to understand specific ecological phenomena such as information sharing (Couzin et al., 2005), the effect of group size in obstacle avoidance (Croft et al., 2013), pooling uncertain estimates of the environment for improved navigation, sometimes referred to as the *'many wrongs principle'* (Codling et al., 2007) and how variation among individuals impacts the overall cohesion of the group (Delgado et al., 2018). These properties of collective movement are thought to be evolutionary adaptations leading to an increased fitness at the population level (Dalziel et al., 2016).

Now, with a wealth of tracking technologies it is possible to analyse real data of aggregations without relying on simulation models or laboratory experiments which have less generalisable conclusions. However, recent studies using real data typically employ a metric based approach to quantify aspects of collective movement such as synchronisation in acceleration or tortuosity (Polansky and Wittemyer, 2011) and measures of proximity and sociability (Delgado et al., 2018) rather than explicitly providing a model of movement. What's more,

most studies are restricted to dyadic interactions (Joo et al., 2018; Long et al., 2014; Polansky et al., 2010). The methods of Dalziel et al. (2016) and Mueller et al. (2011) do extend to multiple individuals, with the latter normalising their results in such a way that cross species comparisons may be made but again, both use metrics to screen the data for collective behavioural traits rather than parametrising a model of movement.

Some approaches do offer a stochastic model for multi-individual movement and may even account for behavioural heterogeneity (Calabrese et al., 2018) but still operate with metric based analysis whose behavioural transitions are dictated by the sampling scheme and whose inference can produce different results on different time-scales. Haydon et al. (2008) uses social structures of large groups to infer population dynamics, mortality rates and fecundity. This is demonstrated with a unique dataset of elk where all individuals are tracked. Whilst movement in this approach is modelled explicitly each animal's movement is modelled individually using a CRW and the group structure is quantified by spatial-temporal proximity from other individuals. Nonetheless, this "socially informed" model gives enlightening results about the population growth rate in relation to fission-fusion processes. They found that solitary individuals (those outside of the proximity threshold) have a higher risk of mortality than their grouped counterparts. Other work such as Strandburg-Peshkin et al. (2015) has shown animal movement to be a result of joint decision making, specifically highlighting that the movement of baboons is influenced by a democratic decision amongst troop members rather than a single dominant species. This stresses the importance of forming coherent models which capture the sophistication of collective motion and social interactions of gregarious animals.

The work of Scharf and Buderman (2020) gives an up-to-date review of the movement models which incorporate dependence between individuals. They suggest that modelling approaches may be separated into two possible avenues, direct and implicit dependence. Direct dependence is where an individual's movement depends directly on interactions with its neighbours via some underlying social network for example, attraction, repulsion or alignment to proximal animals. In contrast, implicit dependence may occur through shared behavioural patterns amongst individuals e.g. a tendency to behave in a similar way. Here, the mechanisms that drive the movement and behaviour are constant across the population and conditional on those the individuals are considered to move independently. The authors further this description by partitioning the implicit dependence approaches into three subtypes, two-stage, complete pooling and partial pooling. The two-stage approach involves separately fitting individual-level models then undergoing some form of post *ad hoc* analysis of dependence, however this approach is not statistically robust. In contrast, complete

and partial pooling both provide a joint movement model for the individuals. In complete pooling, behaviour parameters are constant across individuals whereas, in partial pooling the individuals may deviate form the population level parameters. The group movement models of Langrock et al. (2012) (Section 3.1) and Niu et al. (2016) (Section 3.2), as well as the novel work provided in this thesis can be classified as complete pooling with the dependence being driven by the abstract leader.

Moreover, when developing models of collective movement an important consideration to make is the possibility of incomplete observations of the group. Some species may live in large collectives, possibly hundreds strong and so tagging each individual within that group may be economically or practically infeasible. Methods based on statistical summaries may have varying results depending on who or how many animals were tagged. This lack of robustness can lead to difficulty in inference and forecasting how individuals may respond under different movement patterns.

Existing realistic models of movement, which typically combine continuous locations in space with a discrete representation of behaviour, are generally limited to modelling single individuals. However, Scharf et al. (2016) provide a mechanistic model of group movement which employs direct dependence in movement via an underlying social network. Scharf et al. (2018) further this work by modelling movement as a process convolution, that is, convolving a random process (such as a Weiner process) with a smoothing function. However, in both cases, the underlying social structure is restricted to dyadic connections, thus the number of pairwise interaction parameters grow rapidly as the number of individuals increase, which can lead to difficulty in computation. In contrast, the implicit dependence models of Langrock et al. (2014) give a joint model for the movement of a group of animals explicitly, allowing both dependent and independent behaviours, but their model and inference method are limited to discrete time, and their 'centroid' mechanism to represent attraction is explicitly tied to the time-scale of observations. Niu et al. (2016) give a continuous-time collective movement model (see review in Section 3.2) which assumes consistent group movement at all times, without any variation in behaviour.

This chapter builds on this modelling technique by developing novel methodology which allows exact Bayesian statistical analysis for a class of group movement models with behavioural switching in continuous time, without any need for time-discretisation error. We represent the group movement as a multivariate Ornstein Uhlenbeck process and allow the individuals to switch behaviour, either following the group or moving independently as Brownian motion. The times of changes in behaviour are represented as a thinned Poisson process, allowing exact simulation and Markov chain Monte Carlo inference. The methodology can

be applied to data that are regular or irregular in time, with or without missing or incomplete observations. As well as much greater flexibility in modelling, our approach gives improved computational efficiency by integrating out part of the group movement process using a Kalman filter. In a set of simulation experiments, motivated in part by our analysis of data from simultaneously tracked reindeer (*rangifer tarandus*), we show that our approach can reconstruct unobserved behaviours from location data in a range of scenarios.

The structure of the remainder of this chapter is as follows. We first review the existing modelling approach presented in Niu et al. (2016). Section 3.3 demonstrates a non-stationary version of the model giving biological reasoning for its purposes. Section 3.4 presents an extension to this model to allow for behavioural switching between two modes; following the group and moving independently. Section 3.5 provides inference for the model including details of the trajectories simulation, state space format and application of the inhomogeneous Kalman filter. The chapter is finalised with implementation of the model with multiple simulation data-sets whose underlying processes varied in order to investigate the models capability of recovering parameter values under different settings.

Most of the research described in this Chapter is given in Niu et al. (2020). This work is a collaborative effort of Mu Niu, myself, Jordan Milner, Paul Blackwell and Anna Skarin. Mu provided a switching version of the OU model given in Niu et al. (2016) using similar simulation methods to that of Blackwell et al. (2016) and using the Kalamn filter for inference. Jordan gave insight into the methods needed to derive covariance terms in Section 3.3.1 - Section 3.3.4 and also checked my calculations. Anna Skarin provided reindeer location data and insight in to herding dynamics. The simulation experiments given in Section 3.8.3 - 3.8.2 are a result of comments made by anonymous referees at Biometrics, as are the plots used for data visualisation for example Figure 3.8.

## 3.1 Discrete-time Models of Collective Movement

The models presented in Chapter 2 are all directed towards individual-level movement and do not consider group-level movement. This can be very limiting when working with species who are well regarded as living in and moving as a collective such as fish, birds and ungulates. Fewer models exist that give a coherent model for group movement. In discrete time, Russell et al. (2017) proposed a multi-state discrete-time model for the collective movement of carpenter ants where behavioural states were influenced by covariate information such as environmental conditions or the behavioural state(s) that their neighbour(s) exhibit. Unlike

herd animals, ants do not have similar speed and direction to their neighbours. They often start/stop moving due to interactions with other ants, this prompted the development of a method that models dependent movement with interacting behavioural states. Other models incorporating dependence amongst neighbouring individuals include Croft et al. (2015) who model alignment in movement as an attraction or repulsion to stimulus within a certain proximity to investigate obstacle avoidance in flocks of birds. Similar attraction/repulsion methods are presented in Couzin et al. (2002).

Finally, Langrock et al. (2014) give a group-level model with multiple behavioural modes where individuals within the group are at times either attracted to an abstract point referred to as the group centroid or moving independently. The point may represent a leading animal or the mathematical notion of the term centroid i.e. the centre of mass of the group. The choice of this representation is crucial and should reflect the way in which the individuals make their movement decisions for example, for gregarious animals with no defined leader such as reindeer, the mathematical notion of a centroid may be more appropriate. In contrast, modelling the group movement of matriarchal elephants may use the centroid as a specific leading animal. In this approach they use a discrete time HMM modelling framework. Their multi-state system is composed of two states; the '*exploring*' state is given as a CRW and the '*encamped*' state is represented as a BRW with bias towards the centroid. The likelihood of the HMM is calculated using the forward algorithm given in Equation 2.10 where the density $f_n$ is determined by the type of random walk assumed in state $n$ and the state-dependent distributions considered for step lengths and turning angles. In this case, each state $i$ has movement parameters modelled with

$$\text{Step length} \sim Gamma(\mu_i, \sigma_i), \tag{3.1}$$

$$\text{Turning angle} \sim vonMises(\nu_i, \kappa_i). \tag{3.2}$$

The transition matrix given by

$$\Gamma = \begin{pmatrix} 1-x & x \\ y & 1-y \end{pmatrix} \tag{3.3}$$

This methodology provides a flexible modelling framework for grouping dynamics allowing both dependent and independent behaviours however, their model and inference method are limited to discrete time, and their 'centroid' mechanism to represent attraction is explicitly tied to the time-scale of observations. The next section reviews Niu et al. (2016) who give

methodology analogous to that presented in Langrock et al. (2014) by modelling movement in continuous-time using diffusion processes.

## 3.2 Review of Ornstein-Uhlenbeck Approach

Niu et al. (2016) tackle some of the limitations of current collective modelling techniques by presenting a continuous time movement model for multiple individuals using diffusion processes. Not only does this alleviate poorly understood discretisation methods, it aligns fundamentally with the understanding that animals move in continuous time. The basis of the methodology is similar to the work of Langrock et al. (2014). Heuristically, Niu et al. (2016) represent the interaction between animals as a shared attraction to an abstract point which we refer to as the leader, $L$. To account for group migration to a particular location, the leader is also allowed an attraction towards an inferred point. It is possible to allow this leader to be an actual leading animal which may of course be observed, in which case the model still applies but much of the calculation is greatly simplified; for ease of exposition, we assume here that this is not the case. The observed individuals are conditionally independent, given full information about $L$. Thus animals do not interact directly, but only through their interactions with $L$. This formulation means that the model is robust to incomplete observation of a group of animals, and to variation over time of the number or identity of the observed individuals. The interpretation of the parameters of the model does not depend on the numbers of observed or unobserved animals. This approach is therefore suitable for cases where there may be many unobserved animals, e.g. large herds of herbivores.

The movement of the unobserved leader $L$ is modelled as a stationary Ornstein Uhlenbeck process. Let the random variable $L_t^y$ represent the location of the leader at time $t$ in the $y$ coordinate. A stochastic process $\{L_t^y : t \geq 0\}$ in which $L_t^y$ is attracted to $\theta^y$ is given by the stochastic differential equation (Schach, 1971)

$$\mathrm{d}L_t^y = -\beta(L_t^y - \theta^y)\mathrm{d}t + \rho\,\mathrm{d}V_t^y, \tag{3.4}$$

where $\beta$ is the attraction rate to $\theta^y$; $\theta^y$ is a fixed location which may be unknown; $\rho$ is the coefficient for the noise; $V_t^y$ is standard Brownian motion. By applying the rotational symmetry which is natural in practice (Blackwell, 1997), the model is identical for the $x$ coordinate $L_t^x$, with parameters $\beta$ and $\rho$ in common, and independent Brownian motions $\{V_t^x\}$ and $\{V_t^y\}$ used for $L_t^x$ and $L_t^y$.

A similar stochastic differential equation can model the movement of each of $n$ followers attracted at any instant to the current location of the leader. Let the random variable $F_t^{y,k}$ represent the $y$ coordinate of the $k$th follower's location at time $t$. $\{F_t^{y,k} : t \geq 0\}$ is defined by the following stochastic differential equation with parameters $\alpha$, $\sigma$, $L_t^y$ and Brownian motion $\{W_t^{y,k}\}$, where $F_t^{y,k}$ is attracted to $L_t^y$:

$$\mathrm{d}F_t^{y,k} = -\alpha\left(F_t^{y,k} - L_t^y\right)\mathrm{d}t + \sigma \mathrm{d}W_t^{y,k},$$

with $\alpha$ the attraction rate to $L_t^y$; $\sigma$ the coefficient for the noise. By rotational symmetry as before, $F_t^{x,k}$ and $F_t^{y,k}$ satisfy identical equations.

Throughout this thesis, we express the idea of attraction to the leader by restricting $\alpha$ to be positive. Taking $\alpha$ to be negative would imply repulsion from the moving point at $L_t^y$, which is not a useful model of collective behaviour of the form that we are interested in, although nonetheless interesting. A related model involving repulsion of a single animal from a fixed centre is explored by Blackwell (1997) and Harris and Blackwell (2013).

Combining the above stochastic differential equations for the locations of leader and followers in the $y$ direction gives the multivariate Ornstein-Uhlenbeck process:

$$d\boldsymbol{Y}_t = \mathbf{A}\mathbf{Y}_t dt + \boldsymbol{\Sigma} d\mathbf{B}_t^y, \tag{3.5}$$

where

$$\boldsymbol{Y}_t = \begin{pmatrix} \theta^y \\ L_t^y \\ F_t^{y,1} \\ \vdots \\ F_t^{y,n} \end{pmatrix}, A = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ \beta & -\beta & \ddots & & \vdots \\ 0 & \alpha & -\alpha & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \alpha & 0 & \cdots & -\alpha \end{pmatrix}, \Sigma = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & \rho & \ddots & & \vdots \\ \vdots & \ddots & \sigma & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \sigma \end{pmatrix},$$

$$\boldsymbol{B}^{y_t^T} = \begin{pmatrix} 0 & V_t^y & W_t^{y,1} & \cdots & W_t^{y,n} \end{pmatrix}.$$

$\boldsymbol{Y}_t$ is a vector representing the $y$ coordinates of the attractor, the leader and the followers. The attractor $\theta^y$ is a constant in Niu et al. (2016), but in general it could be modelled by another diffusion process; we include it in the state vector for convenience in describing the inference algorithm later. Note that each $F_t^{y,k}$ is indirectly attracted to $\theta^y$. The matrix $A$ is the attraction rate matrix. We take the stochastic parts (Brownian motion) for the leader and the followers to be uncorrelated, therefore $\Sigma$ is diagonal; each diagonal element of the $\Sigma$, except the initial

zero, represents the coefficient of the individual variance. The solution of this differential equation can be written as

$$Y_t = e^{At} Y_0 + \int_0^t \Sigma e^{A(t-s)} \, \mathrm{d}B_s. \tag{3.6}$$

Conveniently, the distribution of $Y_t$ given $Y_0$ can be described by a multivariate normal distribution

$$\boldsymbol{Y}_t | \boldsymbol{Y}_0 \sim \mathrm{MVN}(\boldsymbol{\mu}, \Xi), \tag{3.7}$$

where

$$\boldsymbol{\mu}^T = \left( \begin{array}{cccc} \theta^y & \mu_{\mathrm{L}}(L_0^y, t) & \mu_{\mathrm{F}}(L_0^y, F_0^{y,1}, t) & \cdots \quad \mu_{\mathrm{F}}(L_0^y, F_0^{y,n}, t) \end{array} \right), \tag{3.8}$$

with

$$\mu_{\mathrm{L}}(L_0^y, t) = \left(L_0^y - \theta^y\right) e^{-\beta t} + \theta^y, \tag{3.9}$$

$$\mu_{\mathrm{F}}(L_0^y, F_0^{y,k}, t) = \left(L_0^y - \theta^y\right) \frac{\alpha}{\alpha - \beta} \left(e^{-\beta t} - e^{-\alpha t}\right) + \left(F_0^{y,k} - \theta^y\right) e^{-\alpha t} + \theta^y, \tag{3.10}$$

and

$$\Xi = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \xi_{\mathrm{L}} & \xi_{\mathrm{LF}} & \cdots & \cdots & \xi_{\mathrm{LF}} \\ \vdots & \xi_{\mathrm{LF}} & \xi_{\mathrm{F}} & \xi_{\mathrm{FF}} & \cdots & \xi_{\mathrm{FF}} \\ \vdots & \vdots & \xi_{\mathrm{FF}} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \xi_{\mathrm{FF}} \\ 0 & \xi_{\mathrm{LF}} & \xi_{\mathrm{FF}} & \cdots & \xi_{\mathrm{FF}} & \xi_{\mathrm{F}} \end{pmatrix}$$

with

$$\xi_{\mathrm{L}}(t) = \frac{\rho^2}{2\beta} \left(1 - e^{-2\beta t}\right), \tag{3.11}$$

$$\xi_{\mathrm{LF}}(t) = \frac{\rho^2 \alpha}{2\beta\,(\alpha+\beta)} - \frac{\rho^2 \alpha}{2\beta\,(\alpha-\beta)} e^{-2\beta t} + \frac{\rho^2 \alpha}{\alpha^2 - \beta^2} e^{-(\beta+\alpha)t}, \tag{3.12}$$

$$\xi_{\mathrm{F}}(t) = \left\{ \frac{\sigma^2}{2\alpha} + \frac{\rho^2 \alpha}{2\beta\,(\alpha+\beta)} \right\} \left(1 - e^{-2\alpha t}\right) - \frac{\rho^2 \alpha^2}{2\beta\,(\alpha-\beta)^2} \left(e^{-\beta t} - e^{-\alpha t}\right)^2$$

$$- \frac{\rho^2\alpha^2}{\beta\left(\alpha^2-\beta^2\right)}\left\{e^{-(\alpha+\beta)t} - e^{-2\alpha t}\right\}, \tag{3.13}$$

$$\xi_{\mathrm{FF}}(t) = \frac{\rho^2\alpha}{2\beta\left(\alpha+\beta\right)}\left(1-e^{-2\alpha t}\right) - \frac{\rho^2\alpha^2}{2\beta\left(\alpha-\beta\right)^2}\left(e^{-\beta t} - e^{-\alpha t}\right)^2$$
$$- \frac{\rho^2\alpha^2}{\beta\left(\alpha^2-\beta^2\right)}\left\{e^{-(\alpha+\beta)t} - e^{-2\alpha t}\right\}. \tag{3.14}$$

For details of the derivation see Niu et al. (2016). The parameter $\alpha$ controls the strength of the attraction of the followers to the leader. One consequence of this is that higher values of $\alpha$ will lead to the followers typically being closer to the leader, although of course their distribution around it depends on the diffusion parameters $\rho$ and $\sigma$ too.

## 3.3 Non-stationary Case

In model described in above, the leader and followers jointly define a multivariate OU process and therefore have a stationary joint distribution. However, in practice the leader may not have a point of attraction, or at least not one that is relevant on the time scale of available data. The most tractable way to allow for this is to simply allow the leader to undergo Brownian motion instead of an OU process, by setting $\beta = 0$ in Equation 3.4. The stochastic process can be described as:

$$\begin{aligned} \mathrm{d}L_t^y &= \rho \mathrm{d}V_t^y, \\ \mathrm{d}F_t^{y,k} &= -\alpha\left(F_t^{y,k} - L_t^y\right)\mathrm{d}t + \sigma \mathrm{d}W_t^{y,k}. \end{aligned} \tag{3.15}$$

Similar to Section 3.2, we combine the equations for the leading point and the followers to give an SDE for (the $y$ coordinates of) both leader and followers. The solution of this multivariate stochastic differential equation can also be written as a multivariate normal distribution

$$\boldsymbol{Y}_t | \boldsymbol{Y}_0 \sim \mathrm{MVN}(\boldsymbol{\mu}^*, \Xi^*), \tag{3.16}$$

where

$$\boldsymbol{\mu}^{*^T} = \left( \begin{array}{cccc} \theta^y & L_0^y & \mu_{\mathrm{F}}^*(L_0^y, F_0^{y,1}, t) & \cdots \quad \mu_{\mathrm{F}}^*(L_0^y, F_0^{y,n}, t) \end{array} \right), \tag{3.17}$$

with

$$\mu^*_{\mathrm{F}}(L^y_0, F^{y,k}_0, t) = L^y_0\left(1 - e^{-\alpha t}\right) + F^{y,k}_0 e^{-\alpha t}.$$

We keep $\theta^y$ in Equation 3.17 in order to be consistent with the existing model in Equation 3.8. $\theta^y$ is fixed to be zero and not used in the inference. The variance matrix can be written as

$$\Xi^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \xi^*_{\mathrm{L}} & \xi^*_{\mathrm{LF}} & \cdots & \cdots & \xi^*_{\mathrm{LF}} \\ 0 & \xi^*_{\mathrm{LF}} & \xi^*_{\mathrm{F}} & \xi^*_{\mathrm{FF}} & \cdots & \xi^*_{\mathrm{FF}} \\ 0 & \vdots & \xi^*_{\mathrm{FF}} & \ddots & \ddots & \vdots \\ 0 & \vdots & \vdots & \ddots & \ddots & \xi^*_{\mathrm{FF}} \\ 0 & \xi^*_{\mathrm{LF}} & \xi^*_{\mathrm{FF}} & \cdots & \xi^*_{\mathrm{FF}} & \xi^*_{\mathrm{F}} \end{pmatrix}. \tag{3.18}$$

However, obtaining the conditional distributions in this case requires additional work, as the derivation of the result given previously in Equation 3.7 relies on stationarity. In this case the solution is as follows.

$$\xi^*_{\mathrm{L}}(t) = \rho^2 t, \tag{3.19}$$

$$\xi^*_{\mathrm{LF}}(t) = \rho^2 t - \frac{\rho^2}{\alpha}\left(1 - e^{-\alpha t}\right), \tag{3.20}$$

$$\begin{aligned} \xi^*_{\mathrm{F}}(t) =& \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}) + \frac{\rho^2}{2\alpha}(2\alpha t - 3) \\ &+ \frac{2e^{-\alpha t}\rho^2}{\alpha} - \frac{e^{-2\alpha t}\rho^2}{2\alpha}, \end{aligned} \tag{3.21}$$

$$\xi^*_{\mathrm{FF}}(t) = \frac{\rho^2}{2\alpha}(2\alpha t - 3) + \frac{2e^{-\alpha t}\rho^2}{\alpha} - \frac{e^{-2\alpha t}\rho^2}{2\alpha}. \tag{3.22}$$

To derive the solutions to the SDE in the non-stationary case, we look at the full expansions of the exponential terms in each of Equation 3.11 to 3.14. Then, we carefully collect terms so that it is possible to set $\beta = 0$ i.e. where $\beta$ does not cause any difficulties with the denominators. The expressions relating to the distribution mean i.e. Equation 3.9 and 3.10 are straightforward to derive. The former is derived from first principles insofar as the leader now follows a Brownian motion process, thus its mean at any time is simply its current location. In the latter it is straightforward to allow $\beta = 0$. However, the derivation of the

covariance terms require more delicate work. We derive each entry in turn, starting from the original form.

### 3.3.1 Leader Variance

We begin with an easy derivation. We start with

$$\xi_{\mathrm{L}}(t) = \frac{\rho^2}{2\beta}\left(1 - e^{-2\beta t}\right).$$

First, expand the exponential to give

$$\begin{aligned}
\xi_{\mathrm{L}}(t) =& \frac{\rho^2}{2\beta}\left(1 - (1 - 2\beta t + \frac{(2\beta t)^2}{2!} - \dots)\right), \\
=& \frac{\rho^2}{2\beta}\left(2\beta t - \frac{(2\beta t)^2}{2!} + \dots\right), \\
=& \rho^2\left(t - \frac{2\beta t}{2!} + \mathscr{O}(\beta^2)\right).
\end{aligned}$$

Then letting $\beta \to 0$ our new variance is

$$\xi_{\mathrm{L}}^* = \rho^2 t. \tag{3.23}$$

### 3.3.2 Leader-Follower Covariance

Starting with the existing equation, we have

$$\begin{aligned}
\xi_{\mathrm{LF}}(t) &= \frac{\rho^2\alpha}{2\beta\,(\alpha+\beta)} - \frac{\rho^2\alpha}{2\beta\,(\alpha-\beta)}e^{-2\beta t} + \frac{\rho^2\alpha}{\alpha^2-\beta^2}e^{-(\beta+\alpha)t}, \\
&= \frac{\rho^2\alpha}{2\beta}\left(\frac{1}{\alpha+\beta}e^{-2\beta t}\right) + \frac{\rho^2\alpha}{\alpha^2-\beta^2}e^{-(\alpha+\beta)t}, \\
&= \frac{\rho^2\alpha}{2\beta}\left(\frac{(\alpha-\beta)-(\alpha+\beta)e^{-2\beta t}}{\alpha^2-\beta^2}\right) + \frac{\rho^2\alpha}{\alpha^2-\beta^2}e^{-(\alpha+\beta)t}, \\
&= \frac{\rho^2\alpha}{2\beta}\left(\frac{(\alpha-\beta)-(1-2\beta t+\frac{(2\beta t)^2}{2}-\dots)(\alpha+\beta)}{\alpha^2-\beta^2}\right) + \frac{\rho^2\alpha}{\alpha^2-\beta^2}e^{-(\alpha+\beta)t},
\end{aligned}$$

$$= \frac{\rho^2\alpha}{2\beta}\left(\frac{-2\beta - 2\beta t(\alpha+\beta) + \mathscr{O}(\beta^2)}{\alpha^2-\beta^2}\right) + \frac{\rho^2\alpha}{\alpha^2-\beta^2}e^{-(\alpha+\beta)t},$$
$$= \rho^2\alpha\left(\frac{-1-(\alpha+\beta)t + \mathscr{O}(\beta^2)}{\alpha^2-\beta^2}\right) + \frac{\rho^2\alpha}{\alpha^2-\beta^2}e^{-(\alpha+\beta)t}$$
$$= \frac{\rho^2\alpha}{\alpha^2-\beta^2}\left(-1 + e^{-(\alpha+\beta)t} - (\alpha+\beta)t + \mathscr{O}(\beta^2)\right).$$

Then letting $\beta \to 0$ we have

$$\xi^*_{\mathrm{LF}} = \rho^2 t - \frac{\rho^2}{\alpha}\left(1 - e^{-\alpha t}\right). \tag{3.24}$$

### 3.3.3 Follower Variance

Starting with the equation

$$\xi_{\mathrm{F}}(t) = \left\{\frac{\sigma^2}{2\alpha} + \frac{\rho^2\alpha}{2\beta(\alpha+\beta)}\right\}\left(1-e^{-2\alpha t}\right) - \frac{\rho^2\alpha^2}{2\beta(\alpha-\beta)^2}\left(e^{-\beta t} - e^{-\alpha t}\right)^2$$
$$- \frac{\rho^2\alpha^2}{\beta(\alpha^2-\beta^2)}\left\{e^{-(\alpha+\beta)t} - e^{-2\alpha t}\right\}, \tag{3.25}$$

rearrange and expand to write in the form

$$= \frac{\sigma^2}{2\alpha}(1-e^{-2\alpha}) + \frac{2e^{-\alpha t}\rho^2}{\alpha} + \underbrace{\frac{\rho^2\alpha}{2\beta}\left(\frac{1}{\alpha+\beta} - \frac{\alpha e^{-2\beta t}}{(\alpha-\beta)^2}\right)}_{1}$$
$$- e^{-2\alpha t}\underbrace{\left(\frac{\rho^2\alpha}{2\beta(\alpha+\beta)} + \frac{\rho^2\alpha^2}{2\beta(\alpha-\beta)^2} - \frac{\rho^2\alpha^2}{\beta(\alpha-\beta)^2}\right)}_{2}.$$

First, let us collect the terms in the underbrace labelled 1.

$$\underbrace{\frac{\rho^2\alpha}{2\beta}\left(\frac{1}{\alpha+\beta} - \frac{\alpha e^{-2\beta t}}{(\alpha-\beta)^2}\right)}_{1} = \frac{\rho^2\alpha}{2\beta}\left[\frac{(\alpha-\beta)^2 - (\alpha+\beta)\alpha e^{-2\beta t}}{(\alpha+\beta)(\alpha-\beta)^2}\right],$$
$$= \frac{\rho^2\alpha}{2\beta}\left[\frac{\alpha^2(1-e^{-2\beta t}) - \alpha\beta(2+e^{-2\beta t}) + \beta^2}{(\alpha^2-\beta^2)(\alpha-\beta)}\right],$$

$$
\begin{aligned}
&= \frac{\rho^2\alpha}{2\beta}\left[\frac{\alpha^2(1-(1-2\beta t+\frac{(2\beta t)^2}{2!}-\ldots))}{(\alpha^2-\beta^2)(\alpha-\beta)}\right] \\
&\quad -\left[\frac{\alpha\beta(2+(1-2\beta t+\frac{(2\beta t)^2}{2!}-\ldots))+\beta^2}{(\alpha^2-\beta^2)(\alpha-\beta)}\right], \\
&= \frac{\rho^2\alpha}{2\beta}\left[\frac{2\alpha^2\beta t-3\alpha\beta+\mathscr{O}(\beta^2))}{(\alpha^2-\beta^2)(\alpha-\beta)}\right].
\end{aligned}
$$

Cancelling the $\beta$ terms and allowing $\beta \to 0$ gives

$$
\frac{\rho^2\alpha}{2}\left[\frac{2\alpha^2 t-3\alpha}{\alpha^3}\right] = \frac{\rho^2}{2\alpha}\left[2\alpha t-3\right]. \tag{3.26}
$$

Now we simplify the second underbrace labelled 2.

$$
\begin{aligned}
\underbrace{\left(\frac{\rho^2\alpha}{2\beta\,(\alpha+\beta)}+\frac{\rho^2\alpha^2}{2\beta\,(\alpha-\beta)^2}-\frac{\rho^2\alpha^2}{\beta\,(\alpha-\beta)^2}\right)}_{2} &= \frac{\rho^2\alpha}{2\beta}\left[\frac{1}{\alpha+\beta}+\frac{\alpha}{(\alpha-\beta)^2}-\frac{2\alpha}{\alpha^2-\beta^2}\right] \\
&= \frac{\rho^2\alpha}{2\beta}\left[\frac{(\alpha-\beta)^2+\alpha(\alpha+\beta)}{(\alpha+\beta)(\alpha-\beta)^2}-\frac{2\alpha}{(\alpha^2-\beta^2)}\right] \\
&= \frac{\rho^2\alpha}{2\beta}\left[\frac{(\alpha^2-\beta^2)((\alpha-\beta)^2+\alpha(\alpha+\beta))}{(\alpha+\beta)(\alpha-\beta)^2(\alpha^2-\beta^2)}\right] \\
&\quad -\left[\frac{2\alpha((\alpha+\beta)(\alpha-\beta)^2)}{(\alpha+\beta)(\alpha-\beta)^2(\alpha^2-\beta^2)}\right] \\
&= \frac{\rho^2\alpha}{2\beta}\left[\frac{\alpha^3\beta-\alpha\beta^3+\alpha 2\beta^2-\beta^4}{(\alpha+\beta)(\alpha-\beta)^2(\alpha^2-\beta^2)}\right].
\end{aligned}
$$

Thus, allowing $\beta \to 0$ we have have the following result,

$$
\frac{\rho^2}{2\alpha}. \tag{3.27}
$$

Summing our above simplified expressions we obtain

$$\xi_{\mathrm{F}}^{*} = \frac{\sigma^2}{2\alpha}(1-e^{-2\alpha t}) + \frac{\rho^2}{2\alpha}(2\alpha t - 3) + \frac{2e^{-\alpha t}\rho^2}{\alpha} - \frac{e^{-2\alpha t}\rho^2}{2\alpha}. \tag{3.28}$$

### 3.3.4 Follower-Follower Covariance

$$\xi_{\mathrm{FF}}(t) = \frac{\rho^2\alpha}{2\beta\left(\alpha+\beta\right)}\left(1-e^{-2\alpha t}\right) - \frac{\rho^2\alpha^2}{2\beta\left(\alpha-\beta\right)^2}\left(e^{-\beta t}-e^{-\alpha t}\right)^2 - \frac{\rho^2\alpha^2}{\beta\left(\alpha^2-\beta^2\right)}\left\{e^{-(\alpha+\beta)t}-e^{-2\alpha t}\right\}.$$

The derivation is similar to the above, with the exception that the first term is not present. This gives the result

$$\xi_{\mathrm{FF}}^{*} = \frac{\rho^2}{2\alpha}(2\alpha t - 3) + \frac{2e^{-\alpha t}\rho^2}{\alpha} - \frac{e^{-2\alpha t}\rho^2}{2\alpha}.$$

The derivations provided in Section 3.3.1 to 3.3.4 described the jointly Gaussian collective movement of multiple individuals. The formulation of the collective movement model we have presented, i.e. through a latent leader, can be seen as a model for correlated movement, with the leader acting as a device to construct a correlated movement model.

## 3.4 Behavioural Switching

Whilst the methodology presented in Niu et al. (2016) provides a model for collective movement in continuous time, it is limited by the homogeneous description of movement. The assumption that animals move with just a single movement process is somewhat unjust. In actuality, animals exist in a state of constant change from both internal and external drivers such as hunger, fatigue, predators and competition of resources. Although some behaviours cause daily shifts in behaviour such as feeding and resting, other behaviours can be displayed seasonally such as rutting, migrating and breeding. These changes in behaviour will indirectly change the way in which they move and the underlying movement process. Datasets with duration long enough that focal individuals will exhibit multiple movement patterns necessitate the inclusion of heterogeneous behaviours (Morales and Ellner, 2002).

Developing a model for collective movement that incorporates this diverse natural world is important for understanding the complexity of animal movement and behavioural ecology.

Motivated by this, we diversify this modelling technique by allowing the individual animals to have two movement processes. In the alternative process, the followers may be free of following the leader and move independently from time to time. A simple but effective way to do this is to have the behaviour of the followers switch between following the leader and independent Brownian motion. The Brownian motion type of movement can be modelled as

$$F_t^{y,k} | F_0^{y,k} \sim \mathrm{N}(F_0^{y,k}, t\sigma_{\mathrm{BM}}^2)$$

where $\sigma_{\mathrm{BM}}^2$ is the diffusion rate of the Brownian motion of the non-following animals.

The final piece left to discuss is the framework of switching between multiple behavioural states coupled with different movement characteristics. In mathematical terms, we can represent this as a Markov process in continuous time with both a diffusion component, location, and a discrete one, behaviour, as in Berman (1994). In the case of individual-level movement, the idea of a switching diffusion process driven by a continuous-time Markov chain was proposed in Blackwell (1997) and formalised in Blackwell (2003). In group movement modelling, a discrete-time version was described by Langrock et al. (2014); here we develop a multivariate Ornstein Uhlenbeck process for a group of animals, driven by a continuous-time Markov chain on a space representing their joint behaviour. In essence this approach conflates themes presented in Langrock et al. (2014), the exact methods from Blackwell et al. (2016) and the multivariate OU model for group movement presented in Niu et al. (2016).

### 3.4.1 Mathematical Formulation of Behavioural Switching

Let $J_t^k$ denote the $k$th animal's behavioural state at time $t$, taking values in $\{1,2\}$, where 1 represents the state of following the leader and 2 represents the state of independent Brownian motion.

Let $\boldsymbol{J}_t$ be a vector containing the behavioural states for the whole group of animals at time $t$, thus taking values in $\{1,2\}^n$. We take each $J_t^k$ independently to be a continuous-time Markov chain on $\{1,2\}$ having transition rates $\lambda_{1,2}$ and $\lambda_{2,1}$, where $\lambda_{1,2}$ is the switching rate of an individual from following the leader (OU process) to Brownian motion and $\lambda_{2,1}$ is the switching rate of an individual switching from Brownian motion to following the leader. The transition rates for $\boldsymbol{J}_t$ are then implied by that structure, although it is possible to allow for

additional structure for example via environmental covariate information as discussed in Chapter 5.

The formulation of behavioural switching for an animal within the group is as follows. The $k$th animal starts in some state $J_0^k = j$ and location $\boldsymbol{Y}_0 = (F_0^{x,k}, F_0^{y,k}) = (x_0, y_0)$, then the animal follows the movement process defined by behaviour $j$; that is, $F_t^{x,k}|F_0^{x,k}$ and $F_t^{y,k}|F_0^{y,k}$ are realisations of the $j$th diffusion process. If $j = 2$ the diffusion process is Brownian motion; if $j = 1$ the animal is following the leader, so that its movement jointly with that of the leader is multivariate Ornstein Uhlenbeck. This continues until the time of the first switch in behaviour, at time $T_1$, when the animal is at $F_{T_1}^{x,k}, F_{T_1}^{y,k}$. If the behaviour switches to $J_{T_1}^k = i$, the next part of the location trajectory is a realisation of the $i$th diffusion process, starting at $F_{T_1}^{x,k}, F_{T_1}^{y,k}$, and this procedure continues iteratively.

The behavioural switching allows a much wider range of observed movement patterns. For example, if switching between behaviours is relatively slow, following individuals will tend to be found closer to the leader, and therefore closer together, the higher the value of $\alpha$, while non-following animals will tend to drift away. However, faster switching between behaviours can complicate this picture, depending on the absolute and relative switching rates. For example, short periods of non-following behaviour will lead to the animals moving independently in the short term while generally remaining close together.

To avoid any potential identifiability and label switching issues such as those described in Frühwirth-Schnatter (2001) we explicitly hard-wire the states into our model formulation such that the identities of the states are distinct. In other words, the Ornstein Uhlenbeck (state 1) group movement state is modelled distinctly from the independent state (state 2).

The times of the switches in behaviour are modelled as a Poisson process that has been 'thinned' details of which are given in the next section.

## 3.5 Inference

### 3.5.1 Exact Simulation

The key to Bayesian inference for the above model is the simulation of trajectories augmented by switching times and locations, appropriately conditioned on the observed data. The simulation is carried out exactly, rather than the typical approach of making a discrete-time approximation as in Langrock et al. (2014) and assuming that a switch can only happen at a

discrete observation time point. Here we want to avoid this unnatural assumption and the poorly understood discretisation error involved.

We utilise a uniformisation approach where the times of switches in behaviour form a Poisson process with rate $\kappa$, which has been 'thinned', that is each potential switching time point is either retained or deleted probabilistically (Guttorp and Minin, 1995), in a way that depends on the movement process. This approach is similar to Markov-modulated Poisson processes of Fearnhead and Sherlock (2006) and the continuous-time Markov chains by Rao and Teh (2013). In a movement context the approach was introduced by Blackwell et al. (2016) (see Section 2.6). However, the inference method we discuss here has been extended to account for multiple individuals so that at every potential switch an individual is randomly selected to switch behaviour. The framework of Blackwell et al. (2016) was initially developed to allow for spatially heterogeneous behavioural switching although, for the time being we are dealing with spatially homogeneous transition rates and so this device is not strictly necessary; however, it is useful to prepare ourselves for developing the model to accept spatial covariates in the future by using a tool which readily allows them. Recall that the rate of the Poisson process, $\kappa$ needs to be an upper bound for all the actual transition rates. Since we are only interested in two behavioural modes (OU and BM) and the switches in behaviour are homogeneous in time and space, the transition rates are all of the form $\mathscr{N}^1\lambda_{1,2} + (n-\mathscr{N}^1)\lambda_{2,1}$, where $n$ is number of animals or members in the group and $\mathscr{N}^1$ is the number in state 1, we take

$$\kappa = n\max(\lambda_{1,2}, \lambda_{2,1}).$$

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 $\kappa$. Starting at some known vector of states $\boldsymbol{J}_{T_0}$ for all members of the group, we can simulate the process forward as follows. Let

$$T \sim \exp(\kappa),$$

be the time of the first event of a process with constant rate $\kappa$. This is the first potential time at which a change in behaviour might occur. We can then determine whether the potential switch at $T$ is an actual switch, an event which has probability

$$P(\text{actual switch}) = \lambda(T)/\kappa,$$

where

$$\lambda(T) = \mathscr{N}^{1,T_0}\lambda_{1,2} + \mathscr{N}^{2,T_0}\lambda_{2,1},$$

is the actual transition rate at time $T$, $\mathscr{N}^{1,T_0}$ is the number of animals following the leader at time $T_0$, and $\mathscr{N}^{2,T_0}$ is the number of animals moving as Brownian motion at time $T_0$. If it is an actual switch, we switch the state of the $k$th animal from the group with probability

$$P(\text{Animal } k \text{ switches}) = \lambda^{k,T_0}/\lambda(T),$$

where $\lambda^{k,T_0}$ is the switching rate of the $k$th animal at time $T_0$; $\lambda^{k,T_0}$ is $\lambda_{1,2}$ if the $k$th animal's previous state $J^k_{T_0}$ is 1 or $\lambda_{2,1}$ if the previous state $J^k_{T_0}$ is 2. If it is not an actual switch, nothing need be changed.

Knowing $\boldsymbol{J}_T$, we can iterate this procedure forwards. This leads to a natural way of extending the simulation over as long interval as we desire. If we denote the events of Poisson process by $T_1, T_2, \ldots$, then for each $T_j$ in turn, we generate location $\boldsymbol{Y}_{T_j}$ by forward simulation.

### 3.5.2 The state space form of the model

Given the behavioural states, we can transform the group dynamic with behaviour switching model into a linear state space model, which can be expressed in the following form:

$$\boldsymbol{Y}_{t_{i+1}} = e^{A_i(t_{i+1}-t_i)}\boldsymbol{Y}_{t_i} + q_i, \quad q_i \sim \text{MVN}(0, \Xi_i), \tag{3.29}$$

$$\boldsymbol{Z}_{t_i} = H_i\boldsymbol{Y}_{t_i} + \varepsilon, \tag{3.30}$$

where $q_i \sim \text{MVN}(0, \Xi_i)$ is the process noise, and $A_i$ and $\Xi_i$ can take different forms based on the behavioural states. The measurement model is constructed by defining $H_i$ through which the model is observed at the discrete time step $t_i$. $\boldsymbol{Z}_{t_i}$ is the observation of the followers' location and the leader location $L_{t_i}$ is unobserved; in the case where the leader is observed, this is straightforward to accommodate by modifying $H_i$ and hence $Z_{ti}$. We assume there is no observation error, and therefore we can set $\varepsilon$ to zero. The state space form of the model is the discrete-time version of the continuous Ornstein Uhlenbeck and Brownian motion behavioural switching model. Here this discretisation is not an approximation; we use all the observations (possibly irregularly spaced in time) and between observations the individuals follow a continuous process corresponding to their current behavioural state. This discretisation can be thought of as a mild solution to the stochastic differential equation (Da Prato and Zabczyk, 2014).

Formatting the model in this way also allows us to deal with frequent irregularities that are inherent in the collection of real observation data. Missing observations can be handled by adapting the measurement matrix $H_i$, for example in a system where we have three followers and one leader, when all animals are observed our matrix will be

$$H_i = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.31}$$

Alternatively, if animal two does not have an observation at this time point the associated third row is omitted and the measurement matrix will be given as

$$H_i = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.32}$$

Given the behavioural states $\boldsymbol{J}_{t_i}$ of the whole group at time $t_i$, the covariance matrix $\Xi_i$ and coefficient matrix $A_i$ need to be changed by setting the corresponding row and column to the Ornstein Uhlenbeck or Brownian motion version of the coefficient. For example, if we have one leader and three followers, and at time $t_i$, the second follower is moving as Brownian motion while the rest follow the leader, $J^2_{t_i} = 2$, $\boldsymbol{J}_{t_i} = [1\ 2\ 1]$. The corresponding $A_i$ and $\Xi_i$ become

$$A_i = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \beta & -\beta & 0 & 0 & 0 \\ 0 & \alpha & -\alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 & -\alpha \end{pmatrix} \quad \Xi_i = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \Xi_{11} & \Xi_{12} & 0 & \Xi_{14} \\ 0 & \Xi_{21} & \Xi_{22} & 0 & \Xi_{24} \\ 0 & 0 & 0 & \sigma^2_{\mathrm{BM}} & 0 \\ 0 & \Xi_{41} & \Xi_{42} & 0 & \Xi_{44}. \end{pmatrix}$$

In this example, the second follower (fourth row and column) is in the Brownian motion behavioural state. When the behavioural state changes from 1 (Ornstein Uhlenbeck) to 2

(Brownian motion) we update the covariance matrix $\Xi_{t_i}$ and $A_{t_i}$ by setting the fourth row and column to zero and leaving the rest unchanged. For the animal moving as Brownian motion, its attraction rate to the leader would be 0 and its movement is independent of the rest of group; recall that $\sigma^2_{\text{BM}}$ is the diffusion parameter of the Brownian motion. Here the first and second row and column of $A_i$ and $\Xi_i$ correspond to the attractor $\theta$ and the leader $L_{t_i}$. We keep the row for $\theta$ to be consistent with the setting in Niu et al. (2016), but since we concentrate here on the non-stationary case, $\theta$ is fixed to be 0 and is not used in the inference.

### 3.5.3 Prior Structure

Throughout the thesis we use uniform priors on $[0,\infty)$ for all parameters unless stated otherwise. In Chapter 6, the priors for the parameters are uniform however, some parameters have alternative constraints and may be negative. In some special cases $\rho$ is fixed (see Section 5.2.1).

The prior structure for the locations is dependent on the modelling framework, that is whether we are using the stationary (Section 3.2) or non-stationary (Section 3.3) version of the model. In the stationary case, the multivariate normal prior is conditional on parameter values and given these parameter values we have a stationary distribution. The stationary distribution is used as the prior distribution for the locations by serving as a starting point for the Kalman filter (Section 3.6). In the non-stationary case, the prior for the locations and thus the initialisation of the Kalman filter is less straightforward and must be treated as a separate case which is described in Section 3.7.

Although we use uniform priors throughout this thesis we can justify the propriety of the posterior densities using standard results. Gelman et al. (2013a) (Section 2.7) show that for a model with known mean and unknown variance the posterior is proper, even with an improper prior, provided that there is at least one observation. When considering the Brownian motion parameter, $\sigma_{BM}$, we can think of this as updating a normal distribution with a known mean of 0 and unknown variance. Of course, technically it is possible that we could have a realisation of the model where all of the animals are in the OU state for the entire time period and thus we have no observations in the BM state however, in practice this does not occur. Similarly, for the OU state each parameter is updated separately meaning that the other parameters are fixed and each update can be thought of as relating to either a normal with known mean and unknown variance, in the case of diffusion parameters, or known variance and unknown mean mean, for the attraction parameters. Again, there would be an issue with propriety if we

did not have a single data point, but in reality this doesn't happen. Lastly, the switching rate parameters are proper due to being bounded and thus lead to a proper posterior distribution.

## 3.6 Markov chain Monte Carlo and the inhomogeneous Kalman filter

### 3.6.1 Sampling the trajectory

Based on the simulation ideas above and the state space form of the model, we can produce an algorithm for Bayesian inference for these models combining Markov chain Monte Carlo techniques with the inhomogeneous Kalman filter. Given data $\boldsymbol{Z}_0, \ldots, \boldsymbol{Z}_t$ we want to sample from the posterior distributions for the parameters of the diffusion process and of the switching rates. The approach involves augmenting the data with the times of all changes of behavioural state, and associated locations. We actually sample times, locations and states for all potential changes, that is at all times of a Poisson($\kappa$) process. Since the true transition rates $\lambda_{1,2}, \lambda_{2,1}$ are unknown, we take their priors to be bounded above by $\kappa_1, \kappa_2$ respectively, and set $\kappa = n \max\{\kappa_1, \kappa_2\}$.

Let $\mathscr{T}_{observe} = \{t_0, \ldots, t_N\}$ be the set of the observation times, $\mathscr{T}_{potential} = \{T_{i,j}, i = 0, \ldots, N-1, j = 1, \ldots, M_i\}$ be the set of all potential switching time points, where $M_i$ is the number of potential switches with $t_i < T_{i,j} < t_{i+1}$ i.e. the potential switches between observation times, and $\mathscr{T}_{actual}$ be the actual switching time, with $\mathscr{T}_{actual} \subset \mathscr{T}_{potential}$. We may have zero, one or multiple switches between two consecutive observation time points. The state of our chain is the collection of all times $\mathscr{T} = \mathscr{T}_{observe} \cup \mathscr{T}_{potential}$, plus associated locations $\boldsymbol{Y}_t$ for the whole group at time $t \in \mathscr{T}$, initial state $\boldsymbol{J}_{t_0}$, the states $\boldsymbol{J}_t$ at potential switching time points, and implied states at the times of observations $\boldsymbol{J}_{t_1}, \ldots, \boldsymbol{J}_{t_N}$.

The key Markov chain Monte Carlo step is to sample the trajectory, that is potential switches, locations and states, over some time interval $(t_a, t_b)$ such that $t_0 \leq t_a < t_b \leq t_N$, conditional on the trajectory outside that interval, on the states $J_{t_a}$, $J_{t_b}$ and on the movement and switching parameters. We define $\mathscr{T}^{ab'}_{potential} = \{T'_{i,j}, i = a, \ldots, b-1, j = 1, \ldots, m_i\}$ with $t_i < T'_{i,j} < t_{i+1}$ and $m_i$ the number of potential switches between $t_i$ and $t_{i+1}$, as the set of all proposed potential switching times in the interval $(t_a, t_b)$, a realisation of a Poisson ($\kappa$) process on $(t_a, t_b)$. Once we propose all the time points of the potential switches in the interval $(t_a, t_b)$, we can determine probabilistically whether these are actual switches and propose behavioural states.

Starting with $\boldsymbol{J}_{t_a}$, the behavioural states for the whole group at $t_a$, the next actual switching time $T'_{a,1}$ and corresponding behavioural states $\boldsymbol{J}_{T'_{a,1}}$ can be proposed as in Section 3.5.1, iterating to obtain $\boldsymbol{J}_{T'_{a,j+1}}$ for $j = 1, \ldots, m_a - 1$. This proposal process is repeated on each subinterval $(t_i, t_{i+1})$ for $i = a, \ldots, b-1$. We require for consistency that the final simulated behavioural states $\boldsymbol{J}_{T'_{b,m_b-1}}$ match the existing augmentation $\boldsymbol{J}_{t_b}$; if not, rejection is automatic. Conditioning on the proposed behavioural states, we can sample the trajectory by simulating the diffusion process forward. The simulation and inference of the model requires the unattainable locations of the 'abstract' leader so an inhomogeneous Kalman Filter is used with the state space model formulation to integrate out the leader's location.

### 3.6.2 Inhomogeneous Kalman filter

The Kalman filter can be used for computing the exact Bayesian posterior distributions of the state in the state space form of the group movement with behaviour switching model. In our case, the transition of the state of Kalman filter depends on the behaviour states $\boldsymbol{J}_t$. Therefore, the system dynamics of the Kalman filter is not linear, giving an inhomogeneous Kalman filter.

Unlike the inference algorithm in Niu et al. (2016) which requires imputing the unobserved leader's location to compute the marginal likelihood, the Kalman filter can integrate out the leader's location. Here, a two-step scheme is presented, which first calculates the marginal distribution of the next step using the known system dynamics, given the behavioural states. In the prediction step, the mean, $m$, and covariance matrix, $P$, can be derived as follows.

$$
\begin{aligned}
m_{t_i|t_{i-1}} &= e^{A_i(t_i - t_{i-1})} m_{t_i|t_{i-1}}, \\
P_{t_i|t_{i-1}} &= e^{A_i(t_i - t_{i-1})} P_{t_i|t_{i-1}} (e^{A_i(t_i - t_{i-1})}.)^{\mathrm{T}} + \Xi_i
\end{aligned}
$$

Here the subscript $t_i|t_{i-1}$ represents the prediction at step $t_i$ conditional on the state at $t_{i-1}$. The recursive iteration is initialised by presenting the prior information in the form $\boldsymbol{Y}_0 \sim \mathrm{MVN}(m_0, P_0)$, where $P_0$ is defined as in Equation 3.18 and $m_0$ is defined as in Equation 3.17. The algorithm then uses each observation to update the distribution to match the new information obtained by the measurement in step $t_i$. The equations for updating the system are given as

$$
k_{t_i} = P_{t_i|t_{i-1}} H_i^{\mathrm{T}} (H_i P_{t_i|t_{i-1}} H_i^{\mathrm{T}})^{-1},
$$

$$m_{t_i|t_i} = m_{t_i|t_{i-1}} + k_{t_i}(\boldsymbol{Z}_{t_i} - H_i m_{t_i|t_{i-1}}),$$
$$P_{t_i|t_i} = P_{t_i|t_{i-1}} - k_{t_i} H_i P_{t_i|t_{i-1}} H_i^{\mathrm{T}} k_{t_i}^{\mathrm{T}},$$

where $(.)^{-1}$ denotes the matrix inverse and $(.)^{\mathrm{T}}$ the matrix transpose. As a result, the filtered forward-time posterior process in step $t_i$ is given by $\boldsymbol{Y}_{t_i} \sim \mathrm{MVN}(m_{t_i|t_i}, P_{t_i|t_i})$. In this iterative computation, $A_i$ and $\Xi_i$ will change according to the behavioural states $\boldsymbol{J}_{t_i}$. $H_i$ will also change according to the availability of the observations at time step $t_i$. The updating step is only run when $t_i \in \mathscr{T}_{observe}$, whereas we need to run the prediction step at every potential switching time and observation time.

Given the behavioural states $\boldsymbol{J}_{T'_{a,1}}, \ldots, \boldsymbol{J}_{T'_{b,m_b-1}}$ in the interval $(t_a, t_b)$, the log likelihood of trajectories in the interval $(t_a, t_b)$ is

$$p(\boldsymbol{Z}|\boldsymbol{\Theta}, \boldsymbol{J}_{t_a,\ldots,t_b}) = -\sum_{i=a+1}^{b} \frac{1}{2}\big\{n\log 2\pi + \log|H_i P_{t_i|t_{i-1}} H_i^{\mathrm{T}}| + (Z_{t_i} - H_i m_{t_i|t_{i-1}})^{\mathrm{T}}(H_i P_{t_i|t_{i-1}} H_i^{\mathrm{T}})^{-1}(Z_{t_i} - H_i m_{t_i|t_{i-1}})\big\}. \tag{3.33}$$

The initialisation of the Kalman filter in the non-stationary case however is not trivial; the details of the formulation are given in the Section 3.7.

### 3.6.3 Inference of Diffusion Parameters

We use standard random-walk Metropolis-Hastings updates for the diffusion parameters. For each set of parameters $\Theta$ we use the Kalmam filter (Equation 3.33) to evaluate the likelihood. We propose new diffusion parameters $\boldsymbol{\Theta'}$ using the symmetric Gaussian proposal distribution centered on the previous values $\boldsymbol{\Theta}$. The new parameters are accepted with probability $\min\{\mathrm{HR}, 1\}$ where HR is the Hastings ratio

$$\frac{p\left(\boldsymbol{\Theta'}|\boldsymbol{J}, \mathscr{T}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{\lambda}\right) q(\boldsymbol{\Theta}|\boldsymbol{\Theta'})}{p\left(\boldsymbol{\Theta}|\boldsymbol{J}, \mathscr{T}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{\lambda}\right) q(\boldsymbol{\Theta'}|\boldsymbol{\Theta})} = \frac{p(\boldsymbol{\Theta'})p(\boldsymbol{Z}|\boldsymbol{\Theta'}, \boldsymbol{J}_{t_0,\ldots,t_N})}{p(\boldsymbol{\Theta})p(\boldsymbol{Z}|\boldsymbol{\Theta}, \boldsymbol{J}_{t_0,\ldots,t_N})}$$

## 3.7 Initialising the Kalman Filter in the Non-stationary Case

For the non-stationary case there is no exact way to initialise the Kalman Filter. The purpose of this section is to derive a meaningful distribution at time $t_0$, just before the first observation.

We suppose that the system has been following the model for some long time, $\Delta t$, and write $t_{-1} = t_0 - \Delta t$. The positions of the animals at $t_{-1}$ are taken to be independent and completely unknown which we can represent with an arbitrary mean $\mu$ and large variance $\delta$. Hence, we let the mean and the covariance matrices be

$$m_{-1} = \begin{pmatrix} 0 \\ \mu \\ \mu \\ \mu \end{pmatrix}, \quad P_{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 \\ 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & \delta \end{pmatrix}.$$

In this case, $m_{-1}$ and $P_{-1}$ represent the scenario with one leader and two followers; note that this case shall be used for demonstration purposes but the principle holds for any dimension. We can show that as $\delta, \Delta t \to \infty$ the limit of the distribution conditional on the initial observation is independent of $\delta$ and $\Delta t$. Recall the Kalman Filter equations.

Prediction step:

$$\begin{aligned} m_{t_i|t_{i-1}} &= e^{A\Delta t} m_{t_i|t_{i-1}}, \\ P_{t_i|t_{i-1}} &= e^{A\Delta t} P_{t_i|t_{i-1}} (e^{A\Delta t})^{\mathrm{T}} + \Xi^*, \end{aligned}$$

Updating Step:

$$\begin{aligned} k_{t_i} &= P_{t_i|t_{i-1}} H_i^{\mathrm{T}} (H_i P_{t_i|t_{i-1}} H_i^{\mathrm{T}})^{-1}, \\ m_{t_i|t_i} &= m_{t_i|t_{i-1}} + k_{t_i} (\boldsymbol{Z}_{t_i} - H_i m_{t_i|t_{i-1}}), \\ P_{t_i|t_i} &= P_{t_i|t_{i-1}} - k_{t_i} H_i P_{t_i|t_{i-1}} H_i^{\mathrm{T}} k_{t_i}^{\mathrm{T}}, \end{aligned}$$

Our attraction matrix $A$ and covariance matrix $\Xi^*$ are

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \alpha & -\alpha & 0 \\ 0 & \alpha & 0 & -\alpha \end{pmatrix}, \quad \Xi^* = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \xi_{\mathrm{L}}^* & \xi_{\mathrm{LF}}^* & \xi_{\mathrm{LF}}^* \\ 0 & \xi_{\mathrm{LF}}^* & \xi_{\mathrm{F}}^* & \xi_{\mathrm{FF}}^* \\ 0 & \xi_{\mathrm{LF}}^* & \xi_{\mathrm{FF}}^* & \xi_{\mathrm{F}}^* \end{pmatrix}.$$

Using our Kalman filter equations for the prediction step we have

$$m_{t_0|t_{-1}} = e^{A\Delta t} m_{t_0|t_{-1}}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1-e^{-\alpha\Delta t} & e^{-\alpha\Delta t} & 0 \\ 0 & 1-e^{-\alpha\Delta t} & 0 & e^{-\alpha\Delta t} \end{pmatrix} \begin{pmatrix} 0 \\ \mu \\ \mu \\ \mu \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ \mu \\ \mu \\ \mu \end{pmatrix}.$$

$$\begin{aligned} P_{t_0|t_{-1}} &= e^{A\Delta t} P_{t_0|t_{-1}} (e^{A\Delta t})^{\mathrm{T}} + \Xi^* \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1-e^{-\alpha\Delta t} & e^{-\alpha\Delta t} & 0 \\ 0 & 1-e^{-\alpha\Delta t} & 0 & e^{-\alpha\Delta t} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \delta & 0 & 0 \\ 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & \delta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1-e^{-\alpha\Delta t} & e^{-\alpha\Delta t} & 0 \\ 0 & 1-e^{-\alpha\Delta t} & 0 & e^{-\alpha\Delta t} \end{pmatrix}^{\mathrm{T}} + \Xi^* \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1-e^{-\alpha\Delta t} & e^{-\alpha\Delta t} & 0 \\ 0 & 1-e^{-\alpha\Delta t} & 0 & e^{-\alpha\Delta t} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \delta & (1-e^{-\alpha\Delta t})\delta & (1-e^{-\alpha\Delta t})\delta \\ 0 & 0 & \delta e^{-\alpha\Delta t} & 0 \\ 0 & 0 & 0 & \delta e^{-\alpha\Delta t} \end{pmatrix} + \Xi^* \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \delta & (1-e^{-\alpha\Delta t})\delta & (1-e^{-\alpha\Delta t})\delta \\ 0 & (1-e^{-\alpha\Delta t})\delta & (1-e^{-\alpha\Delta t})^2\delta + e^{-2\alpha\Delta t}\delta & (1-e^{-\alpha\Delta t})^2\delta \\ 0 & (1-e^{-\alpha\Delta t})\delta & (1-e^{-\alpha\Delta t})^2\delta & (1-e^{-\alpha\Delta t})^2\delta + e^{-2\alpha\Delta t}\delta \end{pmatrix} + \Xi^*. \end{aligned}$$

Denote the summation of the above as

$$P_{t_0|t_{-1}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix},$$

where

$$\zeta_{\mathrm{L}}(\Delta t) = \rho^2 \Delta t + \delta, \tag{3.34}$$

$$\zeta_{\mathrm{LF}}(\Delta t) = \rho^2 \Delta t - \left(\frac{\rho^2}{\alpha} - \delta\right)(1 - e^{-\alpha \Delta t}), \tag{3.35}$$

$$\zeta_{\mathrm{F}}(\Delta t) = \frac{\rho^2(2\alpha\Delta t - 3)}{2\alpha} + 2e^{-\alpha\Delta t}\left(\frac{\rho^2}{\alpha} - \delta\right) - e^{-2\alpha\Delta t}\left(\frac{\rho^2}{2\alpha} - \delta\right) + \delta, \tag{3.36}$$

$$\begin{aligned}\zeta_{\mathrm{FF}}(\Delta t) = &\frac{\sigma^2\left(1 - e^{-2\alpha\Delta t}\right)}{2\alpha} + \frac{\rho^2(2\alpha\Delta t - 3)}{2\alpha} + 2e^{-\alpha\Delta t}\left(\frac{\rho^2}{\alpha} - \delta\right) \\ &- e^{-2\alpha\Delta t}\left(\frac{\rho^2}{2\alpha} - \delta\right) + \delta,\end{aligned} \tag{3.37}$$

The next step is to update our distribution conditional on the first observation. Note that it is possible to observe any subset of the individuals. Therefore, to avoid separate derivations for each case we consider observing only one animal and then instantaneously observe the remaining where applicable. We present the case of updating the distribution having observed one animal below. The transformation matrix, $H$, is

$$H = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}.$$

We can write the Kalman gain as

$$\begin{aligned} k_{t_0} &= P_{t_0|t_{-1}} H_{t_0}^{\mathrm{T}} (H_{t_0} P_{t_0|t_{-1}} H^{\mathrm{T}})^{-1} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \left\{ \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}^{-1} \\ &= \begin{pmatrix} 0 \\ \zeta_{\mathrm{LF}} \\ \zeta_{\mathrm{F}} \\ \zeta_{\mathrm{FF}} \end{pmatrix} \left\{ \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \zeta_{\mathrm{LF}} \\ \zeta_{\mathrm{F}} \\ \zeta_{\mathrm{FF}} \end{pmatrix} \right\}^{-1} \end{aligned}$$

$$= \begin{pmatrix} 0 \\ \frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}} \\ 1 \\ \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \\ \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \end{pmatrix}.$$

Substituting the Kalman gain into our updating steps gives the following mean, $m_{t_0}$ and covariance matrix $P_{t_0}$.

$$\begin{aligned} m_{t_0} &= m_{t_0|t_{-1}} + k_{t_0}(\boldsymbol{Z}_{t_0} - H m_{t_0|t_{-1}}) \\ &= \begin{pmatrix} 0 \\ \mu \\ \mu \\ \mu \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}} \\ 1 \\ \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \end{pmatrix} \cdot \left\{ \boldsymbol{Z}_{t_0} - \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \mu \\ \mu \\ \mu \end{pmatrix} \right\}, \\ &= \begin{pmatrix} 0 \\ \mu \\ \mu \\ \mu \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}}(\boldsymbol{Z}_{t_0} - \mu) \\ (\boldsymbol{Z}_{t_0} - \mu) \\ \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}}(\boldsymbol{Z}_{t_0} - \mu) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}}(\boldsymbol{Z}_{t_0} - \mu) + \mu \\ \boldsymbol{Z}_{t_0} \\ \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}}(\boldsymbol{Z}_{t_0} - \mu) + \mu \end{pmatrix}. \end{aligned} \tag{3.38}$$

$$\begin{aligned} P_{t_0} &= P_{t_0|t_{-1}} - k_{t_0} H P_{t_0|t_{-1}} H^{\mathrm{T}} k_{t_0}^{\mathrm{T}}, \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix} \\ &\quad - \left\{ \begin{pmatrix} 0 \\ \frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}} \\ 1 \\ \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}} & 1 & \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \end{pmatrix} \right\}, \end{aligned}$$

$$
\begin{aligned}
&= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix} \\
&\quad - \left\{ \begin{pmatrix} 0 \\ \frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}} \\ 1 \\ \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{\zeta_{\mathrm{LF}}}{\zeta_{F}} & 1 & \frac{\zeta_{\mathrm{FF}}}{\zeta_{F}} \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\}, \\
&= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix} - \left\{ \begin{pmatrix} 0 \\ \frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}} \\ 1 \\ \frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\zeta_{\mathrm{LF}}{}^2}{\zeta_{\mathrm{F}}} & \zeta_{\mathrm{LF}} & \frac{\zeta_{\mathrm{LF}}\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \frac{\zeta_{\mathrm{LF}}\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} & \zeta_{\mathrm{FF}} & \frac{\zeta_{\mathrm{FF}}{}^2}{\zeta_{\mathrm{F}}} \end{pmatrix} \right\}, \\
&= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{LF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{FF}} & \zeta_{\mathrm{F}} \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\zeta_{\mathrm{LF}}^2}{\zeta_{\mathrm{F}}} & \zeta_{\mathrm{LF}} & \frac{\zeta_{\mathrm{LF}}\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \\ 0 & \zeta_{\mathrm{LF}} & \zeta_{\mathrm{F}} & \zeta_{\mathrm{FF}} \\ 0 & \frac{\zeta_{\mathrm{LF}}\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} & \zeta_{\mathrm{FF}} & \frac{\zeta_{\mathrm{FF}}^2}{\zeta_{\mathrm{F}}} \end{pmatrix}, \\
&= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{L}} - \frac{\zeta_{\mathrm{LF}}^2}{\zeta_{\mathrm{F}}} & 0 & \zeta_{\mathrm{LF}} - \frac{\zeta_{\mathrm{LF}}\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} \\ 0 & 0 & 0 & 0 \\ 0 & \zeta_{\mathrm{LF}} - \frac{\zeta_{\mathrm{LF}}\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} & 0 & \zeta_{\mathrm{F}} - \frac{\zeta_{\mathrm{FF}}^2}{\zeta_{\mathrm{F}}} \end{pmatrix}.
\end{aligned}
\tag{3.39}
$$

For more followers, if we only observe one animal we will always get a similar form to this. The dimensions of the matrix would be larger with the appropriate row and column of zeros depending on which follower was observed. We must now check that as $\delta, \Delta t \to \infty$ the distribution is finite and independent of $\delta$ and $\Delta t$. This is not immediately obvious, so let us first rewrite the equations (3.34) - (3.37) using Landau $\mathscr{O}$ notation. We choose to take the limit in such a way that $\delta \to \infty$, but not as fast as compared with $\Delta t$. Thus, we simplify the equations using $\mathscr{O}(\delta e^{-\alpha \Delta t})$, which tends to 0 as $\delta, \Delta t \to \infty$.

$$\begin{aligned}
\zeta_{\mathrm{L}}(\Delta t) =& \rho^2 \Delta t + \delta, \\
\zeta_{\mathrm{LF}}(\Delta t) =& \rho^2 \Delta t + \delta - \frac{\rho^2}{\alpha} + \mathscr{O}(\delta e^{-\alpha \Delta t}), \\
\zeta_{\mathrm{F}}(\Delta t) =& \rho^2 \Delta t + \delta + \frac{\sigma^2 - 3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha \Delta t}), \\
\zeta_{\mathrm{FF}}(\Delta t) =& \rho^2 \Delta t + \delta - \frac{3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha \Delta t}).
\end{aligned}$$

We calculate the limit of each equation in $m_{t_0}$ and $P_{t_0}$ as $\Delta t, \delta \to \infty$. Firstly, let

$$\mathscr{A} = \rho^2 \Delta t + \delta.$$

Then we calculate the limits for the mean $m_0$ given in Equation 3.38. We can show that the ratios of the $\zeta$ terms all go to 1 which means that every element of $m_{t_0}$ (except the initial 0) will go to $Z_{t_0}$.

$$\begin{aligned}
\frac{\zeta_{\mathrm{LF}}}{\zeta_{\mathrm{F}}} &= \frac{\mathscr{A} - \frac{\rho^2}{\alpha} + \mathscr{O}(\delta e^{-\alpha \Delta t})}{\mathscr{A} + \frac{\sigma^2 - 3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha \Delta t})} \\
&= \frac{\mathscr{A} - \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha \Delta t})}{\mathscr{A} + \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha \Delta t})} \\
&= 1.
\end{aligned}$$

$$\begin{aligned}
\frac{\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} &= \frac{\mathscr{A} - \frac{3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha \Delta t})}{\mathscr{A} + \frac{\sigma^2 - 3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha \Delta t})} \\
&= \frac{\mathscr{A} - \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha \Delta t})}{\mathscr{A} + \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha \Delta t})} \\
&= 1.
\end{aligned}$$

Thus, as $\delta$ and $\Delta t \to \infty$, the distribution mean is finite and independent of $\delta$ and $\Delta t$ and is

$$m_0 = \begin{pmatrix} 0 \\ Z_{t_0} \\ Z_{t_0} \\ Z_{t_0} \end{pmatrix}.$$

Now, we do the same for the covariance elements given in Equation 3.39.

$$\begin{aligned}
\zeta_{\mathrm{L}} - \frac{\zeta_{\mathrm{LF}}^2}{\zeta_{\mathrm{F}}} &= \frac{\left(\mathscr{A}(\mathscr{A} + \frac{\sigma^2-3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t})) - (\mathscr{A} - \frac{\rho^2}{\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t}))^2\right)}{\mathscr{A} + \frac{\sigma^2-3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t})} \\
&= \frac{\mathscr{A}(\mathscr{A} + \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha\Delta t}) - (\mathscr{A} - \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha\Delta t}))^2)}{\mathscr{A} + \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha\Delta t})} \\
&= \frac{\frac{(\sigma^2+\rho^2)\mathscr{A}}{2\alpha}}{\mathscr{A}} \\
&= \frac{(\sigma^2+\rho^2)}{2\alpha}
\end{aligned}$$

$$\begin{aligned}
\zeta_{\mathrm{LF}} - \frac{\zeta_{\mathrm{LF}}\zeta_{\mathrm{FF}}}{\zeta_{\mathrm{F}}} &= \frac{(\mathscr{A} - \frac{\rho^2}{\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t}))(\mathscr{A} + \frac{\sigma^2-3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t}) - \mathscr{A} + \frac{3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t}))}{\mathscr{A} + \frac{\sigma^2-3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t})} \\
&= \frac{(\mathscr{A} - \frac{\rho^2}{\alpha})(\frac{\sigma^2}{\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t}))}{\mathscr{A} + \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha\Delta t})} \\
&= \frac{\mathscr{A}\frac{\sigma^2}{\alpha} + \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha\Delta t})}{\mathscr{A}} \\
&= \frac{\sigma^2}{\alpha}
\end{aligned}$$

$$\begin{aligned}
\zeta_{\mathrm{FF}} - \frac{\zeta_{\mathrm{FF}}^2}{\zeta_{\mathrm{F}}} &= \frac{\left(\mathscr{A} - \frac{3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t})\right)\left(\mathscr{A} + \frac{\sigma^2-3\rho^2}{2\alpha} - \mathscr{A} + \frac{3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t})\right)}{\mathscr{A} + \frac{\sigma^2-3\rho^2}{2\alpha} + \mathscr{O}(\delta e^{-\alpha\Delta t})} \\
&= \frac{\left(\mathscr{A} - \frac{3\rho^2}{2\alpha}\right)\left(\frac{\sigma^2}{2\alpha}\right) + \mathscr{O}(\delta e^{-\alpha\Delta t})}{\mathscr{A} + \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha\Delta t})} \\
&= \frac{(\frac{\sigma^2}{2\alpha})\mathscr{A} - \frac{3\rho^2\sigma^2}{(2\alpha)^2} + \mathscr{O}(\delta e^{\alpha\Delta t})}{\mathscr{A} + \mathscr{O}(1) + \mathscr{O}(\delta e^{-\alpha\Delta t})} \\
&= \frac{\sigma^2}{2\alpha}
\end{aligned} \tag{3.40}$$

Therefore, as $\Delta t, \delta \to \infty$ the limit of the matrix is

$$P_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\rho^2+\sigma^2}{\alpha} & 0 & \frac{\sigma^2}{\alpha} \\ 0 & 0 & 0 & 0 \\ 0 & \frac{\sigma^2}{\alpha} & 0 & \frac{\sigma^2}{\alpha} \end{pmatrix}.$$

Thus, the limits of the equations are finite and independent of $\delta$ and $\Delta t$. These calculations were verified using Maple (Maplesoft, 2018).

### 3.7.1 Parameter inference

The behavioural states, switching rates and the diffusion parameters can be estimated using Markov chain Monte Carlo with a standard Metropolis Hastings algorithm. We propose new switching rates $\boldsymbol{\lambda}'$ using the symmetric Gaussian proposal distribution centered on the previous values $\boldsymbol{\lambda}$. The acceptance probability for $\boldsymbol{\lambda}'$ depends only on $\boldsymbol{J}$, since movement is independent of the rates given the states, and since $\mathscr{T}$ depends only on $\kappa$. The new switching rates are accepted with probability $\min\{\text{HR}, 1\}$ where HR is the Hastings ratio

$$\frac{p\left(\boldsymbol{\lambda}'|\boldsymbol{J},\mathscr{T},\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{\Theta}\right) q(\boldsymbol{\lambda}|\boldsymbol{\lambda}')}{p\left(\boldsymbol{\lambda}|\boldsymbol{J},\mathscr{T},\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{\Theta}\right) q(\boldsymbol{\lambda}'|\boldsymbol{\lambda})} = \frac{p(\boldsymbol{\lambda}')p(\boldsymbol{J}|\boldsymbol{\lambda}')}{p(\boldsymbol{\lambda})p(\boldsymbol{J}|\boldsymbol{\lambda})}$$

by conditional independence and symmetry.

Given the trajectory and states, we know exactly what type of the movement processes the group of animals were following, so the inference about the movement parameters is straightforward. From the Markov property, the trajectory log-likelihood is calculated by summing over terms of the form given in Equation 3.33 for the whole time interval. All followers are considered jointly. We use uniform priors on $[0,\infty)$ for all parameters except when otherwise stated for example, in Chapter 6 some parameters may be negative. We use standard random-walk Metropolis-Hastings updates for all parameters in this and any other subsequent chapters. We insist that diffusion parameters are non-negative values. The only non-standard aspect is the calculation of the likelihood, and so other details are omitted. Similar, lower-dimensional updates for a model of a single animal are described in detail by Blackwell (2003).

The code for the methodology is written in R (R Core Team, 2020) and can be found in the GitHub repository https://github.com/FayFrost. The code provided is for the covariate extension of the methodology discussed in Chapter 6 however, applications of this model in Chapters 3 - Chapter 5 can be implemented as a special case of the code.

## 3.8 Implementation with Simulated Data

To asses the model's performance, several simulation experiments were carried out. Specifically, we investigated the model's ability to recover the true known parameters values even when the algorithm is initialised with widely dispersed starting values. The first experiment is the simplest case where the leader has no point of attraction and there is no behavioural switching i.e. the model presented in Section 3.3; results are described in 3.8.1. In the remaining experiments behavioural switching is accounted for; we check that the model is able to pick up on a wide range of behavioural patterns by using simulated data with a vast range of parameter values. In 3.8.2, the data are simulated with relatively high values for both $\alpha$ and $\sigma_{BM}$, specifically $\alpha = \sigma_{BM} = 5$. Here we can imagine the animals are tightly grouped when in the OU state and widely separated when in the BM state, perhaps representing individual exploring behaviour. In contrast, the data in 3.8.3 are simulated with $\sigma_{BM} = 0.1$. This leads to movement behaviour where, when the animals are not grouped together, they forage locally, leading to rather stationary behaviour. The final simulation experiment uses parameter values similar to those obtained from the analysis of real data from reindeer tracking in 4.2; results are described in 3.8.4.

### 3.8.1 Simulation without switching

We simulated location data of five followers and one leading point in both $x$ and $y$ directions from the non-stationary intrinsic Ornstein Uhlenbeck model given in Section 3.3. The data consisted of 50 time steps for each individual forwarded by using Equation 3.16 iteratively and taking each generated location as the origin for the next. We used a single state version of the MCMC algorithm presented in Section 3.5. We used the MCMC algorithm with 30,000 iterations which took approximately 30 minutes to complete. The posterior point estimates and standard deviations are given in Table 3.1 along with the true parameter values. The posterior densities for the model parameters are given in Figure 3.1 with the true values of each parameter indicated by the dashed blue line.

| Parameter | True Value | Point estimate | Standard deviation |
|---|---|---|---|
| $\alpha$ | 1 | 0.99 | 0.11 |
| $\rho$ | 3 | 3.36 | 0.33 |
| $\sigma$ | 0.5 | 0.52 | 0.03 |

Table 3.1 Parameter estimates for movement and switching model with the simulated dataset where $\beta = 0$.

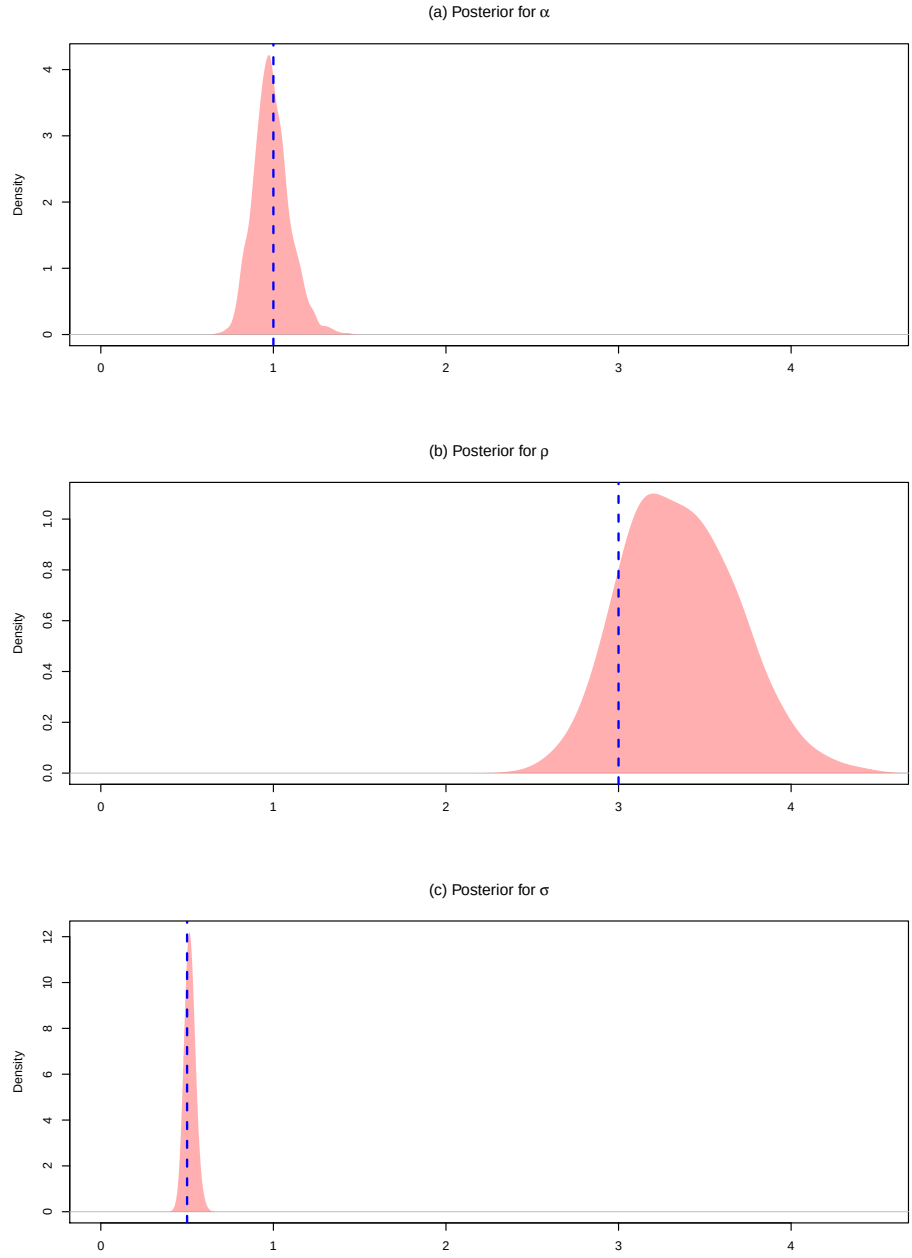


Fig. 3.1 Posterior densities for model parameters with simulation data, based on the Markov chain Monte Carlo runs of 30,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the individual variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower.

The model performs exceptionally well at recovering the true parameter values with all posterior point estimates having a difference from the true value by $< 12\%$.

### 3.8.2 Simulation with high attraction and diffusion

For simulated data with a high attraction parameter and diffusion coefficients, we ran the Markov chain Monte Carlo algorithm for 100,000 iterations after burn-in. This took approximately 180 minutes to complete. The true parameter values are given in Table 3.2, along with the point estimates and standard deviations of the posterior distributions for each parameter. The posterior densities are given in Figure 3.2. The posterior states are given in Figure 3.3. The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (black) represent the mean posterior of the estimated behaviour states.

| Parameter | Point estimate | Standard deviation | True value |
|---|---|---|---|
| $\alpha$ | 4.59 | 0.49 | 5.0 |
| $\rho$ | 3.17 | 0.23 | 3.0 |
| $\sigma$ | 0.46 | 0.03 | 0.5 |
| $\sigma_{\mathrm{BM}}$ | 4.04 | 0.34 | 5.0 |
| $\lambda_{1,2}$ | 0.13 | 0.02 | 0.1 |
| $\lambda_{2,1}$ | 0.49 | 0.08 | 0.4 |

Table 3.2 Parameter estimates for the movement and switching model with simulation data with high attraction and diffusion i.e. where $\alpha = B\sigma = 5$.

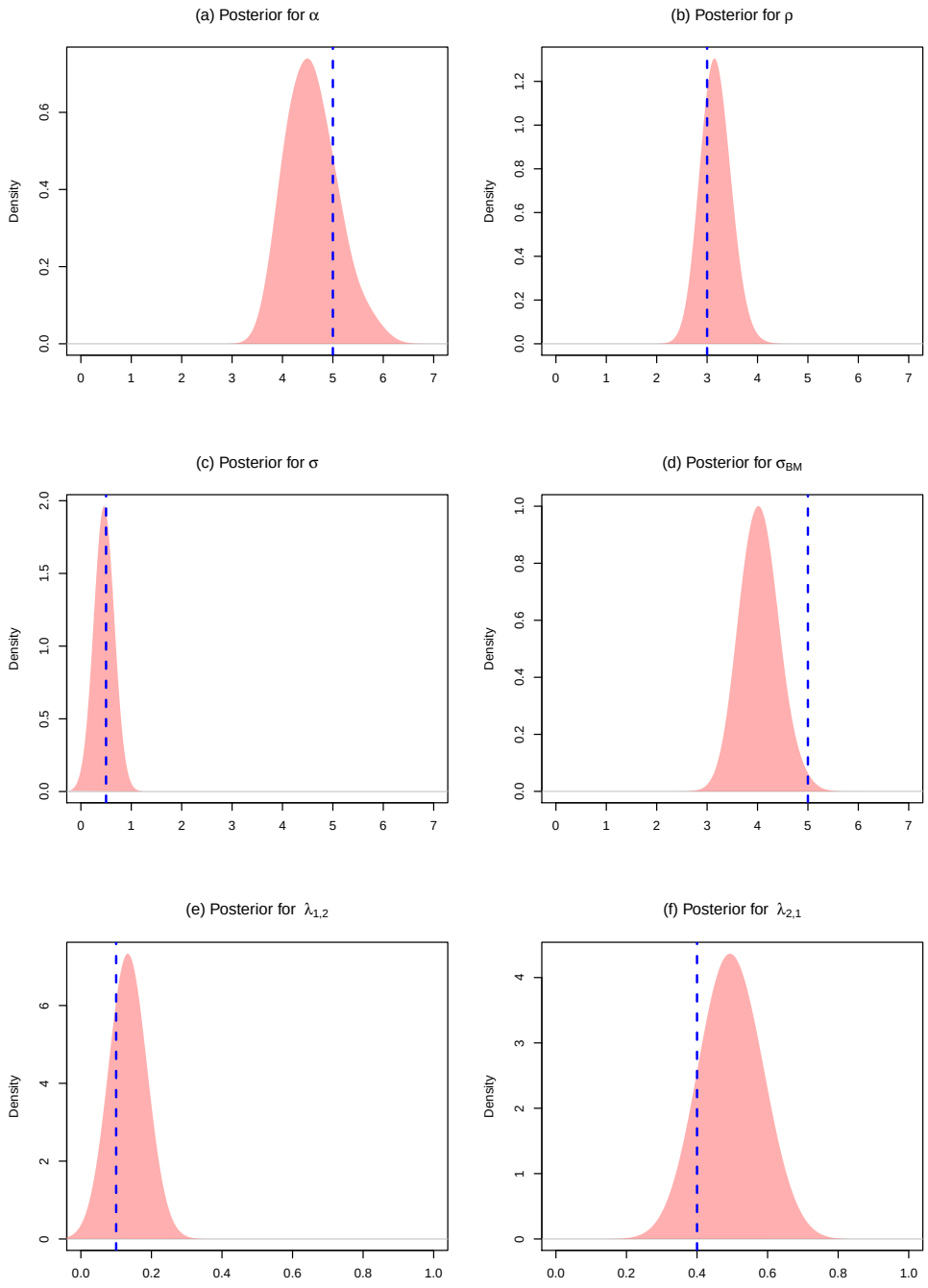


Fig. 3.2 Posterior densities for model parameters with simulation data presented in Table 3.2, based on the Markov chain Monte Carlo runs of 100,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (e) Posterior density for $\lambda_{1,2}$, the switching rate of the follower from OU to BM. (f) Posterior density of $\lambda_{2,1}$, the switching rate of the follower from BM to OU.

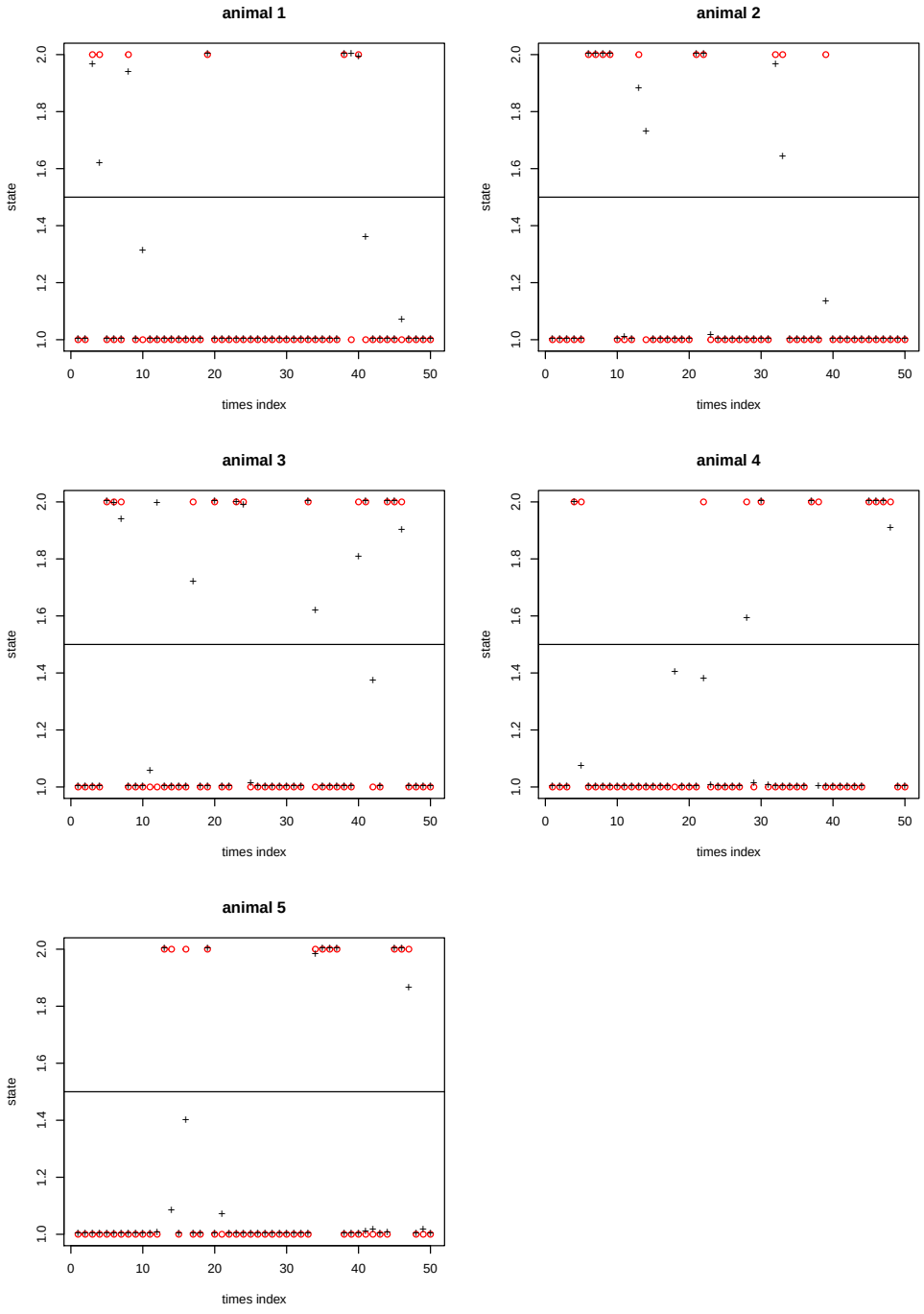


Fig. 3.3 Posterior mean states of all followers for the data with high attraction and diffusion i.e. where $\alpha = B\sigma = 5$. The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (black) represent the mean posterior of the estimated behaviour states. .

The model performed well at retrieving the true values, even with widely dispersed initial values.

We can visualise the movement trajectories by plotting each animal's paths in one dimension against time whilst simultaneously indicating the posterior state estimation at each time step. Trajectory plots of the results are given in Figure 3.4. For completeness the trajectories in two dimensions are presented in Figure 3.5. The point estimates of the posterior states are marked as circles for OU and squares for BM for all five animals at each time step. In this simulation study, we set the true value of $\sigma_{BM} = 5$ , $\lambda_{1,2} = 0.1$ and $\lambda_{2,1} = 0.4$. Since $\lambda_{2,1} > \lambda_{1,2}$, the animal has higher probability to switch from BM to OU.

The state estimation also confirmed animals spent most of the time in OU states. Once they are in BM state, they will quickly switch to OU. The high value of $\sigma_{BM}$ leads to large movement steps when an animal is in the BM state, as is clear from the trajectories plotted in Figure 3.4 and 3.5. Considering the individual trajectories, state estimation is difficult, compared with the results in Section 3.8.3 but carrying out the estimation jointly gives good results.

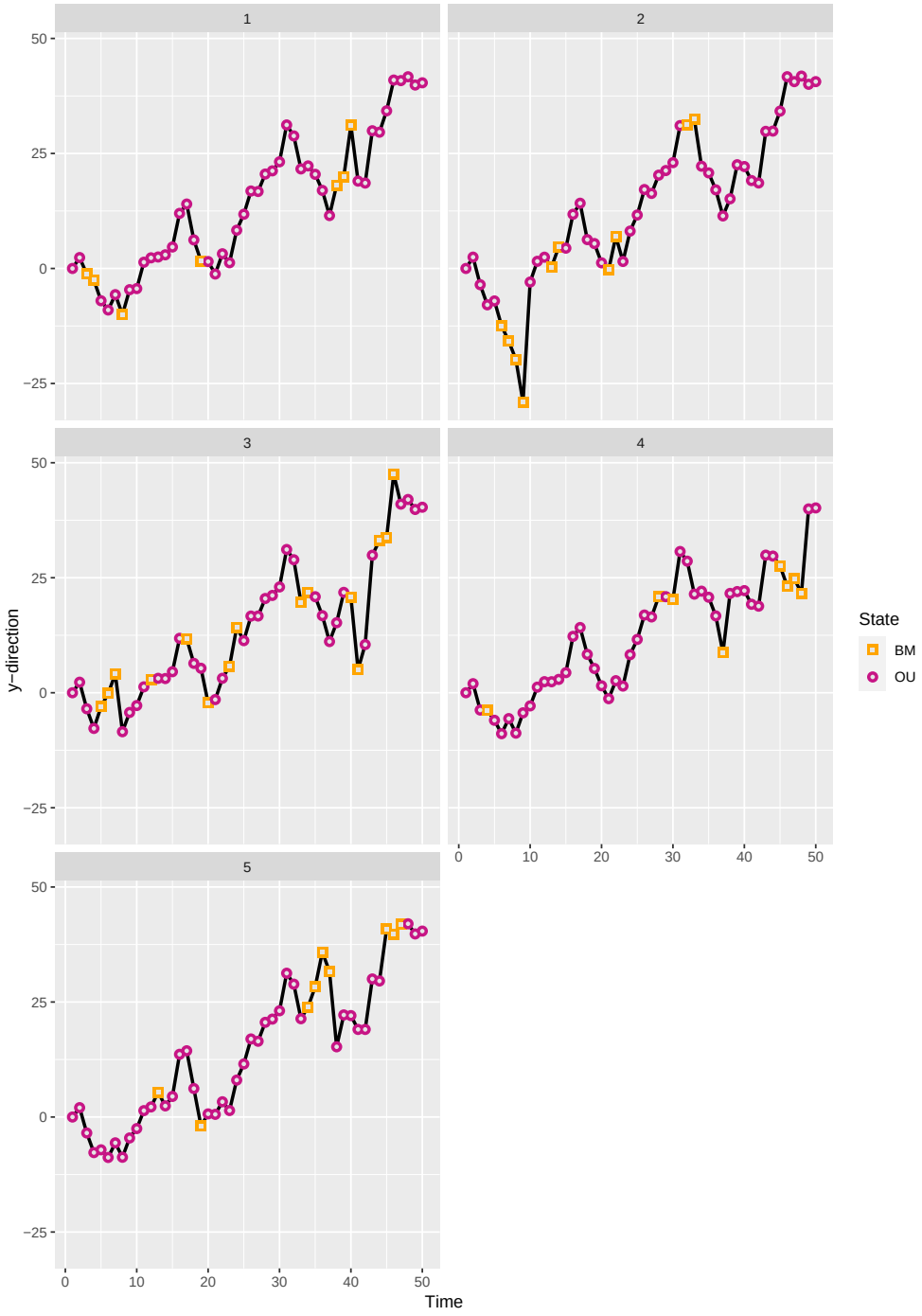


Fig. 3.4 Time trace of locations in $y$-direction for each animal from the simulated data presented in Table 3.2. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange square points indicate an BM state whilst the purple circular points indicate OU states.

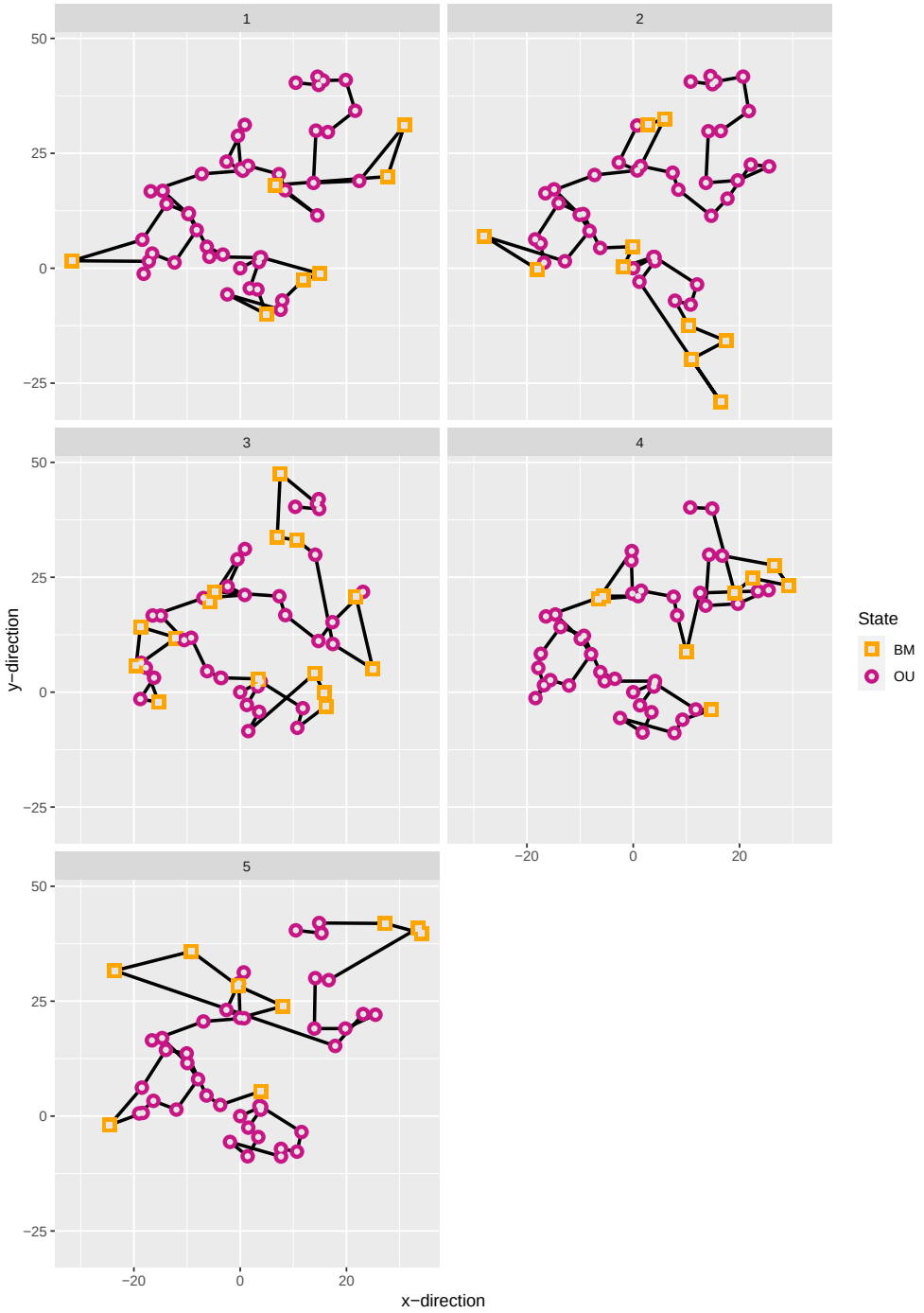


Fig. 3.5 Trajectories for each animal from the simulated data presented in Table 3.2. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange square points indicate an BM state whilst the purple circular points indicate OU states.

### 3.8.3 Simulation with low diffusion

For simulated data with a low diffusion coefficient in the non-following state, we ran the Markov chain Monte Carlo algorithm for 100,000 iterations after burn-in. This took approximately 180 minutes to complete. The true parameter values for each data set are given in Table 3.3, along with the point estimates and standard deviations of the posterior distributions for each parameter. The posterior densities are given in Figure 3.6. The posterior states are given in Figure 3.7.

| Parameter | Point estimate | Standard deviation | True value |
|---|---|---|---|
| $\alpha$ | 5.44 | 0.29 | 5.0 |
| $\rho$ | 3.07 | 0.19 | 3.0 |
| $\sigma$ | 0.52 | 0.03 | 0.5 |
| $\sigma_{\mathrm{BM}}$ | 0.09 | 0.01 | 0.1 |
| $\lambda_{1,2}$ | 0.14 | 0.02 | 0.1 |
| $\lambda_{2,1}$ | 0.46 | 0.08 | 0.4 |

Table 3.3 Parameter estimates for the movement and switching model with simulation data with low diffusion i.e. where $B\sigma = 0.1$.

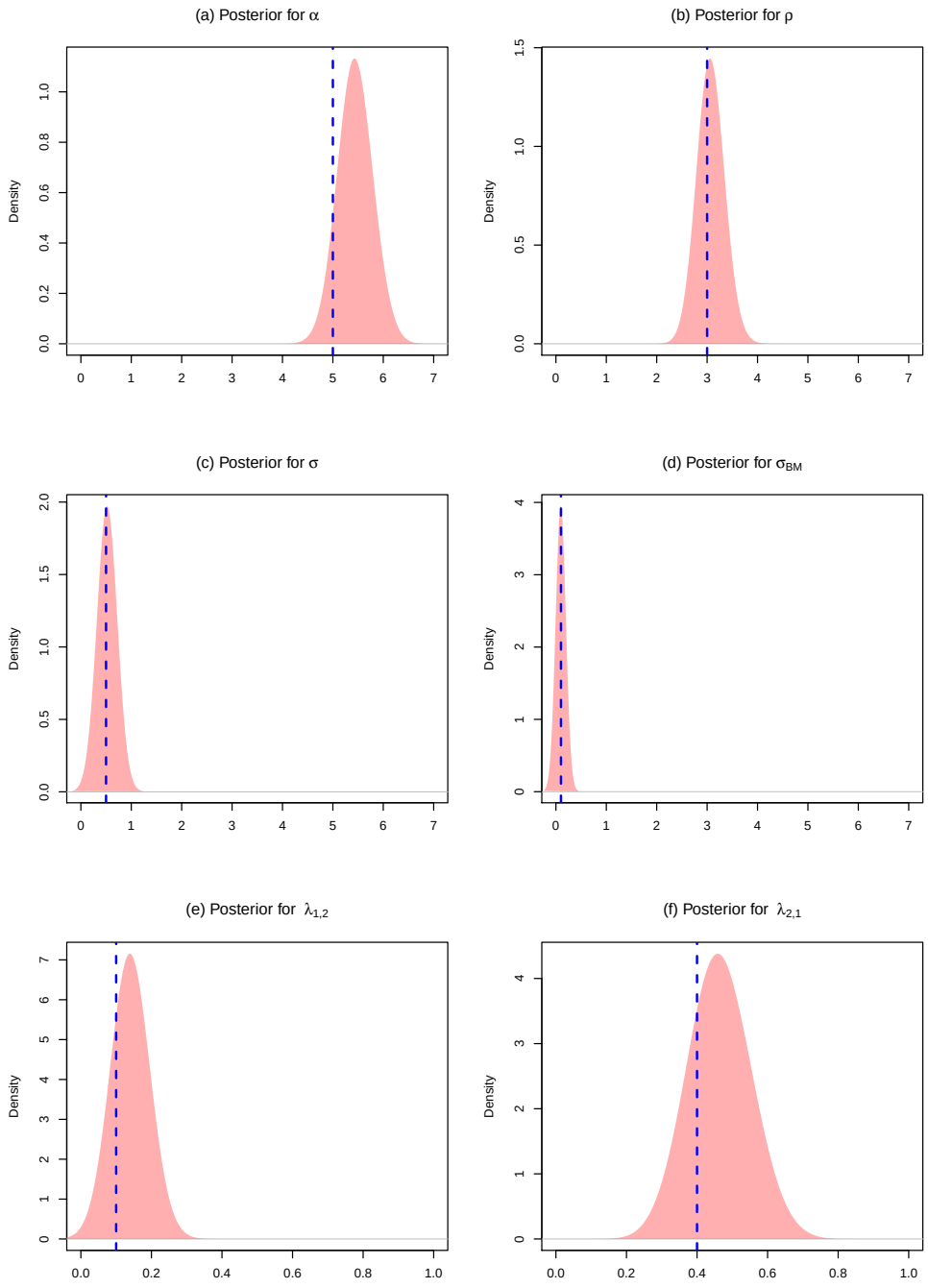


Fig. 3.6 Posterior densities for model parameters with simulation data presented in Table 3.3, based on the Markov chain Monte Carlo runs of 100,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (e) Posterior density for $\lambda_{1,2}$, the switching rate of the follower from OU to BM. (f) Posterior density of $\lambda_{2,1}$, the switching rate of the follower from BM to OU.

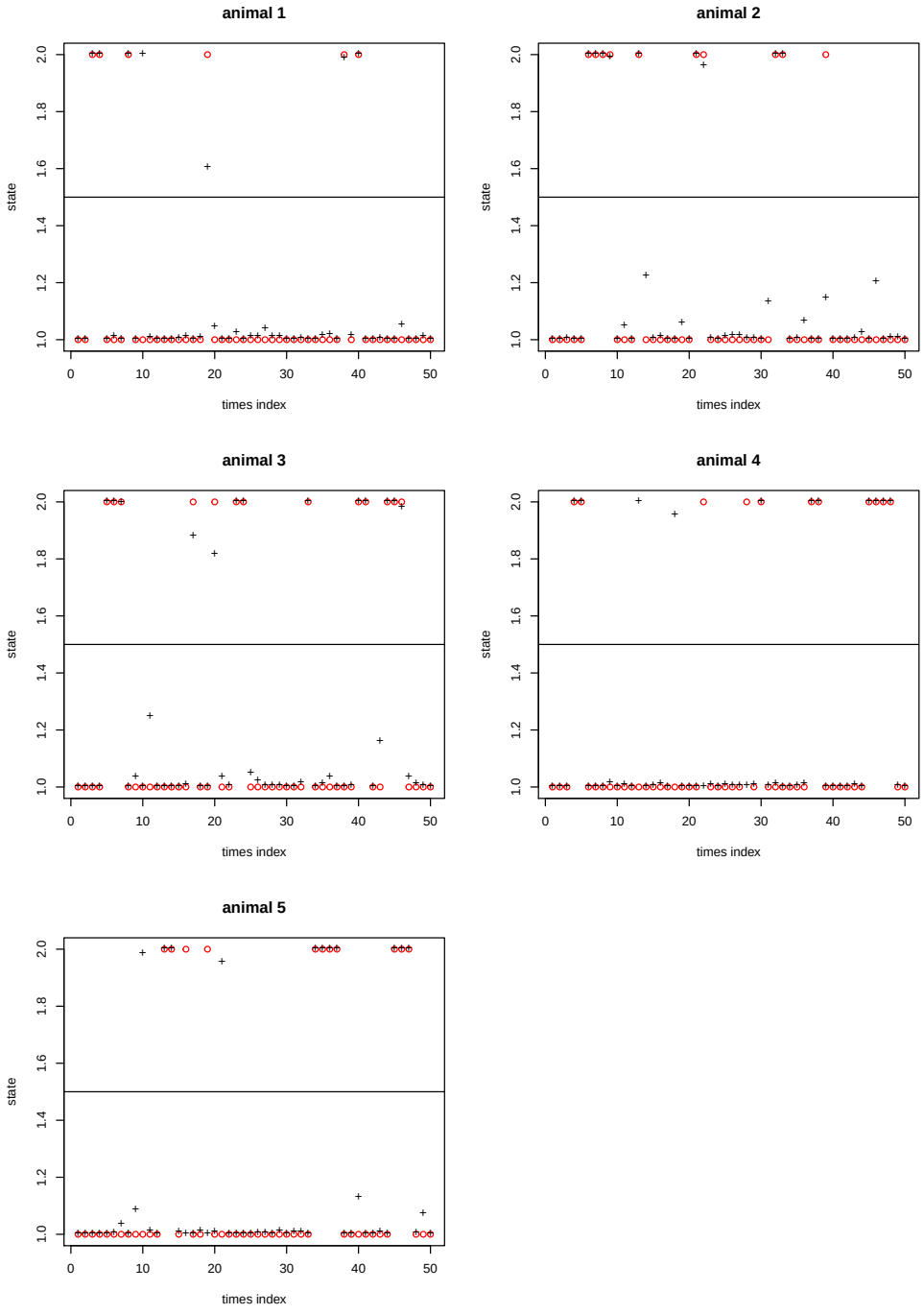


Fig. 3.7 Posterior mean states of all followers for the data set presented in Table 3.3. The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (black) represent the mean posterior of the estimated behaviour states.

For comparison with the previous example, the movement trajectories in one and two dimensions are shown in Figure 3.8 and Figure 3.9 respectively. In this simulation study, the true value of the BM diffusion parameter $\sigma_{BM} = 0.1$ is much smaller than Section 3.8.2. The effects of this small BM diffusion parameter are clearly demonstrated in Figure 3.8 and 3.9, with movement in the BM state being much more localised than before. As expected, state estimation is generally good in this case; the parameter estimation also reflects the true values, and correctly captures the qualitative difference from the previous case.

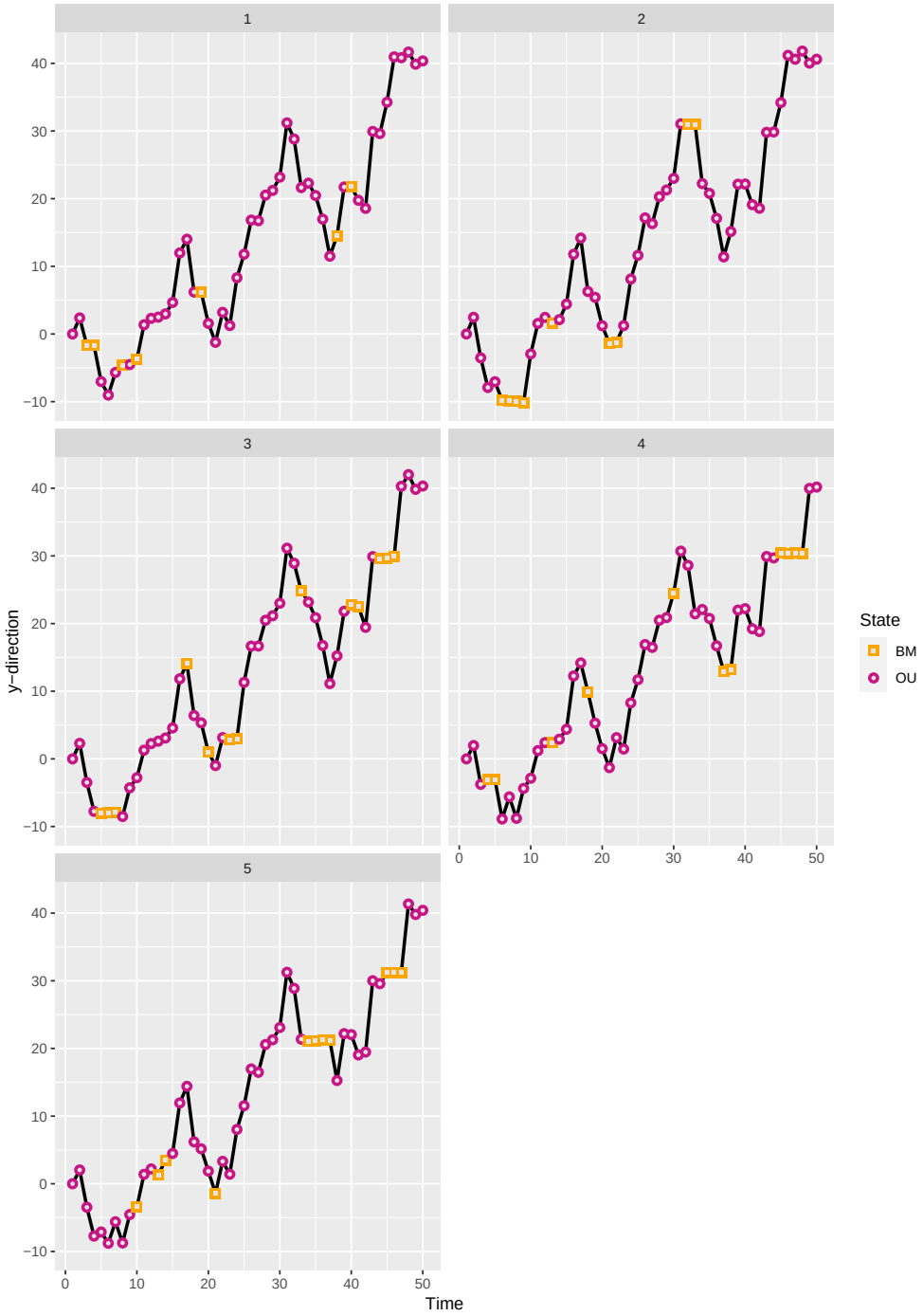


Fig. 3.8 Time trace of locations in $y$-direction for each animal from the simulated data presented in Table 3.3. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange square points indicate an BM state whilst the purple circular points indicate OU states.

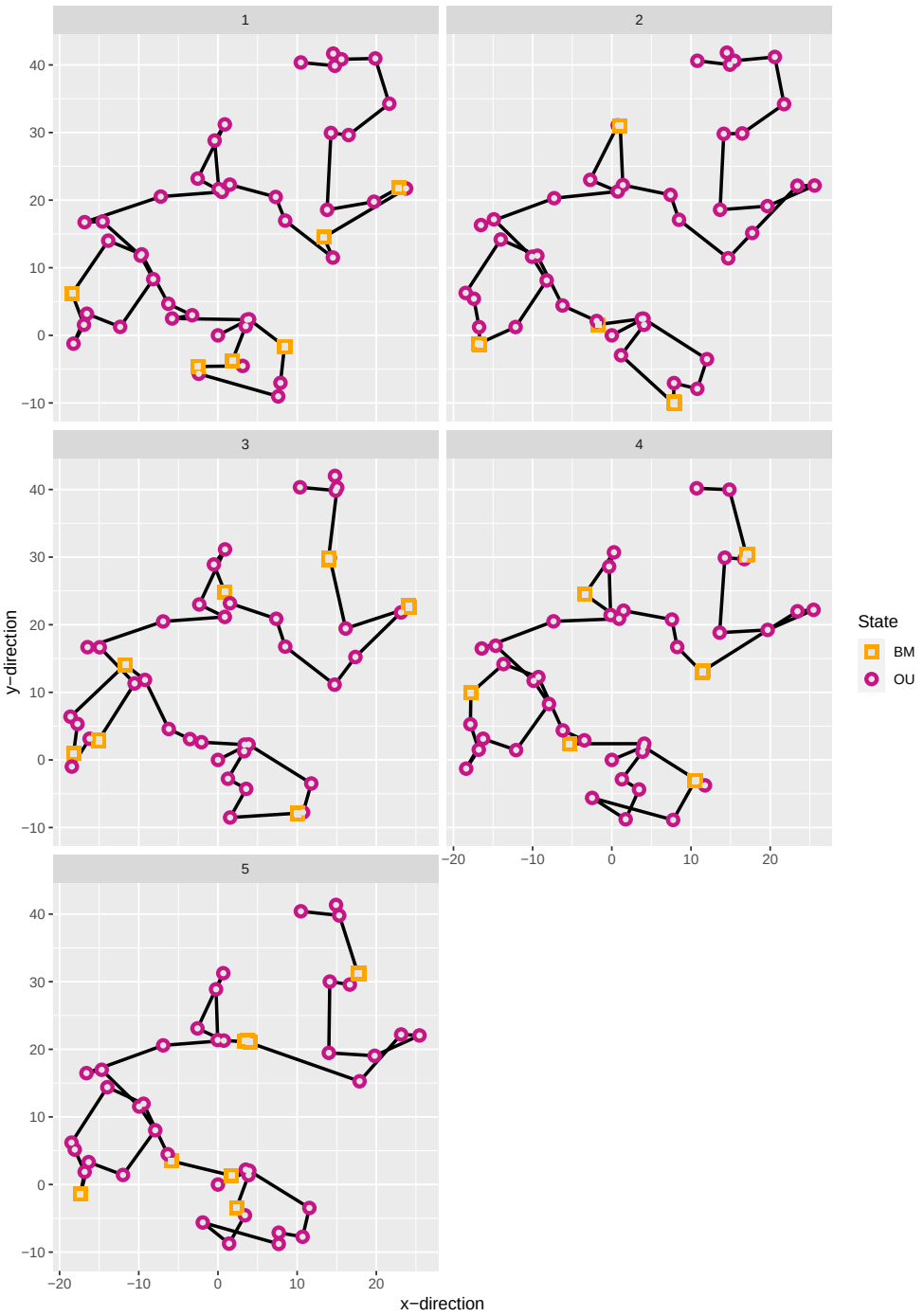


Fig. 3.9 Trajectories for each animal from the simulated data presented in Table 3.3. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange square points indicate an BM state whilst the purple circular points indicate OU states.

### 3.8.4 Reindeer-based Simulation

We simulated location data of five followers where the true diffusion parameters used in the simulation are motivated by the results of real data analysis presented in Chapter 4. The data consisted of 50 time steps. We ran the MCMC algorithm for 50,000 iterations after burn-in. This took approximately 120 minutes to complete.

The posterior point estimates and standard deviations of the model parameters are shown in Table 3.4 along with the true values used in the simulation. The posterior densities for each parameters are given in Figure 3.10. The posterior means of the behavioural states for each follower at every time point are plotted against the true behavioural states in Figure 3.11.

| Parameter | Point estimate | Standard deviation | True value |
|---|---|---|---|
| $\alpha$ | 1.23 | 0.06 | 1.2 |
| $\rho$ | 5.02 | 0.34 | 5.0 |
| $\sigma$ | 0.69 | 0.04 | 0.7 |
| $\sigma_{\text{BM}}$ | 1.59 | 0.17 | 2.0 |
| $\lambda_{1,2}$ | 0.14 | 0.02 | 0.1 |
| $\lambda_{2,1}$ | 0.51 | 0.08 | 0.4 |

Table 3.4 Parameter estimates for the movement and switching model with simulation data motivated by real reindeer data.

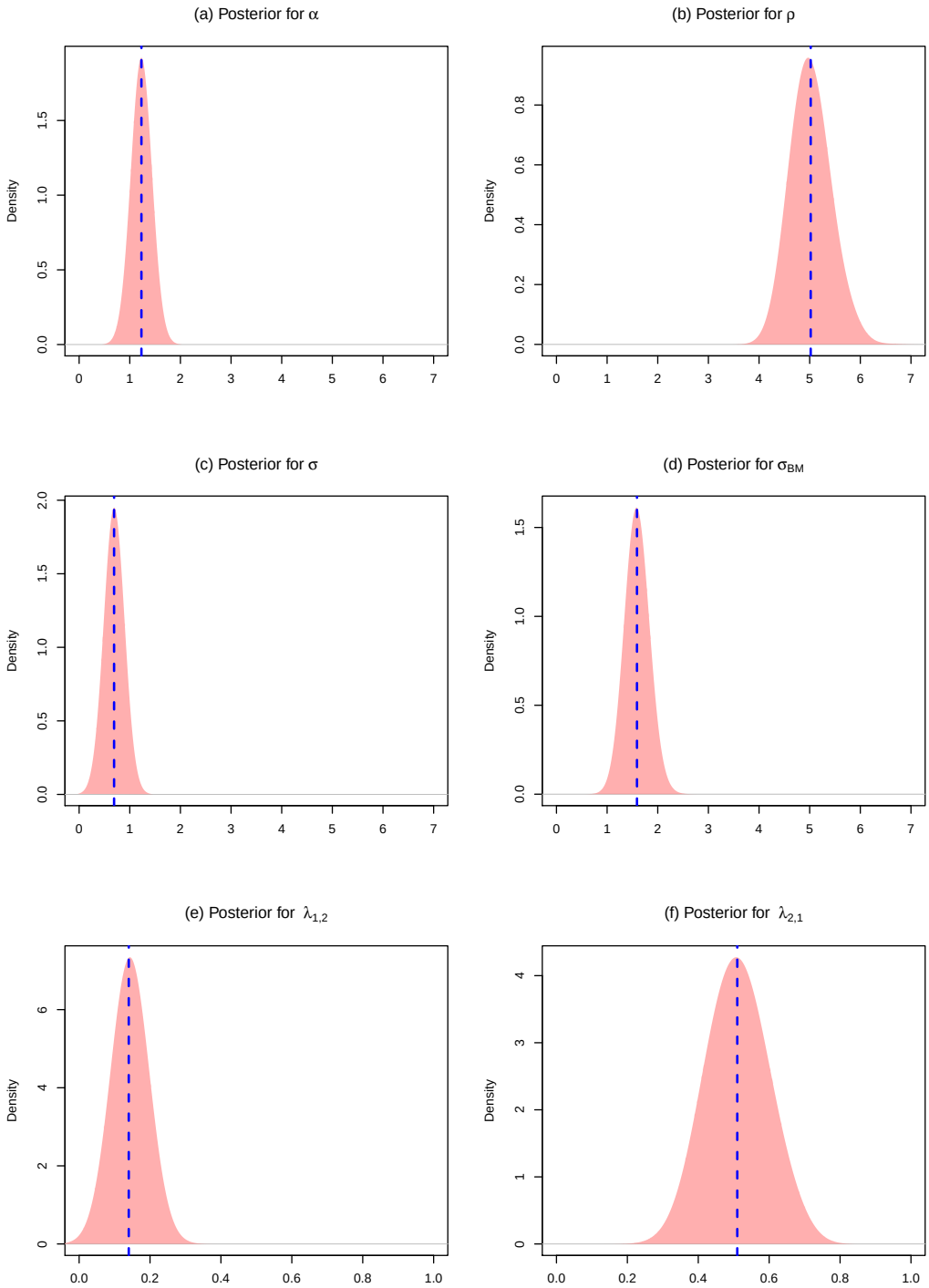


Fig. 3.10 Posterior densities for model parameters with reindeer-based simulated data, based on the Markov chain Monte Carlo runs of 50000 iterations. The dashed blue line represents the true parameter value. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it is in BM. (e) Posterior density for $\lambda_{1,2}$, the switching rate of the follower from OU to BM. (f) Posterior density of $\lambda_{2,1}$, the switching rate of the follower from BM to OU.

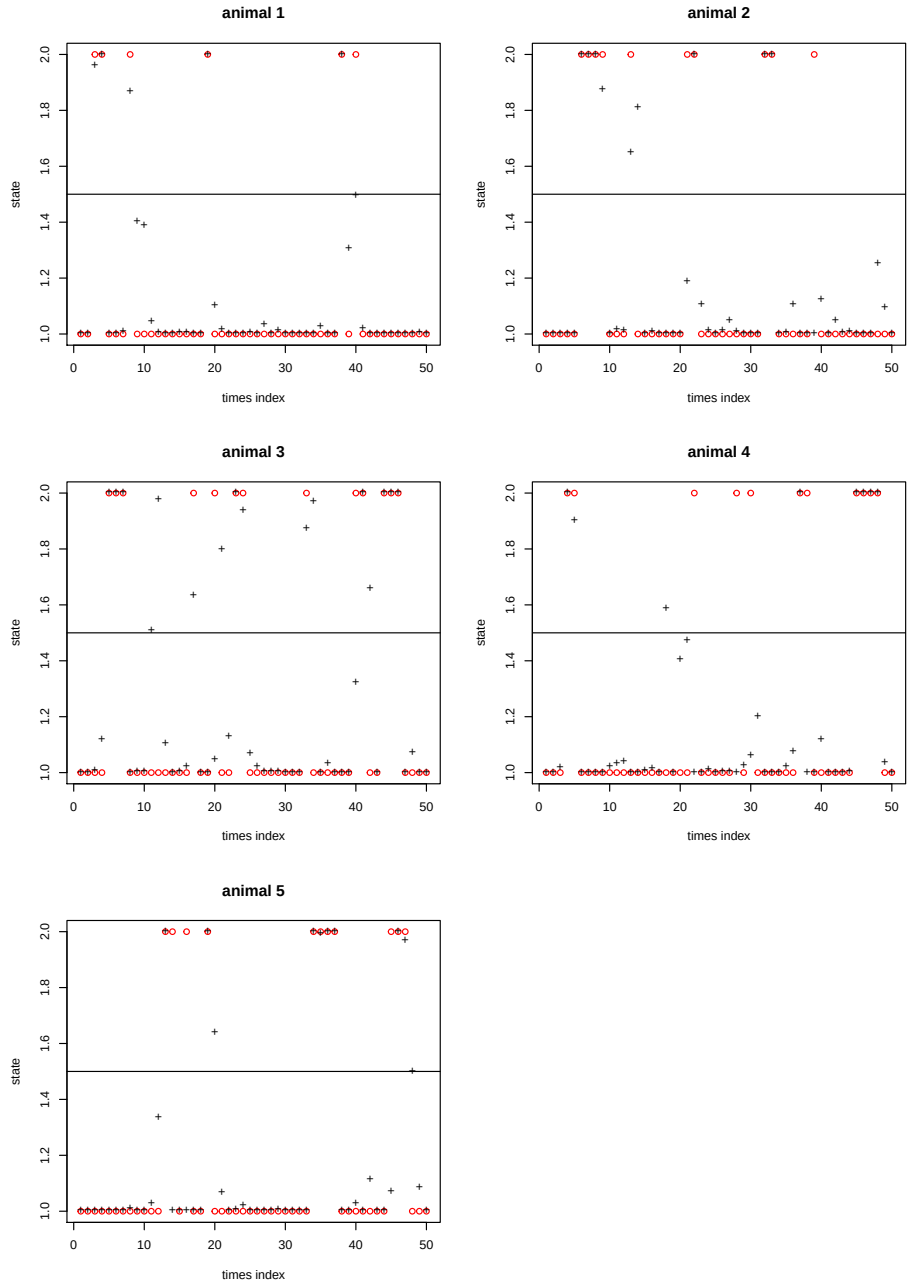


Fig. 3.11 Posterior mean states of all animals from the reindeer-based simulated data. The (red) circles represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The (black) crosses represent the mean posterior of the estimated behaviour states.

It is clear from Figure 3.11 that the model performs well at estimating the the true states. However, in some cases it captures behaviours better than others for example, animal 2's true states at time point 39, and animal 1's true states at time point 38 and 40 seem to have

been poorly estimated. In these two cases, the animals only stay in the Brownian motion behavioural state for very short time and then switch back to the Ornstein Uhlenbeck state following the leader. This makes it harder for the inference algorithm to capture the switching. On the other hand, if the animals move in certain behavioural states for longer time periods such as animal 2 and animal 1 in time interval 1 to 10, the estimated behaviour states match the truth very well.

## 3.9 Discussion

The work in this chapter has aimed to challenge areas of animal movement modelling that have lacked in attention relative to others. We have described the formulation of a group movement model with behaviour switching in continuous time, building on some of the strengths of previous approaches, and an algorithm for fully Bayesian inference. Using simulation experiments we have shown that we can successfully estimate the behavioural states and diffusion parameters. Compared to Niu et al. (2016), we have introduced behaviour switching in the continuous movement model and also extended the model to the non-stationary case by defining the leader's movement process as Brownian motion. By using a state space format and applying the Kalman filter in this case we have also made our calculations computationally more efficient and easily accounted for missing and irregularly spaced data.

Behaviour switching is important in real applications to provide a realistic representation of movement (Blackwell, 1997, 2003; Gurarie et al., 2010; Haydon et al., 2008; Langrock et al., 2014; Morales and Ellner, 2002). Simpler single-behaviour models fail to capture the heterogeneity of movement exhibited by animals as they respond to their environment. When considering multiple animals, these behaviours can represent complex trade-offs between environmental and social factors. For example, some species may exhibit grouping behaviour as protection from predators but in doing so increase their foraging competition (Hart and Mooring, 1992). This approach is unique in allowing this behavioural complexity for group movement while retaining the theoretical and practical benefits of formulation in continuous time.

Of course, if changes in behaviour are rapid compared to the time scale of the information from observations, for example if there are frequently multiple switches between observations, then it becomes impossible to reconstruct the sequence of behaviours, much less their precise timing, with any certainty. That is inevitable in any model of this kind; our approach does at

least allow properly for the different underlying possibilities, and the associated uncertainty, rather than ignoring them as would be necessary in a discrete-time model. Our method also introduces the use of the Kalman filter in the inference, which saves us from the need to impute the location of the leading point as in Niu et al. (2016). By massively reducing the dimension of the space to be explored by the Markov chain Monte Carlo algorithm, this makes the computation feasible even in this more complex model.

In all of the cases presented in this chapter we have fixed $\kappa = 3.5$. This allows the switching rate to be estimated at a maximum of $\kappa$/number of animals i.e. $3.5/5 = 0.7$. It is possible to fix $\kappa$ at a higher value but this comes with a trade-off. Increasing $\kappa$ will increase the amount of potential switches generated and thus proportionally increase the computational time. Since the densities of the switching rates in this chapter do not appear to be fixed around this upper bound of 0.7, then there is no incentive to increase $\kappa$ for these examples. In other cases, where switching rates are larger, there may be a need for increasing $\kappa$. Likewise, for lower switching rates or fewer animals we can save computational time by lowering $\kappa$.

The model does not require every individual to be observed which makes it extremely useful for modelling the movement and behaviours of large collectives where it is practically or economically infeasible to tag all members. This data-driven approach is built within a framework that considers the data collectively, not just in a pairwise or summarised fashion. A key advantage of this methodology over more metric-based approaches is its robustness to incomplete observation of the herd, and to variation over time of the number or identity of observed individuals within the herd. Each observed individual's behaviour within the herd is indirectly quantified through interactions with the leader (represented as an abstract point, whose location is estimated through an averaging over the observed animals). Methods based on statistical summaries may have varying results depending on who or how many animals were tagged. This can lead to difficulty in interpreting the results and understanding how they may respond under different movement patterns. We can think of the formulation we present as a model for the whole herd where we only have '*partial observations'* i.e. the animals which are tagged, and this partial data acts as a proxy for the unobserved individuals.

As the leader is merely an abstraction it is not necessary for there to be some true biological leading animal such as an alpha male or matriarch and so we are not concerned with tagging that particular individual. However, in species where there exists an actual leader, analysis would be much simpler if the leader is observed. If the leader *is* observed, that information would be included through the observation matrix of the state space model, rather than being integrated out. This is also mentioned in Section 3.5.2 when the state-space form of the model is defined.

This approach considers a group represented by a single 'leader' and animals who follow the leader for part of the time. A model which allows switching between multiple separate leaders, suitable for species with more complex social structures, but which relies on a more complete tracking of individuals, is explored by Milner et al. (2020).

Since GPS tacking data usually has low observational error it is often ignored when formulating models of animal movement such as those presented in this chapter. However, the use of the Kalman filter means that it would be straightforward to allow for observation error, taking $\varepsilon$ to be non-zero in Equation 3.30. Another constraint is the invariability of the switching rates; the specific models discussed in detail and applied here have switching rates for each individual which are spatially and temporally homogeneous. However, the method is formulated and implemented within a uniformisation approach which makes it possible to incorporate heterogeneity in switching rates, following Blackwell et al. (2016). This extension is discussed in Chapter 6.

The modelling technique offers some original insight in to collective movement however, as is often the case, the inference methods are limited by their computational cost. That said, I believe that the robust and generalisable conclusions, invariant under which animals are tagged, which can be drawn from such a mechanistic approach far outweighs the computational cost. The model allows for exact Bayesian statistical analysis for movement in continuous time without the need for discretisation error. This flexible approach is unbounded in the diversity of species it may be applied to since it may account for, in principle, any number of animals and in any number of dimensions meaning that there are no restrictions on avian or marine species if the tracking technology allows. In Chapter 4 we apply the model to the location data of semi-domesticated reindeer (*rangifer tarandus*) in the Njaarke herding district of Sweden.

# Chapter 4

# Case Study: Collective Reindeer Movement

Reindeer (*rangifer tarandus*) or Caribou as they are referred to in North America are large gregarious and migratory mammals. They live in a variety of habitats such as Arctic tundra, mountain ranges and forests. In the wild they are native to circum-arctic regions such as north-western US, Canada, Alaska, Greenland and some parts of Scandinavia e.g. Norway. Although, they have been introduced as a semi-domesticated species in other areas such as Iceland. Semi-domesticated (otherwise known as domesticated or tame) reindeer are herded in a pastoral system where they roam freely but with seasonal herding for calving, slaughter and relocation (Skarin and Åhman, 2014). The species are herbivorous, eating the available plants and shrubs in their environment. Reindeer also have the ability to digest lichens, unusual in large herbivores, which they rely on for winter forage (Falldorf et al., 2014) and can often dictate their seasonal movements (Merkle et al., 2016).

In 2015, wild reindeer were classified as a vulnerable species according to the IUCN red list after an observed population decline of 40 % over the course of 21-27 years (Gunn, 2016). This is a dramatic change given that the species were of 'Least Concern' in 2008 (Henttonen and Tikhonov, 2008). Although their domesticated relatives were not assessed in this study it is suggested that these populations are also at risk due to shifts in traditional reindeer herding areas due to competition of land-use (Pape and Löffler, 2012; Vistnes and Nellemann, 2001).

Most reindeer populations use ancestral grazing ranges and migration corridors. Due to this, declines in population can be explained largely through environmental change and habitat loss either directly e.g. as a result of human-animal conflict such as infrastructure development (Polfus et al., 2011; Skarin and Åhman, 2014) or indirectly due to anthropogenic activities

contributing to the warming globe and climate change. The increasing demand for natural resources such as timber, oil, wind and hydro-power has resulted in wider exploitation and encroachment in to historically untouched areas of the globe including arctic and subarctic regions (Hislop, 2013; Klein, 2000; Sandström, 2015).

Despite the footprint of such industrial sites being relatively small the cumulative effect of this and the associated roads, power-lines and cabins can result in avoidance behaviour much larger than the resources site itself. These so called 'zones of avoidance' can result in reindeer exhibiting avoidance behaviour up to 12km which leads to the use of less-suited environments with lower quality forage and higher predation risks (Panzacchi et al., 2013; Skarin and Åhman, 2014; Vistnes et al., 2004). In some extreme cases this can lead to the abandonment of traditional migration corridors and fragmentation of the landscape (Vistnes et al., 2004). Calving mothers have been found to be particularly sensitive to the psychological trauma caused by this and may have lactation trouble and generally poorer reproductive success (Lee et al., 2000; Vistnes and Nellemann, 2001).

Human activities, principally the burning of fossil fuels, can affect species environment as a result of climate change (Forbes et al., 2006; Fyfe et al., 2013). The northernmost countries and arctic regions experience the greatest impact of greenhouse gases with surface temperatures increasing more than twice the global average (Pörtner et al., 2019). The warming atmosphere has an adverse effect on the abundance of parasitic insects which can spread diseases, disturb foraging opportunities and contribute towards herd loss (Ballesteros et al., 2012; Kutz et al., 2014). Increasing temperatures can also have a negative impact on the available forage in the winter and spring. Reindeer are known to follow the growth of new vegetation - often regarded as green wave '*surfing*' (Merkle et al., 2016). The amount of accessible forage in the winter months is strongly influenced by the variation in climate. Warm spells in winter can lead to more rain-on-snow events. As the rain turns to ice the hardened snow-pack makes foraging lichens underneath difficult for the reindeer with heavy energetic costs (Forbes et al., 2016; Hansen et al., 2011; Lee et al., 2000; Putkonen and Roe, 2003).

Reindeer grazing may also play a pivotal role in mitigating the positive feedback loops of climate warming. Arctic shrub cover is increasing with the warming climate and this vegetation absorbs radiation which in turn leads to further warming. Studies have shown that herbivory at higher latitudes can increase surface albedo, the proportion of light reflected, by consuming the shrubbery that covers the land (Beest et al., 2016).

## 4.1 Importance of Semi-domesticated Herds

This thesis is concerned only with reindeer which form part of the semi-domesticated populations in the herding districts of Sweden which exist over most of the northern half of Sweden.

Reindeer husbandry refers to the ownership, management and maintenance of reindeer herds (Tyler et al., 2007). Husbandry in Sweden is an industry reserved for the indigenous Saami population. The herding districts have extensive land use covering around 22.6 million ha; approximately 55% of Swedish land cover. As of 2019, the Swedish domesticated reindeer population was estimated to be around 241,000 with 4600 herders spread over 51 districts (for more details visit the Saami Parliament webpage, https://www.sametinget.se/). Despite the Reindeer Grazing Act (1886) permitting herders to graze their reindeer on all private/public land, the districts are not used exclusively for husbandry and co-exist with other forms of land-use, hence the continuous conflict between herders and other forms of industry such as forestry.

As in other parts of the world, reindeer husbandry has huge economical, cultural and social value (Lee et al., 2000). Reindeer are highly valued for nutrition, clothing, spiritual and cultural reasons (Forbes and Kumpula, 2009). They are said to be a 'keystone species' whose impact on the community and landscape is disproportionately large relative to its abundance (Falldorf et al., 2014; Pape and Löffler, 2012; Power et al., 1996). Husbandry forms an integral part of the Saami livelihood and these populations depend heavily on traditional herding for cultural and physical survival (Forbes et al., 2006; Kofinas et al., 2000; Tyler et al., 2007).

The domesticated reindeer which form the herds in the Saami districts are not isolated from their wild relatives and suffer similar environmental challenges. Reindeer well-being in these herds is vital, not only for their survival but also for the Saami people who depend on them. The information gathered from modelling the movement of semi-domesticated reindeer can be transferred to studies of wild populations and used in management decisions and conservation, especially since the tagging of semi-domesticated reindeer is more convenient because they are already corralled seasonally for calf marking and harvesting.

In conclusion, reindeer face a multitude of threats, especially as a result of human activities. Their presence offers valuable resources such as food, fur and cultural worth whilst potentially dampening the feedback loops of climate warming. Rigorous statistical analysis and coherent modelling of the movement and grouping dynamics of reindeer can yield insightful

knowledge about behaviour, dispersal and land-use in relation to the environment. This information can be used to facilitate decisions about future infrastructure developments, conservation and management of herds to mitigate avoidable physiological and physiological trauma. This can be especially useful in reindeer husbandry which heavily relies on the well-being of a herd as a source of income.

## 4.2 Njaarke Data

The recurring dataset used in this thesis concerns semi-domesticated reindeer herded in the Jamtland province in a northern district of Sweden, Njaarke. The data is collected across three and a half years between 14/11/2009 and 31/12/2013 (Rivrud et al., 2018). The data consists of GPS observations of 78 reindeer where at any one time up to 40 reindeer are observed. The reindeer belong to one of two sub herds, namely "P&D" and "JKP". Those reindeer which form part of the P&D herd are moved from one pasture to another (via lorry transport) in alignment with the change from summer to winter. This is intended to maximise the resources available to the reindeer. The observations are regularly spaced every two hours with the exception that in August and February they are every half hour. However, often GPS fixes do not occur exactly when expected and there can be time lags upon receiving the fix, or in some cases no fix at all.

The tagged reindeer exist as part of a wider herd, consisting of approximately 2000 members (Ahman et al., 2014), where the rest of the individuals are not tagged. To illustrate the models presented so far in this thesis a sub-sample of the data has been selected. In the interest of minimising computational time, a subset of 5 reindeer has been chosen. The specific subset was chosen through exploratory data analysis where the reindeer have some reasonably close proximity and whose movements may be dependent upon on another at least for some of the time. The duration of the subset was chosen to be short enough to alleviate any computational strain but long enough that the reindeer are likely to exhibit multiple behavioural processes. The data consist of up to 50 observations from each individual taken every two hours from 01/12/2009 until 5/12/2009.

Whilst the observations are subject to some of the usual irregularities when dealing with real data, i.e. missing values and observation spacing inconsistencies, the observations are almost regular insofar as they occur up to only 2 minutes before/after the intended timing. Thus, for the simplicity of implementation the time steps of the data were rounded to the nearest hour. However, in principle the methodology accounts for irregular times between observations.

Figure 4.1 shows a map of the Njaarke herding district's location. The left map shows a map of Norway, Sweden and Finland from left to right. Sweden is highlighted in a darker shade of yellow. The perimeters of the herding districts of Sweden are given with black outlines with the Njaarke herding district indicated by the purple polygon. The right hand map gives a zoomed view of Njaarke herding district. The top left polygon denotes the area of land utilised all year by reindeer and the bottom right indicates the winter only pasture. In the figure we can see a vast quantity of environmental heterogeneity. The district has a range of mountainous and boreal forest areas dispersed with water bodies and man-made features such as roads and hydro-power dams.

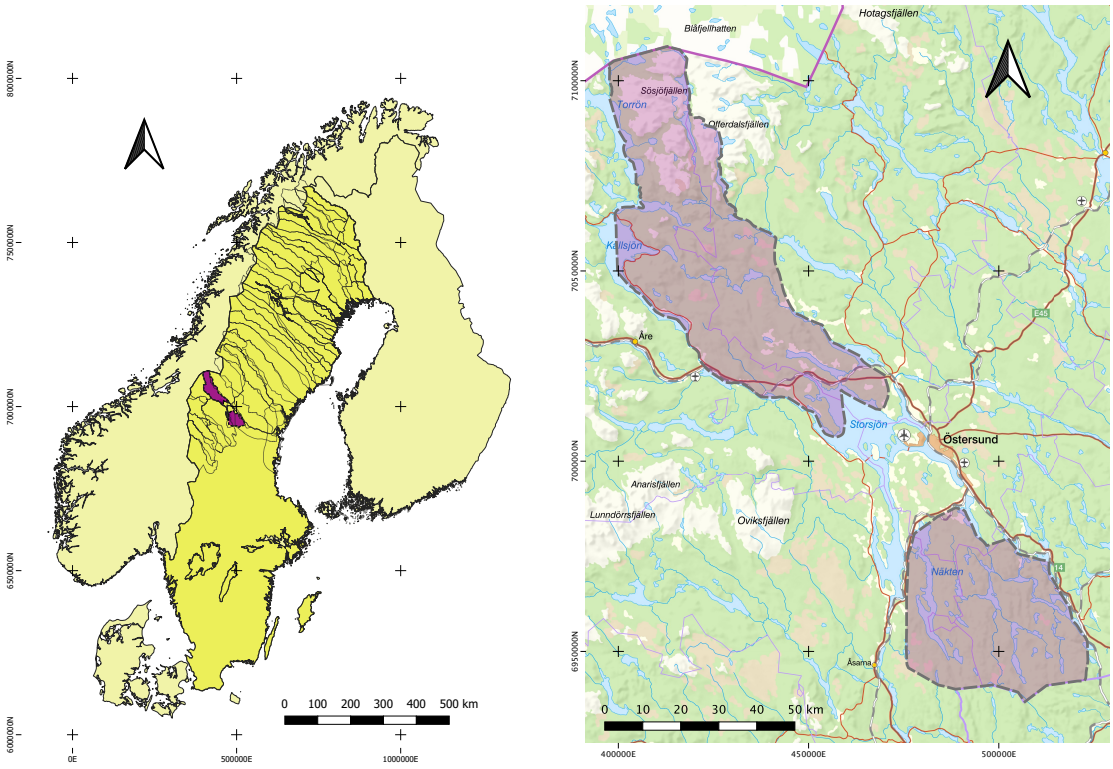


Fig. 4.1 Left: A map of Norway, Sweden and Finland from left to right. Sweden is marked in a darker shade of yellow. The perimeters of the reindeer herding districts are indicated by the black lines and the Njaarke herding district is highlighted by the purple polygon. Right: A zoomed view of Njaarke herding district. The top left polygon denotes the area of land utilised all year by reindeer and the bottom right indicates the winter only pasture.

## 4.3 Application of Non-switching Models

The remainder of this chapter is dedicated to illustrating an application of the various group movement models presented so far with the data set described in Section 4.2. We initially apply the non-switching model presented in Niu et al. (2016) and reviewed in Section 3.2. We then fit the model in the case where $\beta = 0$ and compare the results. The chapter is concluded with the application of the switching model given in Chapter 3 to the same data set. The results and discussion are given at the end of the chapter.

### 4.3.1 Application of Non-Switching Model where $\beta \neq 0$

We apply the existing non-switching and stationary model presented in Niu et al. (2016) and reviewed in Section 3.2 to the real location data described in Section 4.2. Here $\beta$ is nonzero meaning that the leader has a point of attraction given by the variable $\theta$, which we recall is given as a state of the system. Table 4.1 shows posterior means and standard deviations for the parameters of the model. The posterior distribution for the model parameters are given in the density plots in Figure 4.2. The results here are based on 100,000 iterations of Markov chain Monte Carlo runs, with over-dispersed initial values, every tenth iteration being recorded after 30,000 iterations of burn-in. This took approximately 120 minutes to complete.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\theta^x$ | 34.35 | 5.72 |
| $\theta^y$ | 22.78 | 5.58 |
| $\alpha$ | 0.32 | 0.06 |
| $\beta$ | 0.24 | 0.10 |
| $\rho$ | 11.54 | 2.81 |
| $\sigma$ | 1.55 | 0.11 |

Table 4.1 Parameter estimates and standard deviations for model of group movement where $\beta \neq 0$.

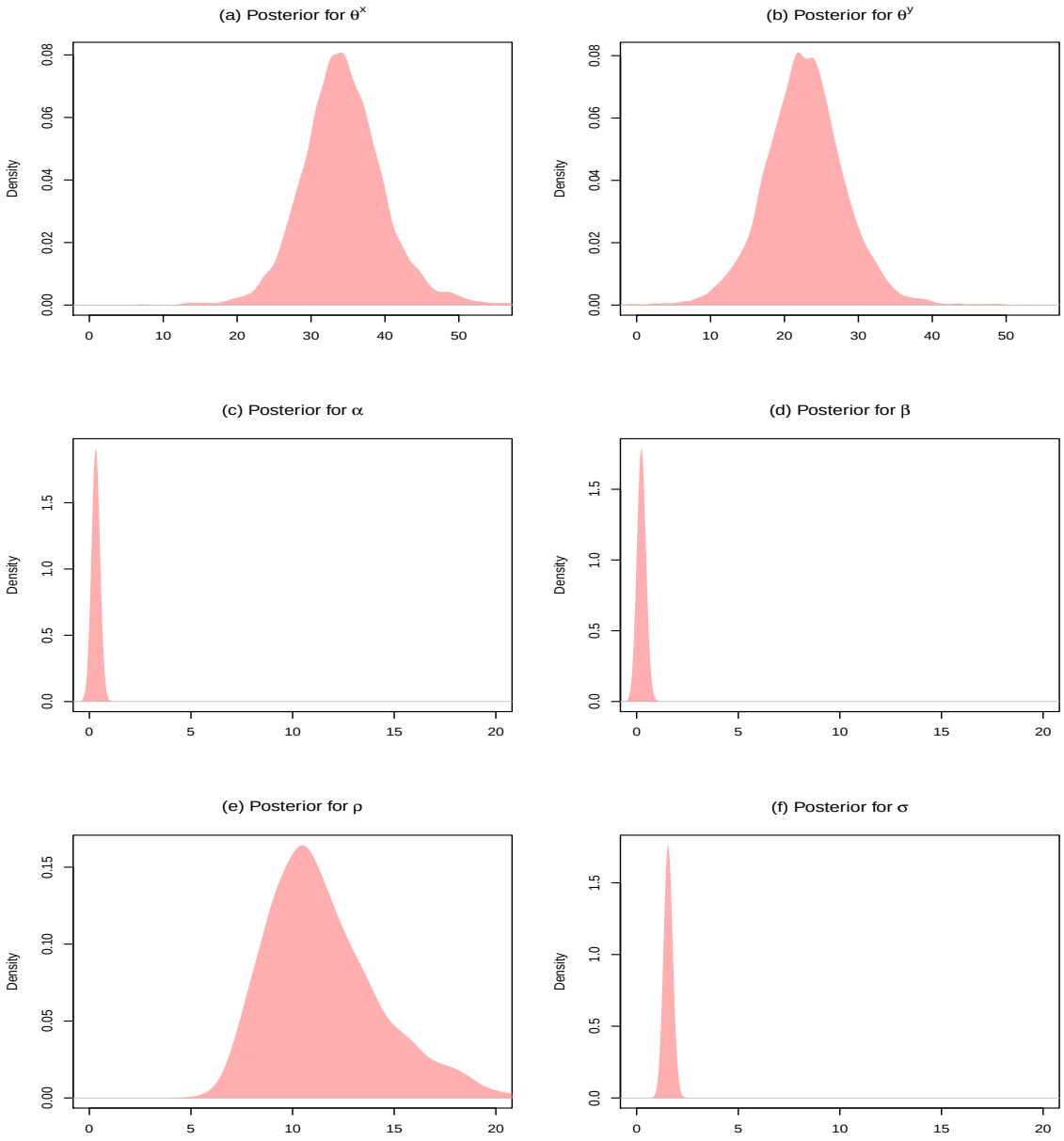


Fig. 4.2 Posterior densities for model parameters with real data, based on the MCMC runs of 100,000 iterations and 30,000 iterations of burn-in. (a) Posterior density for $\theta^x$, the $x$-coordinate of the attraction point for the leader. (b) Posterior density for $\theta^y$, the $y$-coordinate of the attraction point for the leader (c) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (d) Posterior density for $\beta$, the attraction rate of the leader to the attraction point. (e) Posterior density for $\rho$, the individual variance coefficient of the leading point. (f) Posterior density for $\sigma$, the individual variance coefficient of the follower.

### 4.3.2 Application of Non-Switching Model where $\beta = 0$

For comparison with the previous section we apply the non-stationary model presented in Section 3.3 where $\beta = 0$ with the same real location data given in Section 4.2. Table 4.2 shows posterior means and standard deviations for the parameters of the model. The density plots of the posterior distribution of the model parameters are shown in Figure 4.3. The results here are based on 25,000 iterations of Markov chain Monte Carlo runs fitting the switching non-stationary model, with over-dispersed initial values, every tenth iteration being recorded after 5,000 iterations of burn-in. This took approximately 60 minutes to complete.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.42 | 0.06 |
| $\rho$ | 7.70 | 1.40 |
| $\sigma$ | 1.71 | 0.11 |

Table 4.2 Parameter estimates and standard deviations for model of group movement where $\beta = 0$. Note here that since $\beta = 0$ the parameters $\theta^x$ and $\theta^y$ are not relevant to the model and thus their values are absent from the table.

.

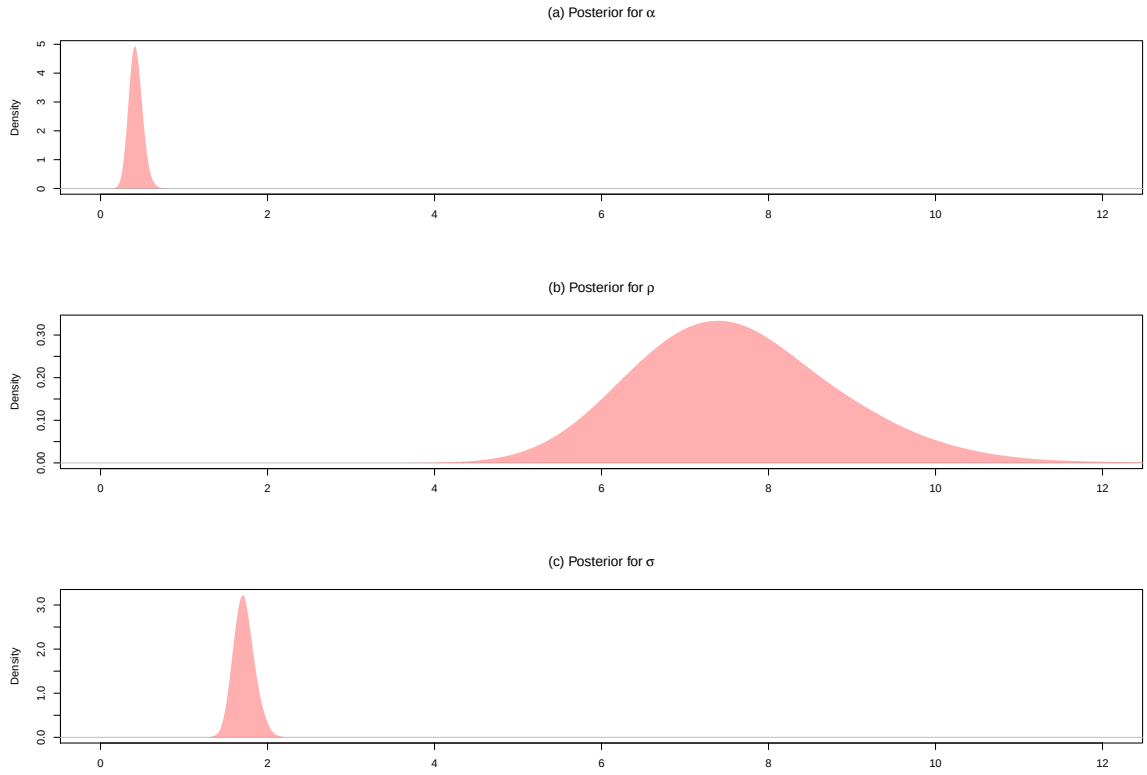


Fig. 4.3 Posterior densities for model parameters with real data, based on the MCMC runs of 25,000 iterations after a burn-in period of 5,000 iterations. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the individual variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower.

### 4.3.3 Comparison of Results

The most noticeable change in the parameter estimates are with $\rho$, the variance coefficient of the leader's location. In the $\beta \neq 0$ case the value is 11.54 whereas when $\beta = 0$ the value of $\rho$ is only 7.70. There is only a small change in the point estimate of $\sigma$ from 1.55 in the stationary case to 1.71 in the non-stationary case. The existing model tries to pinpoint the leader's attraction point $\theta$ and attraction rate to that point $\beta$, which as we discussed may not be appropriate on the time scale of the data. Here our data duration is only a few days and exploratory data analysis suggests that there is no strong attraction point. The consequence of fitting the existing model in this case is that we estimate a larger variance coefficient of the leaders location.

As with all models, there is a trade-off between the complexity of the model and goodness of fit. It is worth noting here that the non-stationary model required fewer iterations to fit and thus was computationally quicker, taking roughly 50% of the time taken to fit the stationary model. This, in part is due to having three fewer parameters to estimate. This experiment is worth bearing in mind, especially for data where there is no assumed point of attraction for the leader.

## 4.4 Application of Switching Model

This section is focussed on an application of the switching model presented in Section 3.4. We use the same subset of real reindeer location data. We illustrate the model's ability to capture behavioural heterogeneity. The duration of the data is 5 days so it is expected that the reindeer will display multiple movement processes in this time. Specifically, we expect the reindeer to switch between following the group and a Brownian motion behaviour.

Table 4.3 shows posterior means and standard deviations for the parameters of the model. The density plots of the posterior distribution of the model parameters are shown in Figure 4.4 whilst the posterior state estimations are shown in Figure 4.5. The results here are based on 100,000 iterations of Markov chain Monte Carlo runs fitting the switching non-stationary model, with over-dispersed initial values, every second iteration being recorded after 30,000 iterations of burn-in. This took approximately 4 hours to complete.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 1.33 | 0.24 |
| $\rho$ | 4.58 | 0.41 |
| $\sigma$ | 0.64 | 0.07 |
| $\sigma_{\mathrm{BM}}$ | 2.49 | 0.39 |
| $\lambda_{1,2}$ | 0.16 | 0.03 |
| $\lambda_{2,1}$ | 0.63 | 0.05 |

Table 4.3 Parameter estimates for reindeer movement and switching model with real dataset.

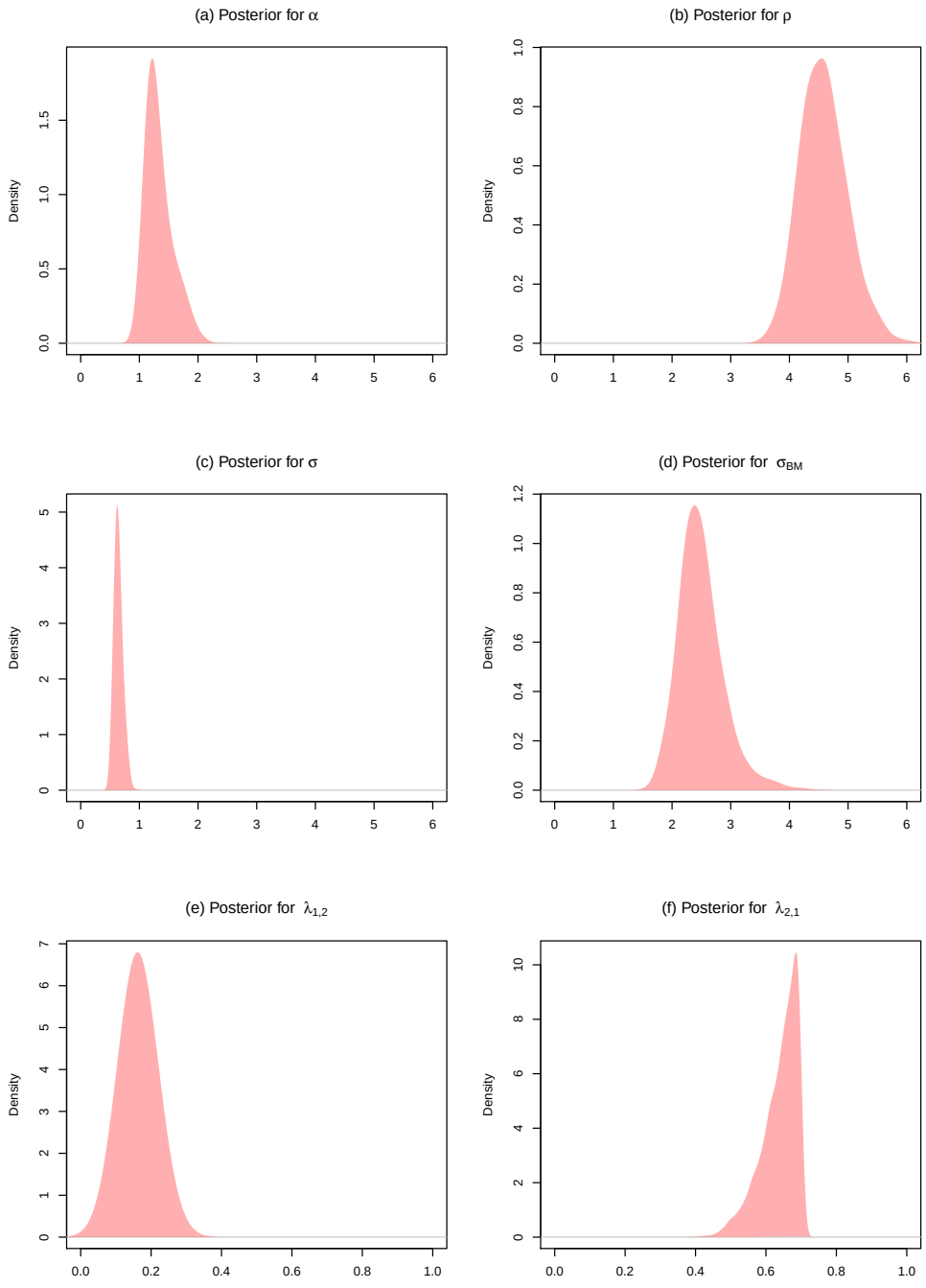


Fig. 4.4 Posterior densities for model parameters with real data, based on the Markov chain Monte Carlo runs of 100,0000 iterations. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (e) Posterior density for $\lambda_{1,2}$, the switching rate of the follower from OU to BM. (f) Posterior density of $\lambda_{2,1}$, the switching rate of the follower from BM to OU.

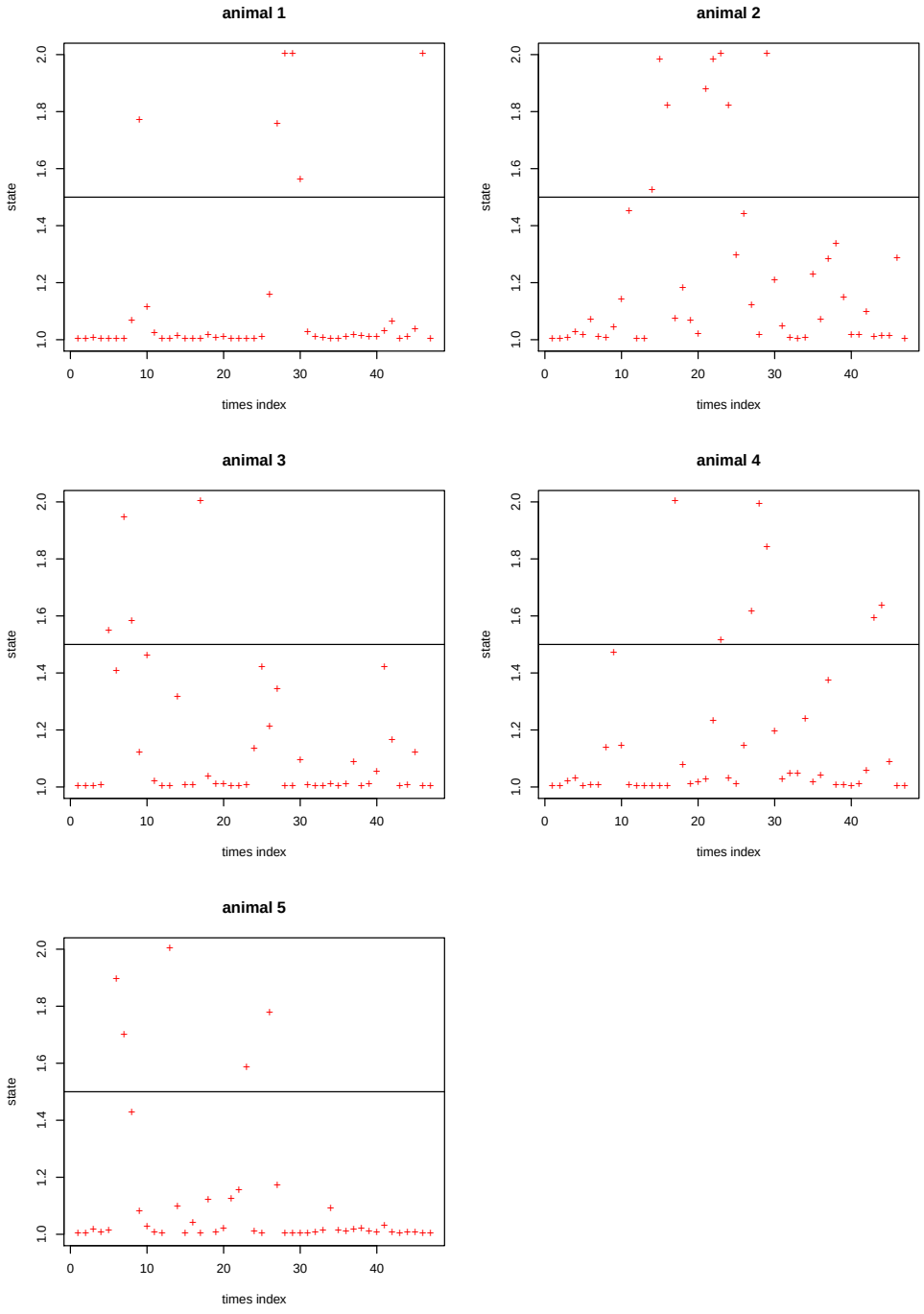


Fig. 4.5 Posterior mean states of all followers for the real data set. The red crosses represent the mean posterior of the estimated behaviour states. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The horizontal line is set at $y = 1.5$.

To visualise the grouping dynamics on a spatial scale Figure 4.6 shows the locations and linearly interpolated trajectories of each animal with the corresponding posterior state esti-

mations. At each location the shape and colour of points indicate whether the individual's posterior state is OU or BM. The orange square points indicate an BM state whilst the purple circular points indicate OU states. The classification of the points are determined by a threshold of 1.5 i.e. if the point estimate of the behavioural state at a particular time point is larger than 1.5 the point is classified as Brownian motion, otherwise it is Ornstein Uhlenbeck. Experimentation was done with a less strict threshold to account for an uncertain category (between 1.4 and 1.6) but this had limited counts as most estimates are confident.

What's more, trajectories may be overlaid on landscape maps to investigate grouping patterns in response to environmental cues such as terrain. Figure 4.7 gives an example of posterior state estimations and location data plotted on a map of the environment indexed by four land cover types namely, anthropogenic areas, mires, forests and water bodies.

Alternatively, if the practitioner is concerned with the behaviour of the reindeer on a temporal scale similar plots can be made by plotting the locations in 1-dimension, say the $y$-coordinate, against time for each individual as in Figure 4.8.

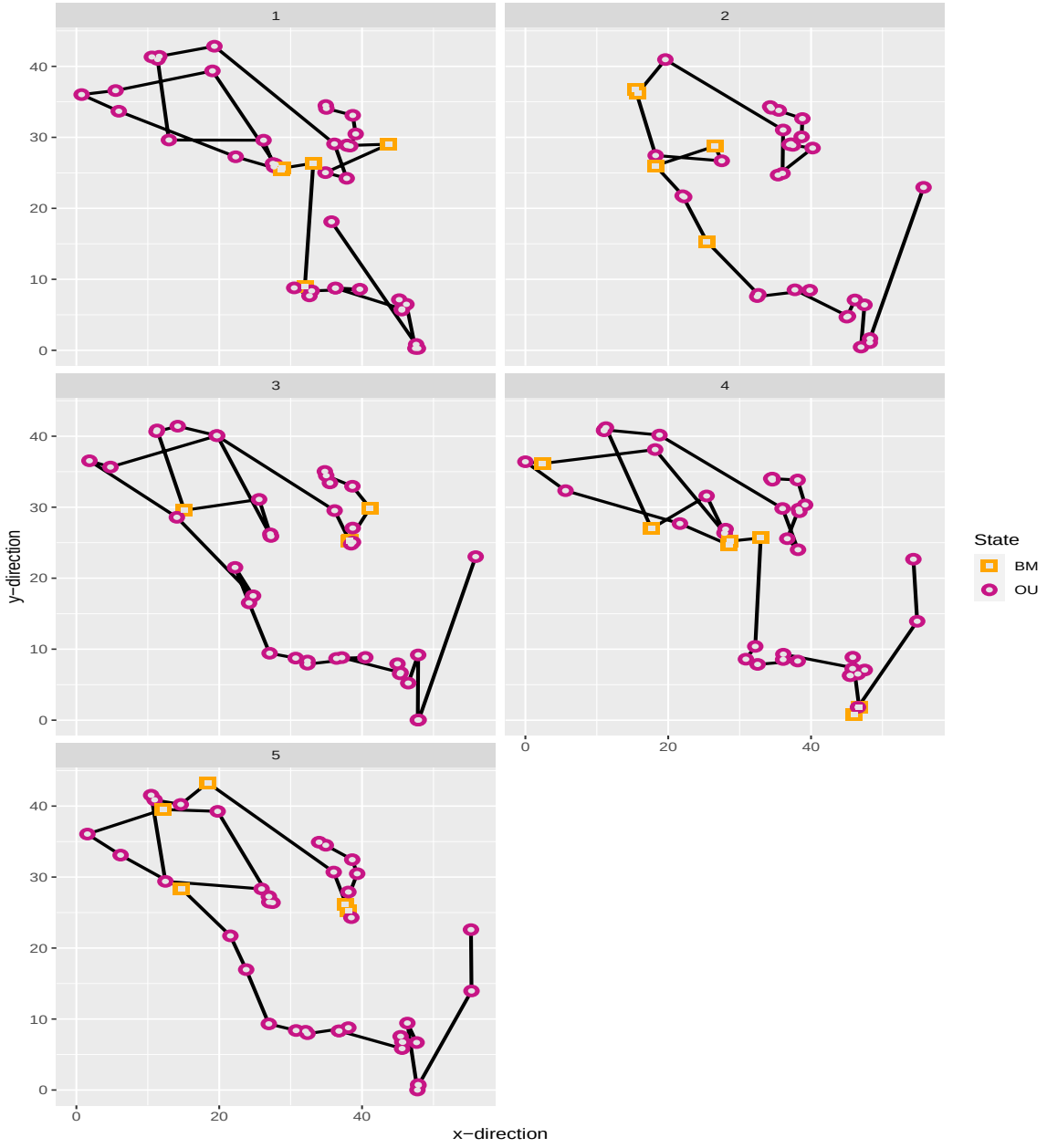


Fig. 4.6 Trajectories for each animal in the real dataset. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange square points indicate an BM state whilst the purple circular points indicate OU states.

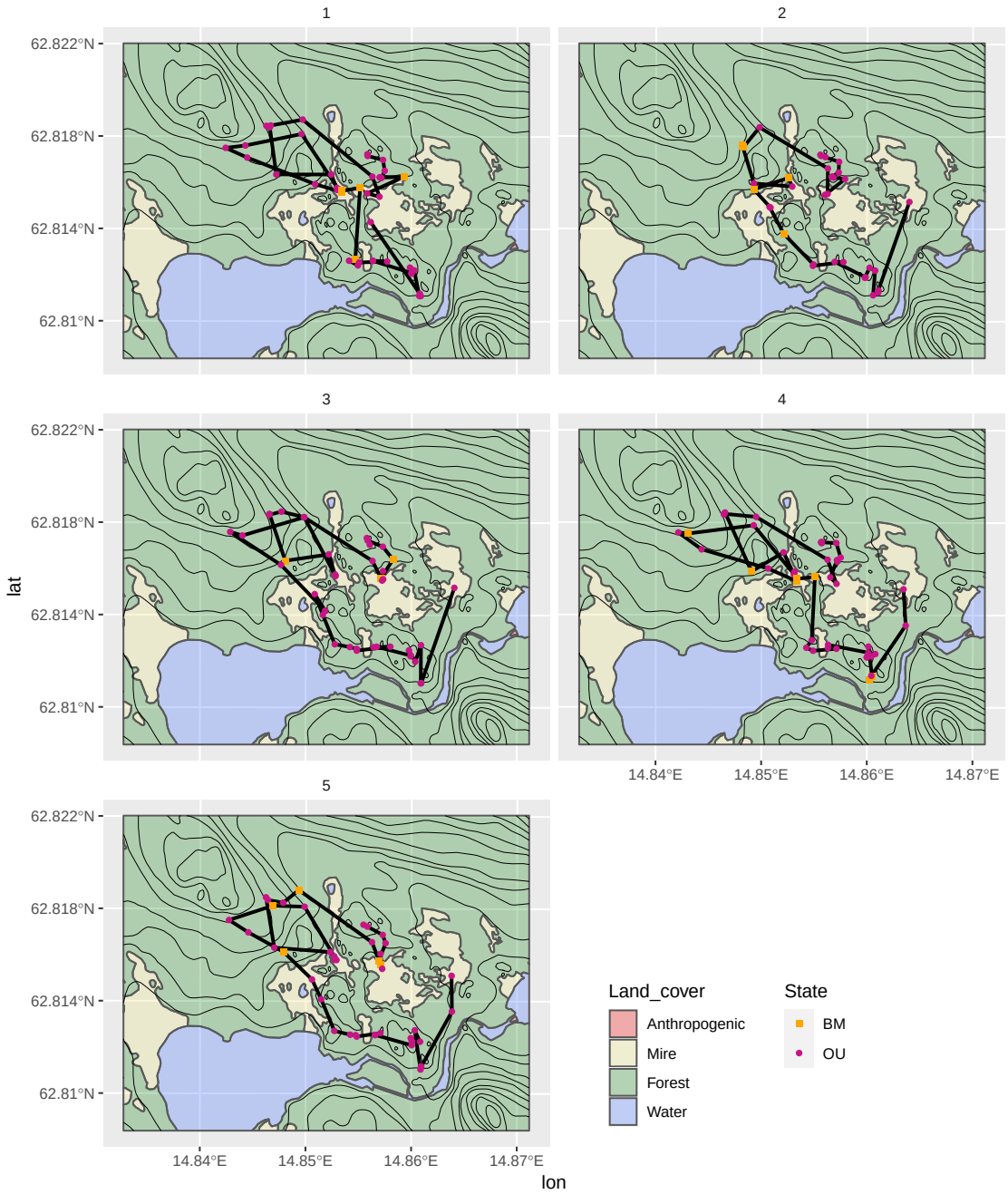


Fig. 4.7 Trajectories for each animal in the real data set projected on to a terrain map. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange square points indicate an BM state whilst the purple circular points indicate OU states. The terrain is split into four categories: anthropogenic, water body, mire and forest given in red, blue, tan and green respectively.

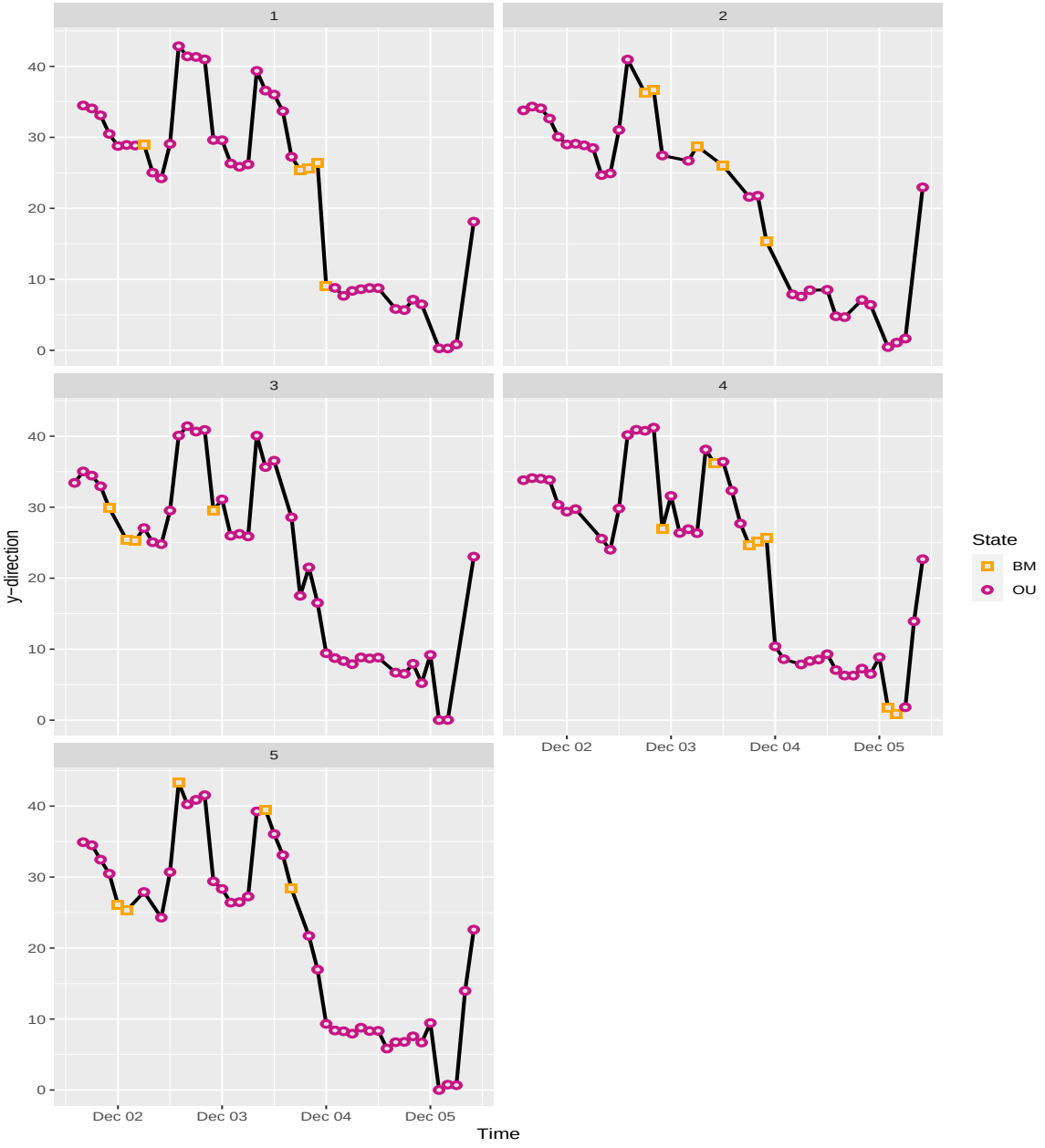


Fig. 4.8 Time trace of locations in the $y$-direction for each animal from the real reindeer data. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange square points indicate an BM state whilst the purple circular points indicate OU states.

The posterior mean of the non-switching variance coefficient of the leader $\rho$ is 7.7 which is much bigger than the parameter in the switching case which is only 4.58. Similar comparison can be seen in the attraction rate parameter $\alpha$, which is 0.42 for the non-switching case and 1.33 for the switching case. First impressions of the results may be that the switching

model is redundant as the animals spend most of their time in grouped state. However, the non-switching model treats the independent movement of followers as the part of the group movement. This leads to the bigger estimated variance of the leader's location and smaller estimated attraction rate, while in the switching group movement model, we successfully distinguished the group movement and independent movement of the followers using the behavioural states. In a simulated example in Chapter 6, the necessity of switching becomes much more apparent as the animals switch behaviour regularly with a sinusoidal pattern.

## 4.5 Discussion

Throughout this chapter we have highlighted the importance of modelling the movement of reindeer, in particular those which are semi-domesticated. We have stressed the significance of particular grouping strategies and their impact on herd survival. For example, although an individual reindeer may reduce its grazing competition by moving away from the herd, it then also stands a greater chance of being killed by predators or, in summer, being harassed by insects, and therefore the choice an individual reindeer makes about how and where to move is balanced between finding enough food for itself but also staying within the safety of the group (Hart and Mooring, 1992). In winter reindeer usually graze in groups digging for lichens underneath the snow. Staying with a group where several animals are digging could be beneficial for the individual reindeer as this saves time and energy from digging. However, this also means competition among the animals for the best lichen forage and individuals may be pushed away and thus need to search for new places to dig (Kojola, 1989).

In Chapter 5 we apply the switching model to two datasets during times which we assume to have contrasting environmental conditions. Specifically, we investigate reindeer grouping strategies for the relief of harassment from mosquitoes in the summer months. We use this application as a case study to demonstrate the model's ability to distinguish between grouping and dispersal behaviour. Using similar visualisation tools we are able highlight preferences of land cover type during assumed presence of mosquitoes. This simple example promotes the functionality of this switching model and the important applications for the conservation and protection of the reindeer herds and the herders who rely on them for their cultural and economic value.

By using the same collection of real reindeer location observations throughout this chapter we are able to understand the differences in parameters estimation between the modelling approaches presented in Chapter 3. We have acknowledged that in cases where it is not

suitable to assume a point of attraction for the leader then the non-stationary model, where $\beta = 0$ is more appropriate and often computationally quicker. However, unless the temporal scale of the data is extremely short it is likely that individual's will exhibit a variety of behavioural modes. In these instances the switching model we have developed is highly applicable and the data visualisation techniques we have demonstrated can be useful for discerning behavioural patterns on a temporal or spatial scale.

Through our exploratory data analysis and literature review we believe that the modelling approach is an appropriate fit for the collective structure in reindeer movement. Although the true diffusion parameters and the behaviour states are unknown we can have confidence about our inferences since the real data is thought to approximately come from the distribution we are fitting and the model performed well at recovering the true values in the simulation experiments with extreme behaviour given in Section 3.8.2, 3.8.3. More specifically, the model provided a good fit when true values were based on the point estimates of the real data in Section 3.8.4.

However, this real example highlights a limitation of having a single leader. In Figure 4.8 we can see time periods in which multiple animals switch to the independent Brownian motion state and move in a similar way e.g animal 4 and 5 around the 2nd December and animal 1 and 4 immediately before the 4th December. Whilst the model presented here treats the whole group as one, at these particular instances it may have been more appropriate to model movement with multiple leaders. However, this is beyond the scope of this thesis and to my knowledge has not been done.

Additionally, this model does not consider intra-group interactions and assumes that the group has no hierarchical structure. In this case, the approach to modelling movement and attraction would be more intricate. Milner et al. (2020) demonstrate this by providing a model of movement for a number of animals within a social hierarchy. They too build on previous work of Niu et al. (2016), using multiple behavioural states with multivariate Ornstein Uhlenbeck diffusion processes to model movement. The main difference of their work is that there is no imputed leader. That is, the leading animal has to be a real, tagged individual rather than an abstraction. As well as this, rather than having a single leader as in our case, the authors allow for attraction between every tagged animal by modelling dyadic interactions. Whilst this novel research is very useful, they acknowledge that this approach is computationally expensive and is limited by the requirement that all individuals within the group must be tagged. There are advantages and disadvantages of the work presented in this thesis and that of Milner et al. (2020) but it is worth noting that ultimately they are built for different purposes. The work in this thesis is designed to pick up on larger scale

group dynamics where there are no defined leaders and which in principle is designed to be invariant on which animals within the group are tagged, whereas Milner et al. (2020) aim to recognise fine scale social interactions amongst known tagged animals.

Another limitation is the assumption of homogeneous switching rates. That is, the switching rates are constant throughout time and space. In reality, grouping or dispersing behaviours are often a result of internal or external stimuli for example, predators, mating seasons, forage competition, hunger or fatigue may drive behavioural choices. In Chapter 5 we discuss how the movement of reindeer may be influenced by the presences of parasitic insects such as mosquitoes. We demonstrate this by comparing grouping behaviour for two different time periods under contrasting levels of insects. In Chpater 6 we formalise this inclusion of covariate information by providing a flexible framework for allowing switching rates to vary temporally. We firstly illustrate this with a simulated example where the switching rates depend on the time of day. Then we revisit our real data application by explicitly allowing the switching rates to depend on covariate information about insect presence.

# Chapter 5

# Reindeer Grouping Strategies for the Relief of Insect Harassment

The purpose of this chapter is to apply the switching model presented in Section 3.4 to two datasets of reindeer locations assumed to be during contrasting environmental settings. We investigate whether the model is able to distinguish between different behavioural patterns and give results consistent with the knowledge of reindeer movement. Specifically, we explore the differences in reindeer grouping strategies in response to the presence or absence of parasitic insects such as mosquitoes and oestrid flies.

Ecosystems inhabited by reindeer are known for large concentrations of parasitic insects such as mosquitoes (*Aedes* sp., Culicidae), horseflies (Tabanidae), blackflies (Simuliidae) and oestrid flies for example, warble and nasal bot flies (*Hypoderma tarandi* L., Oestridae and *Cephenemyia trompe* L., Oestridae respectively) (Hagemoen and Reimers, 2002; Witter et al., 2012b).

The main parasitic species present in the Njaarke herding district described in Section 4.2 are oestrid flies and mosquitoes, so we shall focus our attention toward these. These insects are known to cause discomfort and annoyance to reindeer, forcing relocation and interrupting foraging behaviour. Mosquitoes in particular bite their hosts to feed on their blood whilst both species of oestrid flies use reindeer as hosts for reproduction. Both oestrid fly species have similar life-cycle durations of 1 year. Warble flies leave their hosts as larvae in May-June where they then drop to the ground to burrow and pupate. After a few weeks they emerge as flies to mate. After, the female flies seek reindeer to oviposit on their hide, where the larvae subsequently hatch, penetrate the skin and develop. Nasal bot flies have a similar life-cycle but the female does not oviposit onto the reindeer. Instead she allows the eggs to hatch inside

her then sprays the newly hatched larvae into the muzzle of the reindeer. They then crawl into the nasal cavities of the host where they develop until they are later coughed, sneezed or breathed out (Tryland and Kutz, 2018).

The presence of such insects are thought to be a major factor in shaping reindeer movement, habitat selection and grouping dynamics (Downes et al., 1986; Helle and Aspi, 1983; Skarin et al., 2008; White et al., 1975) which can severely influence reindeer activity budgets. Flying parasites force reindeer to seek topographical relief in more exposed areas at higher altitudes where flying conditions are poorer (Downes et al., 1986); this usually results in decreased foraging opportunity (Morschel and Klein, 1997; Skarin et al., 2010). Not only this, insect harassment can lead to temporary clustering and stationarity in the herd which can result in local overgrazing. Overgrazing may decrease the quality of vegetation, possibly with long term consequences (White et al., 1975).

The increased energy expenditure due to avoidance behaviour, decreased resting time and poorer quality grazing can contribute to nutritional deficits (White et al., 1975). Initially, it was hypothesized that the animals would compensate for this during hours of low insect activity, however, Colman et al. (2003) provided evidence against this. This can be a crucial factor impacting reindeer survival, especially during insect activity peaks around mid-June to the end of August (Skarin et al., 2008; Witter et al., 2012a).

The summer period is a critical time for reindeer to forage and build fat reserves for the winter; it is suggested to be of paramount importance for survival during harsh winters with low forage opportunity (Reimers, 1997; Tryland and Kutz, 2018). The accumulation of daily harassment can lead to decreased body weight which is thought to be strongly correlated with reproductive success, winter survival and calf recruitment (Colman et al., 2003). Moreover, during the summer, reindeer need to account for extra energy expenditure for the rutting season and calving. Any constraints on foraging can have severe detrimental effects on individual reindeer and consequently the population density (Morschel and Klein, 1997).

The behavioural shifts to alleviate parasitic pressure are not limited to small scale movements or micro-habitat relocations. Some reindeer choose large scale migrations to alternative summer pastures after calving to reduce levels of parasite infections. In the study of Folstad et al. (1991), it was shown that the larval abundance of oestrids is negatively correlated with distance between their summer pasture and calving ground suggesting that some reindeer may elect post-calve migration as a means of insect relief.

Both the direct and indirect costs of insect harassment are thought to affect reindeer well-being and overall fitness of the herd (Hagemoen and Reimers, 2002). In extreme cases

reindeer populations have declined in certain regions (Vors and Boyce, 2009). It is postulated that declines in these populations are related to insect harassment, either directly through blood loss and infection or indirectly due to the energetic expenditure related to insect avoidance and disturbed resting cycles (Hagemoen and Reimers, 2002; Morschel and Klein, 1997).

In alignment with the wisdom of herding communities, past observation analyses found that distance between individuals increased with the presence of oestrids and decreased with the presence of only mosquitoes (Morschel and Klein, 1997); group sizes increased with mosquito activity, decreased with both mosquitoes and oestrids and then decreased again during the absence of insects (Downes et al., 1986). Under severe mosquito harassment, the large, tightly grouped congregations make long and rapid movements (White et al., 1975). This suggests that large clumping behaviour can be a strategy of mosquito avoidance whereas dispersion relieves oestrid harassment. This dispersion has been attributed to panicked running in response to oestrids. Helle and Aspi (1983) and White et al. (1975) both give evidence of the advantageousness of grouping behaviour during an insect activity field experiment. They concluded that periphery insect traps were attacked less frequently than central ones meaning that harassment was dependent on an individual's position within herd.

It has been disputed as to which insect has the most influence over reindeer behaviour and grouping dynamics. Witter et al. (2012a) argues that oestrid flies have the most dominance over reindeer behaviour, although the presence of both mosquitoes and oestrids has the greatest effect.

Whilst most studies have observed animals with spotting scopes, which can be laborious and costly, we present an application of the switching model given in Section 3.4 to infer grouping behaviours of reindeer in response to insect harassment using location data alone. We will demonstrate the model's ability to discern varying behavioural patterns between two time periods; one where there is thought to be a high level of insect presence and another where it is thought to be low.

## 5.1 Climatic Conditions for Insect Presence

Often location data of reindeer, from a GPS tag, is collected in the absence of insect activity observations. That said, providing retrospective inference calls for the use of data which can serve as a proxy for insect harassment. A common approach is to use climatic data, which can usually be obtained from nearby weather stations, as a way to index insect activity.

Many studies provide justification for using weather parameters to index insect harassment (Morschel and Klein, 1997; Skarin et al., 2010; Witter et al., 2012b). Past studies have used climatic data and observations of reindeer behaviour, such as foot stomping, nose dropping and tail flicking, which is thought to represent annoyance behaviour, in order to determine weather characteristics with the most influence on insect activity. In particular, temperature and wind speed appear to be the dominant variables affecting insect activity (Downes et al., 1986; Hagemoen and Reimers, 2002; Morschel and Klein, 1997; White et al., 1975). During model selection processes using Akaike Information Criterion (AIC), it was found that models for insect activity which included only temperature and wind speed came second only to a full combination of other less measurable covariates such as light, vegetation, topography (Witter et al., 2012b).

The suggested optimal weather conditions for insect activity varies between species and amongst studies but generally warm temperature and low wind speed are preferred. Some suggest that oestrid fly activity occurs above a temperature threshold of 7°C (Downes et al., 1986), 10°C (Anderson et al., 1994; Mörschel, 1999), or 13°C (White et al., 1975), whilst thresholds of wind speed are reported as less than 6-8 m/s (Anderson et al., 1994). Mosquitoes have been reported to be active at temperatures between 6 and 18°C and wind speeds below 6 m/s (Russell et al., 1993) and to decline in activity above 16°C and 7.5 m/s (Hagemoen and Reimers, 2002).

Past studies hypothesised that the direct effects of weather dominate reindeer behaviour patterns as opposed to the indirect consequences of weather such as insect presence. However, data-driven analyses, with the simultaneous recording of insect abundance and reindeer behaviour, suggest that models of reindeer behaviour that comprised weather variables only did not perform as well as those which also contained covariates related to insect activity, thus suggesting that the indirect effects were larger than direct effects of weather (Hagemoen and Reimers, 2002; Witter et al., 2012a).

## 5.2 Data

To investigate reindeer grouping dynamics in response to varying levels of insect harassment, we fit the switching model described in Section 3.4 to two datasets. Specifically, we focus on two time periods where there is thought to be either a low or high mosquito presence. Both data sets are subsets of the data set described in Section 4.2. The first subset occurs during a time period in early summer where we assume there is low mosquito activity. The

other dataset has been selected in the peak summer period when we believe there are high mosquito levels.

The time period of the datasets were selected through exploratory data analysis using the thresholds for mosquito activity given in Skarin et al. (2010) i.e. mosquitoes are assumed to be mainly active at temperatures between 7 and 17°C and at wind speeds below 7 m/s. The weather data was downloaded from the Swedish Meteorological and Hydrological Institute (http://www.smhi.se/). We chose to collect data from the weather station Korsvattnet A in the north eastern part of the Njaarke district, the closest to the reindeer locations. The observations of temperature and wind speed are recorded hourly and there are no missing data.

The low mosquito activity dataset, which will now be referred to as *early summer*, consists of up to 70 observations of 7 individuals, at two hour intervals between 20/06/2010 and 25/06/2010. A plot of the mosquito activity index for the duration of the early summer data set is given in Figure 5.1. A value of 1 indicates that the wind speed and air temperature are within the thresholds of mosquito activity that are presented in Skarin et al. (2010); a value of 0 indicates that the wind speed and temperature are outside the thresholds and thus mosquitoes are not expected to be present.

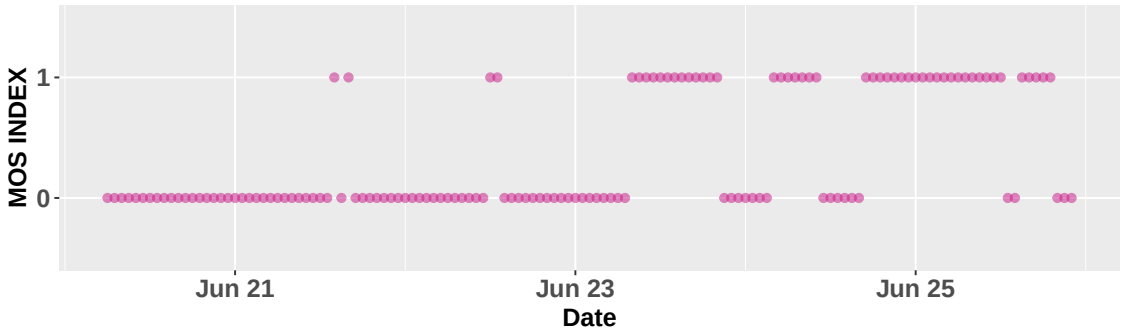


Fig. 5.1 Timeline of mosquito index during the early summer period. A value of 1 indicates that the wind speed and air temperature are within the thresholds of mosquito activity that are presented in Skarin et al. (2010); a value of 0 indicates that the wind speed and temperature are outside the thresholds and thus mosquitoes are not expected to be present.

Despite there being an increased frequency in expected mosquito presence after June 23rd, we have selected a time period up until June 25th. We have done this so as to not use too small a dataset in our analysis.

The high mosquito activity dataset, which will now be referred to as *peak summer*, consists of up to 83 observations of 6 individuals, at two hour intervals between 04/07/2010 and

10/07/2010. A plot of the mosquito activity index for the duration of the peak summer data set is given in Figure 5.2.

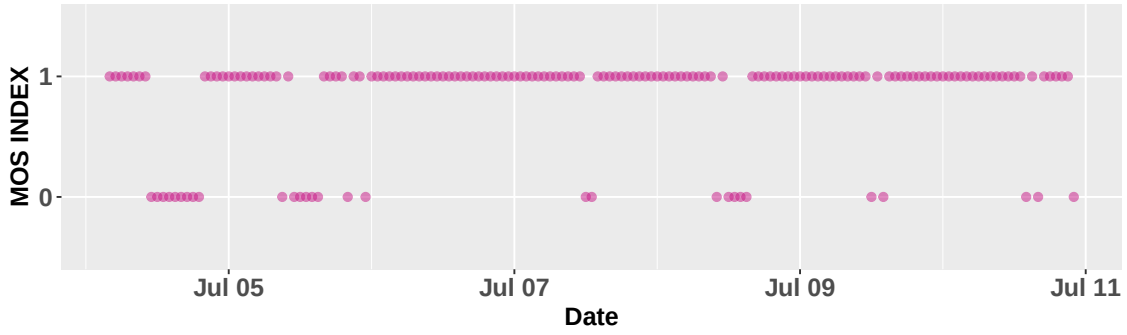


Fig. 5.2 Timeline of mosquito index during the peak summer period. A value of 1 indicates that the wind speed and air temperature are within the thresholds of mosquito activity that are presented in Skarin et al. (2010); a value of 0 indicates that the wind speed and temperature are outside the thresholds and thus mosquitoes are not expected to be present.

Similar to the early summer dataset, we retain observations between July 4th and July 6th despite the fluctuations in mosquito index during this time period.

As before, both datasets are subject to some of the usual irregularities when dealing with real data, i.e. missing values and irregular sampling schedule. However, the observations are almost regular insofar as they occur up to only 2 minutes before/after the intended timing. Thus, for the simplicity of implementation the time steps of the data were rounded to the nearest hour.

### 5.2.1 Identifiability and convergence issues with $\rho$

Before we discuss applying the model to the two datasets there is an important issue which needs addressing. During trial runs of the MCMC algorithm, we encountered difficulties with convergence for the posterior of the parameter $\rho$ insofar as it frequently diverged to large, unreasonable values. We experienced this issue with both the early and peak summer datasets. The wide range of possible values that $\rho$ may take may be considered as an identifiability problem with the parameter. We find that it is not always possible to distinguish between different movement scenarios e.g. a fast leader with low attraction rate (where $\rho$ is large and $\alpha$ is small) may look similar to a slow leader with high attraction ($\rho$ small and $\alpha$ large). To work around this, we fixed $\rho$ to a value we believed to be sensible whilst allowing it to

be large relative to the other Brownian motion parameters, essentially 'scaling' the leader's movement (for more justification, see discussion in Section 5.5). To do this we ran the algorithm with all the parameters unfixed then chose $\rho$ such that it was the least integer greater than twice the maximum of the other Brownian motion parameters, $\sigma$ and $\sigma_{BM}$. In other words, after a trial run, we subsequently choose $\rho$ such that

$$\rho = \left\lceil 2(\max\{\sigma, \sigma_{BM}\}) \right\rceil,$$

where $\lceil x \rceil$ is equal to the least integer greater than or equal to $x$. Using this process, we set $\rho = 3$ for both datasets.

## 5.3 Early Summer Results

We applied the switching model presented in Chapter 3 to the early summer dataset. We ran the Markov chain Monte Carlo algorithm for 50,000 iterations. This took approximately 12 hours to complete. The posterior mean and standard deviation of model parameters are shown in Table 5.1. Posterior state estimations are given in Figure 5.3 whilst posterior densities for the parameters are given in Figure 5.4. For each individual, Figure 5.5 shows the locations in the $y$-direction against time; Figure 5.6 gives the two-dimensional trajectories; Figures 5.7 and 5.8 project these trajectories on to a terrain map.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.047 | 0.01 |
| $\rho$ | 3 | 0 |
| $\sigma$ | 1.25 | 0.077 |
| $\sigma_{\mathrm{BM}}$ | 0.22 | 0.02 |
| $\lambda_{1,2}$ | 0.63 | 0.057 |
| $\lambda_{2,1}$ | 0.19 | 0.033 |

Table 5.1 Parameter estimates of the switching model for the early summer dataset where mosquito activity is assumed to be low.

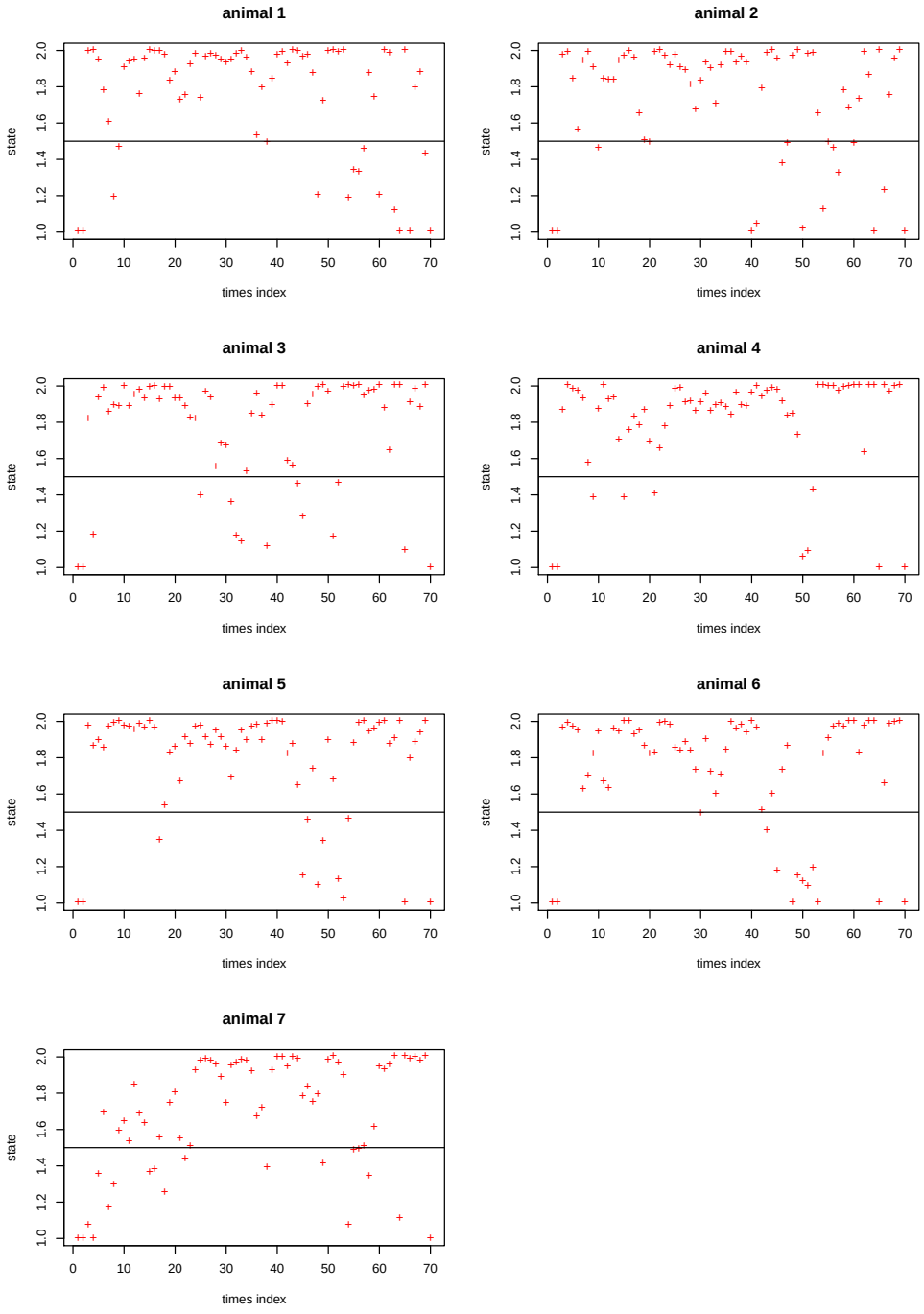


Fig. 5.3 Posterior mean states of all followers for the early summer dataset where mosquito activity is assumed to be low. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (red) represent the mean posterior of the estimated behaviour states.

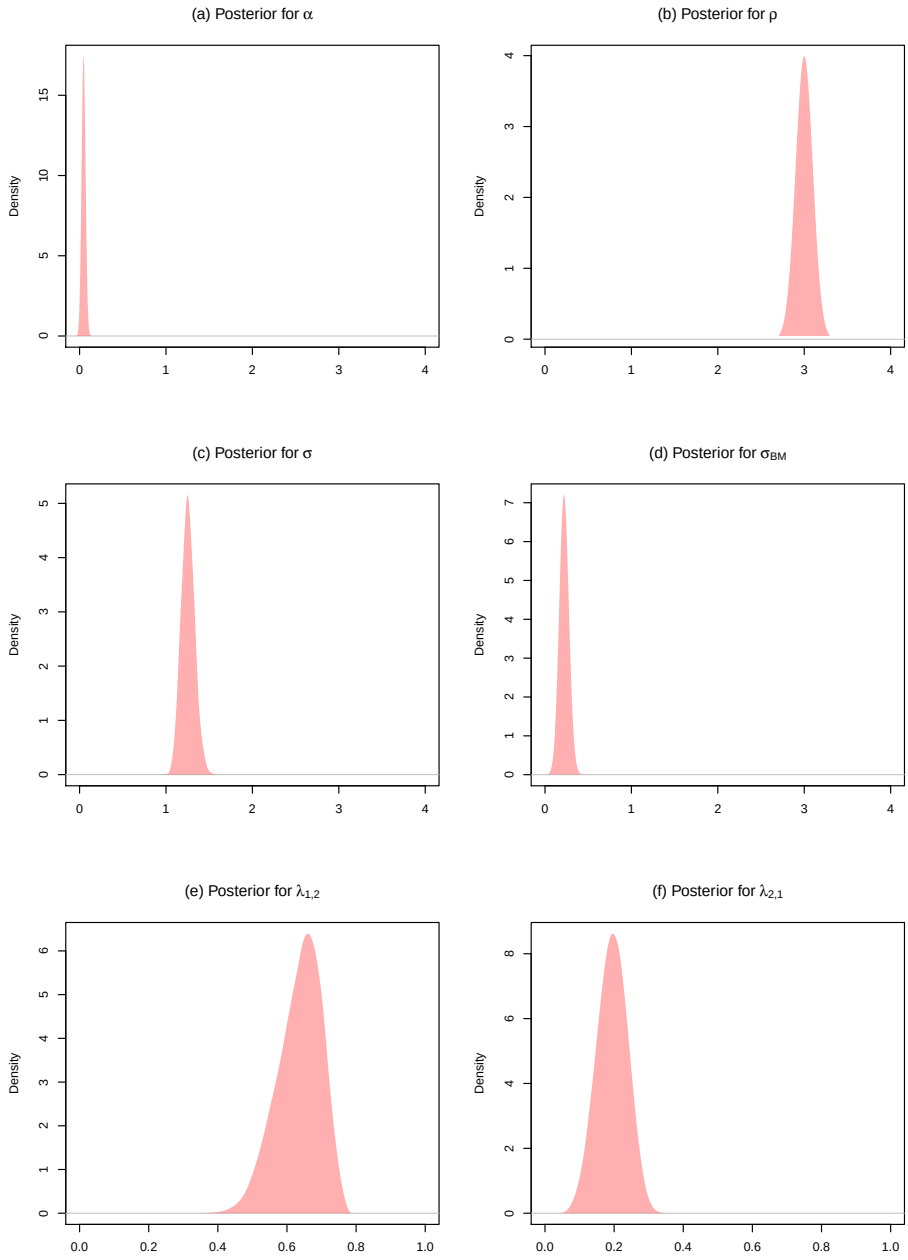


Fig. 5.4 Posterior densities for model parameters for the early summer dataset, based on the Markov chain Monte Carlo runs of 50,000 iterations. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Approximate density for $\rho$, the variance coefficient of the leading point, which is fixed at 3. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (e) Posterior density for $\lambda_{1,2}$, the switching rate of the follower from OU to BM. (f) Posterior density of $\lambda_{2,1}$, the switching rate of the follower from BM to OU.

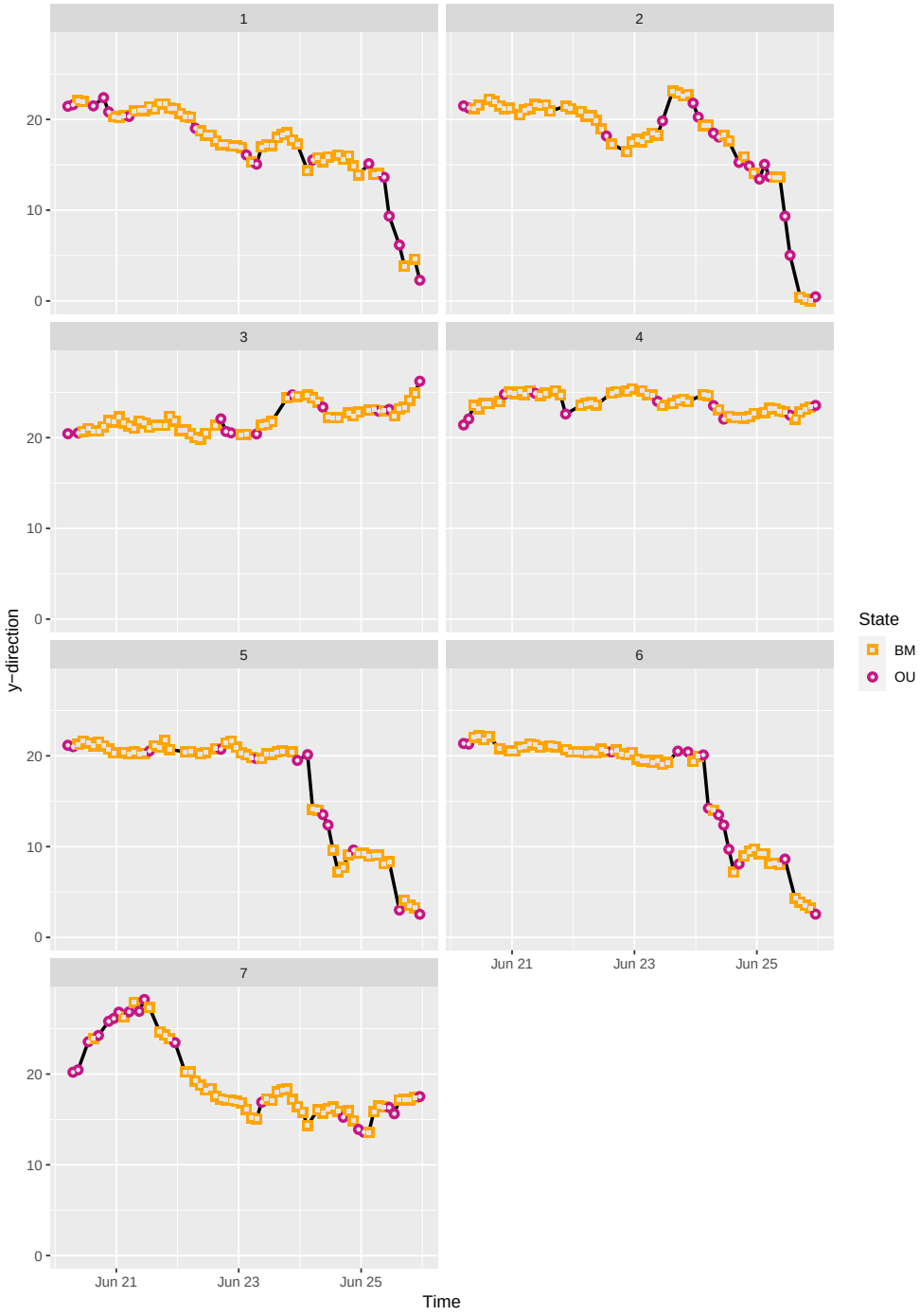


Fig. 5.5 Time trace of locations in the $y$-direction for each animal from the early summer dataset. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange squares indicate a BM state whilst the purple circles indicate an OU state.

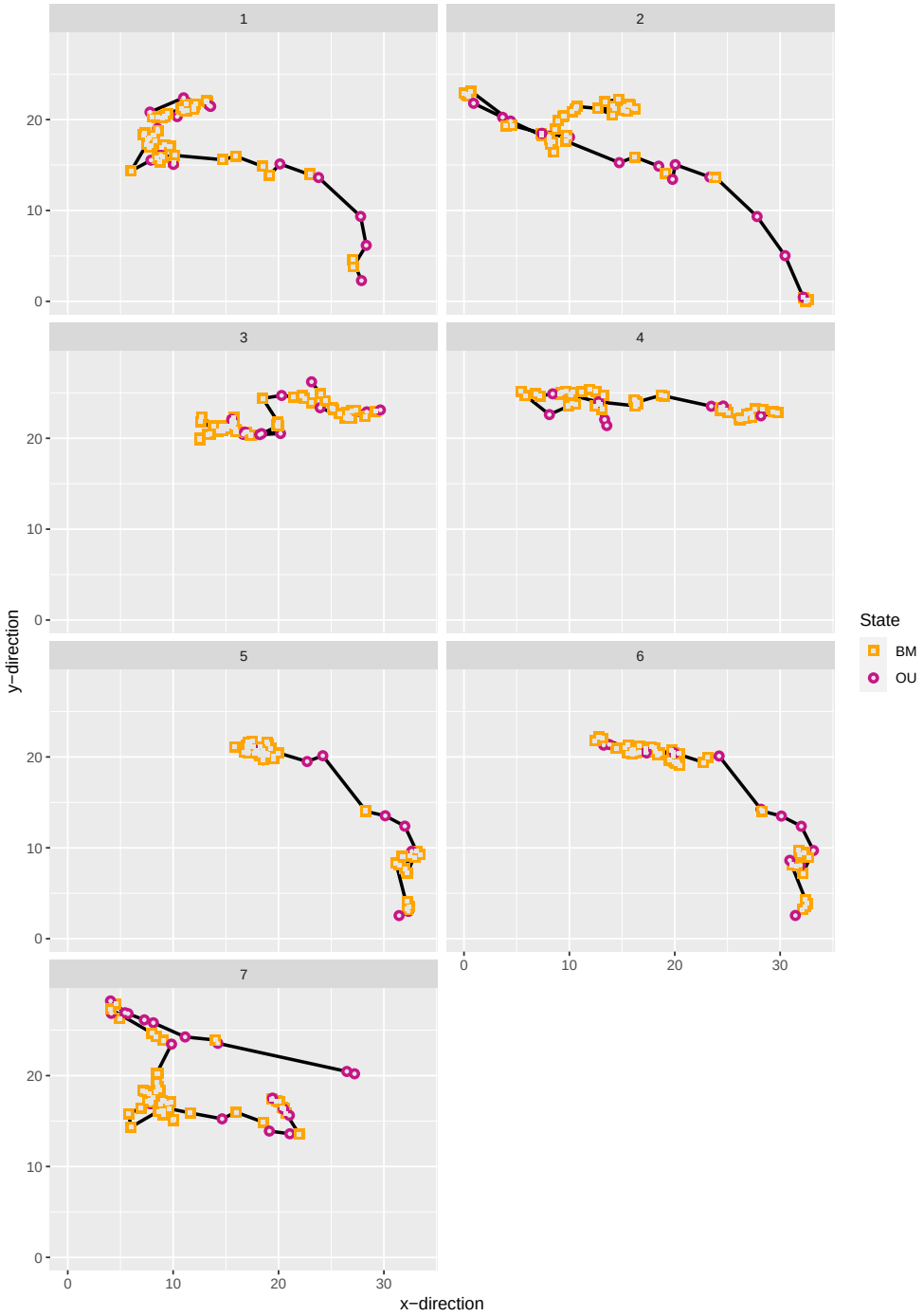


Fig. 5.6 Trajectories for each animal from the early summer dataset. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange squares indicate the BM state whilst the purple circles indicate OU states.

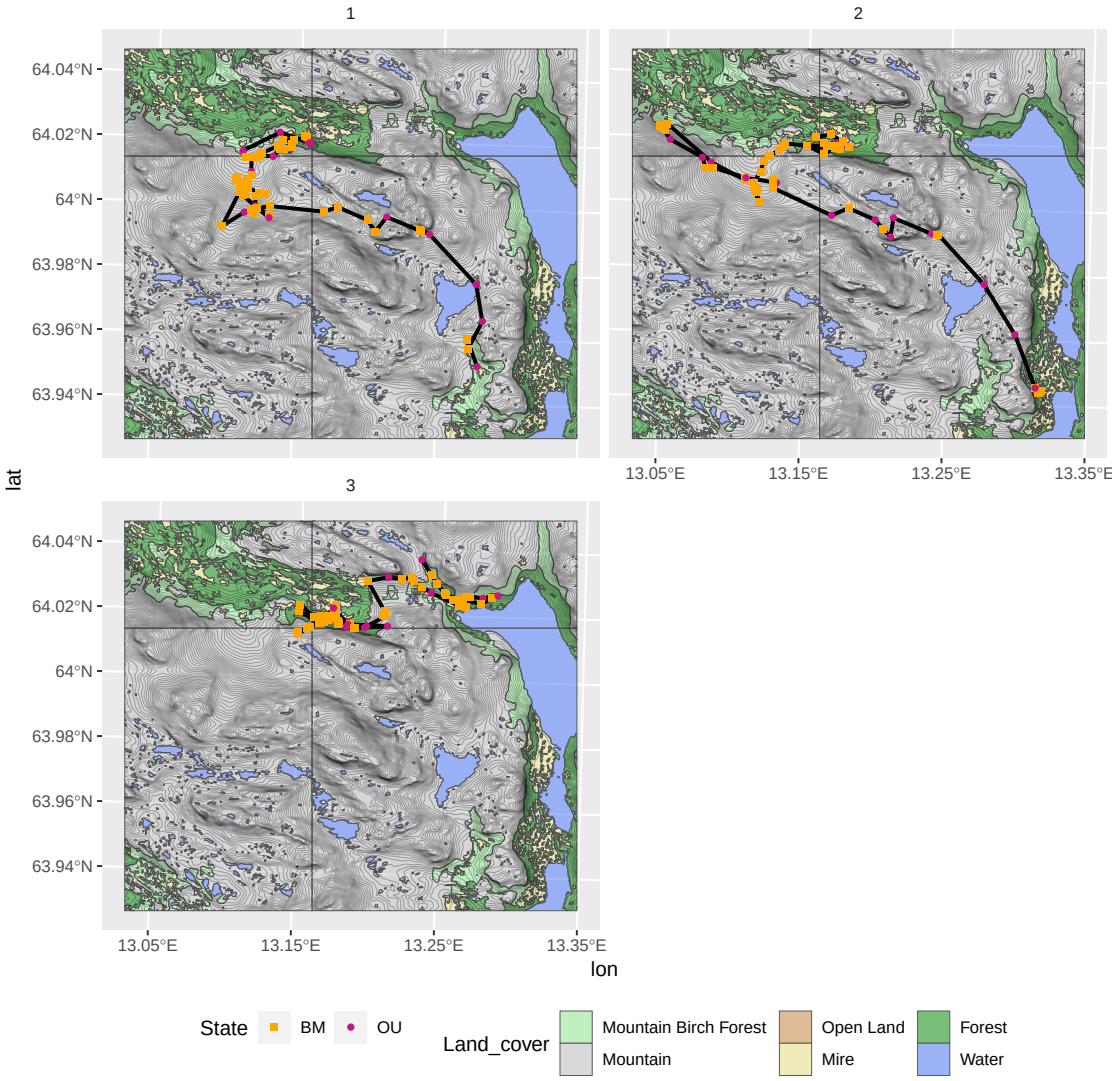


Fig. 5.7 Trajectories for animals 1, 2 and 3 in the early summer data set projected on to a terrain map. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange squares indicate a BM state whilst the purple circles indicate OU states. The terrain is split into six types: mountain birch forest, mountain above tree line, open land, mire, forest and water given in light green, grey, dark tan, tan, green and blue respectively.

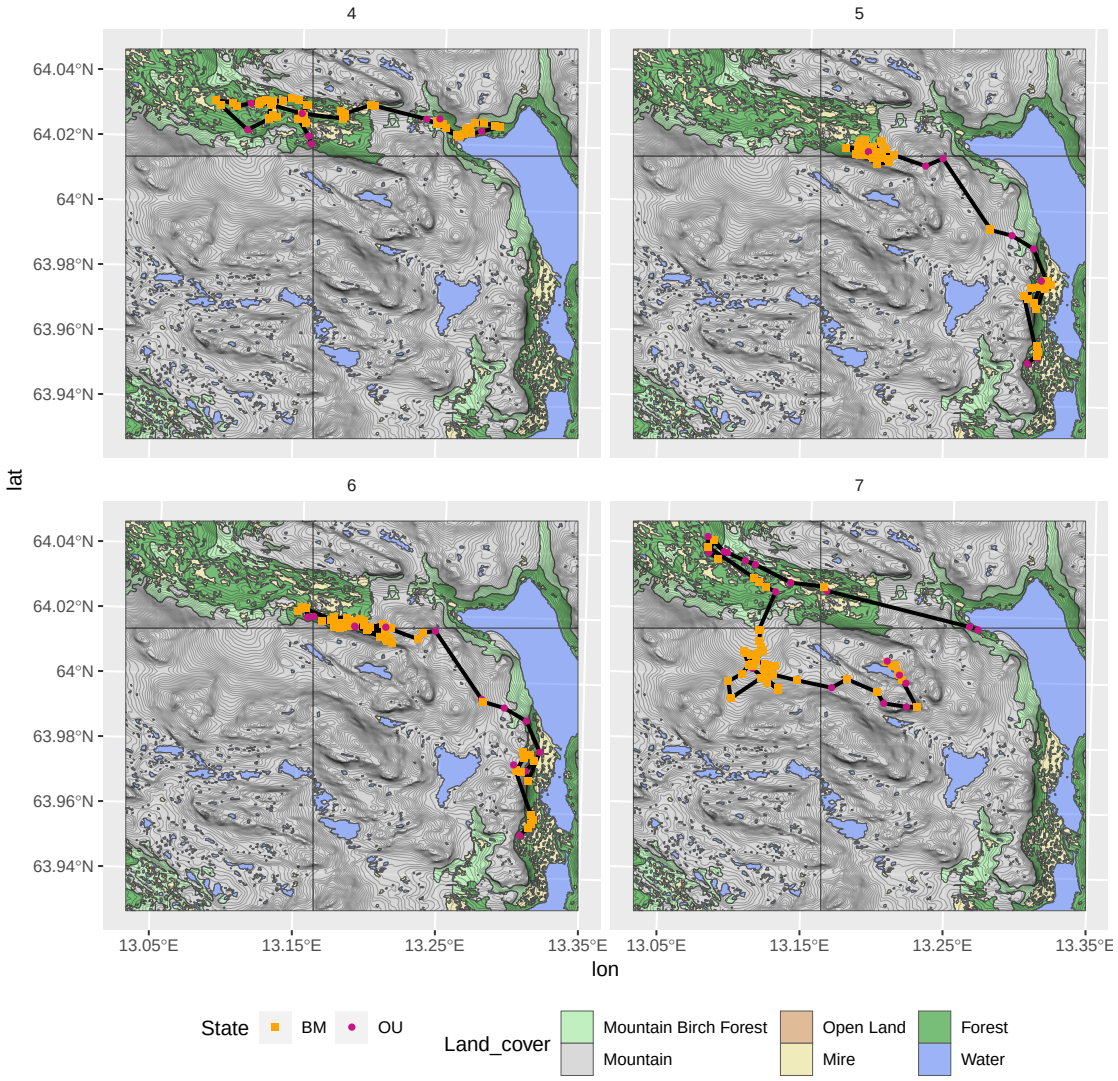


Fig. 5.8 Trajectories for animals 4, 5, 6 and 7 in the early summer data set projected on to a terrain map. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange squares indicate a BM state whilst the purple circles indicate OU states. The terrain is split into six types: mountain birch forest, mountain above tree line, open land, mire, forest and water given in light green, grey, dark tan, tan, green and blue respectively.

## 5.4 Peak Summer Results

We applied the switching model presented in Chapter 3 to the peak summer dataset. We ran the Markov chain Monte Carlo algorithm for 50,000 iterations. This took approximately 12 hours to complete. The posterior mean and standard deviation of model parameters are shown in Table 5.2. Posterior state estimations are given in Figure 5.9 whilst posterior densities for the parameters are given in Figure 5.10. For each individual, Figure 5.11 shows the locations in the $y$-direction against time; Figure 5.12 gives the two-dimensional trajectories; Figures 5.13 and 5.14 project these trajectories on to a terrain map.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.198 | 0.01 |
| $\rho$ | 3 | 0 |
| $\sigma$ | 0.63 | 0.025 |
| $\sigma_{\mathrm{BM}}$ | 0.61 | 0.05 |
| $\lambda_{1,2}$ | 0.08 | 0.017 |
| $\lambda_{2,1}$ | 0.23 | 0.055 |

Table 5.2 Parameter estimates of the switching model for the peak summer dataset where mosquito activity is assumed to be high.

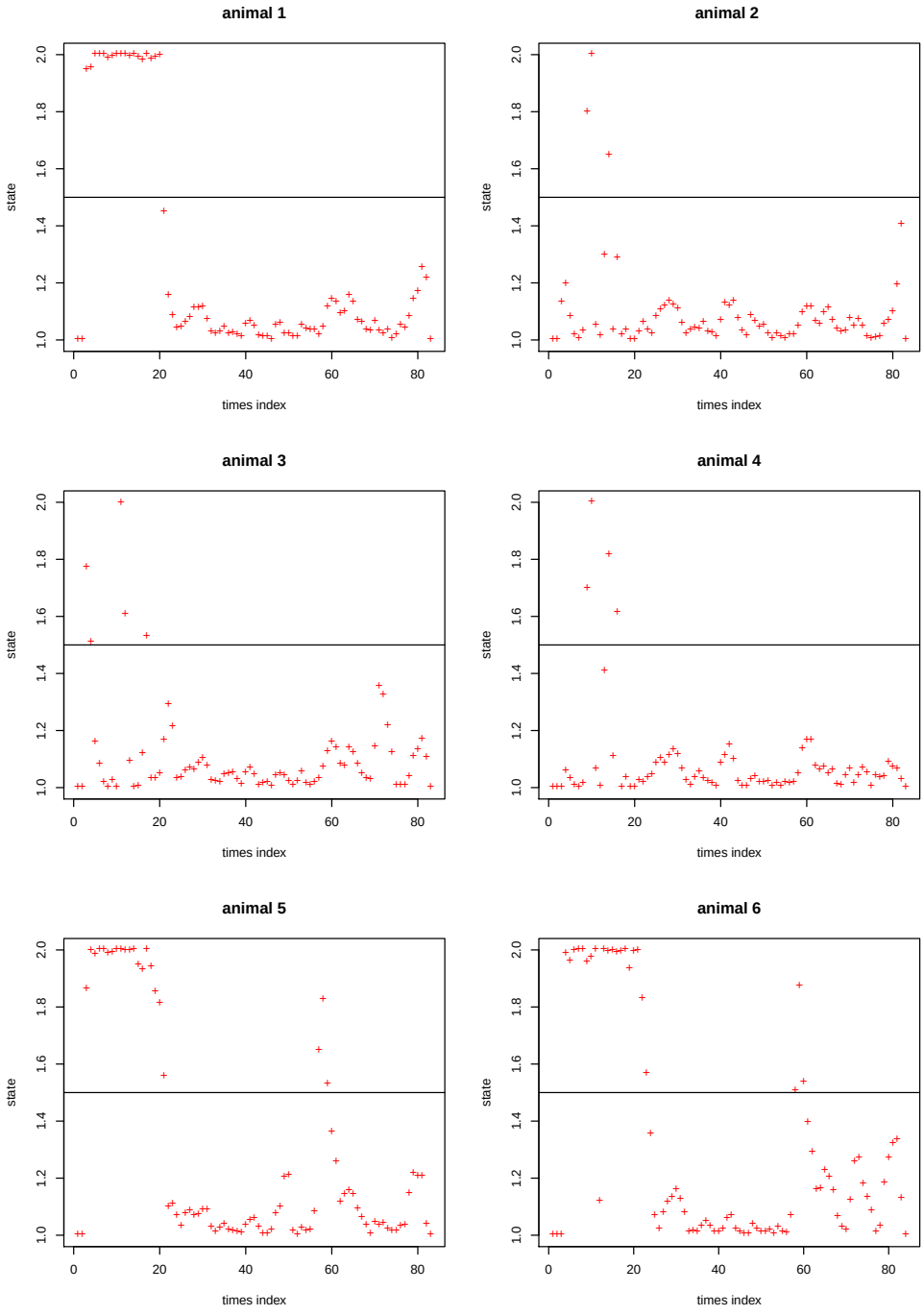


Fig. 5.9 Posterior mean states of all followers for the peak summer dataset where mosquito activity is assumed to be high. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (red) represent the mean posterior of the estimated behaviour states.

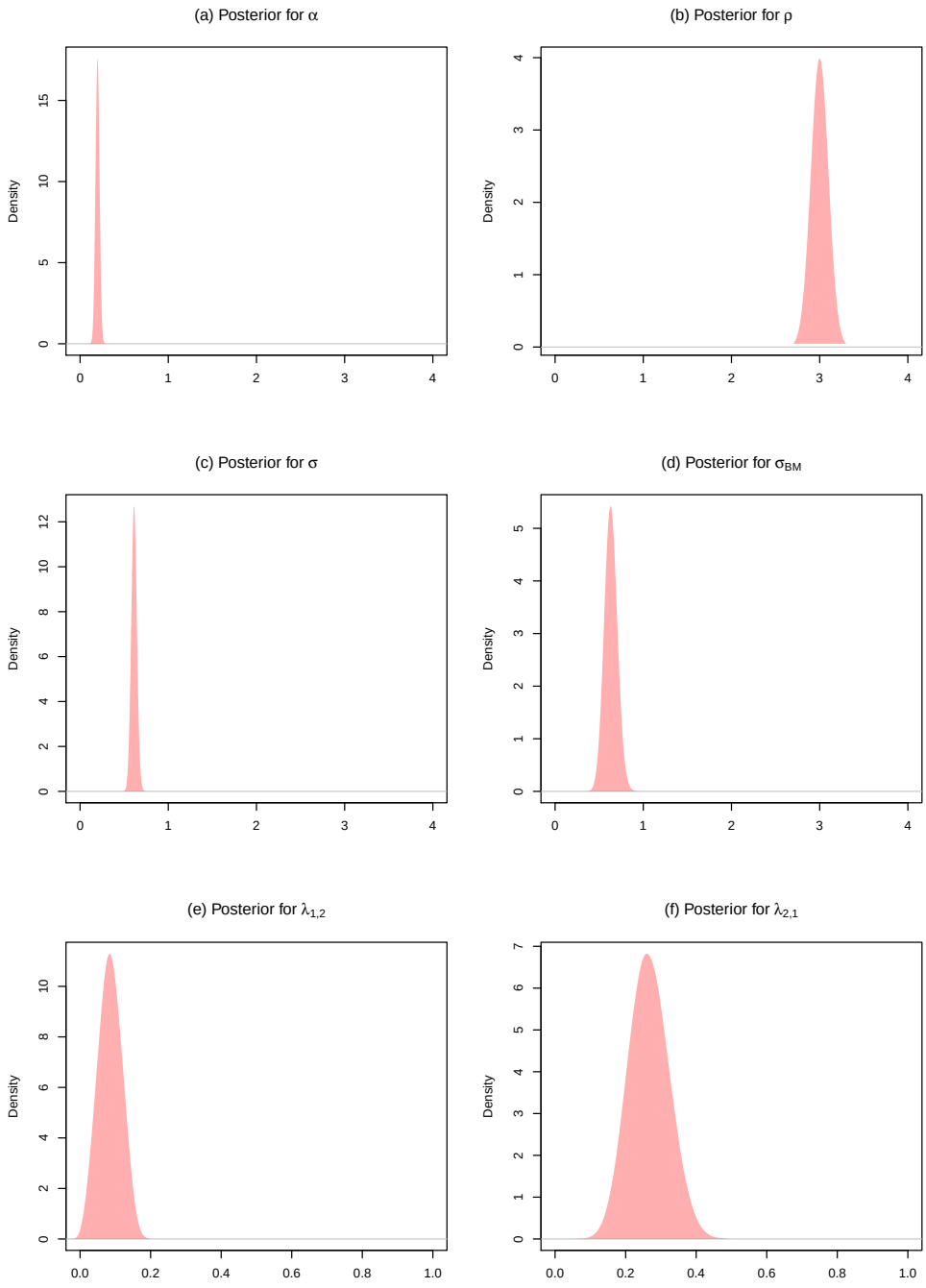


Fig. 5.10 Posterior densities for model parameters for the data set where mosquito activity is high, based on the Markov chain Monte Carlo runs of 50,000 iterations. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Approximate density for $\rho$, the variance coefficient of the leading point, which is fixed at 3. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (e) Posterior density for $\lambda_{1,2}$, the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion. (f) Posterior density of $\lambda_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck.

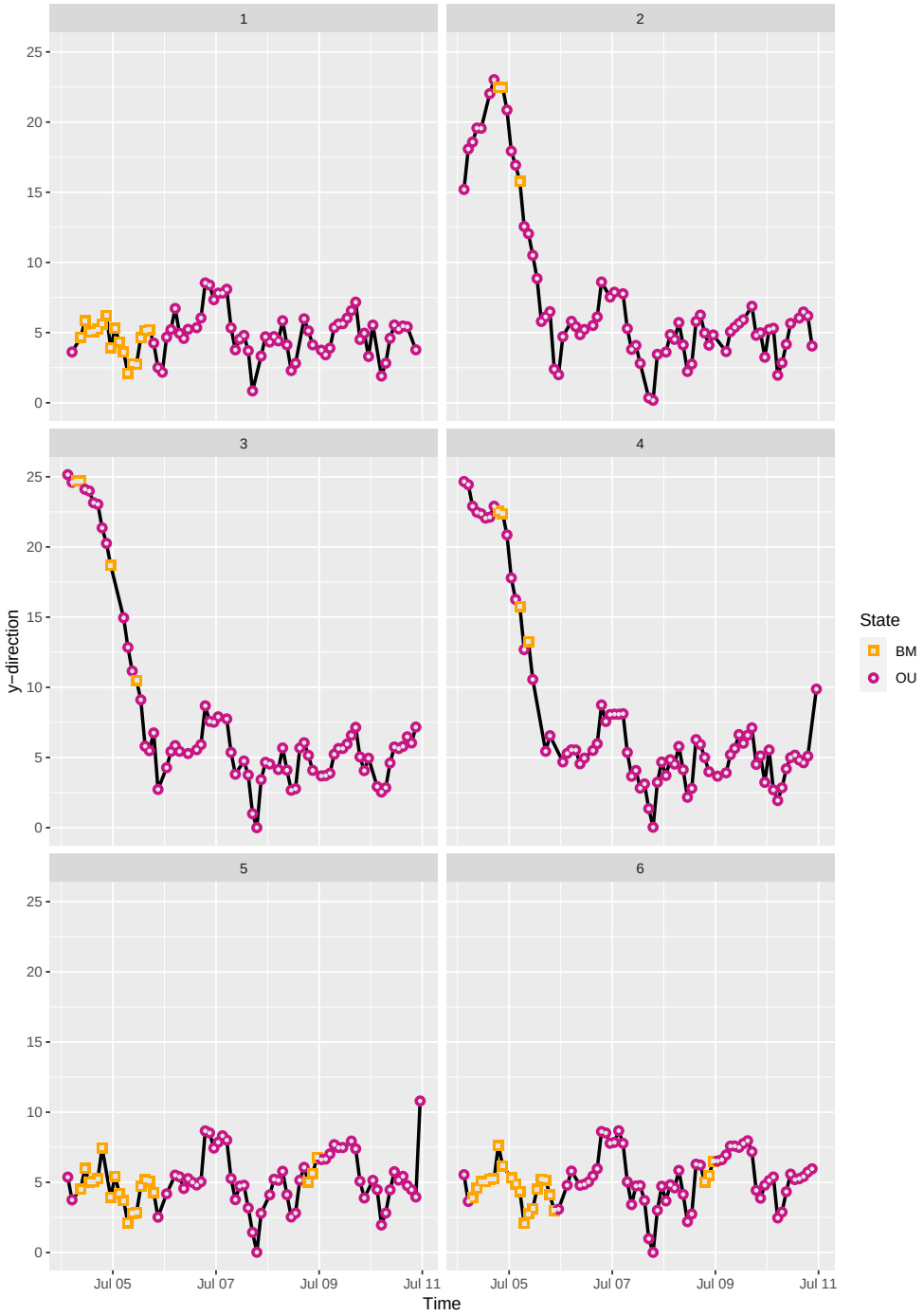


Fig. 5.11 Time trace of locations in the $y$-direction for each animal from peak summer dataset. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange squares indicate the BM state whilst the purple circles indicate OU states.


Fig. 5.12 Trajectories for each animal from the peak summer dataset. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange squares indicate the BM state whilst the purple circles indicate OU states.

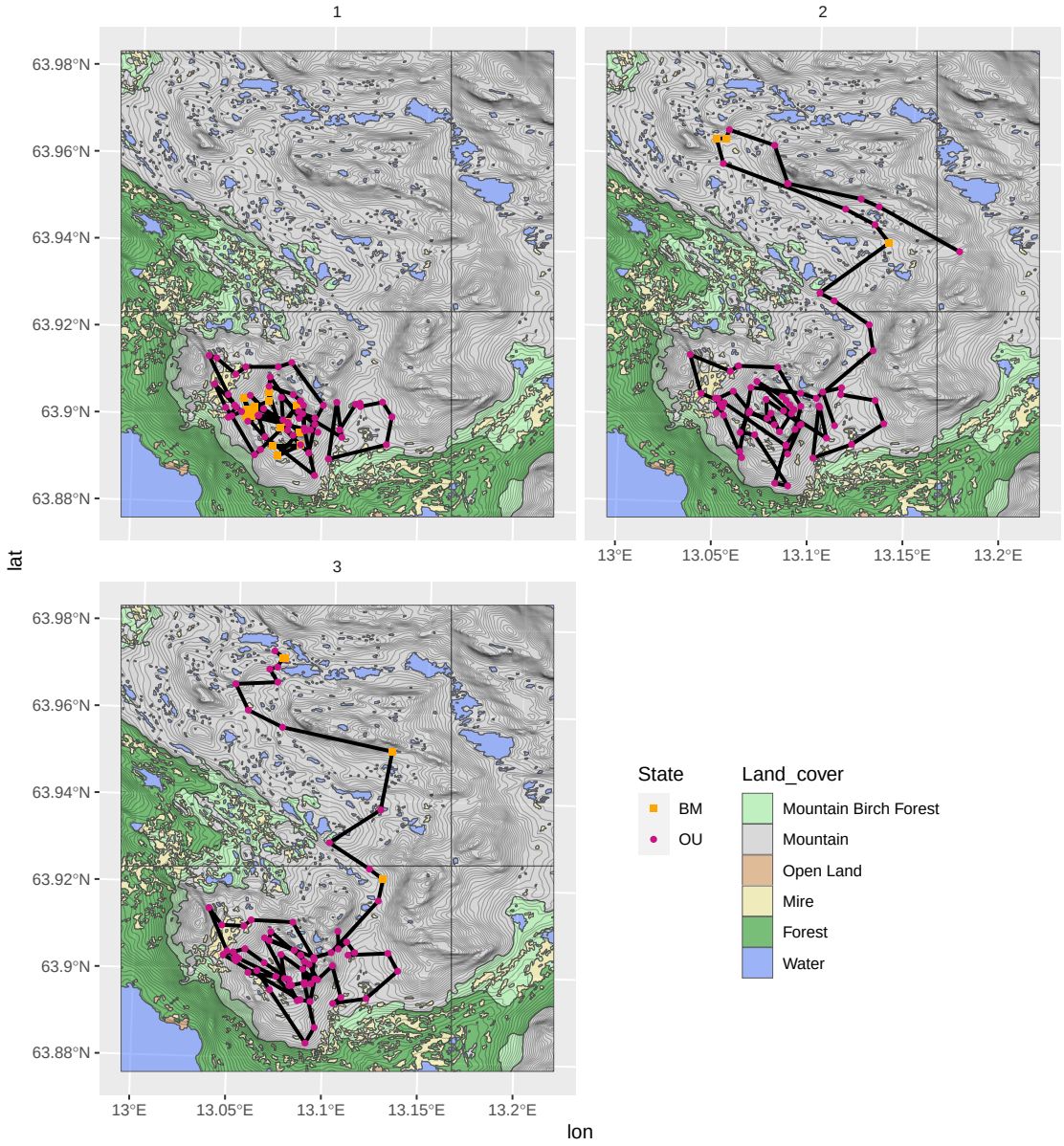


Fig. 5.13 Trajectories for animals 1, 2 and 3 in the peak summer data set projected on to a terrain map. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange squares indicate a BM state whilst the purple circles indicate OU states. The terrain is split into six types: mountain birch forest, mountain above tree line, open land, mire, forest and water given in light green, grey, dark tan, tan, green and blue respectively.

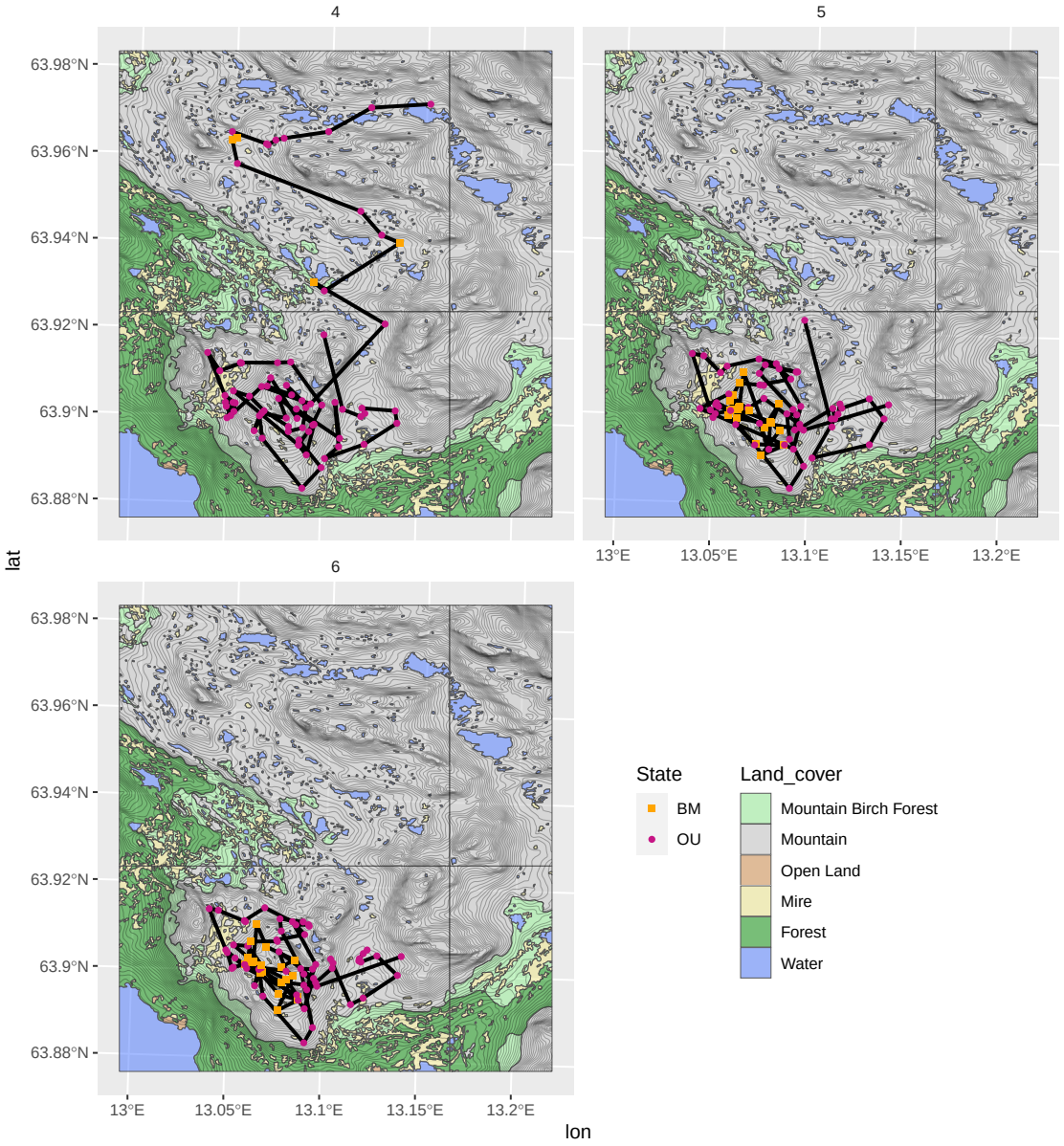


Fig. 5.14 Trajectories for animals 4, 5 and 6 in the peak summer data set projected on to a terrain map. At each time step the points indicate whether the individual's posterior state is OU or BM. The orange squares indicate a BM state whilst the purple circles indicate OU states. The terrain is split into six types: mountain birch forest, mountain above tree line, open land, mire, forest and water given in light green, grey, dark tan, tan, green and blue respectively.

## 5.5 Discussion

Throughout this chapter we have stressed the significance that parasitic harassment can have on the behavioural patterns and grouping dynamics of reindeer. By reviewing past observational studies, we have discussed a variety of reported weather conditions which are optimal for insect activity. Using local weather data as a proxy for insect presence we are able to select two time periods where we would expect contrasting reindeer movement. After applying the switching model to both datasets we found that the parameter values were starkly different between cases. Not only this, the results appear to be consistent with observational studies of reindeer movement in the presence of mosquitoes. That is, we expect less grouping behaviour in early summer due to the lack of mosquitoes and in the peak summer period, when mosquitoes are present, we expect reindeer to be in the OU state most of the time after 06/07/2010 i.e. after the first 20 observations.

The parameter point estimate for $\alpha$, the attraction to the leader, in the early summer case is relatively small, 0.047, in comparison with that of peak summer, 0.198, suggesting that even when the animals are in the OU state it is a very weak attraction. The proportion of time that the animals spend in each state is greatly different between time periods, which is evident from the point estimates of the switching rates. In early summer, the estimates are (0.63, 0.19) for $(\lambda_{1,2}, \lambda_{2,1})$, meaning that the proportion of time they spend in state 1 and 2 can be calculated as 23% and 77% respectively whereas, for peak summer the switching rates are (0.085, 0.27) thus, the proportion is calculated as 76% and 24% respectively.

During the early summer period the noise parameter in the OU state, $\sigma$, is larger (1.25) than that of the peak summer period (0.63) suggesting that there is more variation in the group movement state in the early summer than the peak summer case.

In addition, the parameter controlling the scale of Brownian motion $\sigma_{BM}$, is larger in peak summer than early summer (0.61, 0.22 respectively). One possible biological reason could be that in early summer, when the reindeer move independently, for example, to forage, they may do so without the disruption from insects whereas during peak summer, the reindeer display insect avoidance behaviour whilst attempting to forage independently. Thus, foraging during times of harassment may cause quicker relocation.

The precision of posterior point estimates for the states varies between the datasets. In the early summer dataset there are far more instances where the posterior mean estimates for observations are uncertain, lying somewhere around the 1.5 mark. It is possible that this can be attributed to the switching rates being homogenous in time. Within the early summer

time period there are strong fluctuations in the mosquito index. The switching parameters of the model try to give an overall depiction of the proportion of time spent in a state without describing any underlying mechanics. Hence, they may not actually describe the data and state switches successfully. This issue may alleviated by allowing the switching rates to depend on ancillary data such as mosquito indices, which we will discuss in Chapter 6.

Despite uncertainty in some of the state estimates, when visualising the trajectories the observations are still classified (crudely) as either OU or BM depending on whether they are below or above the 1.5 threshold. This is worth acknowledging before interpreting any of the plots. That said, we are still able to conclude that in the early summer cases (Figure 5.5 and 5.6) the reindeer generally move in the BM state without much directional change relative to the sharp erratic and grouped movement the reindeer exhibit in peak summer (Figure 5.11 and 5.12). This is in line with observational studies of reindeer movement and grouping dynamics in the presence of mosquitoes (Morschel and Klein, 1997; White et al., 1975).

In addition, overlaying trajectories on landscape maps offer great insight into space utilisation and preference during mosquito presence. By comparing Figures 5.7 and 5.8 to Figures 5.13 and 5.14 we can see contrasting resource selection between the two time periods. During early summer, when we expect low mosquito presence, the reindeer move around lower elevations sometimes in forests or open land whereas in peak summer, the reindeer tend to move into a tight group in higher elevation mountain areas, possibly to seek topographical relief.

For these applications, difficulties in parameter estimation meant that it was necessary to fix $\rho$, the parameter controlling the variability of the leader's location. The value of $\rho$ was chosen to be the least integer greater than twice that of the other Brownian motion parameters, $\sigma$ and $\sigma_{BM}$. This seems like a sensible approach given that the leading point is a mathematical abstraction which we expect to behave as if it were a real animal. In essence, by fixing this parameter we are insisting that the leading point moves in a localised fashion making the model more identifiable and interpretable.

Whilst applying the model to the two datasets we experimented with different values of $\kappa$ (see Section 3.5 for the definition of $\kappa$). Increasing $\kappa$ comes at a trade-off for computational time since increasing the value of $\kappa$ leads to more potential switches. However, by increasing $\kappa$ we allow the switching rates to reach higher values. In some applications this may be necessary if the proportion of time spent in a particular state is relatively high. This was the case for the early summer dataset. Here, the point estimate for the switching rate $\lambda_{1,2}$ is 0.63, shown in Table 5.1. This required setting $\kappa = 5$. That way the maximum value that either switching rate may take is given by $\max\{\lambda_{1,2}, \lambda_{2,1}\} = \kappa/n$; in this case, $\max\{\lambda_{1,2}, \lambda_{2,1}\} = 5/7 \approx 0.71$.

Of course, so far we have focussed our attention to the presence of mosquitoes but this is not mutually exclusive with the presence of other flying insects such as oestrid flies. Figure 5.15 and 5.16 indicate the presence (or absence) of oestrids and mosquitoes during the early and peak summer datasets respectively. Here, we assume that oestrid activity occurs when wind speed is below 9 m/s and the temperature was above 11°C, motivated by the thresholds for oestrid activity presented in Skarin et al. (2010).

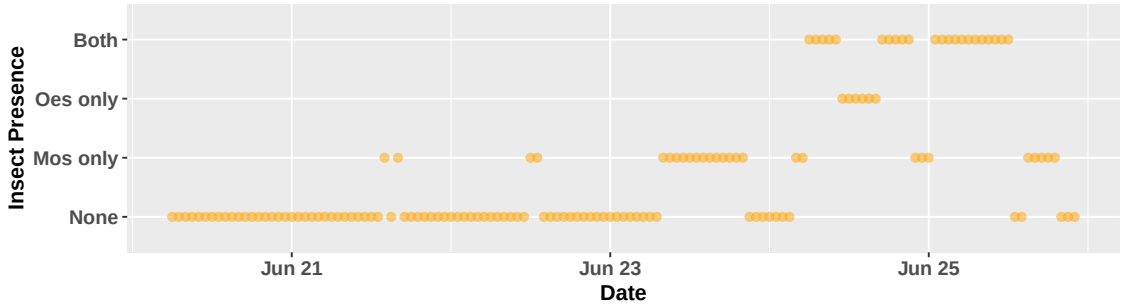


Fig. 5.15 Timeline of insect index during the early summer period. Presence of a particular insect are based on the wind speed and air temperature thresholds for activity as presented in Skarin et al. (2010).

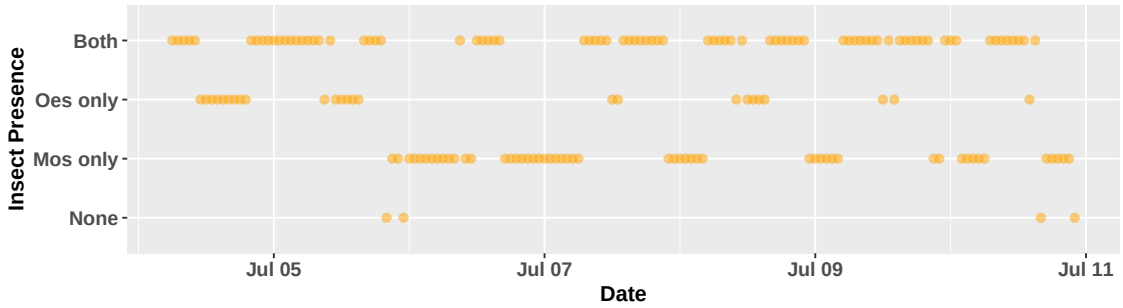


Fig. 5.16 Timeline of insect index during the peak summer period. Presence of a particular insect are based on the wind speed and air temperature thresholds for activity as presented in Skarin et al. (2010).

In Figure 5.15 we see that after 23/06/2010 insects begin to be consistently active, firstly with mosquitoes only and then after 24/06/2010 oestrids begin to appear. At this time the 1-dimensional trajectory plots in Figure 5.5 show the individuals beginning to make relatively large deviations in movement. In some cases the herd move whilst grouped together in an OU state, this could be a characteristic of the presence of mosquitoes; at other times the herd

moves in a BM state which may be attributed to the presence of oestrids. These results are consistent with observational findings of grouping patterns (Helle and Aspi, 1983; White et al., 1975).

During the first 20-30 observations of the peak summer dataset (up to around 06/07/2010), the posterior state estimations given in Figure 5.9 appear to be strongly in a BM state. However, during this time the mosquito index given in Figure 5.16 is not exclusively mosquito free. Therefore, the dispersing of the reindeer may be attributed to other covariates at this time. Figure 5.16 indicates that during the period of time in question the climatic conditions were favourable for oestrid flies and in some cases both mosquitoes and oestrids. Later in the time period there are more occasions where only mosquitoes are assumed to be present which may explain the grouping behaviours at this time. The uncertainty surrounding the posterior mean state estimations may be down to the sporadic and infrequent instances when only oestrids are present.

In the current literature, there has been much discussion about which insects have the largest effect on reindeer movement. At times, the harassment from a particular species of insect may dominate another and this is difficult to deduce from climatic data alone. However, it is possible that by analysing reindeer movement in this way we could gain more insight into parasitic harassment.

We have demonstrated that the model successfully captures different behavioural modes between multiple datasets. Using weather data as a proxy for insect activity, we are able to provide general features of movement and grouping behaviours of reindeer in response to insect presence. By overlaying trajectories with state estimates onto landscape maps we can establish the utilisation of particular habitats in the presence or absence of parasitic pressures. This illustration highlights the possible merits of such techniques for example, their use within conservation and management.

It is clear from this informal comparison that the presence of insects are a major factor in shaping the behaviour of reindeer. However, currently the switching rates within the model are homogenous in time and space which can lead to poorly estimated parameters and states. A more appropriate model would allow the switching rates to depend explicitly on insect levels. In the next chapter, we present a framework to formally include these covariates by allowing the switching rates to be governed by mosquito indices.

# Chapter 6

# A Model for Group Movement with Switching and Covariate Data

An individual's movement and behavioural choices are rarely random; often there exist stimuli and drivers underlying their movement processes. These choices may exist as responses to internal signals or physiological motives such as hunger or fatigue which can lead to behaviour such as foraging or resting. Other movement patterns may be reactions to external stimuli for example, weather, time of day, forage opportunity or distance to human activities. Motivations for group-level movement amongst collective species are not dissimilar to their solitary counterparts. Their movements and grouping dynamics are often actuated by particular needs and the relative cost and benefits of aggregations (Delgado et al., 2018). Grouping or dispersing behaviours may come as responses to predators, mating seasons, forage competition and parasite harassment as discussed in Chapter 5, as well other stimuli.

As a result, pitfalls exist when inferring behaviour based on trajectory data alone and we run the risk of mischaracterising behavioural modes. Utilising covariate information such as weather data or satellite imagery may improve our ability to distinguish behavioural states. In more recent technological advancements, a wealth of biotelemetry ancillary data has been used to extract more out of location data. For instance, McClintock et al. (2013) used salinity and depth sensors to measure dive data of grey harbour seals to differentiate between foraging and travelling behaviour whilst Shepard et al. (2008) used accelerometers to enhance the identification of behaviours from body motion.

Some work has been done to incorporate environmental information into movement models. One popular class of models are step selection functions (SSF). Essentially, these are

resource selection functions (RSF), tools used for estimating the utilisation distribution of an environment, but with added constraints of movement via models such as step and turn. To do this, an animal's observed steps are compared with random steps drawn from a probability distribution given by the movement model to calculate the proportionality of a habitat's utilisation with its availability. Whilst these techniques have proved useful, they have a number of drawbacks; the user must have a priori considerations such as the number of random steps and the scale of movement to be used. The article by Thurfjell et al. (2014) provides a good overview of SSF and their limitations.

Other approaches have chosen to allow the transition probabilities of state switching models to be given explicitly as functions of environmental covariates. The influential work of Morales et al. (2004) gave a flexible suite of models for movement where switches between multiple behavioural modes have the possibility of being influenced by external factors. Their *switching with covariates* model presented here allowed the switching probabilities between two behavioural states to be modelled as a function of distance to habitat type. In a similar vein, McClintock et al. (2012) allowed movement to be modelled as correlated random walks and switching between multiple states were modelled as functions of distance to centres of attractions. In this case, the centres were haul-out sites and foraging areas for grey seals. They found that bias towards centres of attractions gave a better explanation of movement than CRW alone. Further to this, McClintock et al. (2013) used the results of movement models coupled with ancillary dive data from multiple seals to investigate activity budgets at the regional population level. Whilst these studies gave insight into modelling movement with respects to covariate information, all the literature above explained movement using a discrete-time framework, whose switching times are restricted to the same sampling frequency of the observation.

In contrast, Ovaskainen (2004) approximate their correlated random walk models with diffusion processes for individual-level movement in heterogeneous space. Unlike Morales et al. (2004) and McClintock et al. (2012), the study area is categorised into several distinct habitat types and movement in each habitat has its own associated processes. A similar approach is taken in Harris and Blackwell (2013), although they give a more flexible framework by allowing their diffusion equation to have drift and relaxing assumptions about movement around the boundary of regions (for a more detailed comparison see Harris and Blackwell (2013)). Methodology of Bayesian inference for this modelling technique is given in Blackwell et al. (2016) and illustrated with a case study using the real location data of an individual fisher (*Martes pennanti*). The implementation of temporal covariates are also explored in this research. The authors allow switching rates between numerous behavioural states and

bias towards centres of attraction to be modelled as a function of the time of day. To give biological perspective they use this framework with the real location data of wild boar whose switch from 'resting at a nest site' to 'foraging' is assumed to take place at a similar but not identical time each day.

To my knowledge there exists a gap in the literature for explicitly modelling the movement of a group of animals which operates in a continuous-time framework and has the possibility of including covariate information. The following section provides a flexible structure for implementing covariate information. Although most taxa are mobile beings whose behavioural processes will be influenced by spatial heterogeneity, this chapter restricts our focus to the inclusion of temporal covariates such as time of day or weather, as these are simpler to implement than spatial covariates. We demonstrate this methodology in Section 6.2 by introducing a covariate model for group-level movement which exhibits a sinusoidal behavioural processes, where the animals may switch between following the group and moving independently. This could represent diurnal grouping or dispersing behaviour which is based on the time of day. In Section 6.3 we revisit the grouping dynamics of reindeer in response to insect harassment by formally including insect presence as a covariate. That is, we allow the switching rates between states to be dependent on insect harassment data.

## 6.1 Implementing Covariates

Following a similar approach to the wild boar application in Blackwell et al. (2016), we allow the switching rate parameters $\lambda_{1,2}$ and $\lambda_{2,1}$ to be modelled as a function of temporal covariate information $Z_t$, the covariate at time $t$. To recap, without the inclusion of covariates the probability that a potential switch is an actual switch depends only on the current behaviours i.e. it is given as

$$P(\text{actual switch}) = \frac{n_1\lambda_{12} + n_2\lambda_{21}}{\kappa}, \tag{6.1}$$

where

$$\kappa \geq n \max\{\lambda_{12}, \lambda_{21}\} \tag{6.2}$$

and $n_1, n_2$ are the number of animals currently in state 1 and 2 respectively; $n$ is the total number of animals and $\lambda_{12}, \lambda_{21}$ are the switching rates from state 1 to 2 and 2 to 1 respectively.

Now, for some covariate $Z_t$, we allow each switching rate from state $i$ to $j$ to be given as some function of the covariate i.e.

$$\lambda_{ij}(t) = f_{ij}(Z_t). \tag{6.3}$$

Our choice for $f_{ij}$ will depend on our model formulation for the covariate, but the general principle remains the same. Subsequently, we must alter Equation 6.2 to account for these dynamic switching rates in the following way

$$\kappa \geq n \max\{\lambda_{12}^{max}, \lambda_{21}^{max}\}. \tag{6.4}$$

In its flexibility, this methodology allows for the inclusion of a variety of additional covariate information. We begin by exploring a simulated yet biologically realistic example where behavioural patterns depend on the time of day.

### 6.1.1 Inference and Code Modification

The parameters of the switching rates will vary depending on the scenario we are interested in modelling for example, four parameters are used to define the sinusoidal behaviour in Section 6.2. To implement such covariate information we must adapt our inference method. In short, we must now make inference on the underlying parameters which determine the switching rates i.e. the parameters of the function in Equation 6.3 as opposed to using a standard MCMC for $\lambda_{1,2}$ and $\lambda_{2,1}$ themselves. For all switching rate parameters, we use a uniform prior on the appropriate interval of the real line e.g. for the sinusoidal example it is required that two of the parameters are between 0 and 24 as these represent the time of day in hours.

Not only this, the switching rates now vary temporally so the times of switches are crucial. This amounts to keeping track of time in more detail and inputing the switching probabilities for each potential switch at those specific times. From a coding perspective this was not completely trivial, however, we have now developed such a system that, in principle, may employ any form of temporal covariate model. From now on, unless stated otherwise, we will be referring to this updated version of the inference method when mentioning the MCMC algorithm.

The code for the methodology is written in R (R Core Team, 2020) and can be found in the GitHub repository https://github.com/FayFrost.

## 6.2 Modelling Sinusoidal Grouping Patterns

This section aims to model sinusoidal behaviour that can occur in some diurnal animals. We explore a general model for grouping behaviour in which each individual may switch between Ornstein Uhlenbeck and Brownian motion, where the switching rates are modelled using a cosine function of the time of day. Possible real world instances of this behavioural pattern could be night time grouping for warmth or safety from predators and then separating during the day for individual foraging; for nocturnal animals this could model independent exploring behaviour during the night then grouping at a nest site throughout the day. To imitate this behaviour the switching rates are given as a function of the time of day in hours, $t$, as follows.

$$\begin{aligned} \lambda_{1,2} &= \frac{h_{ou}}{2}\left(1+cos\left((t-t_{ou})(2\pi/24)\right)\right), \\ \lambda_{2,1} &= \frac{h_{bm}}{2}\left(1+cos\left((t-t_{bm})(2\pi/24)\right)\right). \end{aligned} \tag{6.5}$$

where $h_{ou}$ and $h_{bm}$ determine the rate of switching from OU to BM and BM to OU respectively. The most likely times of switching away from OU and BM are given by $t_{ou}$ and $t_{bm}$ respectively. As an example, we fix $t_{ou} = 20$, $t_{bm} = 8$ and $h_{ou} = h_{bm} = 0.5$. Thus, the animals are most likely to switch into Brownian motion behaviour in the evening and group during the day. Figure 6.1 shows the rate of switching in either direction over a period of 50 hours, where 0 hours represents midnight. We can see that the switching rates undergo two full daily cycles within this time period.

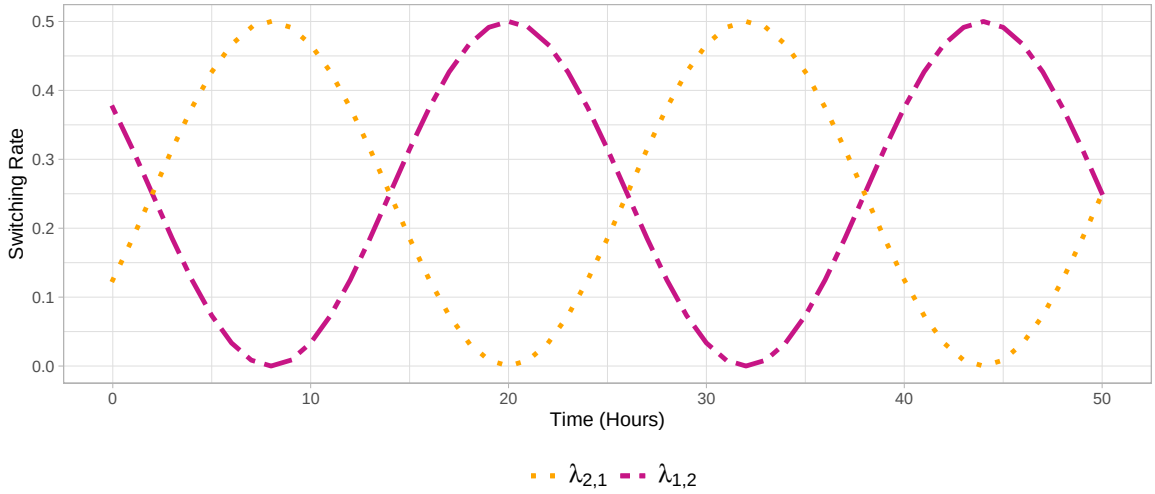


Fig. 6.1 Timeline of switching rates over the course of 50 hours whilst undergoing sinusoidal behavioural patterns. The purple double dashed line represents the switching parameter $\lambda_{1,2}$ whereas the orange dotted line gives the switching parameter $\lambda_{2,1}$.

### 6.2.1 Simulation of two daily cycles

To illustrate the model we simulate location data of 5 individuals using the switching rate functions given in Equation 6.5. The simulated data consists of 48 hourly observations for each individual generated by using Equation 3.16 iteratively and taking each generated location as the origin for the next.

We ran the Markov chain Monte Carlo algorithm (with the updates as discussed in Section 6.1.1) for 20,000 iterations with the initial 70% counting towards a burn-in period. This took approximately 12 hours to complete. The true parameter values for each data set are given in Table 6.1, along with the point estimates and standard deviations of the posterior distributions for each parameter. The posterior states are given in Figure 6.2. The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (black) represent the mean posterior of the estimated behaviour states. The posterior densities for the diffusion parameters and switching parameters are given in Figure 6.3 and 6.4 respectively.

| Parameter | Point estimate | Standard deviation | True value |
|---|---|---|---|
| $\alpha$ | 0.47 | 0.08 | 0.5 |
| $\rho$ | 2.76 | 0.48 | 3 |
| $\sigma$ | 1.03 | 0.06 | 1.0 |
| $\sigma_{\text{BM}}$ | 0.44 | 0.03 | 0.5 |
| $h_{ou}$ | 0.40 | 0.09 | 0.5 |
| $h_{bm}$ | 0.43 | 0.09 | 0.5 |
| $t_{ou}$ | 15.01 | 0.45 | 20 |
| $t_{bm}$ | 7.34 | 1.19 | 8 |

Table 6.1 Parameter estimates for the sinusoidal model using simulated hourly data with 48 observations.

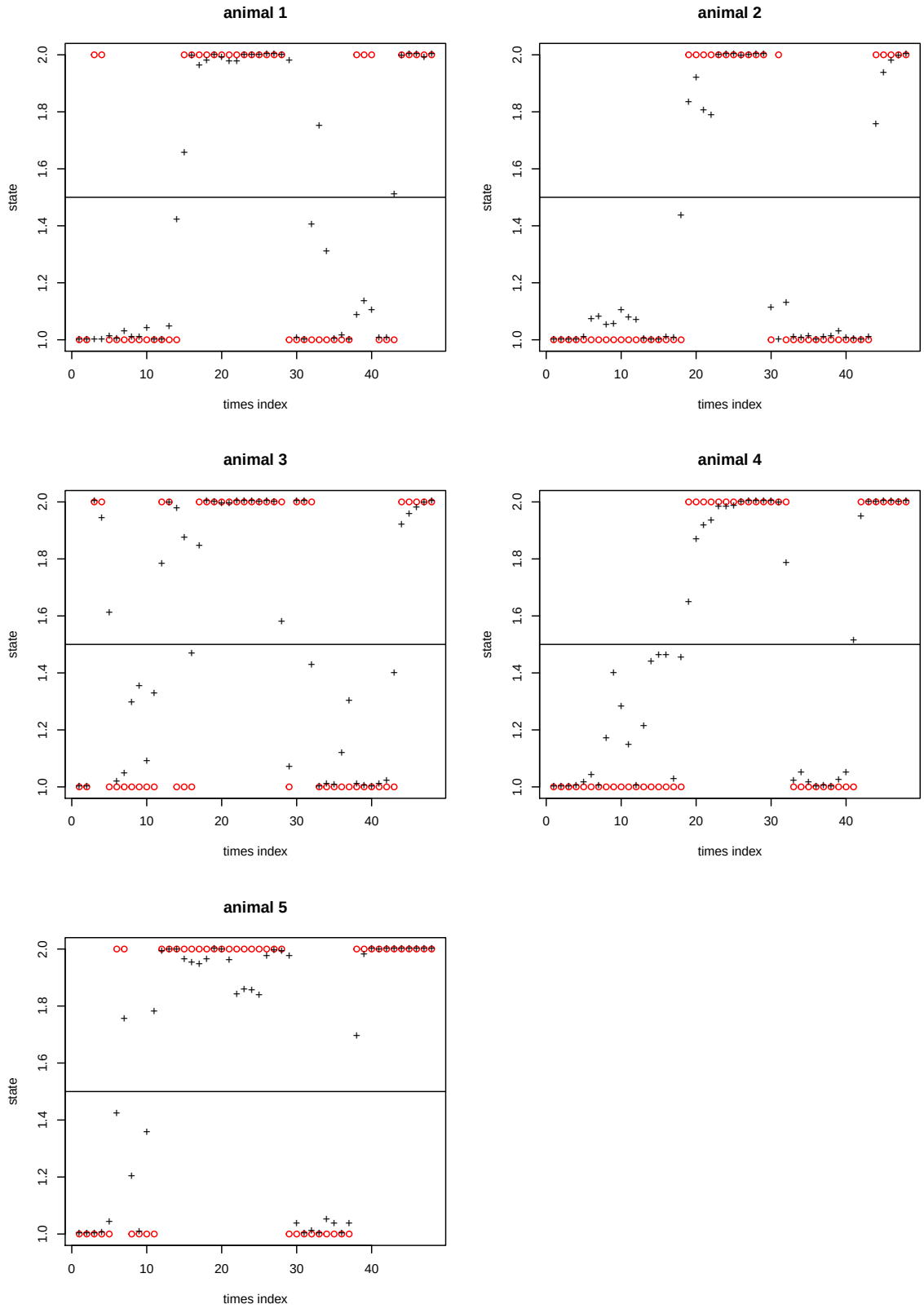


Fig. 6.2 Posterior mean states of all followers for the simulated sinusoidal data with 48 observations. The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (black) represent the mean posterior of the estimated behaviour states.

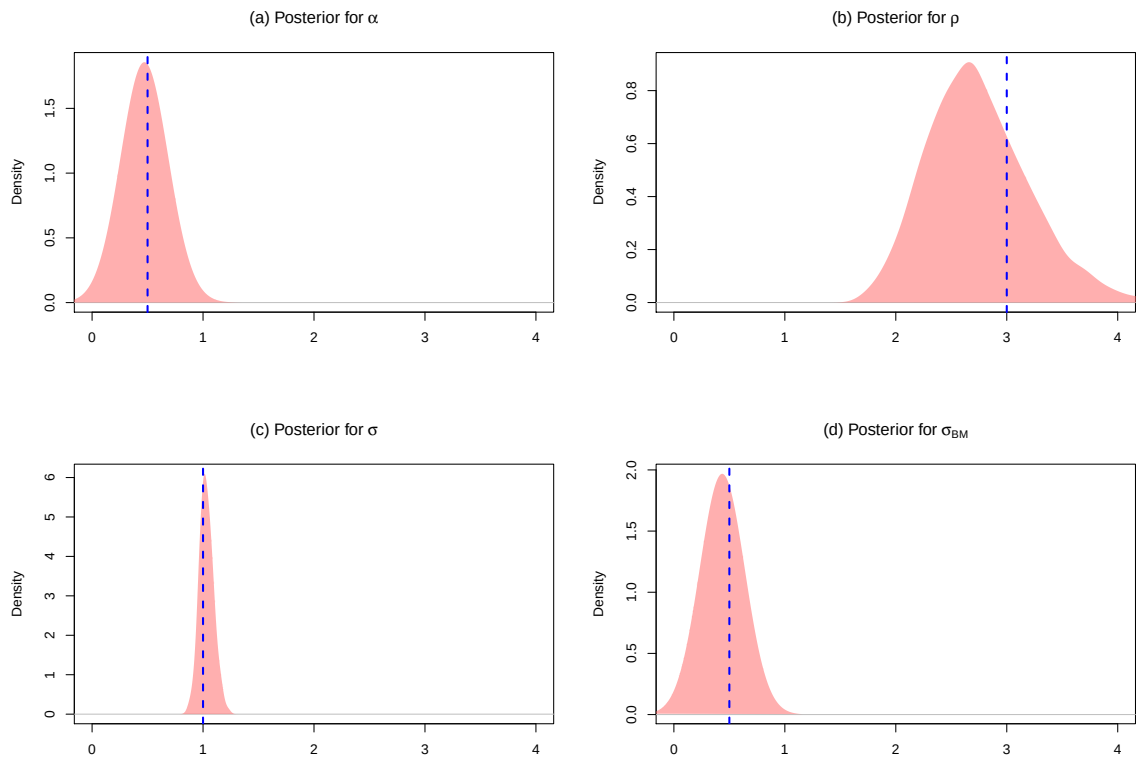


Fig. 6.3 Posterior densities of diffusion parameters for the simulated sinusoidal data with 48 observations, based on the Markov chain Monte Carlo runs of 20,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion).

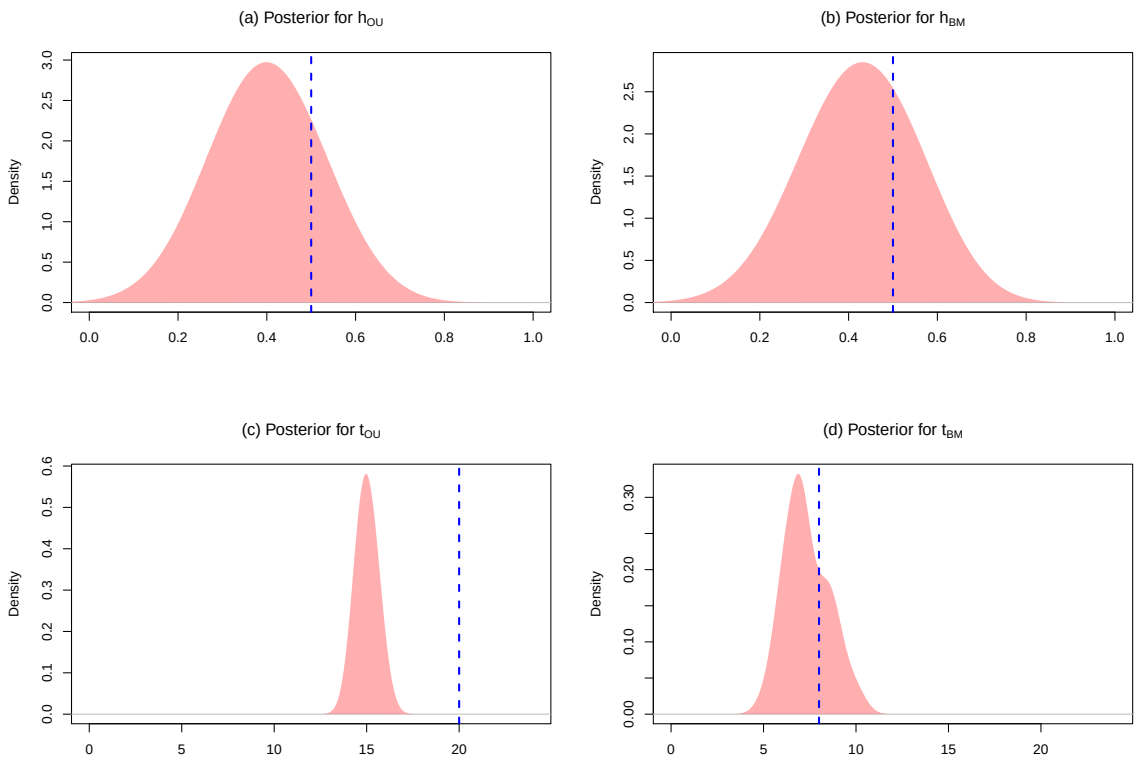


Fig. 6.4 Posterior densities for model's switching parameters for the simulated sinusoidal data with 48 observations, based on the Markov chain Monte Carlo runs of 20,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $h_{ou}$, the rate of switching from OU to BM. (b) Posterior density for $h_{bm}$, the rate of switching from BM to OU. (c) Posterior density for $t_{ou}$, the most likely time of switch out of OU. (d) Posterior density for $t_{bm}$, the most likely time of switch out of BM.

On the whole, the model performs reasonably well at detecting the pattern of movement and behaviour. Although there are at times uncertainty in the state estimations, the general cyclic nature of the behaviour is picked up on. The diffusion parameters are estimated very well with all of the posterior point estimates being close the true value.

However, from Figure 6.4 we see that the switching parameters are less well estimated. This is especially evident with the parameter $t_{ou}$. Not only is the point estimate of $t_{ou}$ far from the true value (15.01 and 20 respectively), the standard deviation is quite small suggesting that the model is confident in its estimation. Since we have the benefit here of knowing what the true value is, we can say that it is likely that this is not the true posterior of this parameter. It appears that the MCMC has not successfully explored the parameter space and become localised at another maximum. In Section 6.2.2 and 6.2.3 we pay particular attention to this

parameter and the improvements we make in its posterior estimation. That said, the point estimates of $h_{ou}$ and $h_{bm}$ do perform well in comparison to $t_{ou}$ and the standard deviations of the $h$ parameters are reasonably small.

Poor estimation, in part, may be down to the small number of actual switches that take place in the duration of two daily cycles. Since this number is likely to be on the order of four (two daily switches for each day) the model lacks enough information to accurately discern the parameters $t_{ou}$ and $t_{bm}$ which represent the most likely time of day that switching occurs.

The duration of this simulation was chosen to be computationally quick. In reality, modern location data is rarely as short as two days and if it is, the sampling scheme may have a much higher resolution. In an effort to test the limitations of this model we investigate whether data of longer duration or higher resolution improves parameter estimation.

### 6.2.2 Simulation with longer duration

Using the same true parameter values as in 6.2.1, we simulate the hourly location data of 5 individuals over the course of three days, that is 72 observations in total.

We ran the covariate Markov chain Monte Carlo algorithm for 20,000 iterations with the initial 70% counting towards a burn-in period. This took approximately 24 hours to complete. The true parameter values for each data set are given in Table 6.2, along with the point estimates and standard deviations of the posterior distributions for each parameter. The posterior states are given in Figure 6.5. Finally, the posterior densities for the diffusion parameters and switching parameters are given in Figure 6.6 and 6.7 respectively.

| Parameter | Point estimate | Standard deviation | True value |
|---|---|---|---|
| $\alpha$ | 0.54 | 0.06 | 0.5 |
| $\rho$ | 2.49 | 0.32 | 3 |
| $\sigma$ | 1.04 | 0.05 | 1.0 |
| $\sigma_{\text{BM}}$ | 0.43 | 0.03 | 0.5 |
| $h_{ou}$ | 0.43 | 0.10 | 0.5 |
| $h_{bm}$ | 0.44 | 0.08 | 0.5 |
| $t_{ou}$ | 16.17 | 2.61 | 20 |
| $t_{bm}$ | 6.95 | 1.30 | 8 |

Table 6.2 Parameter estimates for the sinusoidal model using simulated hourly data with 72 observations.

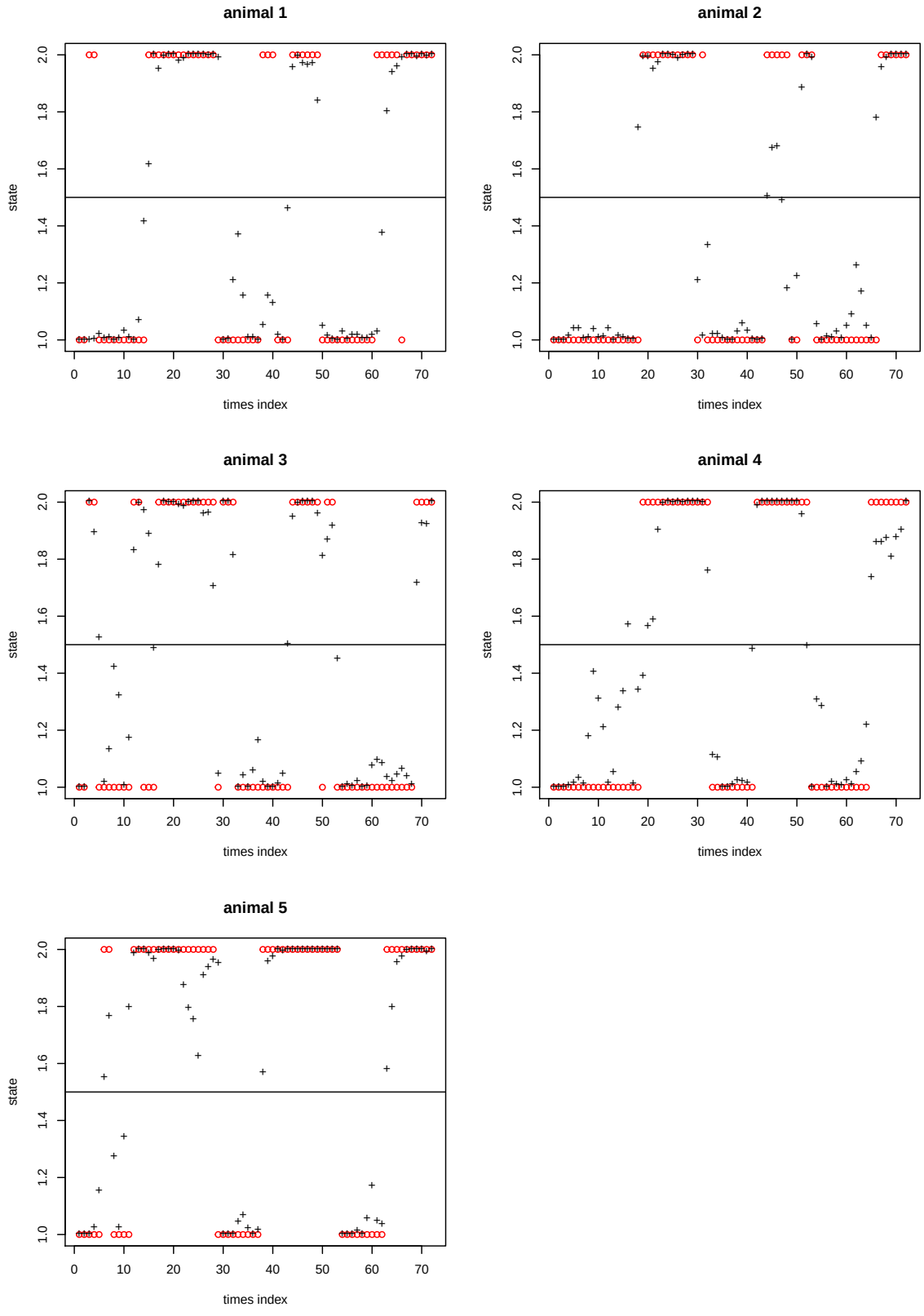


Fig. 6.5 Posterior mean states of all followers for the simulated sinusoidal data with 72 observations. The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (black) represent the mean posterior of the estimated behaviour states.

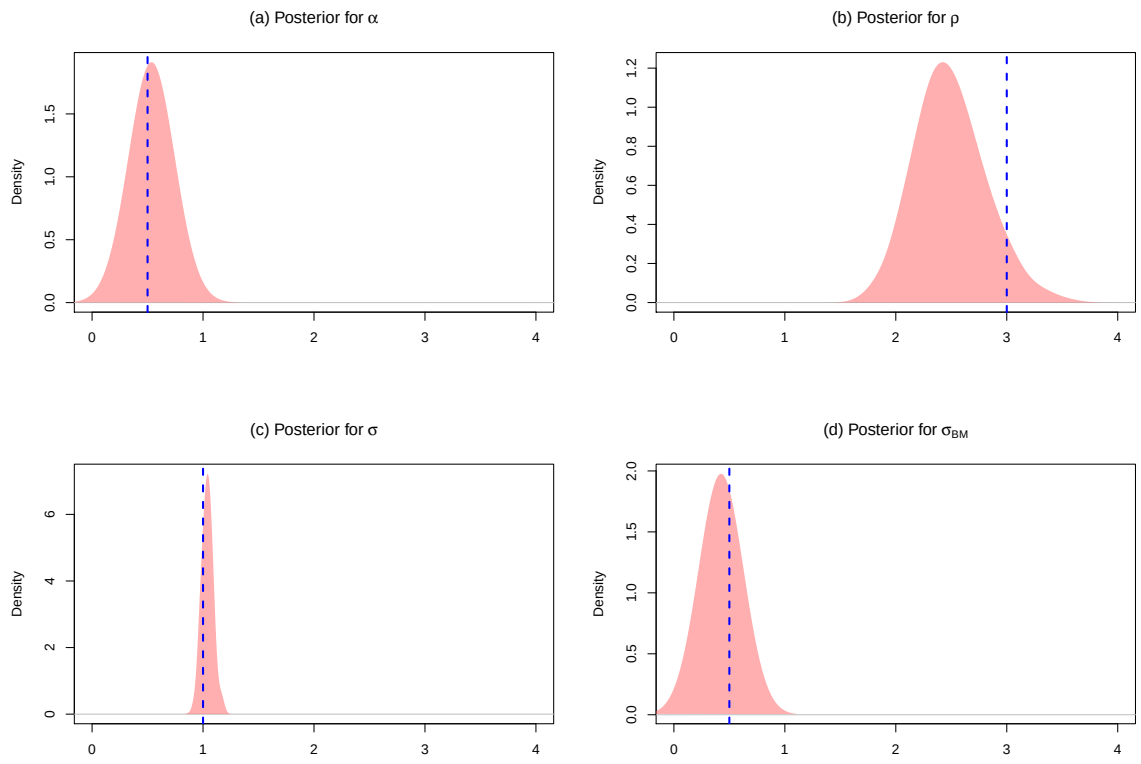


Fig. 6.6 Posterior densities of the diffusion parameters for the simulated sinusoidal data with 72 observations, based on the Markov chain Monte Carlo runs of 20,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion).

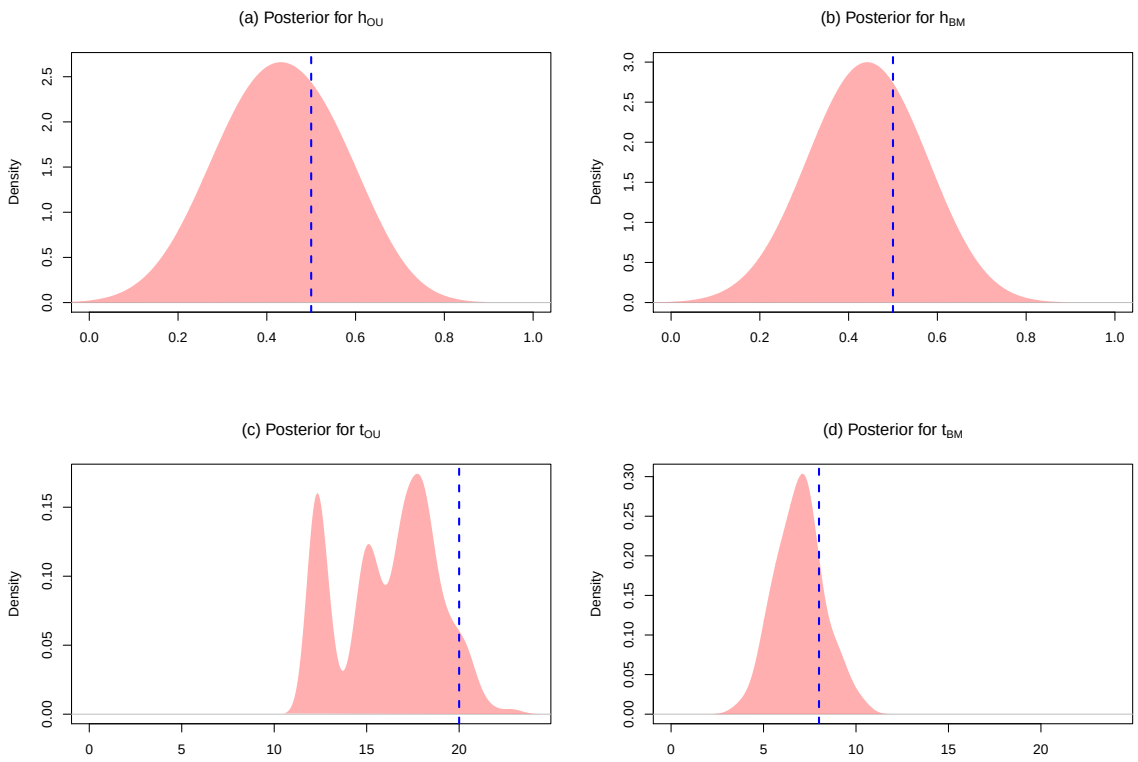


Fig. 6.7 Posterior densities of the switching parameters for the simulated sinusoidal data with 72 observations, based on the Markov chain Monte Carlo runs of 20,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $h_{ou}$, the rate of switching from OU to BM. (b) Posterior density for $h_{bm}$, the rate of switching from BM to OU. (c) Posterior density for $t_{ou}$, the most likely time of switch out of OU. (d) Posterior density for $t_{bm}$, the most likely time of switch out of BM.

By simulating three days of data we notice that the switching parameter estimates have experienced some improvement. Specifically, the posterior density for $t_{ou}$ given in Figure 6.7 gives more weight towards values closer to the true value than the corresponding plot in Figure 6.4 where only two days of data are used. However, in the same vein as in Section 6.2.1, this does not appear to be the true posterior and there may be underlying issues with the MCMC, which may be as a result of using poorly estimated states in the likelihood calculations.

This may be alleviated by using higher resolution data. In general, data with more frequent observations leads to a more precise idea of the animals' movement. This ought to improve the estimation of the movement parameters and the states. Thus, pinning down the timings of state transitions more accurately.

It is also possible that the reason for some uncertainty lies with instances where only a single animal is in the OU state. Biologically, this is not meaningful as a solo animal may not follow a 'group' that only they are part of. Hence, when all animals are in the BM state, an initial switch in state to OU may be overlooked and treated as though it is still moving independently. That said, estimations of the states and model parameters may improve when these events occur proportionally less often. Section 6.2.3 investigates whether increasing the frequency of observations and decreasing instances of solo OU state movement improves parameter estimation.

### 6.2.3 Simulation with more frequent observations

Using the same true parameter values as in Section 6.2.1 and 6.2.2 , we simulate the location data of 5 individuals over the course of three days with sampling frequency every half hour, that is 144 observations in total. Figure 6.8 displays the proportion of the number of animals in the BM state at each observation for this dataset and the one with hourly observations presented in Section 6.2.2. We can see that the dataset with more frequent observations, plot (b), has proportionally fewer times where all the animals are in the BM state.

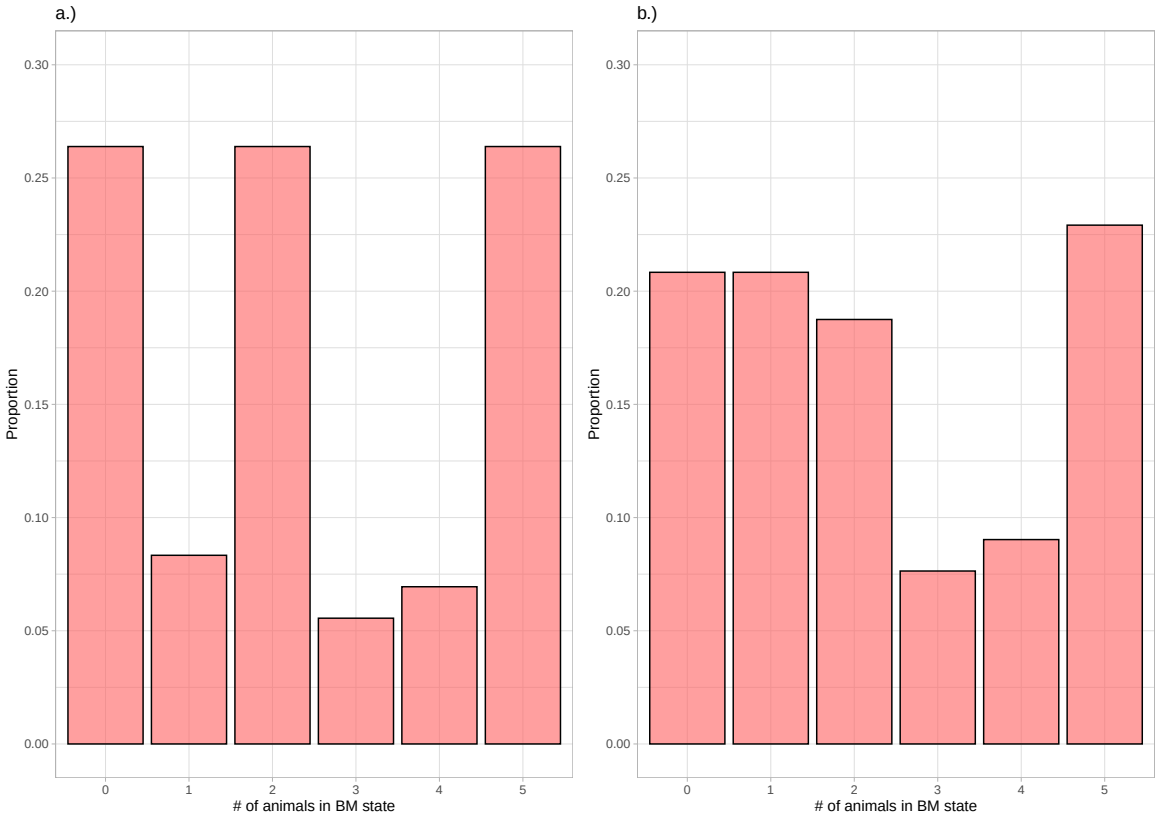


Fig. 6.8 Overall proportion of the number of animals in the BM state at each observation. (a) Proportion of the number of individuals in BM state for data with hourly observations. (b) Proportion of the number of individuals in BM state for data with 30 minute observations.

The difference in proportions in this case is somewhat marginal. So it is likely that any improvements in inference are a result of increased resolution of the data rather than the decrease in the proportion of times that all animals are in the BM state. It is also worth noting that whilst doing this experiment multiple times, we noticed that the distribution of the proportions can vary widely between simulations with different seeds. This adds to our conclusions that although the proportion of times that all the animals are in BM is a factor affecting the quality of estimation, it is likely to be negligible relative to the improvement that comes from higher resolution data alone.

Using the higher resolution data, we ran the covariate Markov chain Monte Carlo algorithm for 20,000 iterations with the initial 70% counting towards a burn-in period. This took approximately 72 hours to complete. The true parameter values for each data set are given in Table 6.3, along with the point estimates and standard deviations of the posterior distributions for each parameter. The posterior states are given in Figure 6.9. Finally, the posterior

densities for the diffusion parameters and switching parameters are given in Figure 6.10 and 6.11 respectively.

| Parameter | Point estimate | Standard deviation | True value |
|---|---|---|---|
| $\alpha$ | 0.47 | 0.03 | 0.5 |
| $\rho$ | 3.32 | 0.31 | 3 |
| $\sigma$ | 0.96 | 0.03 | 1.0 |
| $\sigma_{\text{BM}}$ | 0.32 | 0.01 | 0.5 |
| $h_{ou}$ | 0.53 | 0.05 | 0.5 |
| $h_{bm}$ | 0.54 | 0.05 | 0.5 |
| $t_{ou}$ | 19.27 | 0.96 | 20 |
| $t_{bm}$ | 6.85 | 1.27 | 8 |

Table 6.3 Parameter estimates for the sinusoidal model using simulated data with 144 observations and 30 minute sampling frequency.

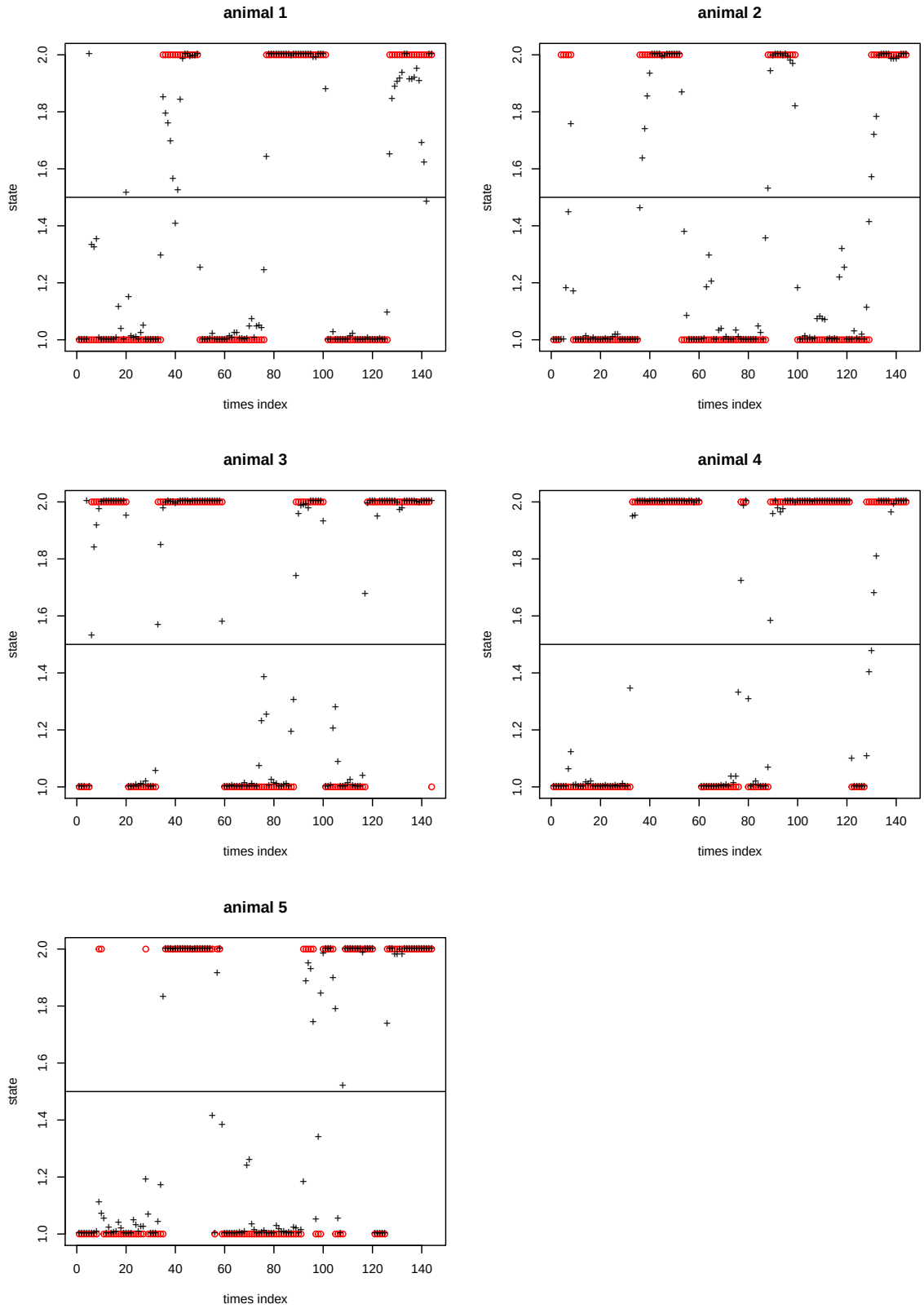


Fig. 6.9 Posterior mean states of all followers for the simulated sinusoidal data with 144 observations. The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (black) represent the mean posterior of the estimated behaviour states.

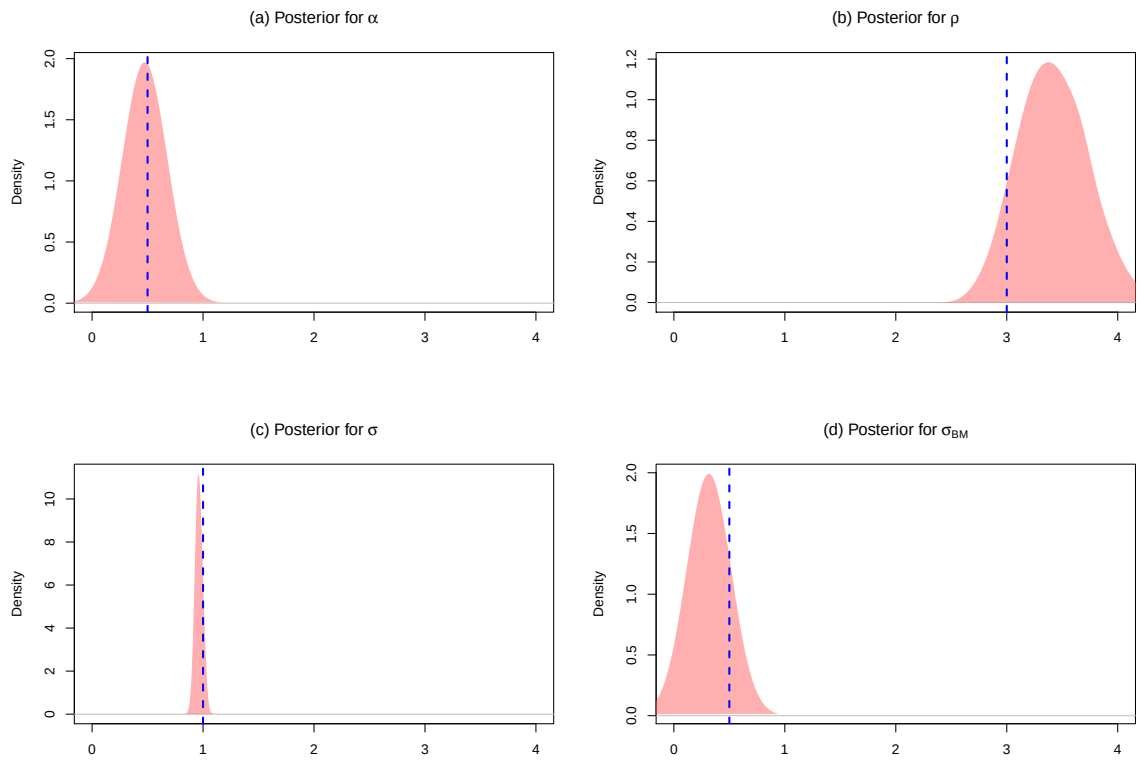


Fig. 6.10 Posterior densities for model parameters for the simulated sinusoidal data with 144 observations, based on the Markov chain Monte Carlo runs of 20,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Posterior density for $\rho$, the variance coefficient of the leading point. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion).

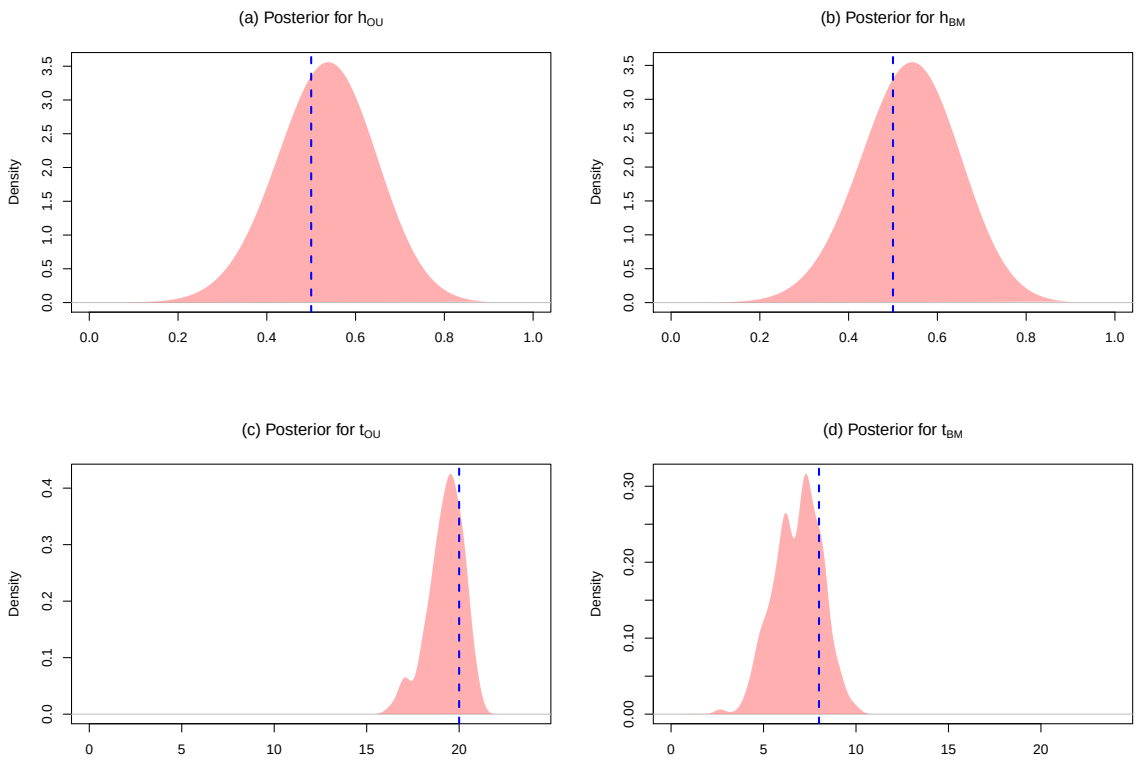


Fig. 6.11 Posterior densities for model's switching parameters for the simulated sinusoidal data with 144 observations, based on the Markov chain Monte Carlo runs of 20,000 iterations. The dashed blue line in each case represents the true parameter value; (a) Posterior density for $h_{ou}$, the rate of switching from OU to BM. (b) Posterior density for $h_{bm}$, the rate of switching from BM to OU. (c) Posterior density for $t_{ou}$, the most likely time of switch out of OU. (d) Posterior density for $t_{bm}$, the most likely time of switch out of BM.

We find that a higher sampling frequency does greatly improve the parameter estimation. From Table 6.3 we can see that all of the parameter point estimates are close to their true values with a reasonably small amount of uncertainty. There is a significant improvement in the posterior estimate of $t_{ou}$, shown in Figure 6.11, whose point estimate, 19.27, is very close to the true value of 20. There are still exist instances where the model fails to capture the correct behavioural state with strong certainty but largely the model performs well.

The main purpose of including ancillary data is to learn about the effects that covariates have on movement. However, including covariate information in the model may also improve inference, especially in the light of dynamic switching and seasonality which may be difficult to estimate using the standard non-covariate model given in Chapter 3. Section 6.2.4 gives

a comparison between the performance of the covariate model presented here and the non-covariate model of Chapter 3.

### 6.2.4 Comparison with non-covariate Model

To investigate how the covariate model performs relative to the non-covariate switching model given in Chapter 3, we fit the non-covariate model to the same simulated data of Section 6.2.3. We compare the density of state estimations faceted by the true state. In other words, we plot the posterior densities for the states separated by whether the true state was either 1 or 2 (OU or BM). Figure 6.12 shows the density of the state estimation faceted by the true state when the non-covariate switching model was used. Similar plots for the covariate switching model are give in Figure 6.13.

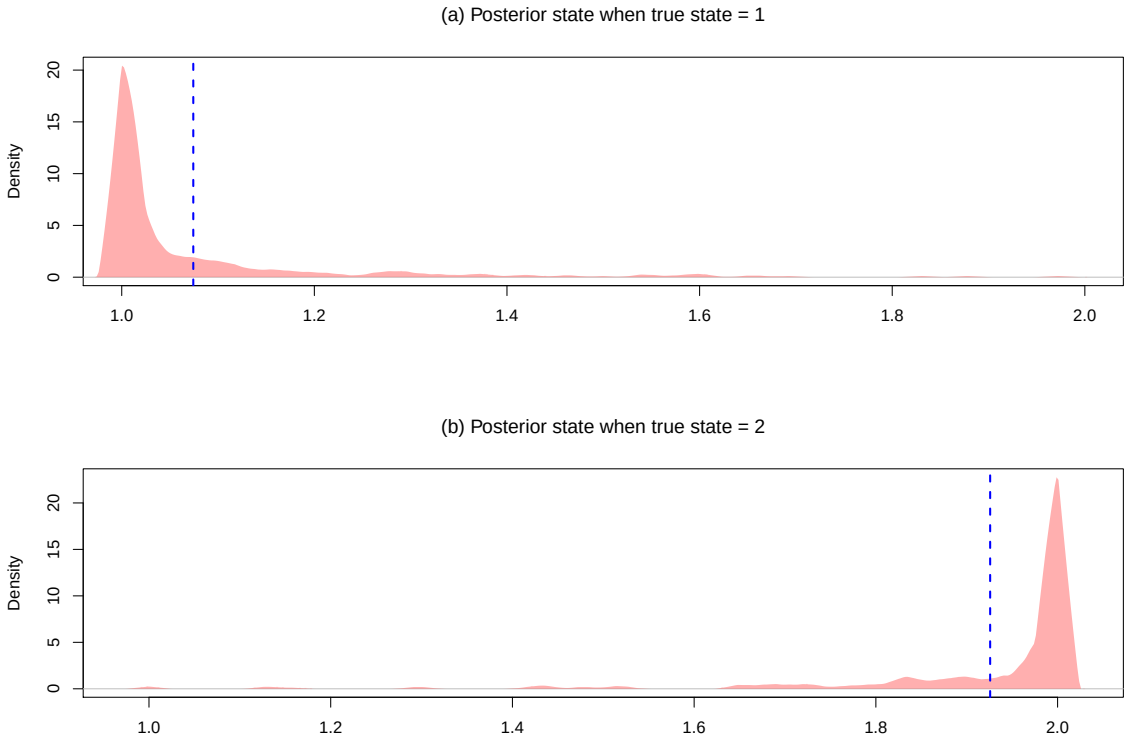


Fig. 6.12 Posterior densities for the state estimates of the simulated sinusoidal data using the non-covariate switching model. (a) Shows the posterior densities for the state estimations when the true state is 1. (b) Shows the posterior densities for the state estimations when the true state is 2. The dashed blue line signifies the mean of the posterior density.

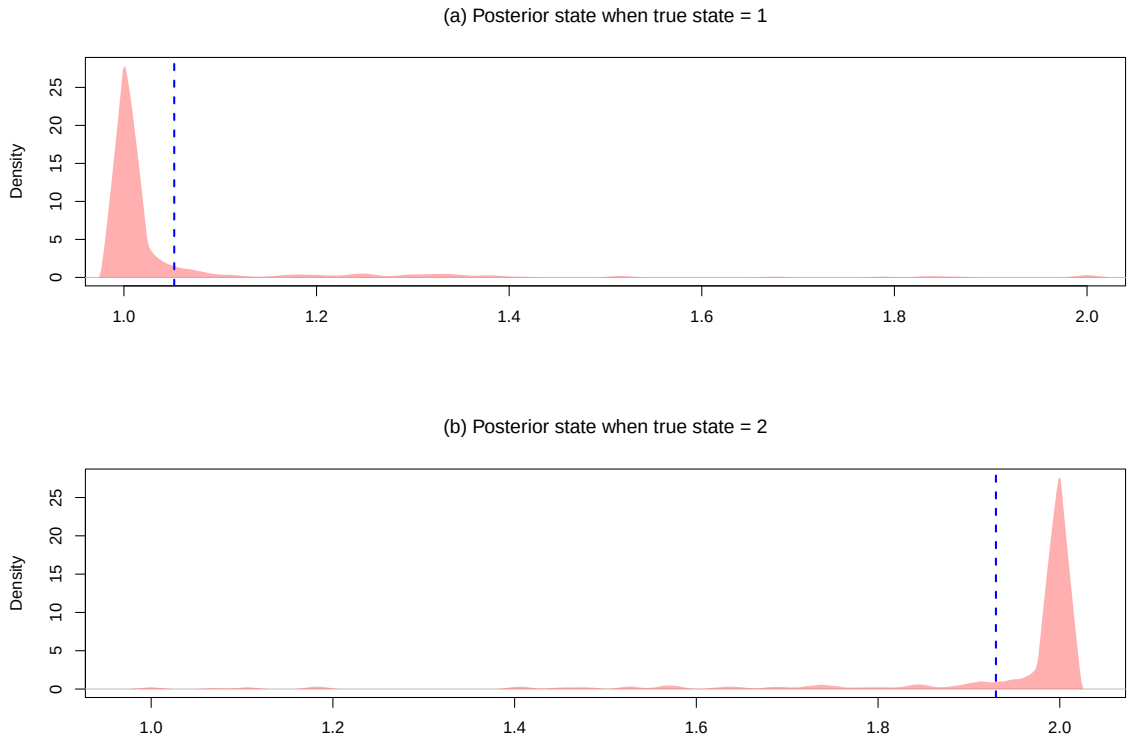


Fig. 6.13 Posterior densities for the state estimates of the simulated sinusoidal data using the covariate switching model. (a) Shows the posterior densities for the state estimations when the true state is 1. (b) Shows the posterior densities for the state estimations when the true state is 2. The dashed blue line signifies the mean of the posterior density.

By comparing the density plots in Figure 6.12 and Figure 6.13 we can see that the covariate model outperforms the non-covariate model. In the covariate case, the density is higher around the true value with less weight around other values. The posterior mean for each state is also closer to the true value in the the covariate case.

However, the original non-covariate model still does rather well. The model successfully estimates the movement parameters whose point estimates and standard deviations are given in Table 6.4. The table also displays the estimated switching rates. Note that the data was simulated using the covariate model given in Equation 6.5 and so the 'true values' in this case are not meaningful.

| Parameter | Point estimate | Standard deviation | True value |
|---|---|---|---|
| $\alpha$ | 0.48 | 0.03 | 0.5 |
| $\rho$ | 3.46 | 0.36 | 3 |
| $\sigma$ | 0.95 | 0.03 | 1.0 |
| $\sigma_{\text{BM}}$ | 0.29 | 0.01 | 0.5 |
| $\lambda_{1,2}$ | 0.36 | 0.05 | |
| $\lambda_{2,1}$ | 0.42 | 0.06 | |

Table 6.4 Parameter estimates for the movement and switching model with simulated covariate dataset.

If your data is plentiful, as in this example with frequent observations over long durations, then it is possible that the non-covariate model will do a good job at estimating the states and model parameters. However, this is useful if you are only interested in fitting a model to the data rather than understanding the mechanisms driving movement, in this case the time of day. To illustrate this we simulate data from the point estimates of Table 6.4 and Table 6.3.

Figure 6.14 shows the states of three days worth of 30 minute observations by way of forward simulation using the point estimates of the covariate model whilst Figure 6.15 shows a similar plot for the non-covariate model.

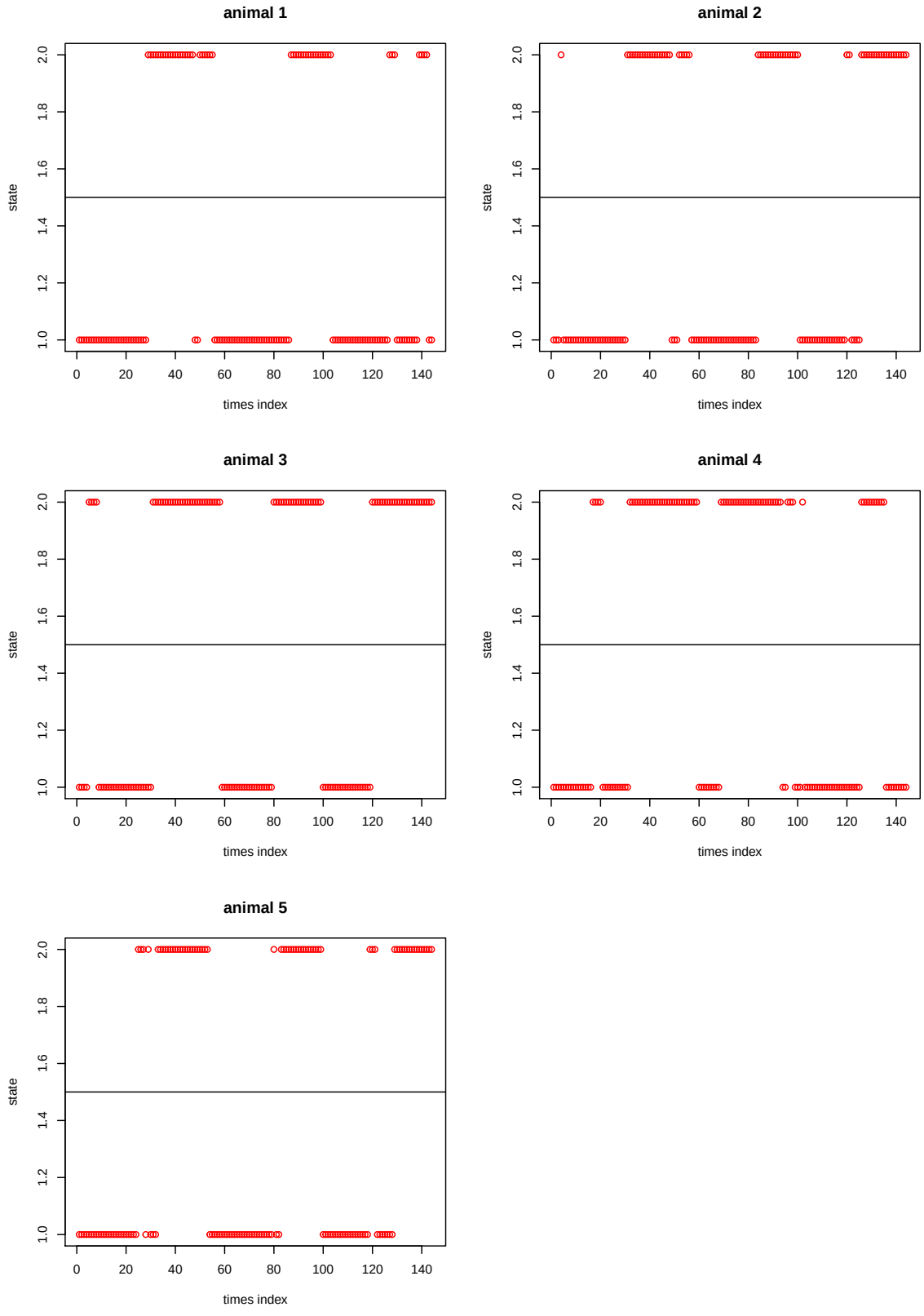


Fig. 6.14 Simulated states using point estimates from the sinusoidal covariate model (Table 6.3). The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion.

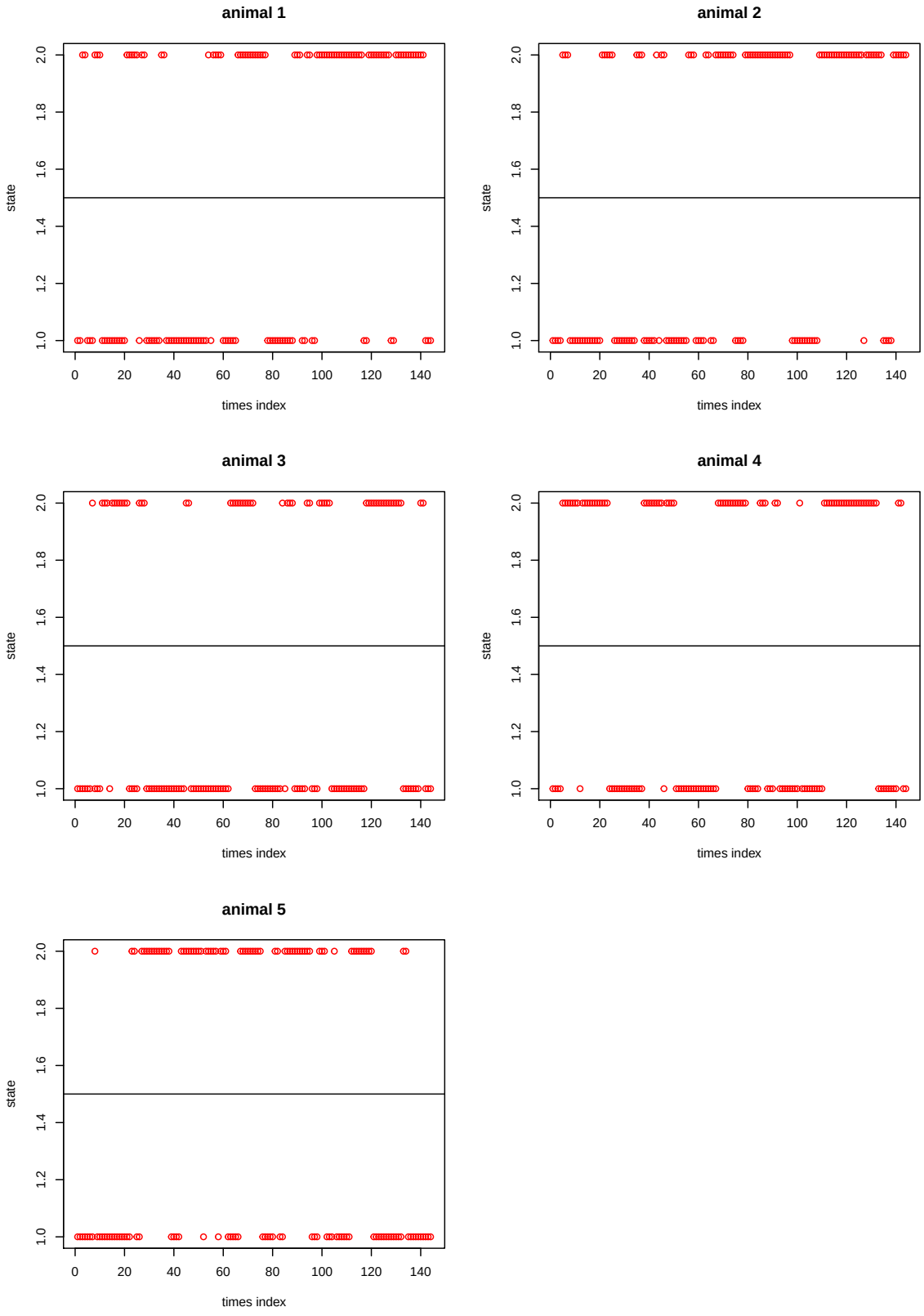


Fig. 6.15 Simulated states using point estimates from the non-covariate switching model (Table 6.4). The circles (red) represent the true states of the follower. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion.

In Figure 6.14 we can see that simulation via the covariate model and its parameter estimates produce strong daily cycles which drive the switching between states. In comparison, Figure 6.15 illustrates that whilst the non-covariate model is able to estimate the overall proportion of time spent in each state, it assumes that the switching rates are static. Thus, it fails to capture the daily patterns and over estimates the amount of switching.

### 6.2.5 Conclusions

So far in this chapter we have highlighted the importance of including covariate information in movement models. We have discussed that an animal's behavioural choices are often governed by stimuli such as weather, terrain, time of day and proximity to human activity. As a consequence, the rate in which animals change behaviour can be a dynamic process, possibly exhibiting patterns or seasonality.

Here we have focussed on developing a model where the rate of switching varies depending on the time of day. This has realistic ecological application since we can imagine the possible scenarios of collective movement such as animals grouping during the night for safety then foraging individually in daylight.

We have explored a variety of simulation experiments to assess the limitations of our model. As expected, the main conclusions are that the model performs better if there are more observations, that is, either longer in duration or more frequent. This can probably be said of any model, that the more information it has the better it can estimate the model parameters.

The functionality of the covariate model was promoted by a comparison between model fitting in the covariate and non-covariate case. Whilst the non-covariate model performed surprisingly well in estimating the movement parameters and states, it failed to capture the true daily cycles which were present in the data. By using the non-covariate model we limit ourselves to accurate model fitting only when the switching rates are static which in many scenarios is not realistic. By allowing the switching rates to vary depending on external information we potentially expand our understanding to the mechanisms driving movement.

However, implementing covariate information is not without computational cost. We found that inference with the covariate model of the data in Section 6.2.3 took around twice the time as that of the covariate model with the same data. In this case the additional time is not too extreme but we foresee this becoming a limitation of the model in the light of much larger datasets. One possible solution for this is to initially estimate the states by fitting the non-covariate model to the data, then use those estimated states as fixed in the covariate

model to estimate the switching parameters. By not updating the trajectory again we are able to gain estimates of the switching parameters whilst saving on time. Although, for obvious reasons this is should be avoided if possible since by doing this we lose out on the benefits of the covariate model to estimate the states.

We have presented general methodology for implementing temporal covariate information. As with most models, accurate estimations of parameters are more difficult when data is limited but we have shown that with just a couple of additional days of data the estimations improve greatly. Although the model fitting process is slower than without covariate information we feel that the extra knowledge and understanding gained is worth the additional computational effort.

The method is very flexible and in principle allows for many different biological applications. In this section we demonstrated its use in modelling seasonality and cyclic behaviour but we will show its applications to weather and insect harassment covariates in the next section.

## 6.3 Revisiting Reindeer Grouping Strategies for Insect Harassment

In this section we revisit the real reindeer location data given in Section 5.2. The aim of this section is to illustrate the implementation of real covariate data, specifically mosquito harassment. Following the method of Section 6.1, we explicitly allow the switching rates to be modelled as functions of mosquito harassment indices, which we derive from proxy weather data described in Section 5.2. We supplement this research by comparing two different models of mosquito harassment, one discrete and one continuous. The discrete case uses weather thresholds described in Skarin et al. (2010) and explored in Chapter 5 whilst the continuous model is proposed from observational findings in Russell et al. (1993).

### 6.3.1 Discrete Model of Mosquito Harassment Covariate

As a discrete model of mosquito harassment at time $t$, we allow the covariate $Z_t$ to take a binary value in $\{0,1\}$ based on the weather data thresholds presented in Skarin et al. (2010) and Chapter 5. A value of 0 would indicate that the weather factors are outside of the thresholds and we assume there is negligible mosquito harassment, whereas, a value of 1 would indicate likely mosquito presence and thus harassment. Mathematically, we can write the covariate $Z_t$ as the indicator function

$$Z_t = \begin{cases} 1, & \text{if} \quad 7 < T < 17, \quad W < 7, \\ 0, & \text{otherwise}, \end{cases}$$

where $T$ is the ambient temperature given in degrees Celsius $^{\circ}C$ and $W$ is the wind speed measured in metres per second, $ms^{-1}$.

Then, to implement this covariate we parametrise the switching rates as follows.

$$\lambda_{12}(t) = \begin{cases} \lambda_{12}^{P}, & \text{if} \quad Z_t = 1, \\ \lambda_{12}^{A}, & \text{if} \quad Z_t = 0, \end{cases}$$

$$\lambda_{21}(t) = \begin{cases} \lambda_{21}^{P}, & \text{if} \quad Z_t = 1, \\ \lambda_{21}^{A}, & \text{if} \quad Z_t = 0, \end{cases}$$

where the probability of an actual switch is given as

$$P(\text{actual switch}) = \frac{n_1\lambda_{12}(t) + n_2\lambda_{21}(t)}{\kappa},$$

and

$$\kappa \geq n \max\{\lambda_{12}^{P,max}, \lambda_{12}^{A,max}, \lambda_{21}^{P,max}, \lambda_{21}^{A,max}\}.$$

Heuristically, we consider the parameters $\lambda_{i,j}^{P}$ and $\lambda_{i,j}^{A}$ as the switching rate from state $i$ to $j$ in the presence and absence of mosquitoes respectively.

## 6.4 Discrete Covariate Results

In a similar vein to Chapter 5, we will apply the covariate model described in Section 6.3.1 to two data sets, peak summer and early summer, described in Section 5.2. The mosquito harassment covariate for each potential switch is calculated using the indicator function given Section 6.3.1. The temperature and wind speed data are downloaded from the Swedish Meteorological and Hydrological Institute (http://www.smhi.se/), collected from the closest weather station to the study site, Korsvattnet A. Since the observations of temperature and wind speed are recorded hourly, it is unlikely that the timing of each potential switch will be at the exact time of the weather observations. That said, for every potential switching time we estimate the weather variables, temperature and wind speed, by linearly interpolating between the observations immediately before and after the potential switch time.

### 6.4.1 Early summer results with discrete covariate

We apply the discrete covariate model described in Section 6.3.1 to the early summer data set. We use the covariate MCMC algorithm (discussed in Section 6.1.1) with 50,000 iterations and the initial 70% counting as burn-in. This took approximately 72 hours to complete. The posterior point estimates and standard deviations of the model parameters are shown in Table 6.5. The posterior mean of the behavioural states are plotted in Figure 6.16. The posterior densities for each of the diffusion parameters are given in Figure 6.17 whilst the posterior densities for the switching rates are given in Figure 6.18.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.05 | 0.01 |
| $\rho$ | 3 | 0 |
| $\sigma$ | 1.15 | 0.08 |
| $\sigma_{\mathrm{BM}}$ | 0.27 | 0.02 |
| $\lambda_{1,2}^{P}$ | 0.368 | 0.063 |
| $\lambda_{1,2}^{A}$ | 0.414 | 0.055 |
| $\lambda_{2,1}^{P}$ | 0.134 | 0.026 |
| $\lambda_{2,1}^{A}$ | 0.129 | 0.023 |

Table 6.5 Parameter estimates for the movement and switching in the early summer data set using the discrete covariate switching model.

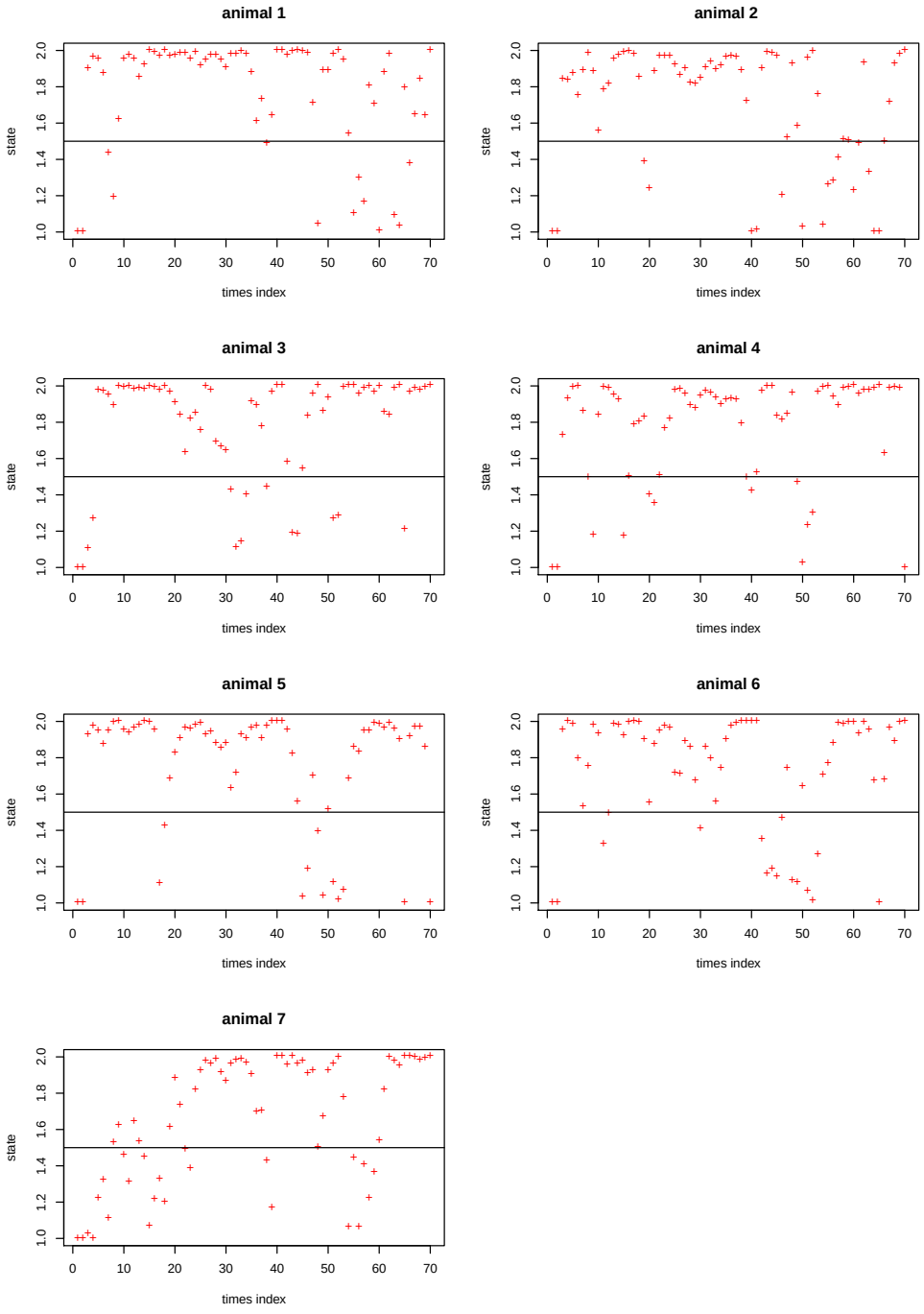


Fig. 6.16 Posterior mean states of all followers for the early summer data set using the discrete covariate switching model. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (red) represent the mean posterior of the estimated behaviour states.

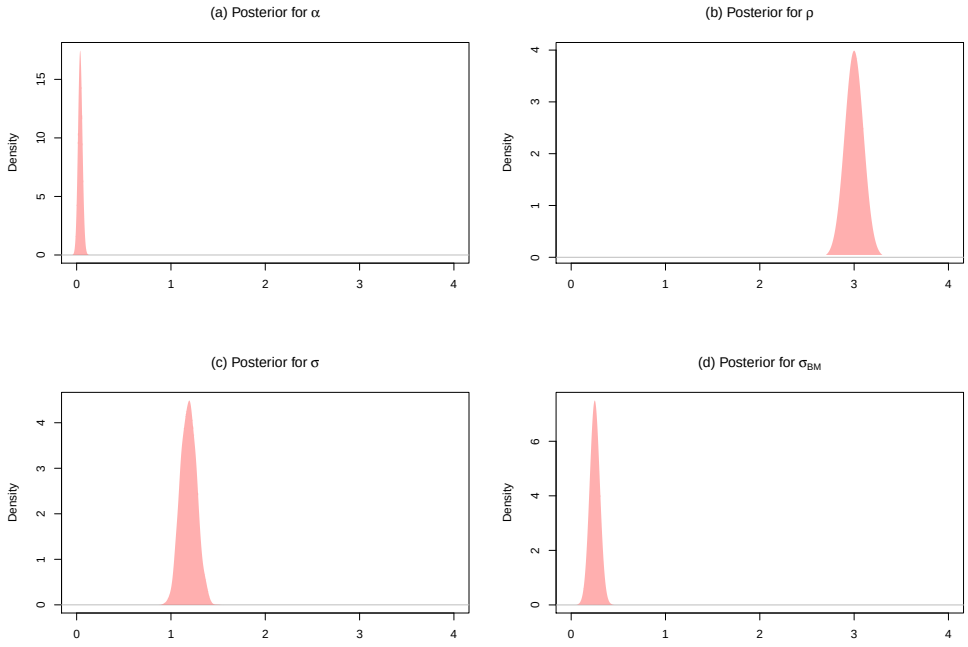


Fig. 6.17 Posterior densities of the model parameters for the early summer data set using the discrete covariate switching model, based on the Markov chain Monte Carlo runs of 50,000 iterations. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Approximate density for $\rho$, the variance coefficient of the leading point, which is fixed at 3. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion).

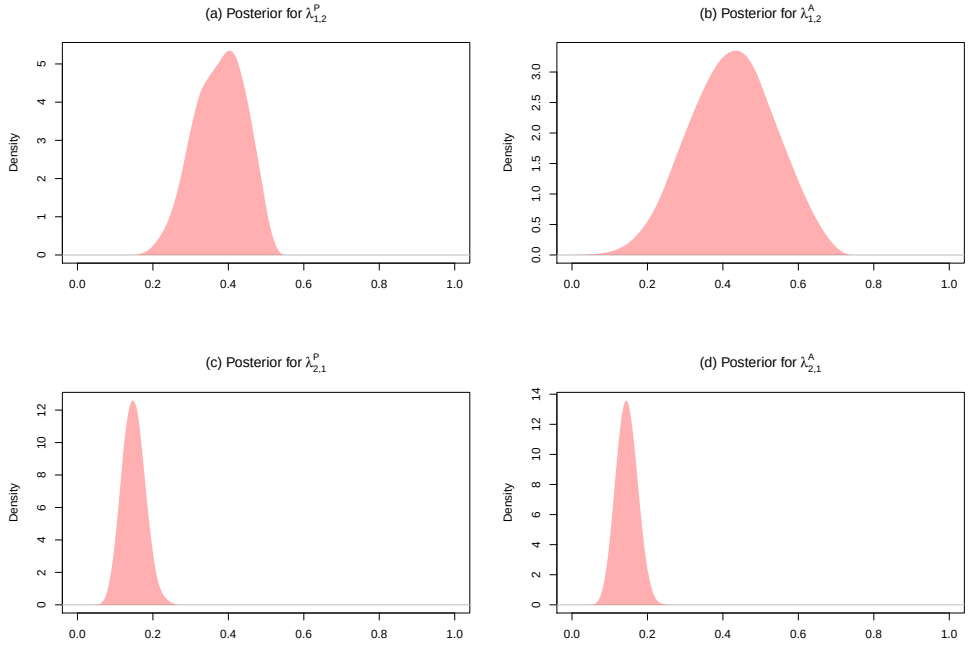


Fig. 6.18 Posterior densities of the switching parameters for the early summer data set using the discrete covariate switching model, based on the Markov chain Monte Carlo runs of 50,000 iterations. (a) Posterior density for $\lambda^P_{1,2}$, the switching rate of the follower from OU to BM in the presence of insects. (b) Posterior density of $\lambda^A_{1,2}$, the switching rate of the follower from OU to BM in the absence of insects. (c) Posterior density for $\lambda^P_{2,1}$, the switching rate of the follower from OU to BM. in the presence of insects. (d) Posterior density of $\lambda^A_{2,1}$, the switching rate of the follower from BM to OU in the absence of insects.

### 6.4.2 Peak summer results with discrete covariate

Here, we apply the discrete covariate model described in Section 6.3.1 to the peak summer data set. We use the covariate MCMC algorithm with 50,000 iterations and the initial 70% counting as burn-in. This took approximately 72 hours to complete. The posterior point estimates and standard deviations of the model parameters are shown in Table 6.6. The posterior means of the behavioural states are plotted in Figure 6.19. The posterior densities for each of the diffusion parameters are given in Figure 6.20 whilst the posterior densities for the switching rates are given in Figure 6.21.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.195 | 0.01 |
| $\rho$ | 3 | 0 |
| $\sigma$ | 0.604 | 0.025 |
| $\sigma_{\mathrm{BM}}$ | 0.645 | 0.066 |
| $\lambda_{1,2}^{P}$ | 0.087 | 0.017 |
| $\lambda_{1,2}^{A}$ | 0.119 | 0.038 |
| $\lambda_{2,1}^{P}$ | 0.237 | 0.042 |
| $\lambda_{2,1}^{A}$ | 0.204 | 0.063 |

Table 6.6 Parameter estimates for the movement and switching rates in the peak summer data using the discrete covariate switching model.

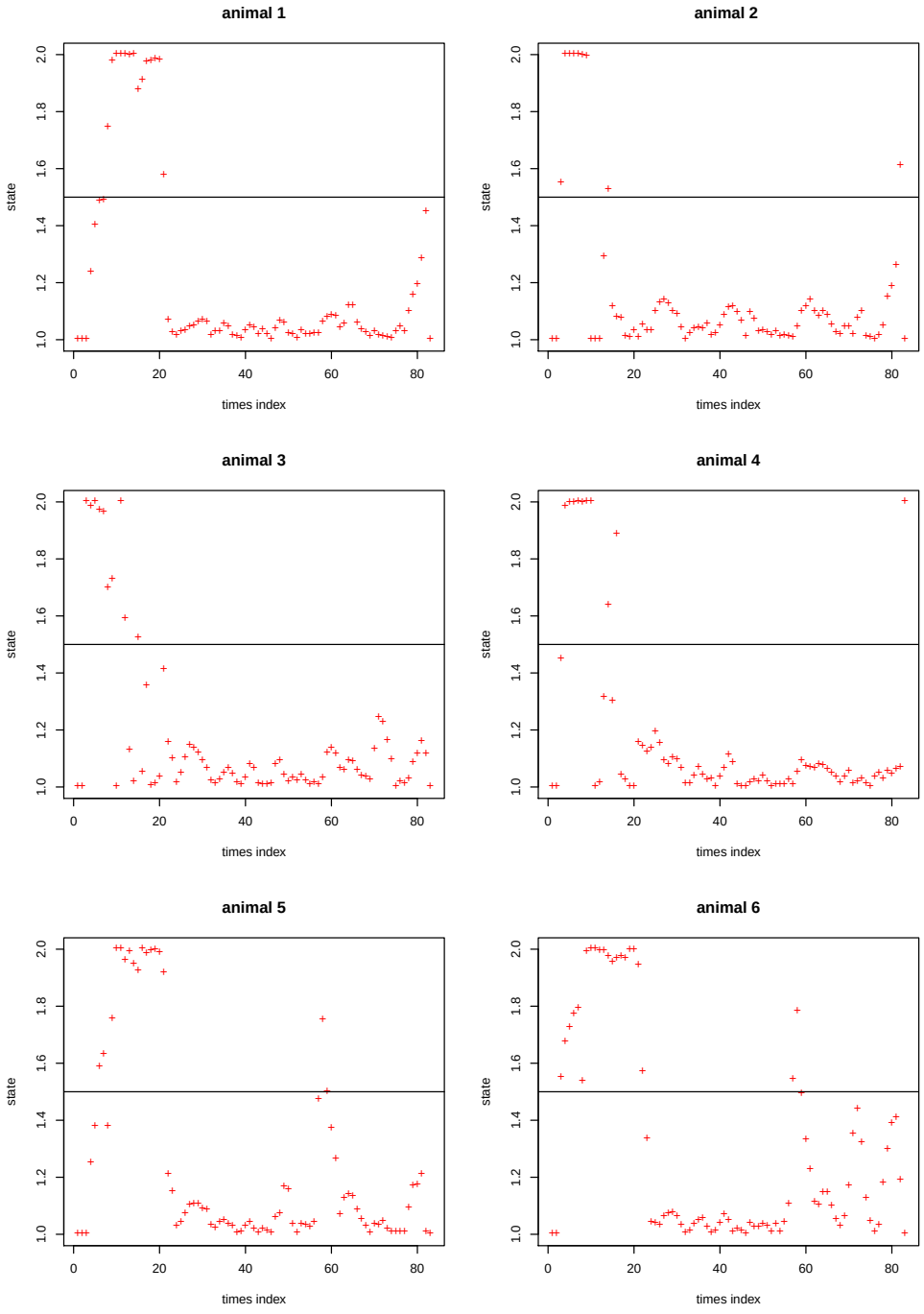


Fig. 6.19 Posterior mean states of all followers for the peak summer data using the discrete covariate switching model. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (red) represent the mean posterior of the estimated behaviour states.

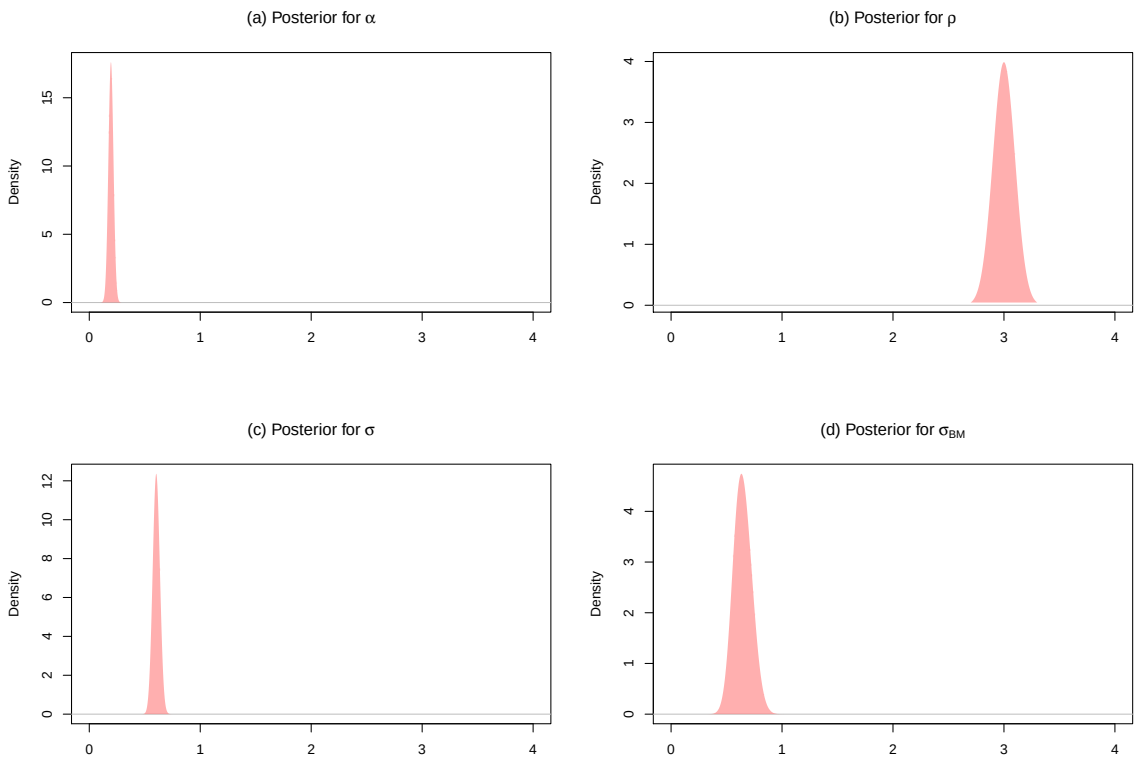


Fig. 6.20 Posterior densities of the model parameters for the peak summer data using the discrete covariate switching model, based on the Markov chain Monte Carlo runs of 50,000 iterations. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Approximate density for $\rho$, the variance coefficient of the leading point, which is fixed at 3. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion).

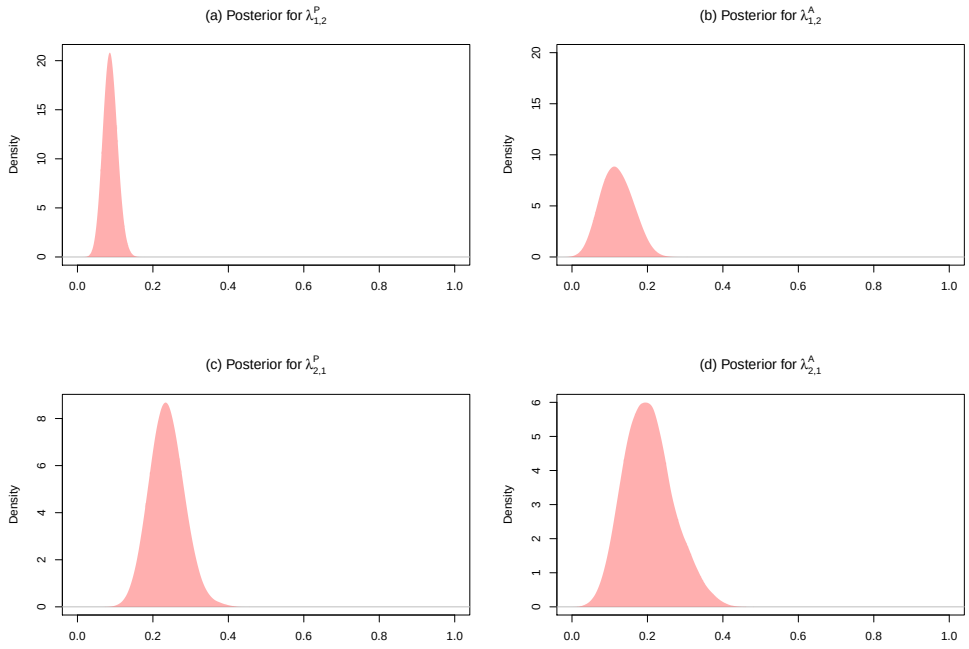


Fig. 6.21 Posterior densities of the switching parameters for the peak summer data using the discrete covariate switching model, based on the Markov chain Monte Carlo runs of 50,000 iterations. (a) Posterior density for $\lambda^P_{1,2}$, the switching rate of the follower from OU to BM in the presence of insects. (b) Posterior density of $\lambda^A_{1,2}$, the switching rate of the follower from OU to BM in the absence of insects. (c) Posterior density for $\lambda^P_{2,1}$, the switching rate of the follower from OU to BM. in the presence of insects. (d) Posterior density of $\lambda^A_{2,1}$, the switching rate of the follower from BM to OU in the absence of insects.

### 6.4.3 Discussion of Discrete Covariate Results

Looking at Table 6.5, we can see that the parameter estimates of $\lambda^P_{1,2}$ and $\lambda^A_{1,2}$, the switching rates from Brownian motion to Ornstein Uhlenbeck are typically larger (0.368 and 0.414 respectively) than that of the opposing parameters, $\lambda^P_{2,1}$ and $\lambda^A_{2,1}$ (0.134 and 0.129). This is consistent with the notion that in the early summer period, the reindeer are less grouped due to low parasitic pressures. However, it is surprising that there is little variation between parameter estimates within each switching direction. Given that we expect the presence of mosquitoes to promote the grouping behaviour of reindeer, we would anticipate more contrasting estimates; within the same switching direction we would expect $\lambda^P_{1,2} < \lambda^A_{1,2}$ and $\lambda^A_{2,1} < \lambda^P_{2,1}$ whereas, between switching directions we would expect $\lambda^P_{1,2} < \lambda^P_{2,1}$ and $\lambda^A_{2,1} < \lambda^A_{1,2}$. Whilst most of these conditions do hold, the differences between parameters are quite often relatively small.

A similar conclusion can be drawn from the peak summer results. In Table 6.6 , we can see that the parameters of the switching rates from grouped behaviour to moving independently, $\lambda^P_{2,1}$ and $\lambda^A_{2,1}$, are estimated to be 0.237 and 0.204 respectively, whereas parameters for switching in the other direction, $\lambda^P_{1,2}$ and $\lambda^A_{1,2}$, are estimated to be 0.087 and 0.119 respectively. In addition, the estimates of all the movement parameters for both datasets are similar to the corresponding results when the non-covariate model was applied in Chapter 5.

Arguably, by using a discrete, binary model of mosquito presence, we lose information about the severity of harassment. Around the thresholds for mosquito activity there may be little or no harassment as the conditions for flying insects are sub-optimal. Alternatively, by using a continuous model we can give a more detailed estimate of the harassment levels using proxy climate data, which can reflect intermediate levels of severity rather than just the extreme values. In the next section we review the use of a continuous model of mosquito harassment which may take any real value between 0 and 1.

## 6.5 Continuous Model of Mosquito Harassment Covariate

In order to alleviate the issues caused by discretely modelling mosquito harassment, it seems natural to investigate adopting a continuous approach. In this section we will use the model of mosquito index presented in Russell et al. (1993). In this case, the mosquito index is formulated as the product of a temperature index, $ti$, and wind speed index, $wi$. These are both continuous, linear interpolations between the values 0 and 1 and are based upon observational findings. Specifically,

$$ti = \begin{cases} 0 & T < 6, \\ \text{1-((18-T)/13)} & 6 \leq T \leq 18, \\ 1 & T > 18, \end{cases} \tag{6.6}$$

$$wi = \begin{cases} 0 & W > 6, \\ \text{(6-W)/6,} & W \leq 6, \end{cases} \tag{6.7}$$

where $T$ is the ambient temperature given in degrees Celsius $^\circ C$ and $W$ is the wind speed measured in metres per second, $ms^{-1}$. Then, the mosquito index is given as

$$MI = ti * wi. \tag{6.8}$$

In comparison to the thresholds of activity presented in Skarin et al. (2010), these constraints are less restrictive. They do not consider an upper temperature bound for mosquito presence; on the contrary, they suggest that higher temperatures in fact promote the presence of mosquitoes. In addition, this continuous modelling approach is able to distinguish between the severity of harassment during optimal and suboptimal conditions whereas Skarin et al. (2010) does not. In other words, using the discrete model, weather data around threshold values such as $7°C$ will take the same covariate value as that of $15°C$, even though it is likely that harassment will be much higher during the latter. Figure 6.22 illustrates differences between the discrete mosquito harassment index presented in Skarin et al. (2010) and the continuous indexing given in Russell et al. (1993). The contour plot represents the mosquito harassment in the continuous case, which may take any real value from 0 (no harassment) to 1 (high harassment). The area contained within the dashed black rectangle indicates the region where the discrete index is taken to be 1 and is 0 elsewhere.

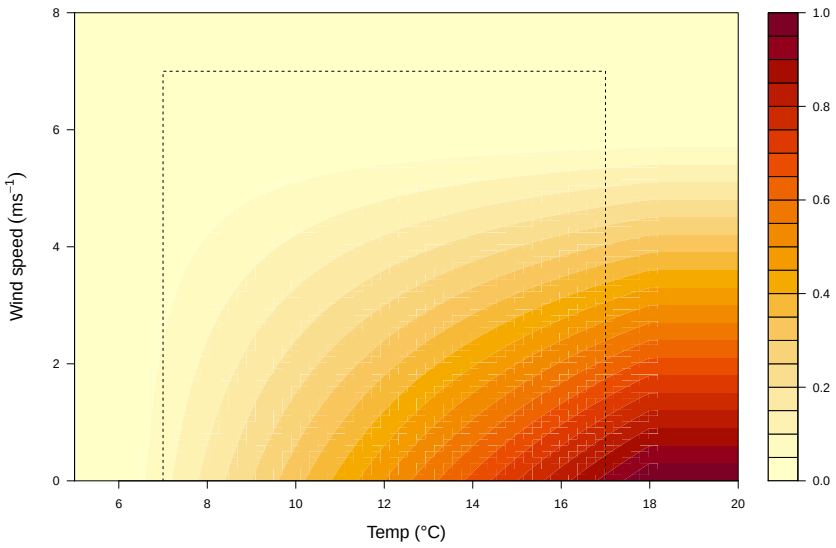


Fig. 6.22 Contour plot of the continuous mosquito harassment index (Russell et al., 1993), a product of temperature and wind speed indices, which may take any real value from 0 (no harassment) to 1 (high harassment). The area contained within the dashed black rectangle indicates the region where the discrete index (Skarin et al., 2010) is taken to be 1 and is 0 elsewhere.

NB: We feel that the denominator in the temperature index (Equation 6.6) should be 12 rather than 13 to avoid discontinuities and allow the index $ti$ to be a linear interpolation from 0 to 1 for the temperature values between $6^{\circ}C$ and $18^{\circ}C$. However, for this thesis we will keep to the literature and use the model presented in Russell et al. (1993) i.e. leaving the denominator as 13. It is worth stressing that we do not expect this to make a huge difference to values of $ti$. In fact, we found that the maximum difference between $ti$ values, when applied to both the early and peak summer data, was 0.08 (to 2 d.p.).

To implement this covariate of mosquito harassment into our framework, we allow the switching rates to be given as a linear function of the covariate at time $t$ as follows:

$$\lambda_{12} = \lambda_{12}^{P} z_t + \lambda_{12}^{A}(1 - z_t),$$
$$\lambda_{21} = \lambda_{21}^{P} z_t + \lambda_{21}^{A}(1 - z_t),$$

where $z_t$ is given as the mosquito index at time $t$, $MI$, calculated using Equation 6.6 - 6.8 and where the probability of an actual switch is given as

$$P(\text{actual switch}) = \frac{n_1\lambda_{12}(t) + n_2\lambda_{21}(t)}{\kappa},$$

where

$$\kappa \geq n \max\{\lambda_{12}^{P,max}, \lambda_{12}^{A,max}, \lambda_{21}^{P,max}, \lambda_{21}^{A,max}\}.$$

Here, the parameters can be assumed to have a consistent interpretation to that discussed in Section 6.3.1.

## 6.6 Continuous Covariate Results

As with Section 6.3.1, we apply the continuous covariate model described in Section 6.5 to the two data sets, peak summer and early summer, described in Chapter 5. The mosquito harassment covariate for each potential switch is calculated as a linear interpolation between the climatic observations immediately before and after the potential switch time.

### 6.6.1 Early Summer Results

We apply the continuous covariate model described in Section 6.6 to the early summer data set. We use the covariate MCMC algorithm (discussed in Section 6.1.1) with 20,000 iterations with the initial 70% counting as burn-in. This took approximately 72 hours to complete. The posterior point estimates and standard deviations of the model parameters are shown in Table 6.7. For this particular application, we choose not to show the posterior densities but instead give trace plots for each parameter in the MCMC algorithm. Trace plots for the MCMC samples of the diffusion and switching parameters are given in Figures 6.23 and 6.24 respectively.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.04 | 0.008 |
| $\rho$ | 3 | 0 |
| $\sigma$ | 1.25 | 0.07 |
| $\sigma_{\text{BM}}$ | 0.23 | 0.01 |
| $\lambda_{1,2}^{P}$ | 0.321 | 0.193 |
| $\lambda_{1,2}^{A}$ | 0.583 | 0.085 |
| $\lambda_{2,1}^{P}$ | 0.249 | 0.102 |
| $\lambda_{2,1}^{A}$ | 0.159 | 0.018 |

Table 6.7 Parameter estimates for the movement and switching in the early summer data set using the continuous covariate switching model.

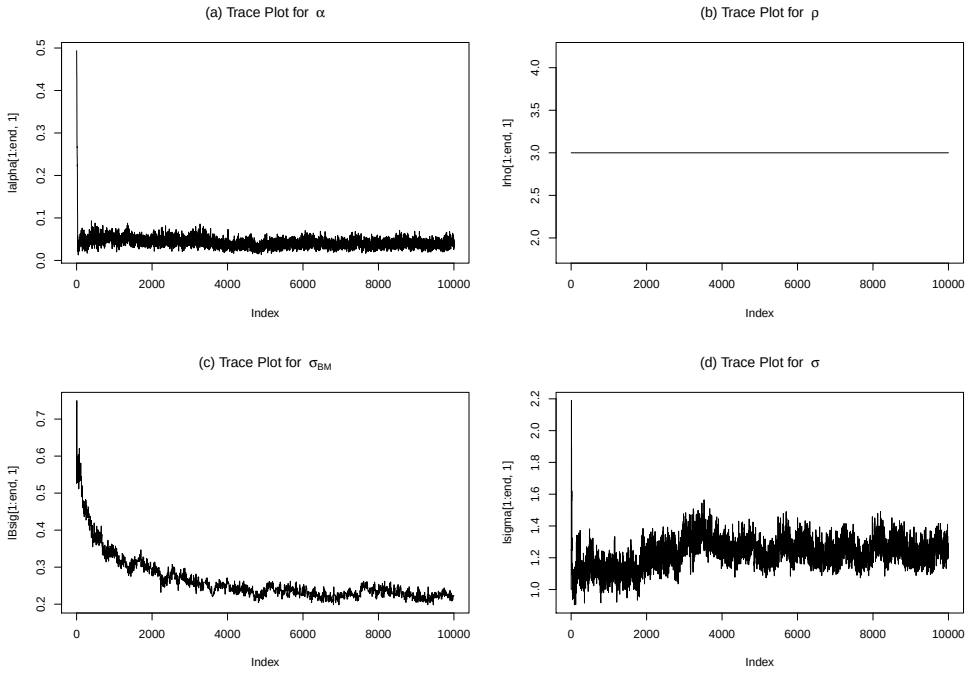


Fig. 6.23 Trace plots for movement parameters of the early summer data using the continuous covariate switching model, based on the MCMC runs of 20,000 iterations. (a) Trace plot for $\alpha$, the attraction rate of the follower to the leading point. (b) Trace plot for $\rho$, the individual variance coefficient of the leading point. This parameter is fixed at 3. (c) Trace plot for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (d) Trace plot for $\sigma$, the individual variance coefficient of the follower.

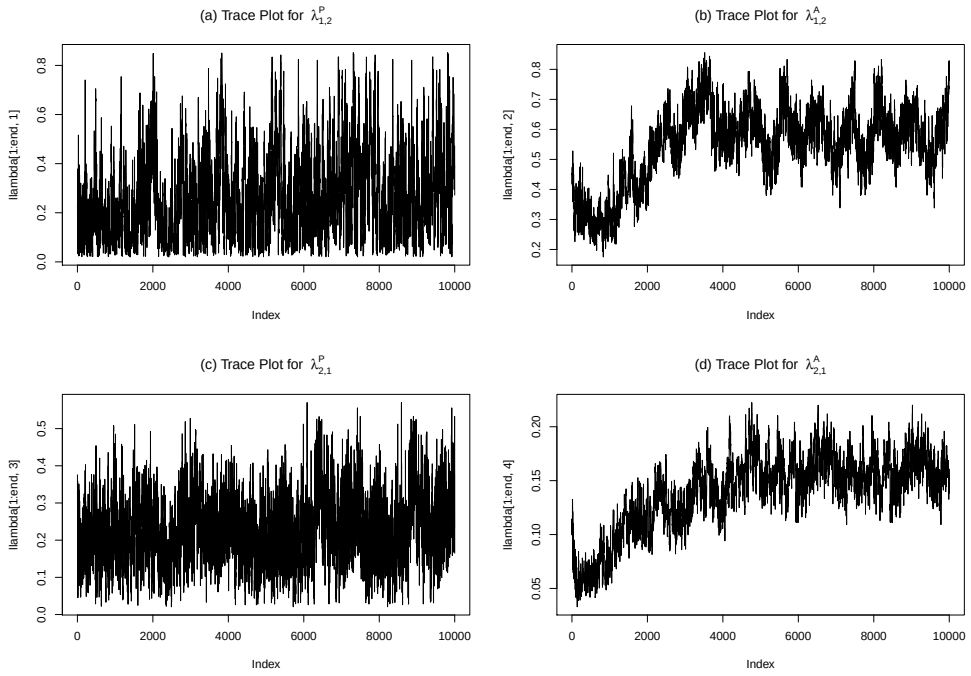


Fig. 6.24 Trace plots for switching parameters of the early summer data using the continuous covariate switching model, based on the MCMC runs of 20,000 iterations. (a) Trace plot for $\lambda^P_{1,2}$, the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion in the presence of insects. (b) Trace plot of $\lambda^A_{1,2}$, the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion in the absence of insects. (c) Trace plot for $\lambda^P_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the presence of insects. (d) Trace plot of $\lambda^A_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the absence of insects.

After experimenting with different proposal standard deviations and initialisations over a number of trial runs of the MCMC algorithm, we found the convergence of some parameters to be quite poor. This is especially evident in the switching parameters given in Figure 6.24. To investigate this further, we carried out a series of convergence diagnostics. A summary table for the autocorrelation function (acf) at time lag 1, 5, 10 and 50, along with the effective sample size (ESS) for each parameter is given in Table 6.8. Plots of the acf for each of the movement and switching parameters are given in Figures 6.25 and 6.25 respectively. Note that the acf for $\rho$ is not given since the parameter is fixed (see Section 5.2.1). Finally, a running mean plot of the switching parameters is given in Figure 6.27. The interpretation of these figures are discussed in Section 6.6.3.

| | acf lag | | | | |
|---|---|---|---|---|---|
| Parameter | 1 | 5 | 10 | 50 | ESS |
| $\alpha$ | 0.46 | 0.11 | 0.09 | 0.07 | 307 |
| $\sigma$ | 0.61 | 0.33 | 0.29 | 0.14 | 90 |
| $\sigma_{BM}$ | 0.92 | 0.73 | 0.63 | 0.43 | 32 |
| $\lambda^P_{1,2}$ | 0.91 | 0.65 | 0.42 | -0.10 | 132 |
| $\lambda^A_{1,2}$ | 0.91 | 0.74 | 0.65 | 0.44 | 23 |
| $\lambda^P_{2,1}$ | 0.78 | 0.42 | 0.31 | 0.17 | 136 |
| $\lambda^A_{2,1}$ | 0.90 | 0.66 | 0.48 | 0.16 | 117 |

Table 6.8 Autocorrelation at lag 1, 5, 10 and 50 and the effective sample size for each parameter in the early summer dataset using the continuous covariate switching model.

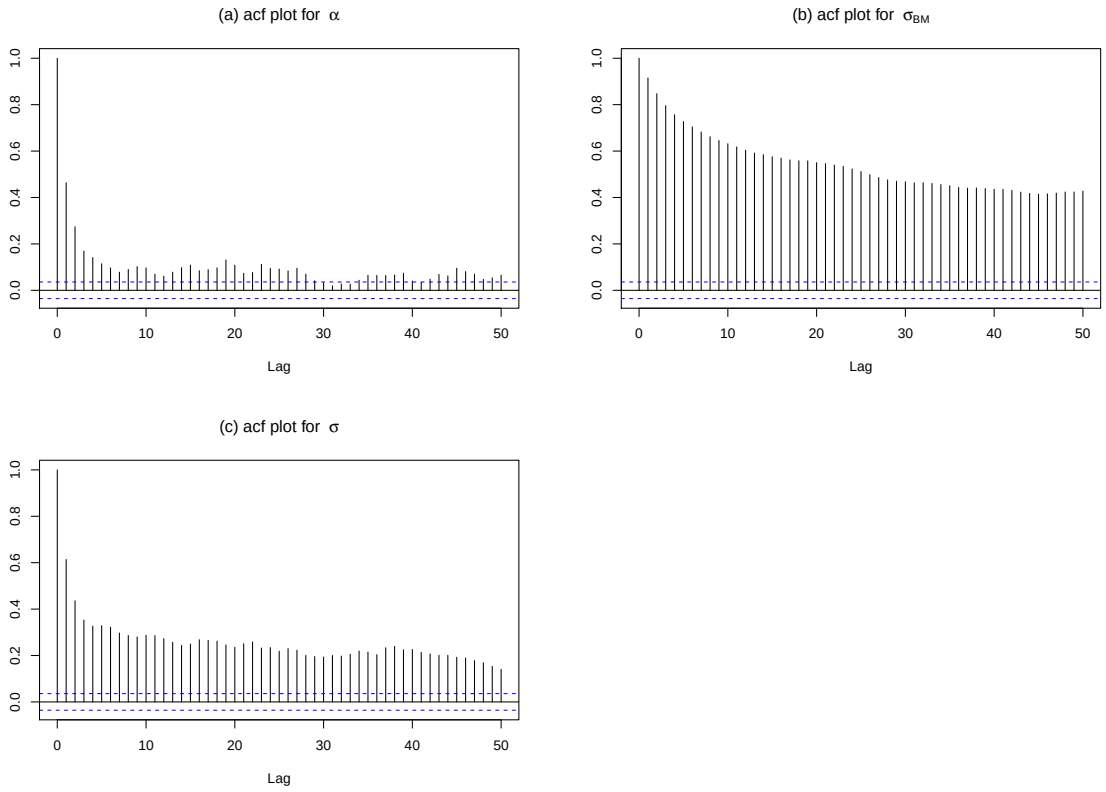


Fig. 6.25 Autocorrelation plots for movement parameters of the early summer data using the continuous covariate switching model, based on the MCMC runs of 20,000 iterations. (a) Autocorrelation plot for $\alpha$, the attraction rate of the follower to the leading point. (b) Autocorrelation plot for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (c) Autocorrelation plot for $\sigma$, the individual variance coefficient of the follower.

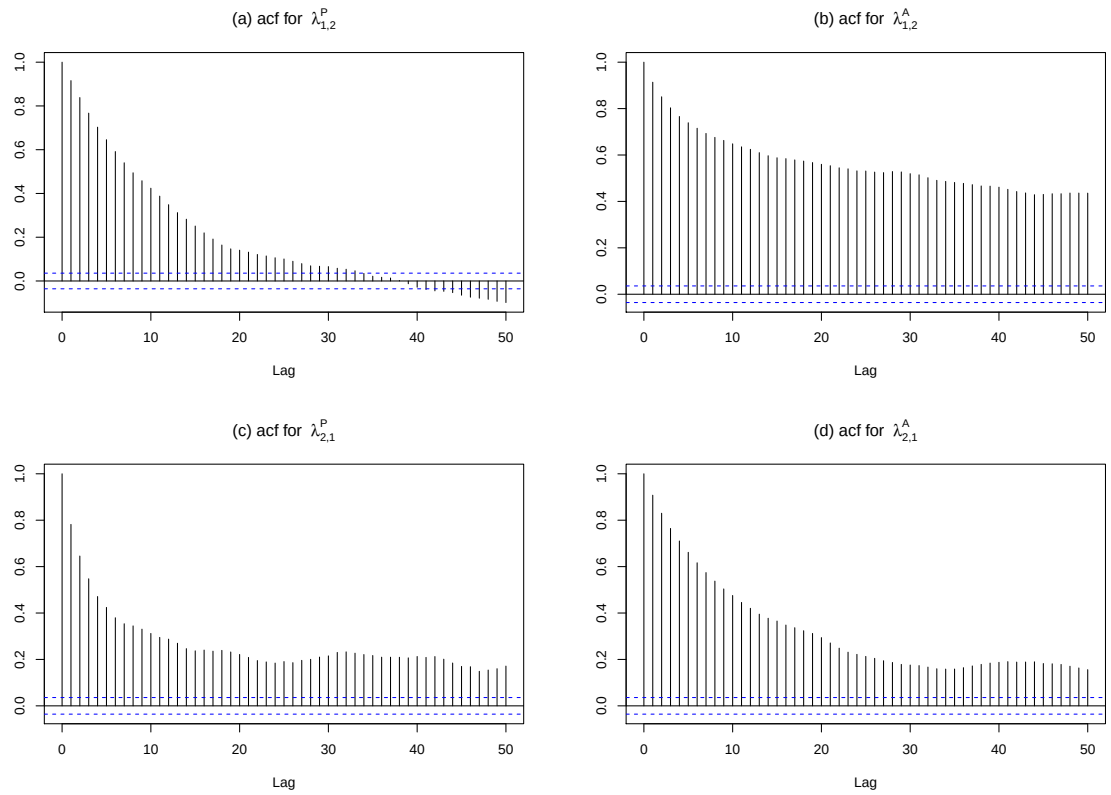


Fig. 6.26 Autocorrelation plots for switching parameters of the early summer data, based on the MCMC runs of 20,000 iterations. (a) Autocorrelation plot for $\lambda^{P}_{1,2}$, the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion in the presence of insects. (b) Autocorrelation plot of $\lambda^{A}_{1,2}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the absence of insects. (c) Autocorrelation plot for $\lambda^{P}_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeckin the presence of insects. (d) Autocorrelation plot of $\lambda^{A}_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the absence of insects.

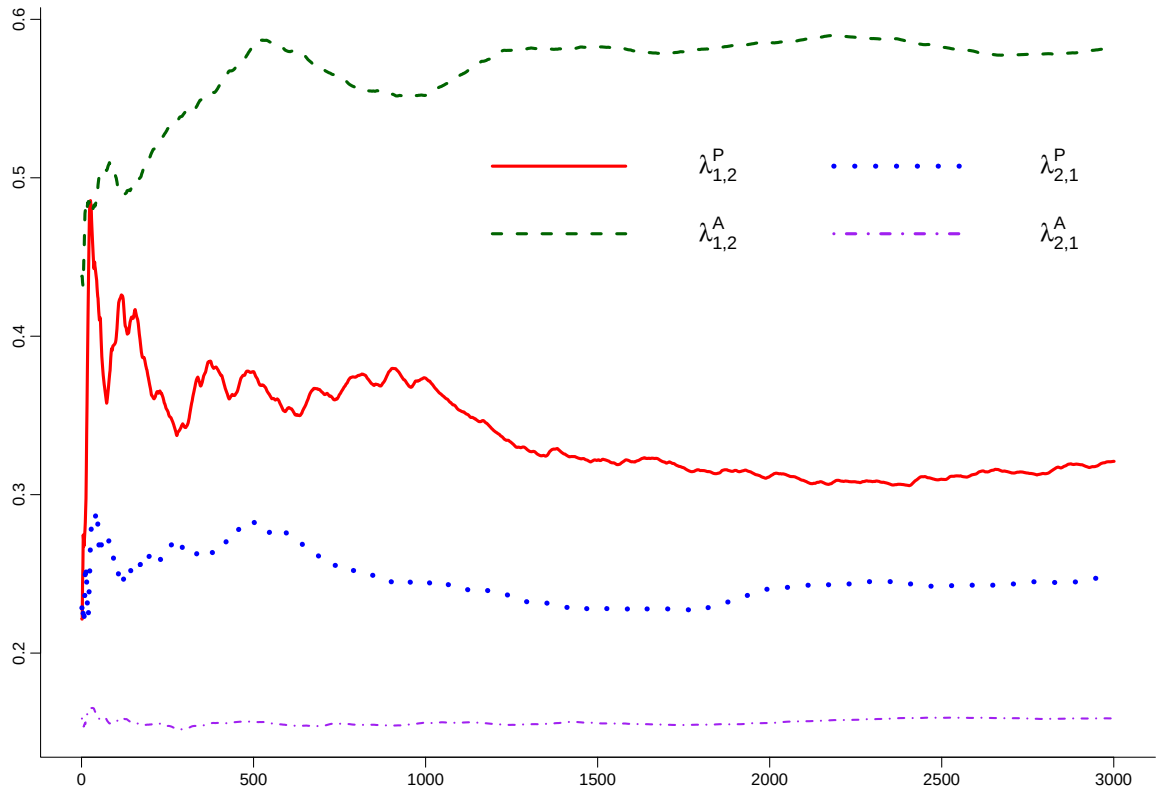


Fig. 6.27 Running mean for the switching parameters of the early summer data using the continuous covariate switching model. The red (solid) line gives the running mean of $\lambda^P_{1,2}$. The green (dashed) line gives the running mean of $\lambda^A_{1,2}$. The blue (dotted) line gives the running mean of $\lambda^P_{2,1}$. Finally, the purple (dot-dash) line gives the running mean of $\lambda^A_{2,1}$.

### 6.6.2 Peak Summer Results

We applied the continuous covariate model described in Section 6.6 to the peak summer data set. We used the covariate MCMC algorithm with 20,000 iterations and the initial 70% counting as burn-in. This took approximately 72 hours to complete. The posterior point estimates and standard deviations of the model parameters are shown in Table 6.9. Trace plots for the MCMC samples of the diffusion and switching parameters are given in Figures 6.28 and 6.29 respectively.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.19 | 0.01 |
| $\rho$ | 3 | 0 |
| $\sigma$ | 0.60 | 0.02 |
| $\sigma_{\mathrm{BM}}$ | 0.67 | 0.08 |
| $\lambda^{P}_{1,2}$ | 0.136 | 0.056 |
| $\lambda^{A}_{1,2}$ | 0.089 | 0.025 |
| $\lambda^{P}_{2,1}$ | 0.179 | 0.124 |
| $\lambda^{A}_{2,1}$ | 0.303 | 0.075 |

Table 6.9 Parameter estimates for the movement and switching rates in the peak summer data using the continuous covariate switching model.

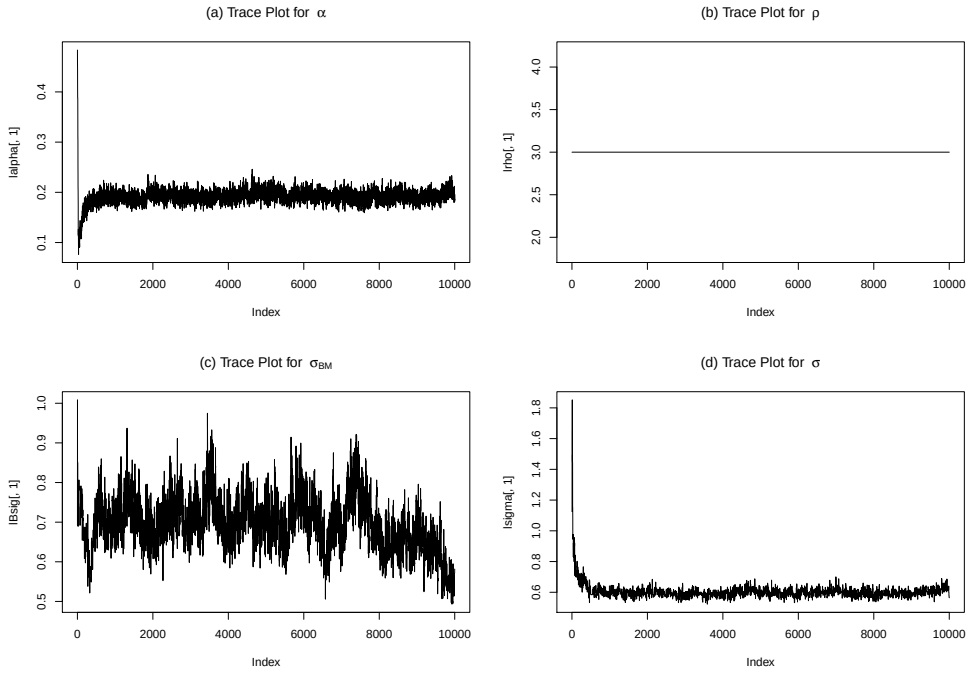


Fig. 6.28 Trace plots for movement parameters of the peak summer data using the continuous covariate switching mode, based on the MCMC runs of 20,000 iterations. (a) Trace plot for $\alpha$, the attraction rate of the follower to the leading point. (b) Trace plot for $\rho$, the individual variance coefficient of the leading point. This parameter is fixed at 3. (c) Trace plot for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (d) Trace plot for $\sigma$, the individual variance coefficient of the follower.

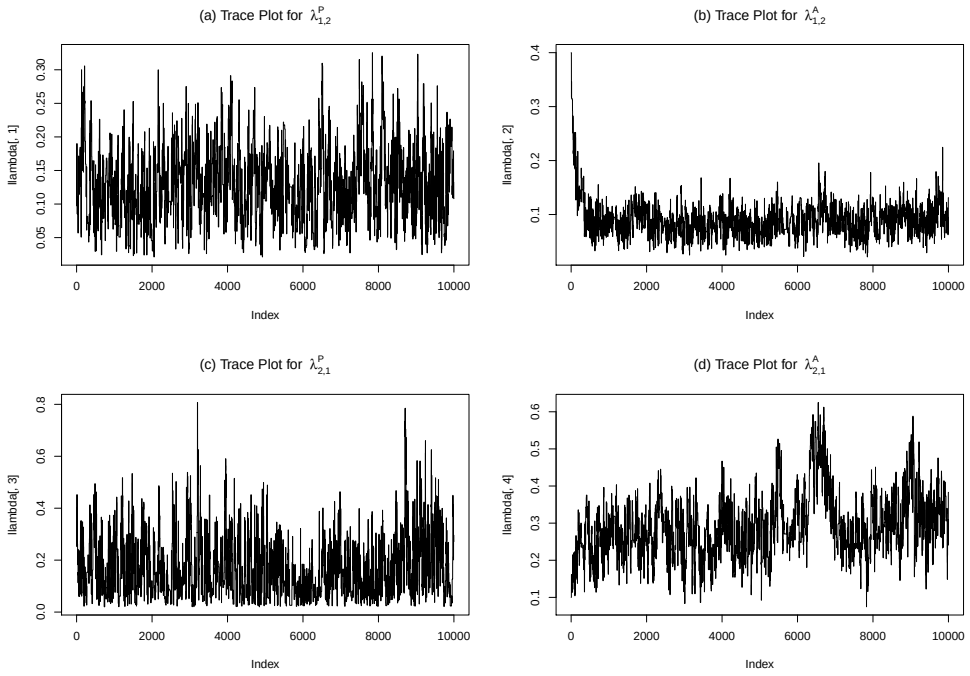


Fig. 6.29 Trace plots for switching parameters of the peak summer data, based on the MCMC runs of 20,000 iterations. (a) Trace plots for $\lambda^P_{1,2}$, the switching rate of the follower from Ornstein Uhlenbeck to Brownian motionin the presence of insects. (b) Trace plot of $\lambda^A_{1,2}$, the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion in the absence of insects. (c) Trace plot for $\lambda^P_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the presence of insects. (d) Trace plot of $\lambda^A_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the absence of insects.

Similar to Section 6.6.1, after a number of trial runs of the MCMC algorithm, we found the convergence of some parameters to be quite poor. In this particular instance, the trace plot for $\sigma_{BM}$ shows some movement towards the very end of the chain which indicates that the MCMC may need additional computing time. A summary table for the autocorrelation function (acf) at time lag 1, 5, 10 and 50, along with the effective sample size (ESS) for each parameter is given in Table 6.10. Plots of the acf for each of the movement and switching parameters are given in Figures 6.30 and 6.31 respectively. Finally, a running mean plot of the switching parameters is given in Figure 6.32.

| | acf lag | | | | |
|---|---|---|---|---|---|
| Parameter | 1 | 5 | 10 | 50 | ESS |
| $\alpha$ | 0.53 | 0.21 | 0.16 | 0.14 | 149 |
| $\sigma$ | 0.81 | 0.44 | 0.31 | 0.20 | 61 |
| $\sigma_{BM}$ | 0.91 | 0.80 | 0.77 | 0.65 | 13 |
| $\lambda^P_{1,2}$ | 0.92 | 0.70 | 0.50 | 0.09 | 101 |
| $\lambda^A_{1,2}$ | 0.85 | 0.52 | 0.38 | 0.18 | 119 |
| $\lambda^P_{2,1}$ | 0.91 | 0.64 | 0.47 | 0.20 | 146 |
| $\lambda^A_{2,1}$ | 0.93 | 0.73 | 0.59 | 0.30 | 71 |

Table 6.10 Autocorrelation at lag 1, 5, 10 and 50 and the effective sample size for each parameter in the peak summer dataset using the continuous covariate switching model.

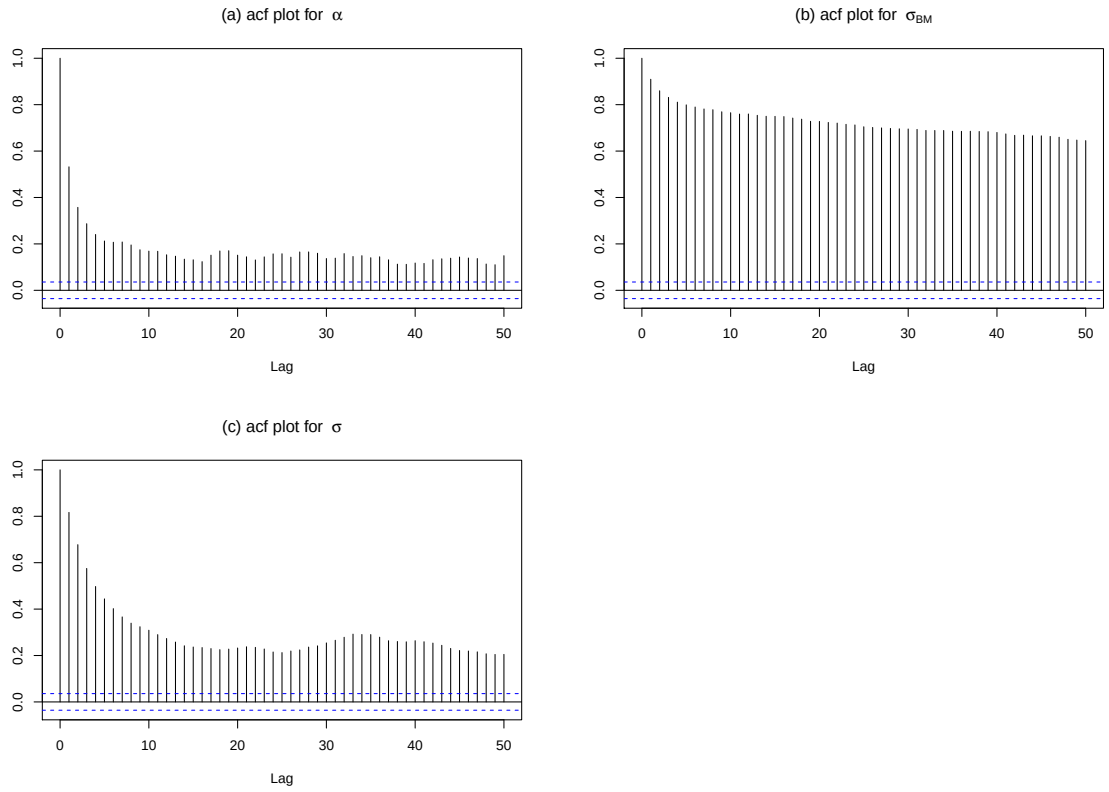


Fig. 6.30 Autocorrelation plots for movement parameters of the peak summer data using the continuous covariate switching model, based on the MCMC runs of 20,000 iterations. (a) Autocorrelation plot for $\alpha$, the attraction rate of the follower to the leading point. (b) Autocorrelation plot for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (c) Autocorrelation plot for $\sigma$, the individual variance coefficient of the follower.

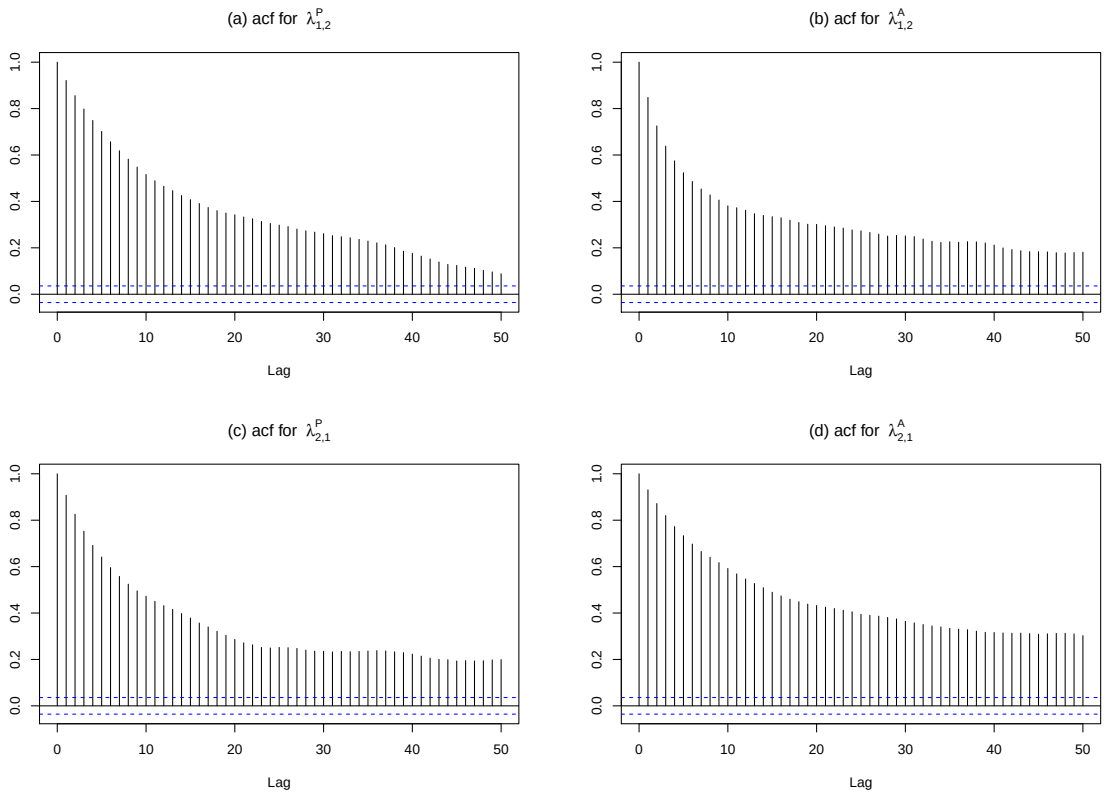


Fig. 6.31 Autocorrelation plots for switching parameters of the peak summer data using the continuous covariate switching model, based on the MCMC runs of 20,000 iterations. (a) Autocorrelation plot for $\lambda^P_{1,2}$, the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion in the presence of insects. (b) Autocorrelation plot of $\lambda^A_{1,2}$, the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion in the absence of insects. (c) Autocorrelation plot for $\lambda^P_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeckin the presence of insects. (d) Autocorrelation plot of $\lambda^A_{2,1}$, the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the absence of insects.

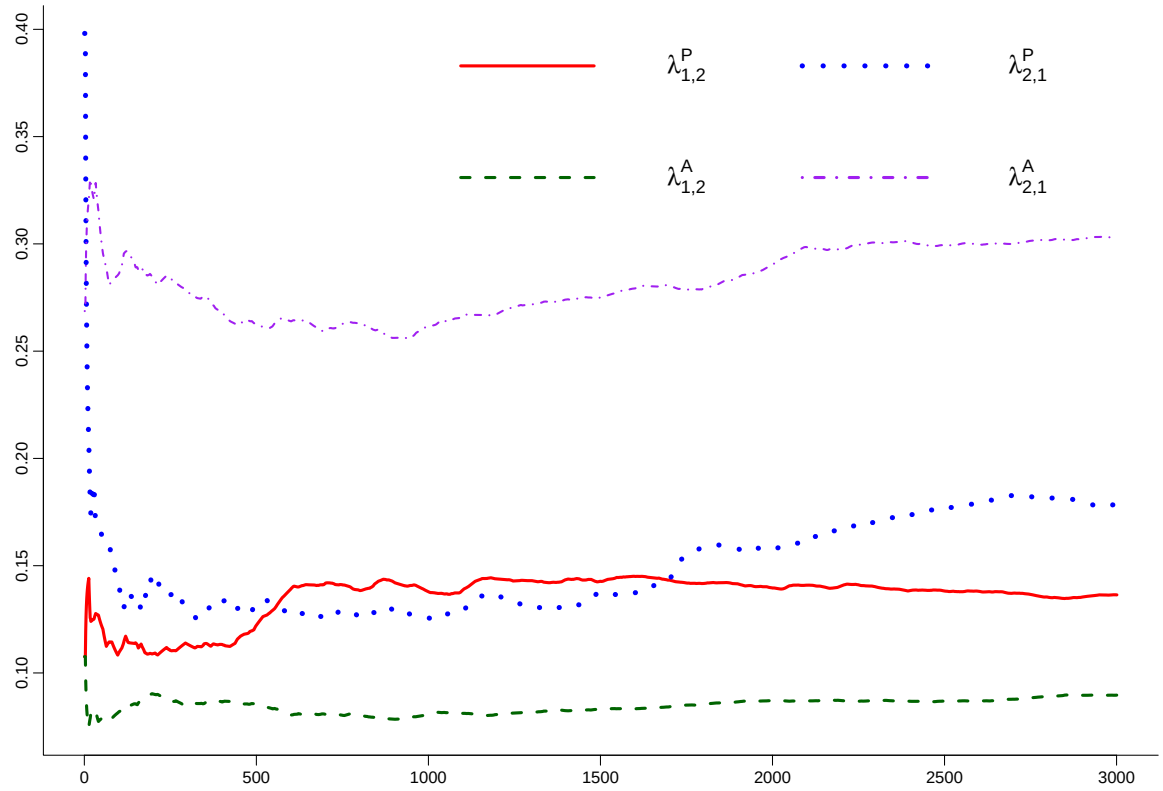


Fig. 6.32 Running mean for the switching parameters of the peak summer data using the continuous covariate switching model. The red (solid) line gives the running mean of $\lambda^P_{1,2}$. The green (dashed) line gives the running mean of $\lambda^A_{1,2}$. The blue (dotted) line gives the running mean of $\lambda^P_{2,1}$. Finally, the purple (dot-dash) line gives the running mean of $\lambda^A_{2,1}$.

### 6.6.3 Discussion

Whilst the use of continuous indexing was a natural progression of the model, its results still did not meet expectations. The estimates of the movement and switching parameters in both cases were similar to that of the discrete model. However, one main difference was the larger variation of the switching rate estimates for the early summer dataset in comparison with those of the discrete case; in the discrete case the estimates were (0.368, 0.414, 0.134, 0.129) and by comparison the continuous are (0.321, 0.583, 0.249, 0.159). Although, this difference is not as substantial as we anticipated.

What's more, we found the mixing of the MCMC to be rather slow and thus convergence of some parameters to be poor. For the early summer data, we can seen from Table 6.10 where slow mixing is most evident. Here, $\lambda^A_{1,2}$, takes the smallest ESS (23) and largest acf at lag 50

(0.44); the reason for poor convergence of this parameter may be due to the lack of switches that occur from state 1 to 2, since the animals spend most of their time in state 2. This is closely followed by the $\sigma_{BM}$, whose ESS is 32 and acf at lag 50 is 0.43.

For the peak summer data, we see a similar situation. However, $\sigma_{BM}$ has the smallest ESS of 13 and largest acf at lag 50 of 0.65. In the switching rates, the poorest performance occurs in the estimation of $\lambda^A_{2,1}$; by a similar intuition, this may be attributed to the animals spending the majority of their time in state 1, with infrequent occurrences of switching from state 2 to 1 in the absence of mosquitoes.

In conclusion, from a statistical analysis perspective the use of a continuous model seems natural and echoes the overall concern of the this thesis. However, its downfall is that in practice the continuous model seems to give little improvement and leads to slow mixing, especially with the switching parameters as seen in the running means plots in Figures 6.27 and 6.32. This may be alleviated by an alternative parametrisation the model.

In essence, we are aiming to estimate the linear relationship between mosquito harassment and the switching rates. In theory, there may be many ways to parametrise this. However, some may be easier to estimate numerically than others. When using the parametrisation given in Section 6.5, a problem arises. When $MI$ takes one of the extreme values, some parameters fail to be estimated. For example, when $MI = 0$, the parametrisation simplifies to

$$\lambda_{12} = \lambda^A_{12},$$
$$\lambda_{21} = \lambda^A_{21}.$$

In both the datasets this occurs frequently hence, can explain why we may have difficulty estimating particular parameters.

## 6.7 Alternative Parametrisation of the Covariate Model

In an effort to combat this issue, we can re-parametrise the model such that the range of values the mosquito index may take is $[-1,1]$ rather than $[0,1]$. Now, let

$$\lambda_{12} = a_1 + b_1 mi, \quad a_1 \geq 0, b_1 > -a_1, b_1 < a_1,$$
$$\lambda_{21} = a_2 + b_2 mi, \quad a_2 \geq 0, b_2 > -a_2, b_2 < a_2,$$

with

$$\kappa \geq n \max\{a_1^{max}, a_2^{max}, a_1^{max} + b_1^{max}, a_2^{max} + b_2^{max}, a_1^{max} - b_1^{max}, a_2^{max} - b_2^{max}\}.$$

Finally, we take the covariate, *mi*, as the transformation

$$mi = 2(MI) - 1, \tag{6.9}$$

where *MI* is the mosquito index as calculated in Russell et al. (1993) (Equation 6.6 - 6.8).

With this new parametrisation, we can think of the $a$ parameters as measuring an average level of transition in each direction, since when $mi = 0$ the switching rates are simply equal to $a$. The $b$ parameters measure the effect that *mi* has on the transition rate. In addition, in the event that the mosquito index has very little effect i.e. $b = 0$, then it possible to still get an estimate of $a$. Parametrising the model in this way essentially standardises the covariate around 0, which may help to stabilise the estimates.

### 6.7.1 Early Summer Results

We applied the alternative parametrisation model described in Section 6.7 to the early summer data set given in Section 5.2. We use the covariate MCMC algorithm with 20,000 iterations and the initial 70% counting as burn-in. This took approximately 72 hours to complete. The posterior point estimates and standard deviations of the model parameters are shown in Table 6.11. The posterior mean of the behavioural states (red crosses) for each individual at every time point are plotted in Figure 6.33. The posterior densities for each of the diffusion parameters are given in Figure 6.34 whilst the posterior densities for the switching rates are given in Figure 6.35.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.08 | 0.015 |
| $\rho$ | 3 | 0 |
| $\sigma$ | 1.134 | 0.07 |
| $\sigma_{\mathrm{BM}}$ | 0.287 | 0.017 |
| $a_1$ | 0.41 | 0.082 |
| $b_1$ | -0.137 | 0.102 |
| $a_2$ | 0.157 | 0.037 |
| $b_2$ | 0.038 | 0.042 |

Table 6.11 Parameter estimates for the movement and switching in the early summer data using the re-parametrised continuous covariate switching model.

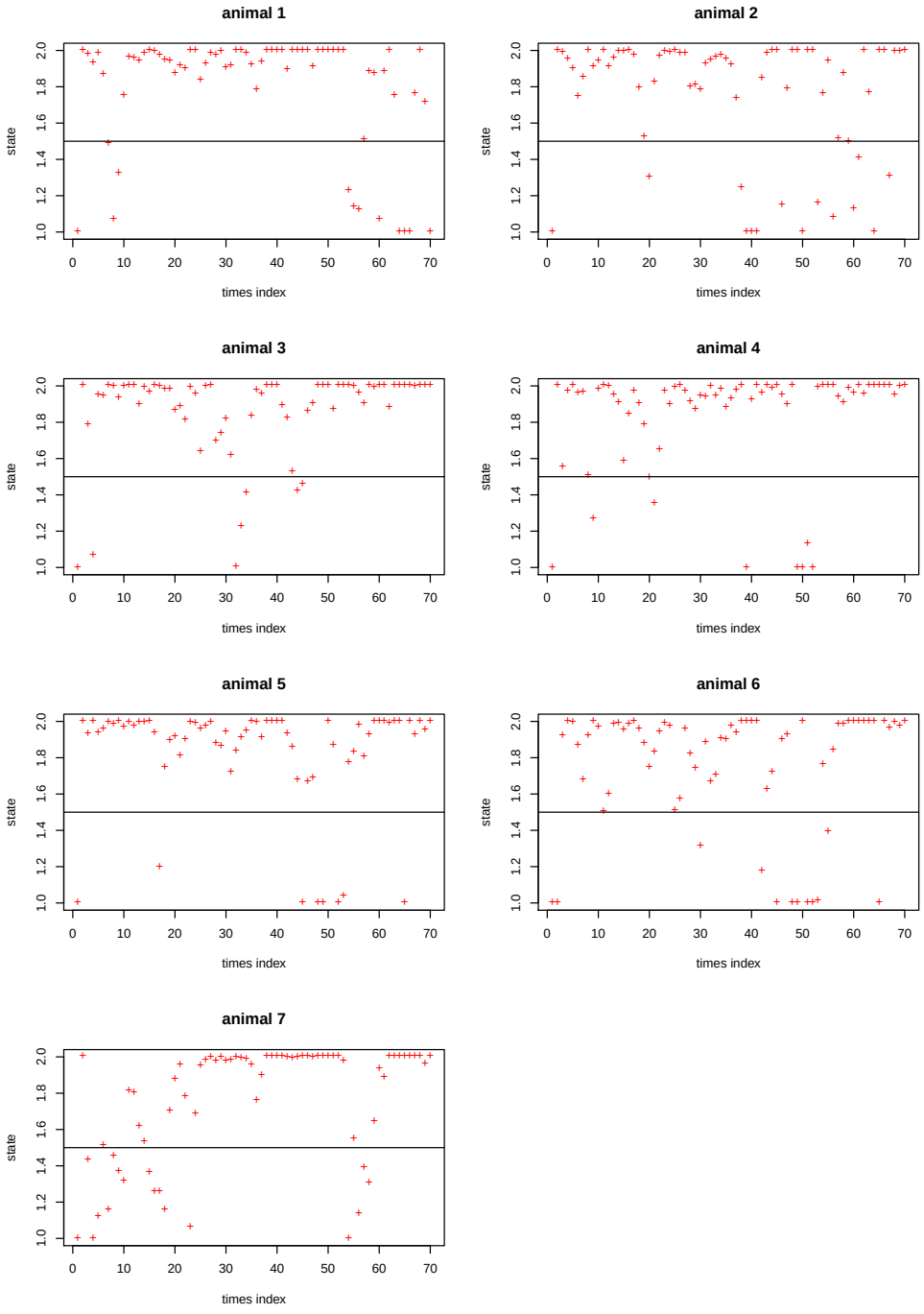


Fig. 6.33 Posterior mean states of all followers for the early summer data using the alternative parametrisation. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (red) represent the mean posterior of the estimated behaviour states.

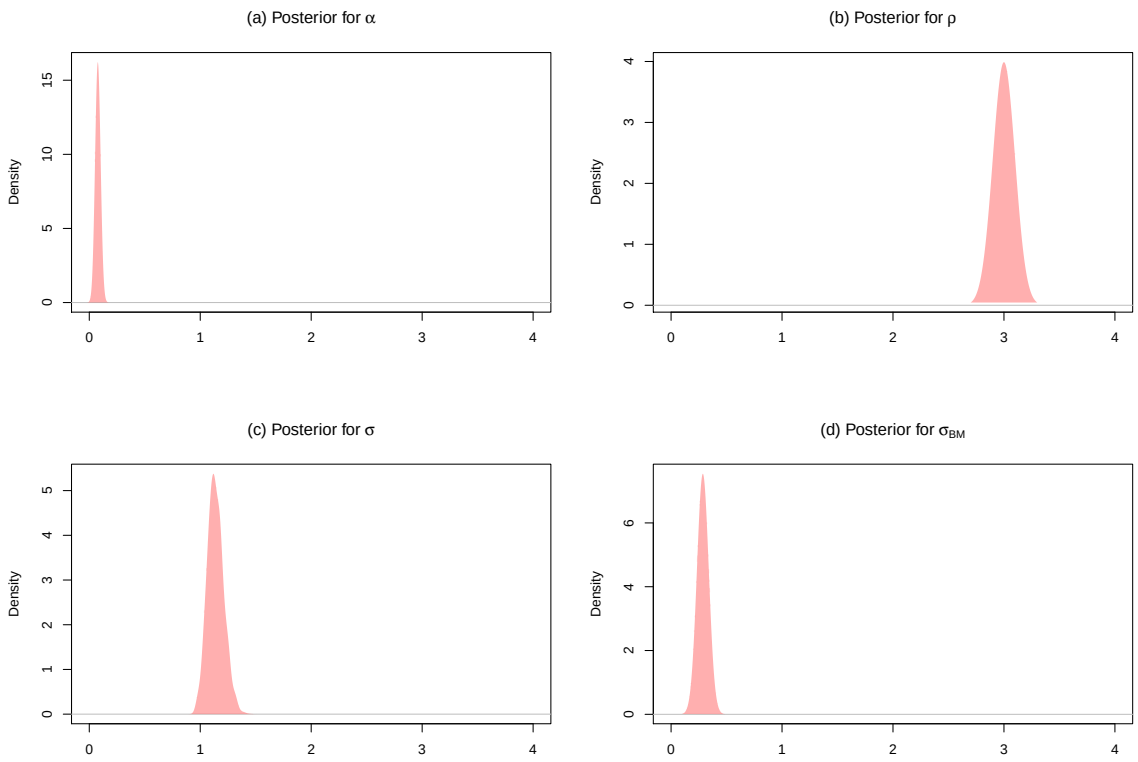


Fig. 6.34 Posterior densities of movement parameters for the early summer data using the alternative parametrisation, based on the Markov chain Monte Carlo runs of 20,000 iterations. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Approximate density for $\rho$, the variance coefficient of the leading point, which is fixed at 3. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion).

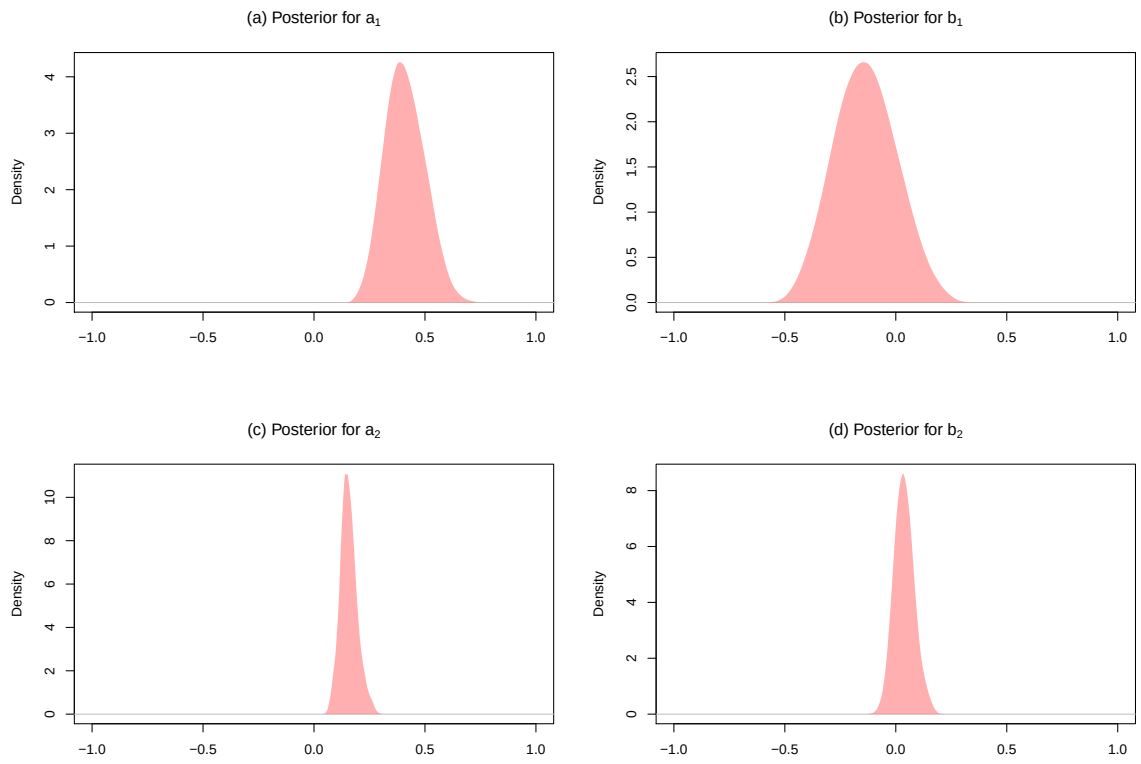


Fig. 6.35 Posterior densities of switching parameters for the early summer data using the alternative parametrisation, based on the Markov chain Monte Carlo runs of 20,000 iterations. (a) Posterior density for $a_1$, the average switching rate of the follower from OU to BM. (b) Posterior density of $b_1$, the relationship between mosquito harassment and the switching rate of the follower from OU to BM. (c) Posterior density for $a_2$, the average switching rate of the follower from BM to OU. (d) Posterior density of $b_2$, the relationship between mosquito harassment and the switching rate of the follower from BM to OU in the absence of insects.

| | acf lag | | | | |
|---|---|---|---|---|---|
| Parameter | 1 | 5 | 10 | 50 | ESS |
| $\alpha$ | 0.76 | 0.37 | 0.27 | 0.17 | 93 |
| $\sigma$ | 0.81 | 0.42 | 0.29 | 0.16 | 127 |
| $\sigma_{BM}$ | 0.93 | 0.78 | 0.72 | 0.52 | 24 |
| $a_1$ | 0.93 | 0.75 | 0.59 | 0.01 | 80 |
| $b_1$ | 0.95 | 0.81 | 0.67 | 0.25 | 62 |
| $a_2$ | 0.96 | 0.83 | 0.74 | 0.25 | 39 |
| $b_2$ | 0.96 | 0.84 | 0.73 | 0.30 | 53 |

Table 6.12 Autocorrelation at lag 1, 5, 10 and 50 and the effective sample size for each parameter in the early summer dataset using the alternative parametrisation model.

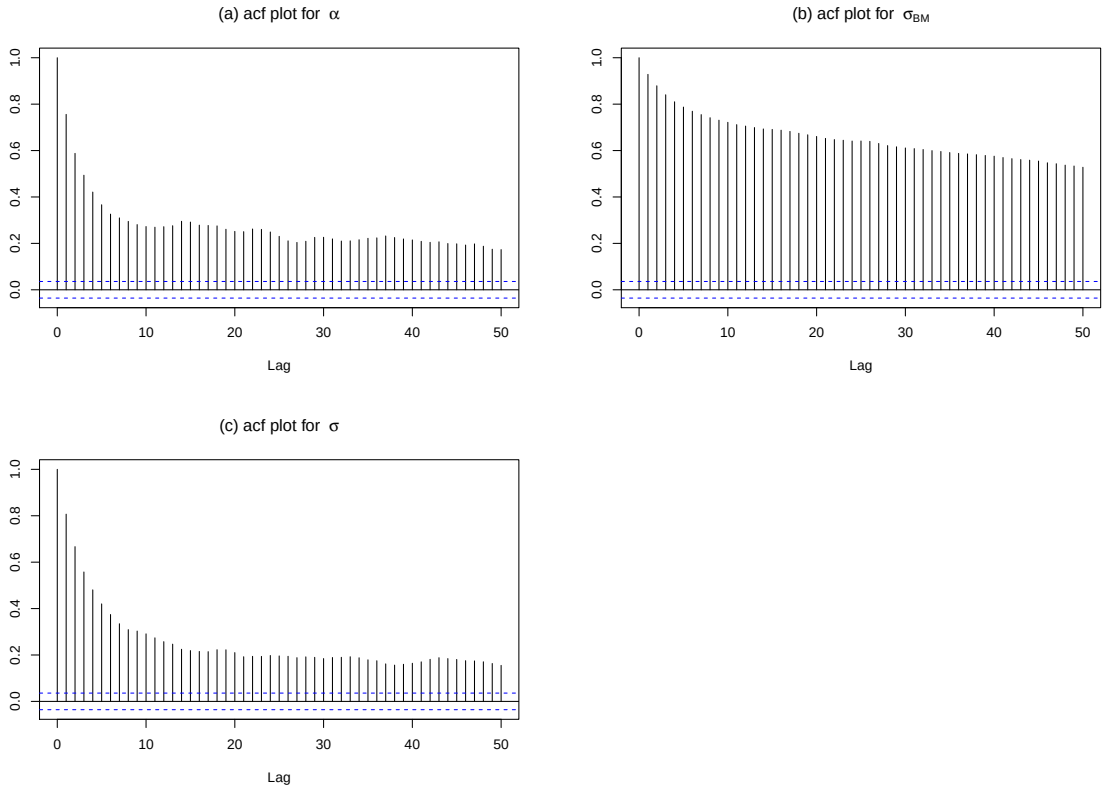


Fig. 6.36 Autocorrelation plots of the movement parameters for the early summer data, using the alternative parametrisation model. (a) Autocorrelation plot for $\alpha$, the attraction rate of the follower to the leading point. (b) Autocorrelation plot for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (c) Autocorrelation plot for $\sigma$, the individual variance coefficient of the follower.

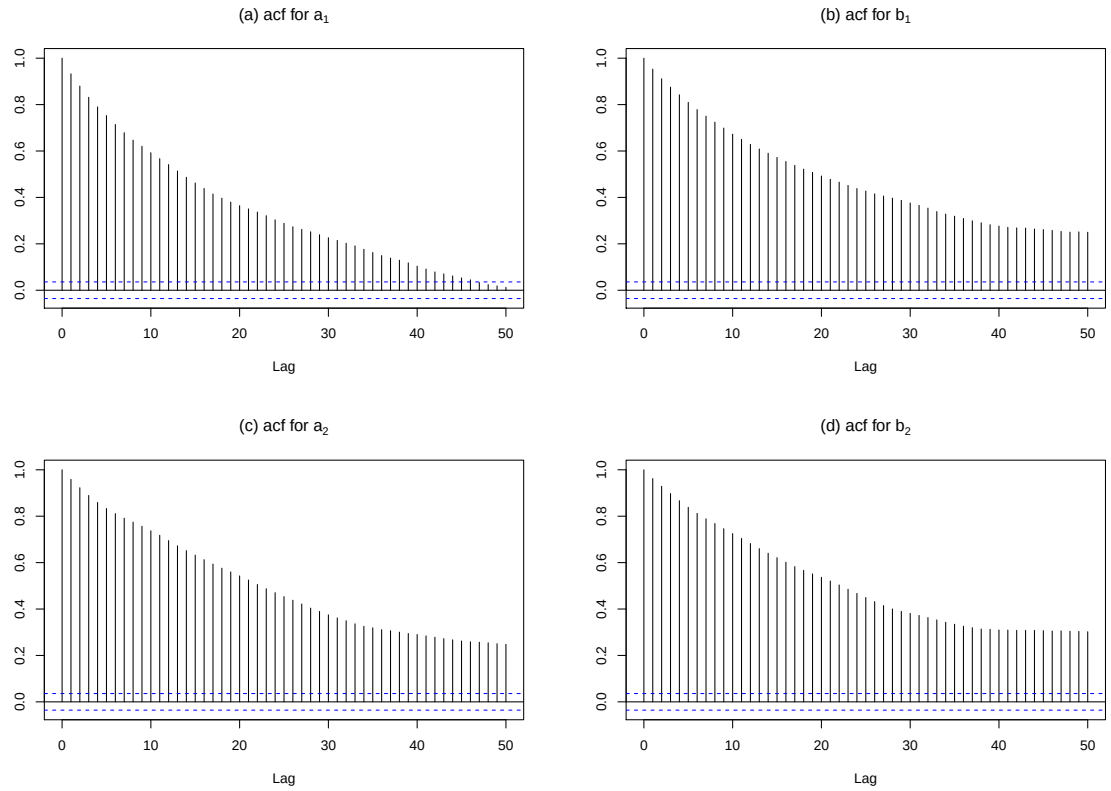


Fig. 6.37 Autocorrelation plots of the switching parameters for the early summer data, using the alternative parametrisation model. (a) Autocorrelation plot for $a_1$, the average switching rate of the follower from Ornstein Uhlenbeck to Brownian motion. (b) Autocorrelation plot of $b_1$, the relationship between mosquito harassment and the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion. (c) Autocorrelation plot for $a_2$, the average switching rate of the follower from Brownian motion to Ornstein Uhlenbeck. (d) Autocorrelation plot of $b_2$, the relationship between mosquito harassment and the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the absence of insects.

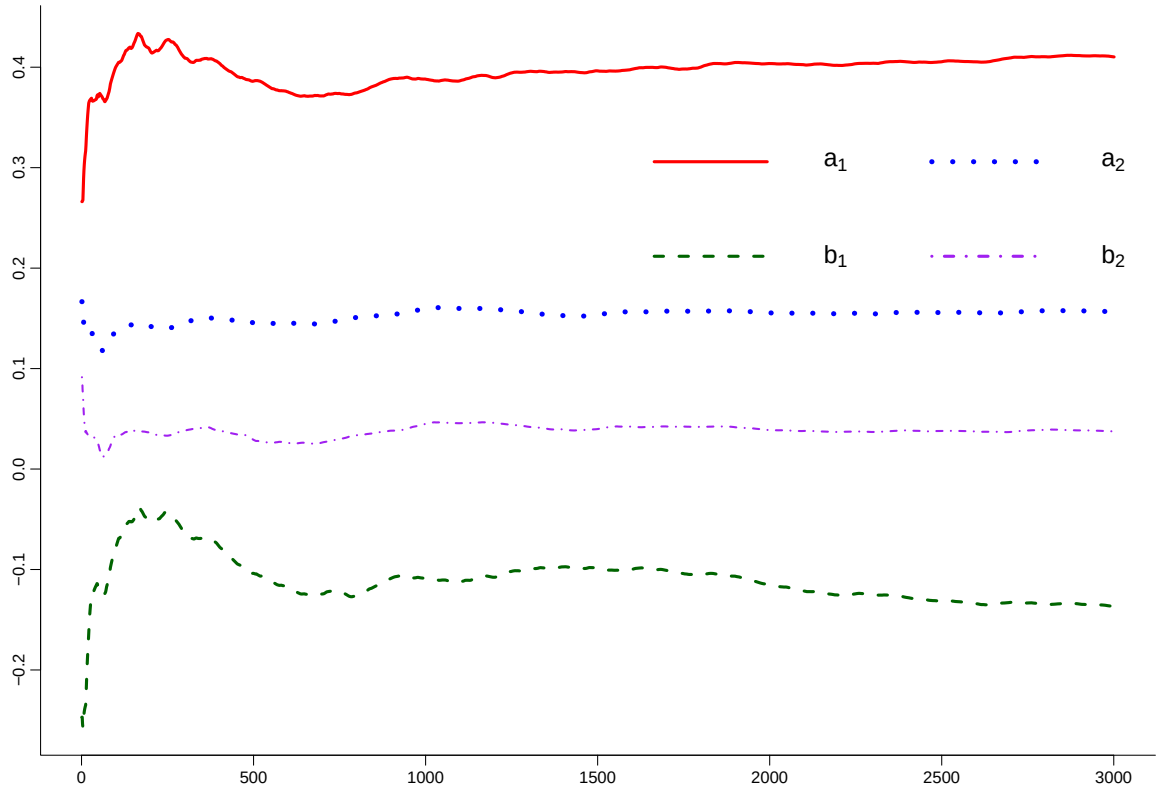


Fig. 6.38 Running mean for the switching parameters of the early summer data, using the alternative parametrisation model. The red (solid) line gives the running mean of $a_1$. The green (dashed) line gives the running mean of $b_1$. The blue (dotted) line gives the running mean of $a_2$. Finally, the purple (dot-dash) line gives the running mean of $b_2$.

### 6.7.2 Peak Summer Results

We applied the alternative parametrisation model described in Section 6.7 to the peak summer data set given in Section 5.2. We use the covariate MCMC algorithm with 20,000 iterations after burn-in. This took approximately 72 hours to complete. The posterior point estimates and standard deviations of the model parameters are shown in Table 6.13. The posterior mean of the behavioural states (red crosses) for each individual at every time point are plotted in Figure 6.39. The posterior densities for each of the diffusion parameters are given in Figure 6.40 whilst the posterior densities for the switching rates are given in Figure 6.41.

| Parameter | Point estimate | Standard deviation |
|---|---|---|
| $\alpha$ | 0.26 | 0.02 |
| $\rho$ | 3 | 0 |
| $\sigma$ | 0.50 | 0.02 |
| $\sigma_{\text{BM}}$ | 0.77 | 0.06 |
| $a_1$ | 0.122 | 0.025 |
| $b_1$ | 0.0499 | 0.028 |
| $a_2$ | 0.212 | 0.056 |
| $b_2$ | -0.059 | 0.085 |

Table 6.13 Parameter estimates for the movement and switching in the peak summer data using the alternative parametrisation continuous covariate switching model.

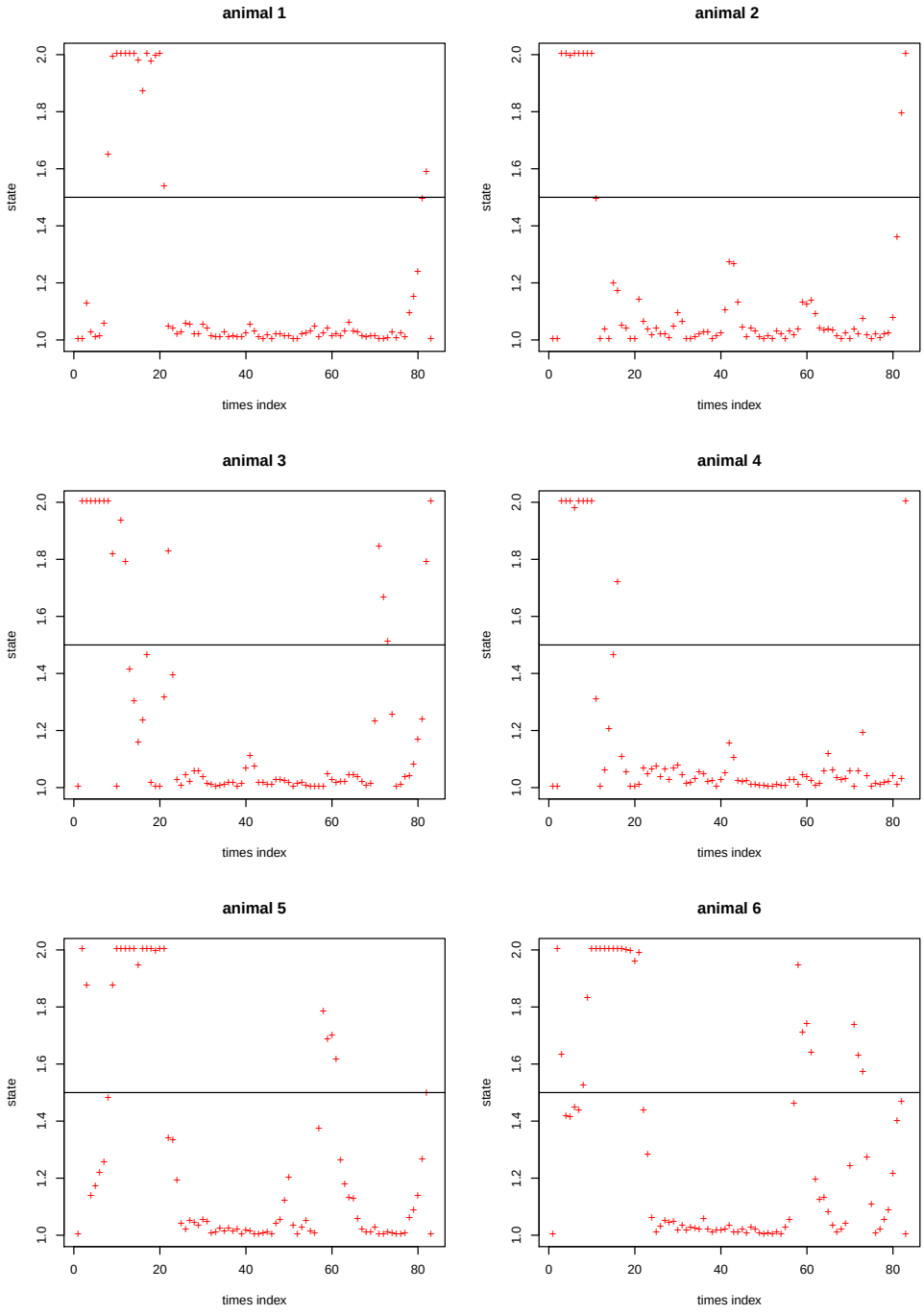


Fig. 6.39 Posterior mean states of all followers for the peak summer data, using the alternative parametrisation model. The vertical axis represents the states, 1 for Ornstein Uhlenbeck and 2 for Brownian motion. The crosses (red) represent the mean posterior of the estimated behaviour states.

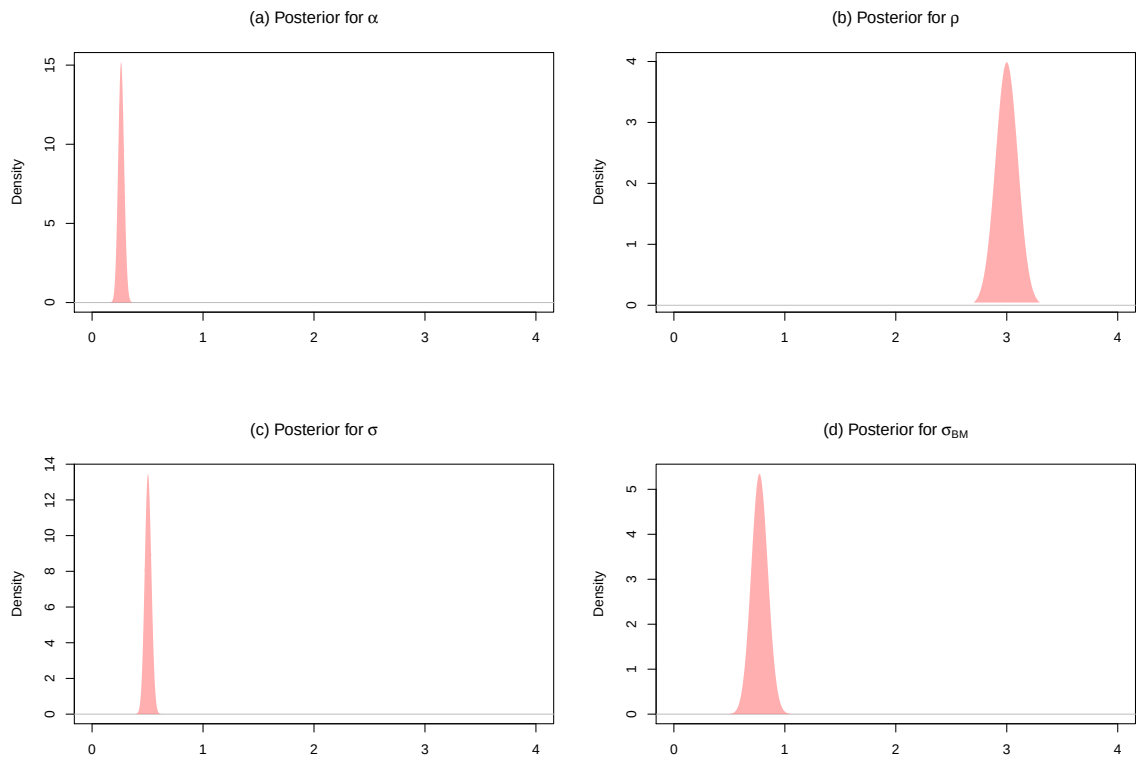


Fig. 6.40 Posterior densities of the movement parameters for the peak summer data using the alternative parametrisation model, based on the Markov chain Monte Carlo runs of 20,000 iterations. (a) Posterior density for $\alpha$, the attraction rate of the follower to the leading point. (b) Approximate density for $\rho$, the variance coefficient of the leading point, which is fixed at 3. (c) Posterior density for $\sigma$, the individual variance coefficient of the follower. (d) Posterior density for $\sigma_{\mathrm{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion).

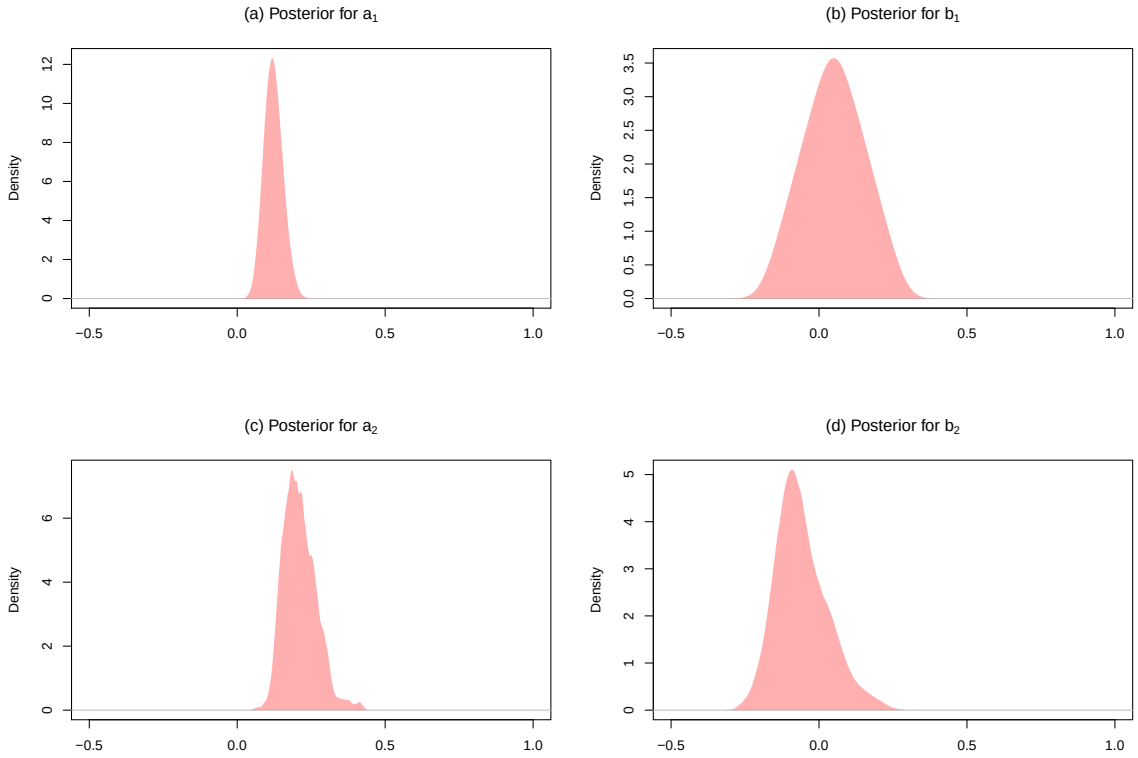


Fig. 6.41 Posterior densities of switching parameters for the peak summer data using the alternative parametrisation model, based on the Markov chain Monte Carlo runs of 20,000 iterations. (a) Posterior density for $a_1$, the average switching rate of the follower from OU to BM. (b) Posterior density of $b_1$, the relationship between mosquito harassment and the switching rate of the follower from OU to BM. (c) Posterior density for $a_2$, the average switching rate of the follower from BM to OU. (d) Posterior density of $b_2$, the relationship between mosquito harassment and the switching rate of the follower from BM to OU in the absence of insects.

| | acf lag | | | | |
|---|---|---|---|---|---|
| Parameter | 1 | 5 | 10 | 50 | ESS |
| $\alpha$ | 0.59 | 0.21 | 0.14 | 0.05 | 343 |
| $\sigma$ | 0.67 | 0.30 | 0.24 | 0.16 | 188 |
| $\sigma_{BM}$ | 0.78 | 0.52 | 0.47 | 0.26 | 80 |
| $a_1$ | 0.93 | 0.74 | 0.58 | 0.14 | 83 |
| $b_1$ | 0.95 | 0.80 | 0.64 | 0.18 | 72 |
| $a_2$ | 0.91 | 0.69 | 0.51 | 0.17 | 101 |
| $b_2$ | 0.92 | 0.69 | 0.49 | 0.07 | 99 |

Table 6.14 Autocorrelation at lag 1, 5, 10 and 50 and the effective sample size for each parameter in the peak summer dataset using the alternative parametrisation model.

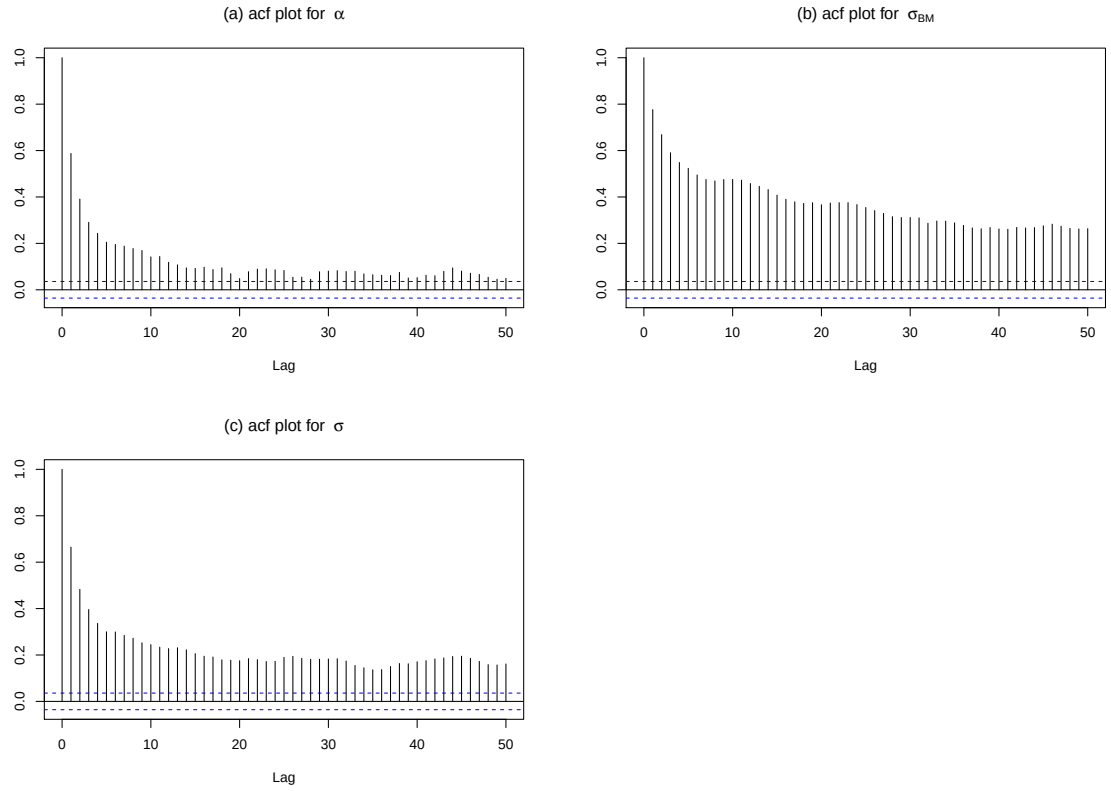


Fig. 6.42 Autocorrelation plots of the movement parameters for the peak summer data, using the alternative parametrisation model. (a) Autocorrelation plot for $\alpha$, the attraction rate of the follower to the leading point. (b) Autocorrelation plot for $\sigma_{\text{BM}}$, the variance coefficient of follower when it doesn't follow the leader (Brownian motion). (c) Autocorrelation plot for $\sigma$, the individual variance coefficient of the follower.

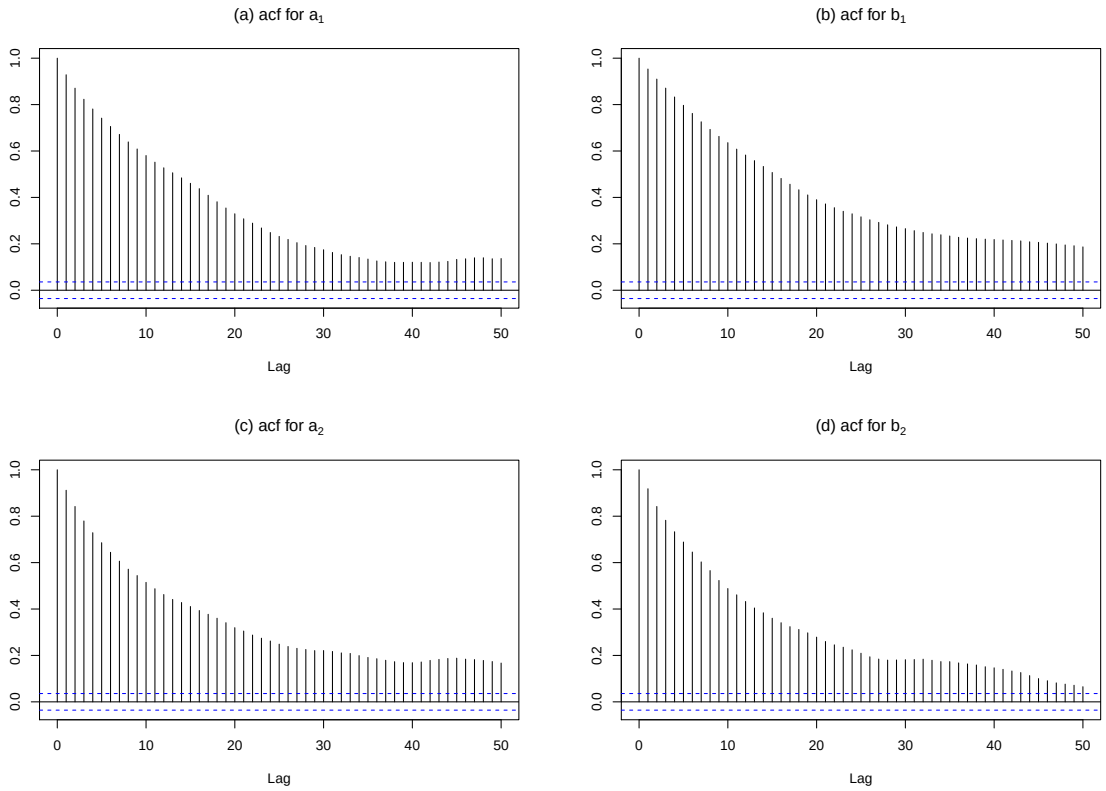


Fig. 6.43 Autocorrelation plots of the switching parameters for the peak summer data, using the alternative parametrisation model. (a) Autocorrelation plot for $a_1$, the average switching rate of the follower from Ornstein Uhlenbeck to Brownian motion. (b) Autocorrelation plot of $b_1$, the relationship between mosquito harassment and the switching rate of the follower from Ornstein Uhlenbeck to Brownian motion. (c) Autocorrelation plot for $a_2$, the average switching rate of the follower from Brownian motion to Ornstein Uhlenbeck. (d) Autocorrelation plot of $b_2$, the relationship between mosquito harassment and the switching rate of the follower from Brownian motion to Ornstein Uhlenbeck in the absence of insects.

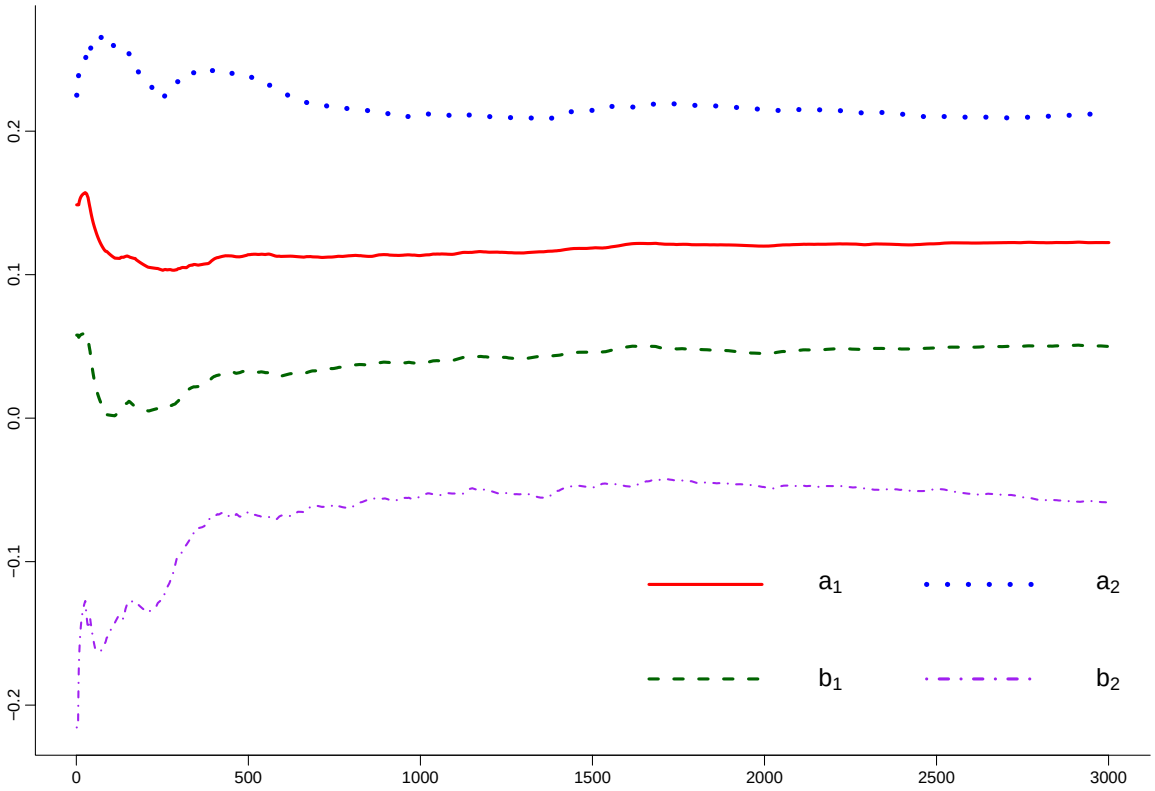


Fig. 6.44 Running mean for the switching parameters of the peak summer data, using the alternative parametrisation model. The red (solid) line gives the running mean of $a_1$. The green (dashed) line gives the running mean of $b_1$. The blue (dotted) line gives the running mean of $a_2$. Finally, the purple (dot-dash) line gives the running mean of $b_2$.

## 6.8 Discussion

The consistent estimates of the movement parameters give us confidence in our modelling approach; however, the improvement in MCMC mixing using the alternative parametrisation varies. For the early summer dataset, the ESS of most parameters given in Table 6.12 are smaller than those previously (Table 6.8). Although, the smallest ESS of the switching parameters is that of $a_2$ which is 39 and its acf at lag 50 is 0.25, an improvement on before. In contrast, for the peak summer data, the mixing is much better using the alternative parametrisation (Table 6.14); the ESS of all the movement parameters are significantly larger and the acf of switching rates smaller than previously (Table 6.10).

By contrasting the running mean plots for the old parametrisation (Figures 6.27 and 6.32) with the alternative parametrisation (Figures 6.38 and 6.44) we can see that in both cases the latter gives more stable estimations of the switching rates.

From Table 6.11 we can see that posterior point estimates of the diffusion parameters are similar to those of the discrete covariate model and the original continuous parametrisation given in Table 6.5 and 6.7 respectively. The estimates of the switching parameters suggest that the level of transitions from state 1 to state 2 declines with an increase of mosquito presence whereas transitions from state 2 to state 1 have a small but positive increase ($b_1 = -0.137, b_2 = 0.038$). Moreover, the average level of transition from state 1 to state 2, $a_1$, is estimated to be 0.41, more than twice that from state 2 to state 1, $a_2 = 0.157$.

Similarly, from Table 6.13 we can see that posterior point estimates of the diffusion parameters are similar to those of the discrete covariate model and the original continuous parametrisation given in Table 6.6 and 6.9 respectively. In contrast to the early summer data given in Table 6.11, as the presence of mosquitoes increases there is a small, positive ($b_1 = 0.0499$) gradient for the switching rates from state 1 to 2 but a negative ($b_2 = -0.059$) gradient from state 2 to state 1. Although this may contradict certain literature regarding the grouping of reindeer, it is worth appreciating that the point estimates of $b_1$ and $b_2$ are quite close to zero suggesting a negligible correlation between mosquito levels and the behavioural states of the reindeer. What is more illuminating is the average amount of transition from state 2 to state 1, $a_2 = 0.212$ is almost twice that from state 1 to state 2 $a_1 = 0.122$.

Although one might expect the absolute values of the estimates $b_1$ and $b_2$ to be greater if the mosquitoes truly have a strong effect on the reindeer movement, in practice the model we have used for mosquito harassment may not be a good predictor of movement. Whilst the Russell et al. (1993) model for harassment is continuous, it is a rather crude product of linearly interpolated temperature and wind indices. What's more, we feel that the denominator in the temperature index (Equation 6.6) should be 12 rather than 13 to avoid discontinuities. There is much dispute in the literature over models of mosquito harassment and using weather data as a proxy rather than direct observations of mosquito presence is likely to contribute to poor predictability.

To aid our interpretation of the switching parameter estimates, we compare ratios of the posterior $\lambda_{1,2}/\lambda_{2,1}$ when the mosquito harassment is at its extreme values, -1 and 1. Table 6.15 and 6.16 provide a summary of the densities given in Figures 6.45 and 6.46, the ratios during the early summer and peak summer dataset respectively.

It is worth noting that the maximum value of posterior ratio is rather large when $mi = 1$, 1848.8, 638.9 respectively; this can be attributed to small estimates of $\lambda_{2,1}$ which are close to 0. However, in these cases 95% of the data lies well under the values 4.38, 12.65 respectively.

| Ratio | Min. | 5%. | 1st Qu. | 3rd Qu. | 95% | Max. |
|---|---|---|---|---|---|---|
| $\lambda_{1,2}$ / $\lambda_{2,1}$, when $mi = -1$ | 2.57 | 3.34 | 4.09 | 5.15 | 6.02 | 8.38 |
| $\lambda_{1,2}$ / $\lambda_{2,1}$, when $mi = 1$ | 0.0002 | 0.15 | 0.69 | 2.19 | 4.38 | 1848.8 |

Table 6.15 Summary statistics for the densities of posterior switching rate ratios in the early summer data. The first row gives the summary where $mi = -1$ whereas the second row summarises the density where $mi = 1$.

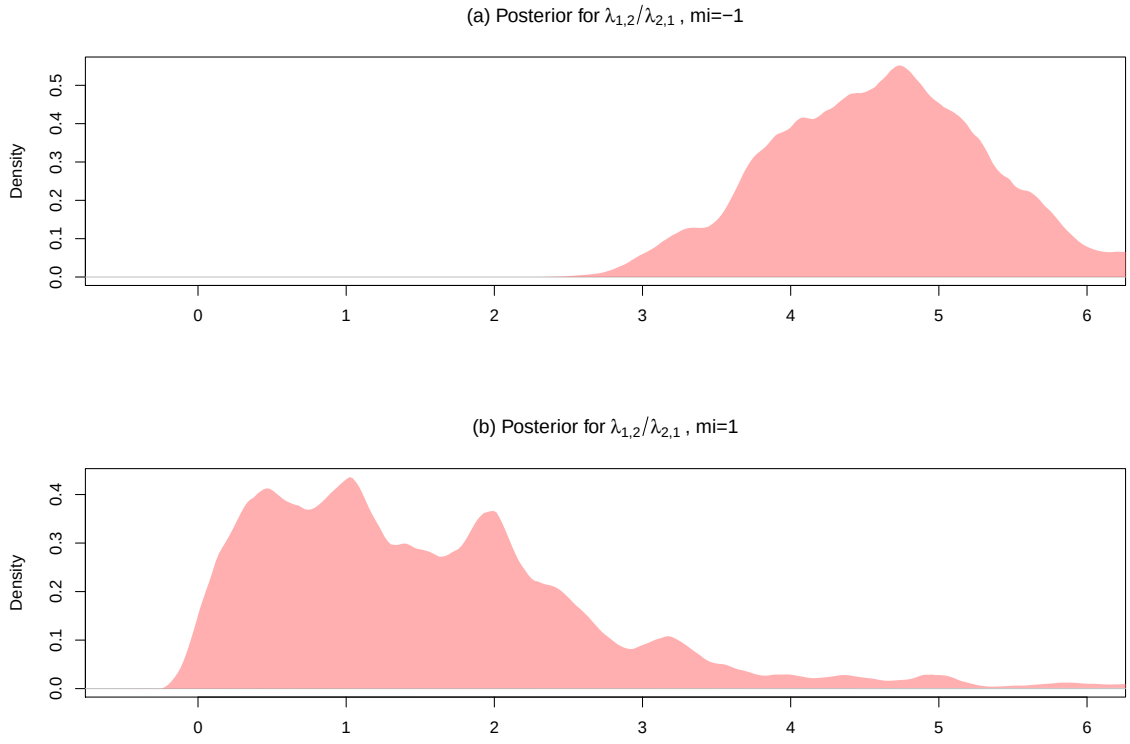


Fig. 6.45 Posterior densities of the switching rate ratio $\lambda_{1,2}/\lambda_{2,1}$ in the early summer dataset. (a.) Posterior density for $\lambda_{1,2}/\lambda_{2,1}$ when $mi = -1$. (b.) Posterior density for $\lambda_{1,2}/\lambda_{2,1}$ when $mi = 1$.

| Ratio | Min. | 5% | 1st Qu. | 3rd Qu. | 95% | Max. |
|---|---|---|---|---|---|---|
| $\lambda_{1,2}$ / $\lambda_{2,1}$, when $mi = -1$ | 0.02 | 0.12 | 0.21 | 0.34 | 0.50 | 1.07 |
| $\lambda_{1,2}$ / $\lambda_{2,1}$, when $mi = 1$ | 0.01 | 0.38 | 0.73 | 2.91 | 12.65 | 638.9 |

Table 6.16 Summary statistics for the densities of posterior switching rate ratios in the peak summer data. The first row gives the summary where $mi = -1$ whereas the second row summarises the density where $mi = 1$.

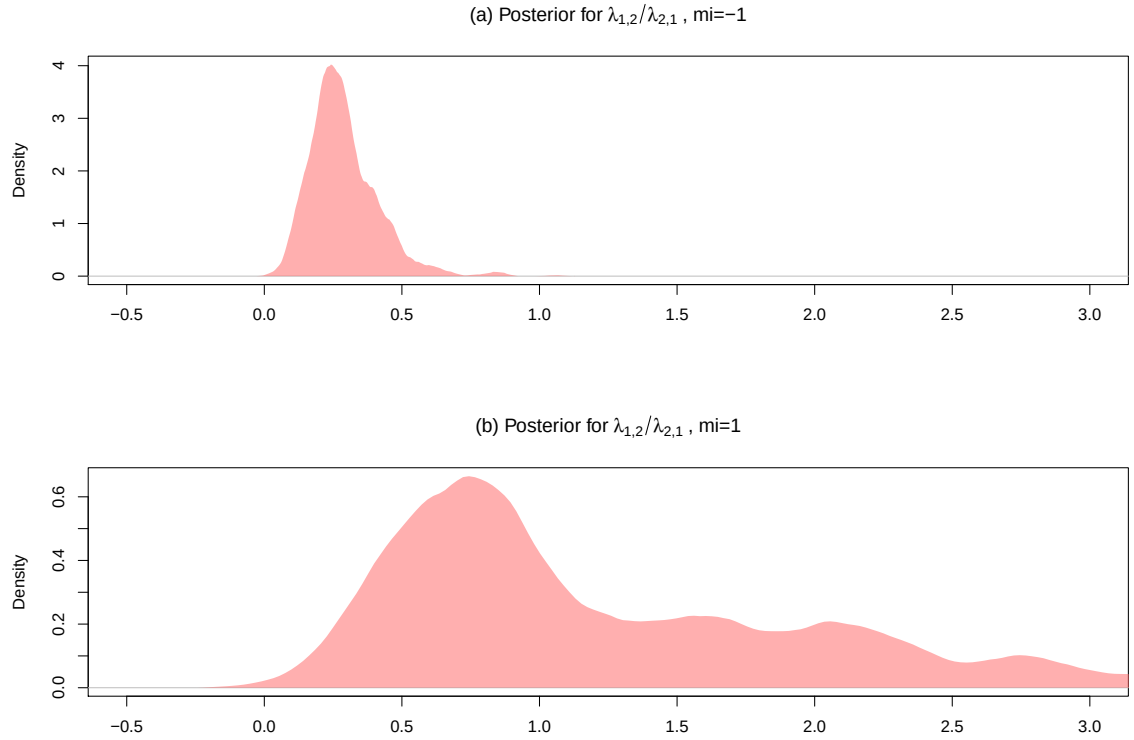


Fig. 6.46 Posterior densities of the switching rate ratio $\lambda_{1,2}/\lambda_{2,1}$ in the peak summer dataset. (a.) Posterior density for $\lambda_{1,2}/\lambda_{2,1}$ when $mi = -1$. (b.) Posterior density for $\lambda_{1,2}/\lambda_{2,1}$ when $mi = 1$.

Although we cannot establish precisely how mosquitoes affect the reindeer, using these ratios we are able to notice clear differences in the proportion of transitions between high and low mosquito presence. In both cases, we can see that the ratios of switching rates are closer to 1 when $mi = 1$. That is, the level of transition in each direction is more similar when the mosquito harassment is high rather than low. Moreover, between datasets we can see from plot (a) in both figures, that in early summer there is a dominance of Brownian motion behaviour but a dominance of Ornstein Uhlenbeck in peak summer when $mi = -1$. This is

not entirely consistent with what we expect; we anticipate that when $mi = -1$ the proportion of time spent in Brownian motion will be greater. One possible reason for inconsistency is that in our analysis we are comparing a separate model fitting for each dataset. The posterior estimates for the movement parameters are starkly different between datasets meaning that the distinction of states between datasets is different. Ideally, this could be combatted by fitting the model to one large dataset which encompassed both time periods but this not currently computationally feasible.

# Chapter 7

# Discussion and Further Work

Despite the rapid growth in tracking technologies and the tools used to analyse such tracks, models of collective animal movement have been under-represented. Instead, practitioners have preferred single animal models due to their ease of fitting, computational efficiency and since historically ecologists were not practically or financially able to mass tag species. Similarly, a lack in progress can be seen in the development of continuous-time models, which in part is due to their inherently complex nature, computational strain and difficulty in parameter interpretation.

In Chapter 3 we provided a continuous-time group movement model which allows for behavioural state switching. This novel and flexible framework hopes to break down the barriers associated with modelling movement collectively or in continuous-time. In essence, the model combines the continuous and collective movement model of Niu et al. (2016) with the exact simulation methods of Blackwell et al. (2016). The underlying notion of the methodology is similar to the work of Langrock et al. (2014), where the interaction between animals are a shared attraction to an abstract leader (which has its own point of attraction) and where the animals may switch between following this leader and moving independently with Brownian motion. The model can be applied to data which may be irregular in time or have incomplete observations, thus the model is not restricted to a particular time scale and may be used to compare datasets with different sampling schedules. Moreover, the approach gives an improvement on the computational efficiency of fitting by use of the Kalman filter.

The applicability of the model is demonstrated by a series of simulation experiments in Section 3.8. Here we applied the non-stationary model discussed in Section 3.3, where the leader does not have a point of attraction, to a variety of simulated movement processes emulating real biological possibilities. In each case the model successfully recovered the

true parameter values. This gives confidence in our results when the model is applied to real data where the true parameter values are unknown as in Chapter 4.

It is common for ecologists to be interested in the movement processes of animals in response to environmental factors and how this affects habitat section. Animals often exhibit trade-offs between a range of demands such as foraging quality, guarding territory, caring for young and being safe from predators. These demands will influence the behaviours the animals exhibit and the habitat they select (Roever et al., 2014). In Chapter 5 we undertook some preliminary work to investigate whether the model can pick up on different behavioural patterns consistent with known responses to insect harassment. We compared two datasets where we expected contrasting levels of insects, which were implied by proxy weather data. Generally the model performed well, resulting in estimates of behavioural and movement processes which were congruous with observational studies.

By overlaying each animal's trajectory and estimated states on a terrain map we were able to see distinct differences in habitat selection. During times of low harassment, the reindeer seemed to spend the majority of their time moving independently in lower elevations or in woodland. In contrast, during high harassment, the reindeer appeared to move up into higher elevation mountainous regions where forage is scarce, possibly to seek topographical relief from the insects. From this case study, we envision direct applications for conservation and wildlife management.

However, as highlighted in Section 5.5, by using homogenous switching parameters the model attempts to give an overall depiction of the proportion of time spent in each state, without describing any underlying mechanics. Since the movement and thus state transitions are thought to be affected by insect harassment, this ought to be included explicitly in the model as we explore in Chapter 6.

In Chapter 6, we follow similar methodology to others by allowing the transition probabilities of our SSM to be given as functions of environmental covariates (Harris and Blackwell, 2013; McClintock et al., 2012, 2013). We discuss a flexible framework for incorporating a range of temporal covariates into the state switching model, then investigate its limitations through a series of simulation studies where movement is influenced by the time of day, as seen in some diurnal species. Unsurprisingly, we find that the model performs better with richer data that is, with more frequent or larger quantities of observations.

In Section 6.3, we revisit the reindeer case study of Chapter 5 by explicitly including covariate information of mosquito presence into the model. Initially, we used a discrete model of mosquito presence; however, we found little variation in the parameter estimates $\lambda_{i,j}^{P}$ and $\lambda_{i,j}^{A}$

for each switching direction. That is, the switching rates did not seem to be influenced greatly by the presence of mosquitoes. Naturally, we instead used a continuous model although, this only gave a marginal improvement and lead to slow mixing of the MCMC for certain parameters.

Due to difficulties in convergence, we discussed an alternative parametrisation. However, the improvement in convergence from this parametrisation varied between datasets. There was little change in the convergence performance of the early summer dataset but an improvement for the peak summer data. Overall, the model seems to perform well in simulated settings so poor convergence in this case may be attributed to mosquito harassment alone being an inadequate predictor of reindeer movement, or that the model for insect harassment has scope for improvement. In any case, we feel that the methodology presented in this thesis is a step in the right direction in facilitating the use of continuous-time collective movement models.

## 7.1 Further Work

### 7.1.1 Model selection and convergence diagnostics

We acknowledge the lack of formal model selection within our work. For comparison between the non-covariate model of Chapter 5 and the covariate models of Section 6.3.1, 6.5 and 6.7, it would be interesting to calculate information criteria such as the Deviance information criterion (DIC) (Spiegelhalter et al., 2002) or more recently the Watanabe–Akaike information criterion (WAIC) (Gelman et al., 2013b; Watanabe, 2010) to assess the relative quality of each model. Similarly, we could do this to rank the covariate circadian model in Section 6.2.3 and the standard switching approach in Section 6.2.4 . However, the nature of such criteria require the likelihood of the data, given each parameter, to be recorded at every iteration of the MCMC algorithm, which can be computationally costly (see Appendix 1 for an explicit formulation of WAIC). Instead, we offer an anecdotal comparison by simulating forward from point estimates of the parameter values for both cases (shown in Figure 6.14 and 6.15). From this we see that the non-covariate model fails to imitate the true underlying mechanics of switching, which gives justification to our preference of the covariate model.

In addition to this, we limit our assessment of convergence to single-run tools such as acf and ESS. However, other convergence diagnostic tools would supplement our investigations. With multiple MCMC sequences, initialised at overdispersed values, it would be possible to

use tools such as the Gelman-Rubin statistic (Brooks and Gelman, 1997; Gelman and Rubin, 1992) to analyse convergence.

### 7.1.2 Model variants

With our flexible modelling approach it is possible to experiment with a range of temporal covariates using both simulated and real data. One encouraged direction of further work would be to investigate the performance of oestrid fly indices as a predictor of movement. For this, Mörschel (1999) provides a continuous logistic function for predicting the presence of oestrid flies based on the climatic variables, temperature and wind speed. This is given as

$$y = \frac{f(T,W)}{1+f(T,W)},$$

where

$$f(T,W) = \exp(-2.9646 + 0.166T - 0.1951W)$$

and where $T$ is the ambient temperature given in degrees Celsius $^{\circ}C$ and $W$ is the wind speed measured in metres per second, $ms^{-1}$.

However, as touched upon in the discussion of Chapter 5, a better approach would simultaneously model the effect that mosquitoes and oestrids have on reindeer behaviour, since they do not occur exclusively. In fact, observational findings of Witter et al. (2012a) suggested that such a combined model would significantly improve predictive ability. In their study, they used AIC to compare a range of models involving biotic and abiotic variables as predictors of reindeer behaviour. The found that out of the two best models, which had nearly identical AIC scores, the variables included black fly or mosquitoes counts plus all of the other following variables: oestrids, time, easting, northing, growing degree days (gdd), year and group size.

Here, the variables black fly, mosquito and oestrid are abundance levels of that insect; time represents the time of day; easting and northing are the Cartesian coordinates of the study site; gdd is the growing degree day, a measure of seasonality relevant to the development of new green vegetation; year is the year of study which considers any annual variations and finally, group size is a count estimate for the number of adult reindeer in the group. For a more detailed description of the variables, see Witter et al. (2012a).

Most of these covariates are temporal variables and would be straightforward to implement into the current model, with the exclusion of the spatial variables, northing and easting, which require additional model development (see Section 7.1.5). However, this would need highly sampled observations of mosquito and oestrid levels. Since this may not be readily available, some proxy for the abundance of both species of insect would still be required and to our knowledge they only exist as separate models for each species.

There are multiple avenues for adapting the movement model itself. Firstly, we have only considered the case where there exists one leader but there are immediate and obvious biological reasons for having multiple leaders. For this framework, each individual has the possibility of switching between following any of the $n$ leaders, forming many sub-herds, or moving with Brownian motion. To provide insight into the set-up of this model we will give the corresponding form of the attraction matrix, $A$, as given in Equation 3.5, for multiple leaders. In the case of multiple leaders, the attraction matrix will have an additional $n$ columns and rows, one for each leader. We will present the case for two leaders and three followers, where the first and third animal are following leader 1 and the second animal is following leader 2; $A$ takes the format as follows.

$$
\begin{array}{c}
\\
\text{Leaders}\left\{\vphantom{\begin{pmatrix}0\\0\\0\\0\\0\\0\end{pmatrix}}\right. \\
\\
\\
\\
\\
\\
\text{Attraction to specific leader}
\end{array}
\begin{array}{c}
\overbrace{\phantom{00}}^{\text{Leaders}}\ \overbrace{\phantom{000}}^{\text{Followers}} \\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
\beta & -\beta & 0 & 0 & 0 & 0 \\
0 & 0 & -\beta & 0 & 0 & 0 \\
0 & \alpha & 0 & -\alpha & 0 & 0 \\
0 & 0 & \alpha & 0 & -\alpha & 0 \\
0 & \alpha & 0 & 0 & 0 & -\alpha
\end{pmatrix}\left.\vphantom{\begin{pmatrix}0\\0\\0\\0\\0\\0\end{pmatrix}}\right\}\text{Followers}
\end{array}
$$

Another possibility is to allow the location of the leader(s) to affect not only the movement of the followers but also their switching probabilities. Ecologically, it is easy to imagine a scenario whereby the followers are more inclined to switch behaviour from Brownian motion to following the group if the leader becomes too far away. In contrast, if the leader becomes too close some individuals may choose to move independently in order to avoid any foraging competition. This covariate information could be implemented into the existing model by allowing the switching rate for an individual to be given as a function of their location relative to the leaders. Following on from this, in the spatially heterogeneous case we may choose to let the switching probabilities be influenced by the absolute location of the leader rather than its relation to the followers. For example, we could allow the switching rates to be dependent

on the the specific habitat that the leader occupies. However, as mentioned previously, the inclusion of spatial variables involves some development of the model (see Section 7.1.5).

### 7.1.3 Computational improvement

Ideally, we would be able to provide inference on collective movement over long time periods of multiple weeks but in reality we are limited by the computational strain of the model. In Chapter 6, data consisting of five days of observations sampled every two hours for six individuals took around 72 hours to complete 20,000 iterations. Thus, with only a couple of additional animals or weeks of data we can foresee issues with computational feasibility.

Obvious approaches to relieve these issues are to recode parts of the algorithm in faster compiled languages such as C/C++ (ISO, 2020). Of course, some profiling may first be done in R to establish where this would be most useful.

Another approach would be to use some approximation to the model. Rather than allowing an exact continuous simulation of the trajectories described in Section 3.5, we could insist that switches occur only at a fixed, pre-determined sampling schedule but that the animals still move with a continuous Ornstein Uhlenbeck or Brownian motion process. The sampling schedule may be chosen as fine as we like and it may even be possible to accommodate irregular observations. Although this may make the algorithm faster, this discretisation goes against the philosophy of the thesis.

### 7.1.4 Spatial Covariates and NDVI

In this thesis we have focussed only on the inclusion of temporal covariates. However, there exist many spatial attributes which may affect the way in which animals move. For example, the infrastructure needed for gas exploitation and wind turbines can create avoidance behaviour in reindeer up to 5km away, much larger than the structure itself (Hall et al., 2006; Skarin and Åhman, 2014). Moreover, animals may alter their behaviour as a function of proximity to particular factors such as nests sites, nearby neighbours or predator territories (Broom et al., 2019; Harris and Blackwell, 2013; Potts and Lewis, 2014; Russell et al., 2017). Specific to reindeer movement, Witter et al. (2012b) suggested the use of spatial variables such as topography and vegetation could be used in indices of insect harassment.

An increasingly explored environmental attribute affecting reindeer behaviour and patch selection is the quality and quantity of nutritional forage. From winter into spring, reindeer

are said to follow a 'green-wave', the melting of snow and the growth of new forage (Merkle et al., 2016).

The use of satellite imagery has been popularised in ecology as a means to quantify the level of nutritional forage. Satellites such as MODIS (MODerate resolution imaging spectroradiometer) regularly take images of the reflectance of red, near-infrared, and blue wavebands of the Earth's surface; these can be used to calculate the normalised difference vegetation index (NDVI), a crude estimate of the quality of green vegetation and forage.

The NDVI of surface can range from -1 to 1 and is calculated as (NIR - Red)/(NIR + Red), where NIR and Red denote the spectral reflectance in the near-infrared and red (visible) regions respectively. Live green vegetation has a positive NDVI, with healthier plants taking values closer to 1, this is because the chlorophyll in plants absorb visible light but their leaf structures strongly reflect near-infrared light (Pettorelli et al., 2005b). By contrast, infrastructure, soil and snow have NDVI close to 0.

In conservation and management, NDVI has been used to detect droughts, flooding, wild fires and degradation in the landscape (Pettorelli et al., 2011). In ungulate research, it has been shown to be a strong predictor in movement, biomass and well-being (Mueller et al., 2011; Pettorelli et al., 2005a,c). More specifically, Bartlam-Brooks et al. (2013) highlight NDVI as a crucial predictor for the onset of migration in zebra.

### 7.1.5 Implementing Spatio-temporal Covariates

The inclusion of spatial covariates such as habitat type or indeed spatio-temporal information such as NDVI require more care than time-varying covariates alone. As the model stands, when using the Kalman filter to evaluate the likelihood we integrate out the leader's locations and also average over the follower's locations outside of observations. For time-varying covariates this is not a problem since we need not know specifically where each animal is at every switching time.

However, for spatial covariates the location of each animal at the time of every switch is essential and we must obtain this by sampling from the movement model of the animal's current state. When we know the respective locations, only then can we determine the switching probabilities from spatial covariates. Whilst this methodology is straightforward in principle, the algorithm requires more sophistication and in practice the additional sampling of locations will inevitably lead to a more computationally intensive algorithm.

To simplify the matter, rather than using a single point process which captures all switches, we could use a separate one for each of the $n$ individuals. In this case, the rate of each process, $\kappa$, does not need to be as large as in the original formulation as it only tries to capture switches from a single animal i.e. for each process we only require

$$\kappa \geq \max(\lambda_{1,2}, \lambda_{2,1})$$

i.e. $n$ times smaller than in original algorithm. We then sample from all $n$ processes simultaneously and order the switching times to be evaluated. Since each switching time is specific to a single individual we only have to sample their location. By avoiding the need to sample every individual's locations the computational time should be quicker.

Models of movement and behaviour which take into consideration spatio-temporal landscape dynamics provide a more realistic perspective on species–habitat relationships. From this, conservation and management initiatives are able to better tailor their actions towards specific goals (Roever et al., 2014). Continuous models of collective movement which include complex covariate information will offer rich applications and as the technology and computational power improves research should be driven towards these 'gold standard' approaches.

# References


Ahman, B., Svensson, K., and Rönnegard, L. (2014). High female mortality resulting in herd collapse in free-ranging domesticated reindeer (Rangifer tarandus tarandus) in Sweden. *PLoS ONE*, 9(10).

Anderson, J. R., Nilssen, A. C., and Folstad, I. (1994). Mating behavior and thermoregulation of the reindeer warble fly, Hypoderma tarandi L. (Diptera: Oestridae). *Journal of Insect Behavior*, 7(5):679–706.

Aoki, I. (1982). A Simulation Study on the Schooling Mechanism in Fish. *Nippon Suisan Gakkaishi*, 48(8):1081–1088.

Auger-Méthé, M., Field, C., Albertsen, C. M., Derocher, A. E., Lewis, M. A., Jonsen, I. D., and Flemming, J. M. (2016). State-space models' dirty little secrets: Even simple linear Gaussian models can have estimation problems. *Scientific Reports*, 6(December 2015):1–10.

Ballesteros, M., Bårdsen, B. J., Langeland, K., Fauchald, P., Stien, A., and Tveraa, T. (2012). The effect of warble flies on reindeer fitness: A parasite removal experiment. *Journal of Zoology*, 287(1):34–40.

Bartlam-Brooks, H. L., Beck, P. S., Bohrer, G., and Harris, S. (2013). In search of greener pastures: Using satellite images to predict the effects of environmental change on zebra migration. *Journal of Geophysical Research: Biogeosciences*, 118(4):1427–1437.

Beest, M. T., Sitters, J., Ménard, C. B., and Olofsson, J. (2016). Reindeer grazing increases summer albedo by reducing shrub abundance in Arctic tundra. *Environmental Research Letters*, 11(12):125013.

Berman, S. M. (1994). A Bivariate Markov Process with Diffusion and Discrete Components. *Communications in Statistics. Stochastic Models*, 10(2):271–308.

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(3):613–627.

Blackwell, P. G., Niu, M., Lambert, M. S., and Lapoint, S. D. (2016). Exact Bayesian Inference for Animal Movement in Continuous Time. *Methods in Ecology and Evolution*, 7:184–195.

Bode, N. W. F., Wood, A. J., and Franks, D. W. (2011). The impact of social networks on animal collective motion. *Animal Behaviour*, 82:29–38.

Breed, G. A., Costa, D. P., Goebel, M. E., and Robinson, P. W. (2011). Electronic tracking tag programming is critical to data collection for behavioral time-series analysis. *Ecosphere*, 2(January):1–12.

Breed, G. A., Costa, D. P., Jonsen, I. D., Robinson, P. W., and Mills-Flemming, J. (2012). State-space methods for more completely capturing behavioral dynamics from animal tracks. *Ecological Modelling*, 235-236:49–48.

Breed, G. A., Jonsen, I. D., Myers, R. A., Don Bowen, W., and Leonard, M. L. (2009). Sex-specific, seasonal foraging tactics of adult grey seals (Halichoerus grypus) revealed by state-space analysis. *Ecology*, 90(11):3209–3221.

Brillinger, D. R., Preisler, H. K., Ager, A. A., Kie, J. G., and Stewart, B. S. (2002). Employing stochastic differential equations to model wildlife motion. *Bulletin of the Brazilian Mathematical Society*, 33(3):385–408.

Brooks, S. P. and Gelman, A. (1997). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455.

Broom, M., Erovenko, I. V., Rowell, J. T., and Rychtar, J. (2019). Models and measures of animal aggregation and dispersal. *Journal of Theoretical Biology*, 484(2020):2.

Buckland, S. T., Newman, K. B., Thomas, L., and Koesters, N. B. (2004). State-space models for the dynamics of wild animal populations. *Ecological Modelling*, 171(1-2):157–175.

Buhl, J., Sumpter, D. J., Couzin, I. D., Hale, J. J., Despland, E., Miller, E. R., and Simpson, S. J. (2006). From Disorder to Order in Marching Locusts. *Science*, 312(5778):1402–1406.

Cagnacci, F., Boitani, L., Powell, R. A., and Boyce, M. S. (2010). Animal ecology meets GPS-based radiotelemetry : a perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, pages 2157–2162.

Calabrese, J. M., Fleming, C. H., Fagan, W. F., Rimmler, M., Kaczensky, P., Bewick, S., Leimgruber, P., and Mueller, T. (2018). Disentangling social interactions and environmental drivers in multi-individual wildlife tracking data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1746).

Calabrese, J. M., Fleming, C. H., and Gurarie, E. (2016). Ctmm: an R Package for Analyzing Animal Relocation Data as a Continuous-Time Stochastic Process. *Methods in Ecology and Evolution*, 7(9):1124–1132.

Codling, E. A., Pitchford, J. W., and Simpson, S. D. (2007). Group Navigation and The 'Many-Wrongs Principle' in Models of Animal Movement. *Ecology*, 88(7):1864–1870.

Cody, M. L. (1971). Finch flocks in the Mohave Desert. *Theoretical Population Biology*, 2(2):142–158.

Colman, J. E., Pedersen, C., Hjermann, D. Ø., and Holand, Ø. (2003). Do Wild Reindeer Exhibit Grazing Compensation during Insect Harassment ? *Journal of Wildlife Management*, 67(1):11–19.

Cooke, S. J., Hinch, S. G., Wikelski, M., Andrews, R. D., Kuchel, L. J., Wolcott, T. G., and Butler, P. J. (2004). Biotelemetry: A mechanistic approach to ecology. *Trends in Ecology and Evolution*, 19(6):334–343.

Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. A. (2005). Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025):513–516.

Couzin, I. D., Krause, J., James, R., Ruxton, G. D., and Franks, N. R. (2002). Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218(1):1–11.

Craighead, F. C. and Craighead, J. J. (1972). Data on Grizzly Bear Denning Activities and Behavior Obtained by Using Wildlife Telemetry. *International Association for Bear Research and Management*, 2:84–106.

Croft, S., Budgey, R., Pitchford, J. W., and Wood, A. J. (2013). The influence of group size and social interactions on collision risk with obstacles. *Ecological Complexity*, 16:77–82.

Croft, S., Budgey, R., Pitchford, J. W., and Wood, A. J. (2015). Obstacle avoidance in social groups: New insights from asynchronous models. *Journal of the Royal Society Interface*, 12(106).

Croxall, J. P., Silk, J. R., Phillips, R. A., Afanasyev, V., and Briggs, D. R. (2005). Global circumnavigations: Tracking year-round ranges of nonbreeding albatrosses. *Science*, 307(5707):249–250.

Da Prato, G. and Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions*. Cambridge University Press.

Dalziel, B. D., Corre, M. L., Côté, S. D., and Ellner, S. P. (2016). Detecting collective behaviour in animal relocation data, with application to migrating caribou. *Methods in Ecology and Evolution*, 7(1):30–41.

Delgado, M. d. M., Miranda, M., Alvarez, S. J., Gurarie, E., Fagan, W. F., Penteriani, V., di Virgilio, A., and Morales, J. M. (2018). The importance of individual variation in the dynamics of animal collective movements. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1746):20170008.

Doncaster, C. P. (1990). Non-parametric estimates of interaction from radio-tracking data. *Journal of Theoretical Biology*, 143(4):431–443.

Downes, C. M., Theberge, J. B., and Smith, S. M. (1986). The influence of insects on the distribution, microhabitat choice, and behavior of the Burwash caribou herd. *Canadian Journal of Zoology*, 64(3):622–629.

Dunn, J. E. and Gipson, P. S. (1977). Analysis of Radio Telemetry Data in Studies of Home Range. *Biometrics*, 33(1):85–101.

Durbin, J. and Koopman, S. (2001). *Time Series Analysis by State Space Methods.* Oxford University Press, Oxford.

Falldorf, T., Strand, O., Panzacchi, M., and Tømmervik, H. (2014). Estimating lichen volume and reindeer winter pasture quality from Landsat imagery. *Remote Sensing of Environment*, 140:573–579.

Farhadinia, M. S., Michelot, T., Johnson, P. J., Hunter, L. T. B., and Macdonald, D. W. (2020). Understanding decision making in a food-caching predator using hidden Markov models. *Movement Ecology*, 8(9).

Fearnhead, P. and Sherlock, C. (2006). An exact Gibbs sampler for the Markov-modulated poisson process. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 68(5):767–784.

Fleming, C. H., Calabrese, J. M., Mueller, T., Olson, K. A., Leimgruber, P., and Fagan, W. F. (2014). From Fine-Scale Foraging to Home Ranges: A Semivariance Approach to Identifying Movement Modes across Spatiotemporal Scales. *The American Naturalist*, 183(5).

Fleming, C. H., Sheldon, D., Gurarie, E., Fagan, W. F., Lapoint, S., and Calabrese, J. M. (2017). Kálmán filters for continuous-time movement models. *Ecological Informatics*, 40:8–21.

Folstad, I., Nilssen, A. C., Halvorsen, O., and Andersen, J. (1991). Parasite avoidance: the cause of post-calving migrations in Rangifer? *Canadian Journal of Zoology*, 69(9):2423–2429.

Forbes, B., Bölter, M., Müller-Wille, L., Hukkinen, J., Müller, F., N.Gunslay, and Konstantinov, Y. (2006). *Reindeer Management in Northernmost Europe*, volume 184. Springer, Berlin, Heidelberg.

Forbes, B. C. and Kumpula, T. (2009). The ecological role and geography of reindeer (rangifer tarandus) in Northern Eurasia. *Geography Compass*, 3(4):1356–1380.

Forbes, B. C., Kumpula, T., Meschtyb, N., Laptander, R., MacIas-Fauria, M., Zetterberg, P., Verdonen, M., Skarin, A., Kim, K. Y., Boisvert, L. N., Stroeve, J. C., and Bartsch, A. (2016). Sea ice, rain-on-snow and tundra reindeer nomadism in Arctic Russia. *Biology Letters*, 12(11):4–8.

Frair, J. L., Fieberg, J., Hebblewhite, M., Cagnacci, F., Decesare, N. J., and Pedrotti, L. (2010). Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365:2187–2200.

Frühwirth-Schnatter, S. (2001). Fully Bayesian analysis of switching gaussian state space models.

Fyfe, J. C., Von Salzen, K., Gillett, N. P., Arora, V. K., Flato, G. M., and McConnell, J. R. (2013). One hundred years of Arctic surface temperature variation due to anthropogenic influence. *Scientific Reports*, 3(September):4–11.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013a). *Bayesian data analysis, third edition*.

Gelman, A., Hwang, J., and Vehtari, A. (2013b). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24:997–1016.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statisitical Science*, 7(4):457–472.

Gilks, W., Richardson, S., and Spiegelhalter, D. (1996). *Markov chain Monte Carlo in practice*. Chapman & Hall, London.

Gunn, A. (2016). Rangifer tarandus. The IUCN Red List of Threatened Species 2016.

Gurarie, E., Andrews, R. D., and Laidre, K. L. (2010). A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, 12(5):395–408.

Guttorp, P. and Minin, V. N. (1995). *Stochastic modeling of scientific data*. CRC Press, 1st edition.

Hagemoen, R. I. M. and Reimers, E. (2002). Reindeer summer activity pattern in relation to weather and insect harassment. *Journal of Animal Ecology*, 71(5):883–892.

Hall, S., Nielson, Ryan, M., Lindzey, F., and McDonald, L. L. (2006). Winter Habitat Selection of Mule Deer Before and During Development of a Natural Gas Field. *Journal of Wildlife Management*, 70(2):396–403.

Hansen, B. B., Aanes, R., Herfindal, I., Kohler, J., Sæther, B. E., and Oli, M. K. (2011). Climate, icing, and wild arctic reindeer: Past relationships and future prospects. *Ecology*, 92(10):1917–1923.

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

Hart, B. L. and Mooring, M. S. (1992). Animal Grouping for Protection From Parasites: Selfish Herd and Encounter-Dilution Effects. *Behaviour*, 123(3):173–193.

Hastings, W. K. (1970). Monte carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.

Haydon, D. T., Morales, J. M., Yott, A., Jenkins, D. A., Rosatte, R., and Fryxell, J. M. (2008). Socially informed random walks: Incorporating group dynamics into models of population spread and growth. *Proceedings of the Royal Society B: Biological Sciences*, 275(1638):1101–1109.

Hebblewhite, M. and Haydon, D. T. (2010). Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1746):20170007.

Helle, T. and Aspi, J. (1983). Does herd formation reduce insect harassment among reindeer? A field experiment with animal traps. *Acta Zoologica Fennica*, 175(January):129–131.

Henttonen, H. and Tikhonov, A. (2008). Rangifer tarandus. The IUCN Red List of Threatened Species 2008.

Herbert-Read, J. E. (2016). Understanding how animal groups achieve coordinated movement. *Journal of Experimental Biology*, 219:2971–2983.

Hislop, L. (2013). *The View from the Top: Searching for responses to a rapidly changing Arctic*. United Nations Environment Programme.

Holgate, P. (1971). *Random walk models for animal behavior.*, volume 2. Penn State University Press, University Park, PA.

Hooten, M. B., Scharf, H. R., Hefley, T. J., Pearse, A. T., and Weegman, M. D. (2018). Animal movement models for migratory individuals and groups. *Methods in Ecology and Evolution*, 2018(March):1–14.

Huth, A. and Wissel, C. (1992). The simulation of the movement of fish schools. *Journal of Theoretical Biology*, 156(3):365–385.

ISO (2020). ISO International Standard ISO/IEC 14882:2020(E) – Programming Language C++.

Johnson, D. S., London, J. M., Lea, M.-A., and Durban, J. W. (2008). Continuous-Time Correlated Random Walk Model for Animal Telemetry Data. *Ecology*, 89(5):1208–1215.

Jonsen, I. D., Flemming, J. M., and Myers, R. A. (2005). Robust state-space modeling of animal movement data. *Ecology*, 86(11):2874–2880.

Jonsen, I. D., Myers, R. A., and Flemming, J. M. (2003). Meta-analysis of animal movement using state-space models. *Ecology*, 84(11):3055–3063.

Joo, R., Etienne, M.-P., Bez, N., and Mahévas, S. (2018). Metrics for describing dyadic movement: a review. *Movement Ecology*, 6(26).

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering, Transactions of the ASME*, 82(1):35–45.

Kays, R., Crofoot, M. C., Jetz, W., and Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, 348(6240):2478–2478.

Klein, D. R. (2000). Arctic grazing systems and industrial development: Can we minimize conflicts? *Polar Research*, 19(1):91–98.

Kofinas, G., Osherenko, G., Klein, D., and Forbes, B. (2000). Research planning in the face of change: The human role in reindeer/caribou systems. *Polar Research*, 19(1):3–21.

Kojola, I. (1989). Mother's dominance status and differential investment in reindeer calves. *Animal Behaviour*, 38(2):177–185.

Kutz, S. J., Hoberg, E. P., Molnár, P. K., Dobson, A., and Verocai, G. G. (2014). A walk on the tundra: Host-parasite interactions in an extreme environment. *International Journal for Parasitology: Parasites and Wildlife*, 3(2):198–208.

Langrock, R., Hopcraft, J. G. C., Blackwell, P. G., Goodall, V., King, R., Niu, M., Patterson, T. A., Pedersen, M. W., Skarin, A., and Schick, R. S. (2014). Modelling group dynamic animal movement. *Methods in Ecology and Evolution*, 5(2):190–199.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., and Morales, J. M. (2012). Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, 93(11):2336–2342.

Lee, S. E., Press, M. C., Lee, J. A., Ingold, T., and Kurttila, T. (2000). Regional effects of climate change on reindeer: A case study of the Muotkatunturi region in Finnish Lapland. *Polar Research*, 19(1):99–105.

Leos-Barajas, V., Gangloff, E. J., Adam, T., Langrock, R., Van Beest, F. M., Nabe-Nielsen, J., and Morales, J. M. (2017). Multi-scale Modeling of Animal Movement and General Behavior Data Using Hidden Markov Models with Hierarchical Structures. *Journal of Agricultural, Biological, and Environmental Statistics*, 22(3):2320–0248.

Long, J. A., Trisalyn, N. A., Webb, S. L., and Gee, K. L. (2014). A critical examination of indicies of dynamic interaction for wildlife telemtry studies. *Journal of Animal Ecology.*, 83(5):1216–1233.

Lord, R. D. J., Bellrose, F. C. ., and Cochran, W. W. (1962). Radiotelemetry of the Respiration of a Flying Duck. *Science*, 137:39–41.

Macdonald, D., Ball, F., and Hough N.G. (1980). *A Handbook on Biotelemetry and Radio Tracking*. Pergamon.

Maplesoft (2018). Maple.

McClintock, B. T., Johnson, D. S., Hooten, M. B., Ver Hoef, J. M., and Morales, J. M. (2014). When to be discrete: the importance of time formulation in understanding animal movement. *Movement Ecology*, 2(1):21.

McClintock, B. T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B. J., and Morales, J. M. (2012). A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs*, 82(3):335–349.

McClintock, B. T. and Michelot, T. (2018). momentuHMM: R package for generalized hidden Markov models of animal movement. *Methods in Ecology and Evolution*, 9(6):1518–1530.

McClintock, B. T., Russell, D. J., Matthiopoulos, J., and King, R. (2013). Combining individual animal movement and ancillary biotelemetry data to investigate population-level activity budgets. *Ecology*, 94(4):838–849.

Mckellar, A. E., Langrock, R., Walters, J. R., and Kesler, D. C. (2015). Behavioral Ecology Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird. *Behavioral Ecology*, 26(1):148–157.

Merkle, J. A., Monteith, K. L., Aikens, E. O., Hayes, M. M., Hersey, K. R., Middleton, A. D., Oates, B. A., Sawyer, H., Scurlock, B. M., and Kauffman, M. J. (2016). Large herbivores surf waves of green-up during spring. *Proceedings of the Royal Society B: Biological Sciences*, 283(1833).

Merkle, J. A., Potts, J. R., and Fortin, D. (2017). Energy benefits and emergent space use patterns of an empirically parameterized model of memory-based patch selection. *Oikos*, 126(2):185–195.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

Michelot, T. and Blackwell, P. G. (2019). State-switching continuous-time correlated random walks. *Methods in Ecology and Evolution*, 10(5):637–649.

Michelot, T., Langrock, R., and Patterson, T. A. (2016). moveHMM An R package for the analysis of animal movement data. *Methods in Ecology and Evolution*, 7(11).

Milner, J. E., Blackwell, P. G., and Niu, M. (2020). Modelling and Inference for the Movement of Interacting Animals. *Methods in Ecology and Evolution*.

Moorcroft, P. R., Lewis, M. A., and Crabtree, R. L. (1999). Home range analysis using a mechanistic home range model. *Ecology*, 80(5):1656–1665.

Morales, J. M. and Ellner, S. P. (2002). Scaling up animal movements in heterogeneous landscapes: The importance of behavior. *Ecology*, 83(8):2240–2247.

Morales, J. M., Haydon, Author, W. D. T., Frair, J., Holsinger, K. E., and Fryxell, J. M. (2004). Extracting More out of Relocation Data: Building Movement Models as Mixtures of Random. *Ecology*, 85(9):2436–2445.

Mörschel, F. M. (1999). Use of Climatic Data to Model the Presence of Oestrid Flies in Caribou Herds. *The Journal of Wildlife Management*, 63(2):588–593.

Morschel, F. M. and Klein, D. R. (1997). Effects of weather and parasitic insects on behavior and group dynamics of caribou of the Delta Herd, Alaska. *Canadian Journal of Zoology*, 75(10):1659–1670.

Mueller, T., Olson, K. A., Dressler, G., Leimgruber, P., Fuller, T. K., Nicolson, C., Novaro, A. J., Bolgeri, M. J., Wattles, D., Destefano, S., Calabrese, J. M., and Fagan, W. F. (2011). How landscape dynamics link individual- to population-level movement patterns: A multispecies comparison of ungulate relocation data. *Global Ecology and Biogeography*, 20(5):683–694.

Munden, R., Börger, L., Wilson, R. P., Redcliffe, J., Loison, A., Garel, M., and Potts, J. R. (2019). Making sense of ultrahigh-resolution movement data: A new algorithm for inferring sites of interest. *Ecology and Evolution*, 9(1):265–274.

Niu, M., Blackwell, P. G., and Skarin, A. (2016). Modeling interdependent animal movement in continuous time. *Biometrics*, 72(2):315–324.

Niu, M., Frost, F., Milner, J. E., Skarin, A., and Blackwell, P. G. (2020). Modelling group movement with behavior switching in continuous time. *Biometrics*.

Odum, E. P. and Kuenzler, E. J. (1955). Measurement of Territory and Home Range Size in Birds. *The Auk*, 72(2):128–137.

Ovaskainen, O. (2004). Habitat-specific movement parameters estimated using mark-recapture data and a diffusion model. *Ecology*, 85(1):242–257.

Panzacchi, M., Van Moorter, B., Jordhøy, P., and Strand, O. (2013). Learning from the past to predict the future: Using archaeological findings and GPS data to quantify reindeer sensitivity to anthropogenic disturbance in Norway. *Landscape Ecology*, 28(5):847–859.

Pape, R. and Löffler, J. (2012). Climate change, land use conflicts, predation and ecological degradation as challenges for reindeer husbandry in northern europe: What do we really know after half a century of research? *Ambio*, 41(5):421–434.

Parton, A. and Blackwell, P. G. (2017). Bayesian Inference for Multistate 'Step and Turn' Animal Movement in Continuous Time. *Journal of Agricultural, Biological and Environmental Statistics*, 22.

Parton, A., Blackwell, P. G., and Skarin, A. (2016). Bayesian Inference for Continuous Time Animal Movement Based on Steps and Turns. *Bayesian Statistics in Action: BAYSM 2016, Florence, Italy*, 194.

Patterson, T. A., McConnell, B. J., Fedak, M. A., Bravington, M. V., and Hindell, M. A. (2010). Using GPS data to evaluate the accuracy of state – space methods for correction of Argos satellite telemetry error. *Ecological Society of America*, 91(1):273–285.

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. (2008). State-space models of individual animal movement. *Trends in Ecology and Evolution*, 23(2):87–94.

Pettorelli, N., Mysterud, A., Yoccoz, N. G., Langvatn, R., and Stenseth, N. C. (2005a). Importance of climatological downscaling and plant phenology for red deer in heterogeneous landscapes. *Proceedings of the Royal Society B: Biological Sciences*, 272(1579):2357–2364.

Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jedrzejewska, B., Lima, M., and Kausrud, K. (2011). The Normalized Difference Vegetation Index (NDVI): Unforeseen successes in animal ecology. *Climate Research*, 46(1):15–27.

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J., and Stenseth, N. C. (2005b). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology and Evolution*, 20(9):503–510.

Pettorelli, N., Weladji, R. B., Holand, Ø., Mysterud, A., Breie, H., and Stenseth, N. C. (2005c). The relative role of winter and spring conditions: Linking climate and landscape-scale plant phenology to alpine reindeer body mass. *Biology Letters*, 1(1):24–26.

Polansky, L. and Wittemyer, G. (2011). A framework for understanding the architecture of collective movements using pairwise analyses of animal movement data. *Journal of the Royal Society Interface*, 8(56):322–333.

Polansky, L., Wittemyer, G., Cross, P. C., Tambling, C. J., and Getz, W. M. (2010). From moonlight to movement and synchronized randomness: Fourier and wavelet analyses of animal location time series data. *Ecology*, 91(5):1506–1518.

Polfus, J. L., Hebblewhite, M., and Heinemeyer, K. (2011). Identifying indirect habitat loss and avoidance of human infrastructure by northern mountain woodland caribou. *Biological Conservation*, 144(11):2637–2646.

Pörtner, H.-O., Roberts, D., Masson-Delmotte, V., Zhai, P., Tignor, M., Poloczanska, E., Mintenbeck, K., Alegria, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., and Weyer, N. (2019). IPCC, 2019: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate. Technical report, UNEP/WMO.

Potts, J. R. and Lewis, M. A. (2014). How do animal territories form and change? lessons from 20 years of mechanistic modelling. *Proceedings of the Royal Society B: Biological Sciences*, 281(1784).

Power, M. E., Tilman, D., Estes, J. A., Menge, B. A., Bond, W. J., Mills, L. S., Daily, G., Castilla, J. C., Lubchenco, J., and Paine, R. T. (1996). Challenges in the Quest for Keystones. *BioScience*, 46(8):609–620.

Preisler, H. K., Ager, A. A., and Wisdom, M. J. (2013). Analyzing animal movement patterns using potential functions. *Ecosphere*, 4(3):1–13.

Putkonen, J. and Roe, G. (2003). Rain-on-snow events impact soil temperatures and affect ungulate survival. *Geophysical Research Letters*, 30(4).

R Core Team (2020). R: A Language and Environment for Statistical Computing.

Rao, V. and Teh, Y. W. (2013). Fast MCMC sampling for Markov jump processes and extensions. *Journal of Machine Learning Research*, 14:3295–3320.

Reimers, E. (1997). Rangifer population ecology: a Scandinavian perspective. *Rangifer*, 17(3):105.

Rivrud, I. M., Sivertsen, T. R., Mysterud, A., Åhman, B., Støen, O. G., and Skarin, A. (2018). Reindeer green-wave surfing constrained by predators. *Ecosphere*, 9(5).

Roever, C. L., Beyer, H. L., Chase, M. J., and Van Aarde, R. J. (2014). The pitfalls of ignoring behaviour when quantifying habitat selection. *Diversity and Distributions*, 20:322–333.

Roquet, F., Williams, G., Hindell, M. A., Harcourt, R., McMahon, C., Guinet, C., Charrassin, J. B., Reverdin, G., Boehme, L., Lovell, P., and Fedak, M. (2014). A Southern Indian Ocean database of hydrographic profiles obtained with instrumented elephant seals. *Scientific Data*, 1:1–10.

Rowland, M. M., Wisdom, M. J., Johnson, B. K., and Kie, J. G. (2000). Elk Distribution and Modeling in Relation to Roads. *The Journal of Wildlife Management*, 64(3):672.

Russell, D. E., Martell, A. M., and Nixon, W. A. C. (1993). Range Ecology of the Porcupine Caribou Herd in Canada. *Rangifer*, 13(5):1.

Russell, J. C., Hanks, E. M., Modlmeier, A. P., and Hughes, D. P. (2017). Modeling Collective Animal Movement Through Interactions in Behavioral States. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3):313–334.

Sandström, P. (2015). *A toolbox for Co-production of Knowledge and Improved Land Use Dialogues - The Perspective of Reindeer Husbandry*. Department of Forest Resource Management, Swedish University of Agricultural Sciences.

Särkkä, S. (2013). *Bayesian filtering and smoothing*. Cambridge University Press, Cambridge, UK.

Schach, S. (1971). Weak Convergence Results for a Class of Multivariate Markov Processes. *The Annals of Mathematical Statistics*.

Scharf, H. R. and Buderman, F. E. (2020). Animal movement models for multiple individuals. *Wiley Interdisciplinary Reviews: Computational Statistics*, 12(6):1–16.

Scharf, H. R., Hooten, M. B., Fosdick, B. K., Johnson, D. S., London, J. M., and Durban, J. W. (2016). Dynamic social networks based on movement. *Annals of Applied Statistics*, 10(4):2182–2202.

Scharf, H. R., Hooten, M. B., Johnson, D. S., and Durban, J. W. (2018). Process convolution approaches for modeling interacting trajectories. *Environmetrics*, 29(3):1–18.

Schlägel, U. E., Signer, J., Herde, A., Eden, S., Jeltsch, F., Eccard, J. A., and Dammhahn, M. (2019). Estimating interactions between individuals from concurrent animal movements. *Methods in Ecology and Evolution*, 10(8):1234–1245.

Shepard, E. L., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., Halsey, L. G., Gleiss, A., Morgan, D. T., Myers, A. E., Newman, C., and Macdonald, D. W. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research*, 10(1):47–60.

Siniff, D. and Jessen, C. (1969). A simulation model of animal movement patterns. *Advances in Ecological Research*, 6:185–217.

Skarin, A. and Åhman, B. (2014). Do human activity and infrastructure disturb domesticated reindeer? The need for the reindeer's perspective. *Polar Biology*, 37(7):1041–1054.

Skarin, A., Danell, Ö., Bergström, R., and Moen, J. (2008). Summer habitat preferences of GPS-collared reindeer Rangifer tarandus tarandus. *Wildlife Biology*, 14(1):1–15.

Skarin, A., Danell, Ö., Bergström, R., and Moen, J. (2010). Reindeer movement patterns in alpine summer ranges. *Polar Biology*, 33(9):1263–1275.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 64(4):583–616.

Strandburg-Peshkin, A., Farine, D. R., Couzin, I. D., and Crofoot, M. C. (2015). Shared decision-making drives collective movement in wild baboons. *Science (New York, N.Y.)*, 348(6241):1358–61.

Thomas, B., Holland, J. D., and Minot, E. O. (2011). Wildlife tracking technology options and cost considerations. *Wildlife Research*, 38(8):653–663.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collinghan, Y. C., Erasmus, B. F. N., de. Siqueira, M. F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van. Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, M. A., Peterson, A. T., Phillips, O. L., and Williams, S. E. (2004). Extinction risk from climate change. *Nature*, 427:294–296.

Thurfjell, H., Ciuti, S., and Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, 2(1):1–12.

Tomkiewicz, S. M., Fuller, M. R., Kie, J. G., and Bates, K. K. (2010). Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2163–2176.

Torney, C. J., Berdahl, A. M., Lamont, M., Debell, L., Angohiatok, R. J., Leclerc, L.-M. M., and Berdahl, A. M. (2018). Inferring the rules of social interaction in migrating caribou. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1746):20170385.

Tryland, M. and Kutz, S. (2018). *Reindeer and Caribou: Health and Disease*. CRC Press.

Turchin, P. (1999). Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants. In *The Quarterly Review of Biology*, volume 74. Stony Brook University.

Tyler, N. J. C., Turi, J. M., Sundset, M. A., Strøm Bull, K., Sara, M. N., Reinert, E., Oskal, N., Nellemann, C., McCarthy, J. J., Mathiesen, S. D., Martello, M. L., Magga, O. H., Hovelsrud, G. K., Hanssen-Bauer, I., Eira, N. I., Eira, I. M. G., and Corell, R. W. (2007). Saami reindeer pastoralism under climate change: Applying a generalized framework for vulnerability studies to a sub-arctic social-ecological system. *Global Environmental Change*, 17(2):191–206.

Uhlenbeck, G. E. and Ornstein, L. S. (1930). On the Theory of the Brownian Motion. *American Physical Society*, 36(5):823–841.

Vistnes, I. and Nellemann, C. (2001). Avoidance of Cabins , Roads , and Power Lines by Reindeer during Calving. *The Journal of Wildlife Management*, 65(4):915–925.

Vistnes, I., Nellemann, C., Jordhøy, P., and Strand, O. (2004). Effects of Infrastructure on Migration and Range Use of Wild Reindeer. *The Journal of Wildlife Management*, 68(1):101–108.

Vistnes, I. I., Nellemann, C., Jordhøy, P., and Støen, O. G. (2008). Summer distribution of wild reindeer in relation to human activity and insect stress. *Polar Biology*, 31(11):1307–1317.

Vors, L. S. and Boyce, M. S. (2009). Global declines of caribou and reindeer. *Global Change Biology*, 15(11):2626–2633.

Wang, Y. S., Blackwell, P. G., Merkle, J. A., and Potts, J. R. (2019). Continuous time resource selection analysis for moving animals. *Methods in Ecology and Evolution*, 10(10):1664–1678.

Watanabe, S. (2010). Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *Journal of Machine Learning Research*, 11(2010):3571–3594.

Westley, P. A. H., Berdahl, A. M., Torney, C. J., Biro, D., Westley, P. A. H., Torney, C. J., and Biro, D. (2018). Collective movement in ecology: from emerging technologies to conservation and management. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1746):20170004.

White, R., Thomson, B., Skogland, T., Person, S., Russell, D., and Holleman, D. (1975). Ecology of caribou at Prudhoe Bay, Alaska. Ecological Investigations of the Tundra Biome in the Prudhoe Bay Region, Alaska (ed. J. Brown). *Biological Papers of the University of Alaska, Fairbanks, AK, USA.*, pages 151–187.

Witter, L. A., Johnson, C. J., Croft, B., Gunn, A., and Gillingham, M. P. (2012a). Behavioural trade-offs in response to external stimuli: Time allocation of an Arctic ungulate during varying intensities of harassment by parasitic flies. *Journal of Animal Ecology*, 81(1):284–295.

Witter, L. A., Johnson, C. J., Croft, B., Gunn, A., and Poirier, L. M. (2012b). Gauging climate change effects at local scales: Weather-based indices to monitor insect harassment in caribou. *Ecological Applications*, 22(6):1838–1851.

Zimmerman, J. and Powell, R. A. (1995). Confidence ellipses. *Canadian Journal of Zoology*, 73(6):1123–1133.

Zucchini, W., MacDonald, I. L., and Langrock, R. (2016). *Hidden Markov models for time series : an introduction using R*. Chapman and Hall/CRC, 2nd edition.

# Appendix A

# Watanabe–Akaike information criterion (WAIC)

Here we discuss the Watanabe–Akaike information criterion (WAIC) (Watanabe, 2010) as formulated in the text by Gelman et al. (2013b). The WAIC is an information criterion used in Bayesian model selection. Heuristically, we can think of this as comprising of two components; the first, a measure of fit given by the log pointwise posterior predictive density and secondly, a penalisation for the effective number of parameters to adjust for overfitting.

For practical purposes we will give the computed formulations of WAIC as given in Gelman et al. (2013b), that is what can be computed from draws of MCMC simulations as we have in this thesis. Given data $y_1, \dots y_n$ which are modelled with parameters $\boldsymbol{\theta}$, WAIC is calculated using draws from the posterior $p(\boldsymbol{\theta})$, $\boldsymbol{\theta}^s$, where $s = 1, \dots, S$ i.e. each of the accepted samples of the MCMC. The WAIC is computed as

$$-2(lppd - p_{waic}), \tag{A.1}$$

where $lppd$, the log pointwise predictive density, is given as

$$lppd = \sum_{i=1}^{n} \log \left( \frac{1}{S} \sum_{s=1}^{S} p(y_i|\boldsymbol{\theta}^s) \right), \tag{A.2}$$

and $p_{waic}$, the adjust for overfitting, can be given in two forms. The first, $p_{waic^1}$, uses averages over the $S$ posterior draws and is computed as

$$p_{waic^1} = 2\sum_{i=1}^{n}\left(\log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_i|\boldsymbol{\theta}^s)\right) - \frac{1}{S}\sum_{s=1}^{S}\log(p(y_i|\boldsymbol{\theta}^s)\right). \tag{A.3}$$

The second, $p_{waic^2}$, uses the posterior variance of the log predictive density for each $y_i$ and is calculated as

$$p_{waic^2} = \sum_{i=1}^{n} V_{s=1}^{S}(\log p(y_i|\boldsymbol{\theta}^s)), \tag{A.4}$$

where, $V_{s=1}^{S}$ represents the sample variance.