Deriving patterns from animal movement decisions: a comparison of approximation techniques and a continuous-time resource selection method

Yi-Shan Wang

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Science
School of Mathematics and Statistics

November 2018

ACKNOWLEDGEMENT

Firstly, I would like to express my gratitude to my supervisor, Dr Potts, for the continuous support of my PhD study, for his guidance and patience. His detailed comments on my writing helped me transform my obscure sentences into understandable paragraphs.

Besides my supervisor, I would like to thank Prof Blackwell for his support on statistical inference, and thank Dr Merkle for providing the mule deer data. I also would like to thank Ministry of Education of Taiwan for financial support.

My sincere thanks also goes to friends from Christ Church Fulwood in Sheffield. They constantly asked me about my progress and prayed for me.

Finally, I would like to thank my family: my husband, Caleb, our kids, my mum and my brother for their love and support.

Abstract

This thesis consists of two parts. First, I investigate the effect of using three partial differential equation (PDE) techniques on analysing some simple animal movement models. Results of examining a biased random walk show that an old approach from Patlak's work in 1953 can give a very poor approximation even in this very simple case, while more recent methods correctly describe the movement process. By analysing central-place foraging models and movement in heterogeneous landscapes, I show that more recent PDE techniques can provide more accurate approximations of space use patterns when the kernel describing the movement is sufficiently smooth. However, for non-smooth movement kernels, all methods can result in quantitatively misleading approximations. This analysis provides an insight into the conditions under which the PDE methods might perform better.

Second, I present two continuous-time modelling frameworks for analysing animal movement depending on selection of resources over the whole landscape or in the surrounding area. The models are parameterised by a Markov chain Monte Carlo (MCMC) algorithm, allowing for movement decisions made at any time. Based on these frameworks, I generate simulations in various situations, including migration and foraging in patchy or rasterised landscapes. Analysis of simulated trajectories reveals that the inference algorithm can successfully capture the parameter values used in simulations in most cases. I also fit the migration model to spring migration data of some mule deer (*Odocoileus hemionus*). The results imply that migration might be explained by the trade-off between resources and travel distance. This work addresses some limitations of methods relying on discrete-time movement models and therefore provides an advanced tool for understanding movement driven by environmental factors.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Animal movement plays a central role in understanding relationships and patterns in ecosystems. For instance, population distribution or space use patterns can be regarded as the consequence of movements (Börger et al., 2008). Movements can also stem from conspecific and interspecific interactions as well as interactions between animals and the environments (Lewis and Murray, 1993; Chetkiewicz and Boyce, 2009; Vanak et al., 2013).

By analysing animal movements, researchers have attempted to derive spatial patterns, including the formation of home ranges and territories (Moorcroft et al., 1999; Börger et al., 2008; Bateman et al., 2015; Potts and Lewis, 2016b; Merkle et al., 2017), utilisation distributions (Benhamou, 2011; Signer et al., 2017; Wilson et al., 2018), the space use patterns resulting from competition (Potts and Petrovskii, 2017) and the influence of spatial attributes (Forester et al., 2009).

Furthermore, studies of movement are also key to our insight into other aspects of the living world because movement governs not only the life of individuals but also patterns at scales from population, community to ecosystem (Nathan et al., 2008). In addition, the advance in tracking technology has enabled the collection of high-resolution data, which considerably improves our knowledge of the underlying mechanisms, causes and consequences of movement (Kays et al., 2015).

Behind the great importance of the analysis of movement in understanding the

living world, some challenges make it difficult to analyse movement data. Some major problems arise from behavioural changes, autocorrelation between observations and observation errors in data (Gurarie et al., 2009). The change of movement behaviour can stem from the heterogeneity of the environment and different activities in life and may have a greater impact on scaling up individual movement to space patterns than environmental factors do (Morales and Ellner, 2002). Much research has been devoted to the identification of behavioural modes in order to capture the scenario of animals' life such as migration more precisely (Gurarie et al., 2009; Bunnefeld et al., 2011; Pedersen et al., 2011; Fleming et al., 2014a; Bastille-Rousseau et al., 2016). The autocorrelation in position and velocity comes from the fine scale of data collection and can be incorporated in movement models such as a discrete-time correlated random walk (CRW) and a continuous-time stochastic movement model (e.g. the Ornstein-Uhlenbeck process). Beyond the consideration of autocorrelation in a movement model, Fleming et al. (2014b) and Fleming et al. (2017) developed methods to facilitate the use of continuous-time models to efficiently deal with autocorrelated data. For tackling the complexity caused by observation errors, a promising tool is a state-space model (Patterson et al., 2008; Albertsen et al., 2015), which uses a model to explicitly incorporates observation errors in addition to a model for movement process. Although this thesis will not focus on resolving these issues, they could be taken into account in future research based on the work of this thesis.

This thesis is composed of two topics related to the derivation of spatial patterns from animal movements. The first is to compare the accuracy of approximation methods for predicting population space use patterns from individual movement rules. The second introduces two modelling frameworks and an algorithm for inference to infer the preferences of animals for resources from movement data in two separate situations. In one case, an animal is assumed to assess the resources across the landscape when making movement decisions, similar to Ford (1983) and Mitchell and Powell (2004), whereas the other modelling method considers only the resources in the immediate vicinity of the animal's location, similar to Preisler et al. (2013). Real situations may reside in between these two extremes.

§1.1   A comparison of approximation techniques

The first topic discussed in this thesis focuses on methods for analysing individual movement mechanisms, represented by a function termed a *movement kernel*. In this part of the thesis, I will only consider models in a 1-dimensional space and not involving direction. A movement kernel $k_\tau(z|x)$ describes the probability of an animal moving to a place z in a (typically small) period of time τ, given its current location x. This probability can be affected by factors such as distance to the destination from the animal's position, environmental conditions and interactions between animals and these factors can be integrated into an extended movement kernel (Potts, Mokross and Lewis, 2014). Thus it is convenient to use a movement kernel to describe individual movement rules, as the studies in Rhodes et al. (2005), where movements depend on the location of the animal's home range.

However, a movement kernel only represents the probability of selecting a position in a relatively short period of time compared to an animal's lifetime. On the other hand, understanding an ecosystem often involves long-term patterns at population level, described by some key information such as the distribution and abundance of a species. Therefore, techniques are necessary to scale up the decision-making process at individual level, described by a movement kernel, to space use patterns at population level. To scale up individual movement to long-term patterns, it is conceptually straightforward to use the Master Equation (ME), which propagates the movement kernel forward in time and is commonly used (Moorcroft and Barnett, 2008; Potts, Bastille-Rousseau, Murray, Schaefer and Lewis, 2014; Merkle et al., 2018):

$$u(x, t + \tau) = \int_{-\infty}^{\infty} k_\tau(x|y)u(y, t)\mathrm{d}y, \tag{1.1}$$

where the function $u(x, t)$ is the probability density of the animal's position x at time t. In practice, it requires the iteration of Equation (1.1) until the difference between distributions $u(x, t)$ and $u(x, t+\tau)$ is sufficiently small to obtain the long-term distribution at steady state. However, calculating this integral repeatedly can be computationally demanding, so more efficient approximation methods are

favoured.

Some techniques have been developed to approximate the long-term distribution by converting the ME to a partial differential equation (PDE) (e.g. Codling et al. (2008), Section 2.2). Using the solution to a PDE at steady state to estimate the long-term distribution is much more efficient than iterating the ME. Nevertheless, these approximation techniques usually require particular assumptions such as omitting higher order moments in the system to make it tractable. The extent to which these approximations are reliable is unknown. For instance, a second-order moment closure assumes that moments at orders higher than two can be expressed by only the first and second moments. However, such a closure can make a poor approximation of spatially structured populations if the spatial features involve much information at higher orders (Murrell et al., 2004). Therefore, my first object was to investigate the accuracy of the approximations using three such techniques, all of which are formulated in PDEs, and draw a comparison between them. This investigation was performed by examining either long-term or steady-state distributions in various example cases, described in detail below.

The oldest method was developed in the mid-20th century by Patlak (1953) and popularised in the context of animal movement by Turchin (1991). Hillen and Painter (2013) reviewed two more recently developed methods, namely the Hyperbolic Scaling and Moment Closure methods. Potts et al. (2016) used these three methods to examine animal distributions near a habitat edge in a one-dimensional interval composed of two segments featuring different spatial attributes. In this thesis, I consider the application of these PDE methods on the whole 1D real line as well as a finite interval in several situations.

First, I compared the three approximate methods by using them to examine a biased random walk, biased towards a fixed direction and simple enough that the PDEs involved can be solved analytically. This shows Patlak's method gives a poor approximation, while the other methods provide an exactly correct prediction. Since Patlak's approach can fail even a very simple example, it may also produce misleading results when analysing a more complicated movement. Therefore, it needs further consideration to understand the effect of the three methods on the steady-state distribution when examining more general cases.
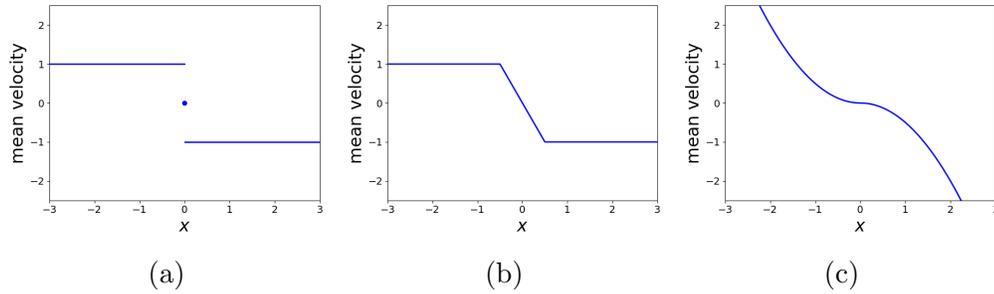
Figure 1.1: Mean velocity functions of central-place foraging models with different levels of continuity. (a) A discontinuous mean velocity function. (b) A continuous mean velocity function, having non-differentiable points (c) A differentiable mean velocity function.

Subsequently, I investigated cases where the PDEs derived could be solved analytically at steady state to obtain the long-term distributions. Such examples considered include three types of central-place foraging movement models. With the central place being located at the origin, the mean velocity functions of these models have different levels of continuity: the first type has a discontinuous point at the central place (e.g. Figure 1.1a), the second is continuous over the real line but has two non-differentiable points near the central place (e.g. Figure 1.1b) and the last is differentiable everywhere (e.g. Figure 1.1c). These three central-place foraging models were examined to reveal how changes in velocity affect the results of using the approximate methods.
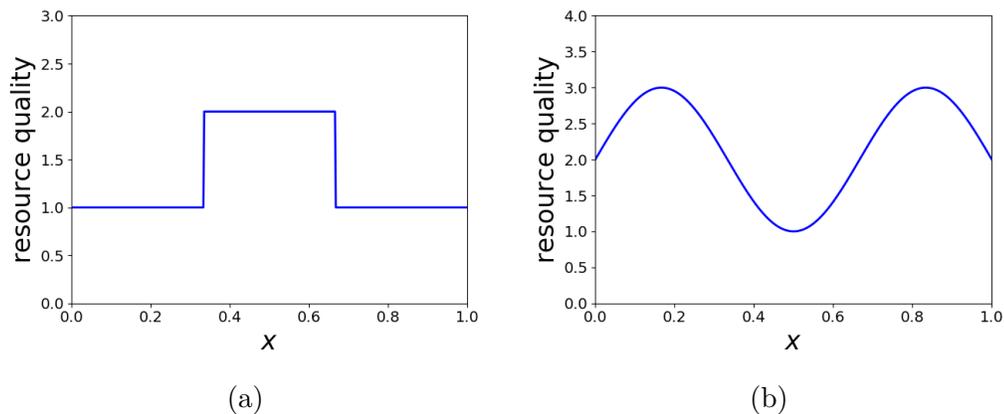
Figure 1.2: Examples of heterogeneous resources. (a) Resource quality with discontinuous jumps. (b) Resource quality which changes smoothly.

Finally, I used the PDE methods to analyse movement kernels combined with

heterogeneous resources of two types, one of which has discontinuous jumps (e.g. Figure 1.2a) and the other changes smoothly (e.g. Figure 1.2b). These two types of resources were considered to understand the impact of changes in resources on the approximation of long-term distribution. This is similar to the strategy which examines central-place foraging models with different level of smoothness.

In general, the results of examining these cases indicate that the PDE methods can provide poor approximations if the movement kernel is non-smooth and better estimates if the movement kernel is sufficiently smooth.

§ 1.2 A continuous-time resource selection method

The second part of the thesis introduces tools for analysing animals' movement responses to resources changing over time. As animals use resources in space selectively, understanding how they make selection decisions is central to gaining insight into underlying movement mechanisms (Cagnacci et al., 2010). A widely used tool for representing the selection of resources is the resource selection analysis (RSA), often relying on a resource selection function (RSF) (Manly et al., 2002), which takes values proportional to the probability of using a resource unit. When applying an RSF, a used resource unit is often compared to some other resource units which are available but not used. However, it may not be straightforward to define the availability of a resource unit and may not be realistic to assume that every unit across the land is equally available.

This problem of defining the availability of a resource unit has been resolved by step selection analysis (SSA) (Fortin et al., 2005; Forester et al., 2009; Thurfjell et al., 2014). Instead of considering the selection of locations in space, SSA examines the selection of 'steps', defined by linking two consecutive observed points. Each used step is compared to some random steps starting from the same beginning position of the used step. In this way, the availability of a step is naturally constrained by the animal's mobility, which can be described by the probability of moving from the starting to end points of the step. Furthermore, SSA has been extended to allow simultaneous estimation of the parameters of resource selection and movement, since movement traits such as velocity may be

influenced by resource selection. This extension of SSA is termed integrated step selection analysis (iSSA) (Avgar et al., 2016).

Although SSA and iSSA have successfully enhanced the analysis of resource selection by considering mobility, there are some drawbacks of SSA and iSSA because they rely on a discrete-time movement modelling framework. Since a discrete-time movement model requires data to be collected with fixed time intervals, this makes it difficult to manage data irregular in time (McClintock et al., 2014). Moreover, SSA and iSSA are based on the assumption that the animal decides to move exactly at points observed and never changes its mind between two observations. As a result, when applying SSA or iSSA, the time scale needs to correspond to decisions to move to reach a correct conclusion (Thurfjell et al., 2014). Only a few types of discrete-time movement models are robust enough to be adjusted to match the temporal scale to movement processes (Schlägel and Lewis, 2016).
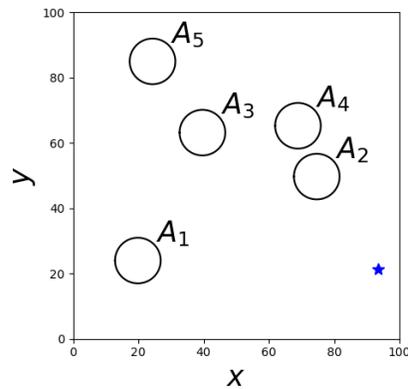
Figure 1.3: A landscape with five food patches to choose from. The blue star is the animal's location and circles A_1 to A_5 are food patches. The animal is assumed to choose its target patch by comparing the attractiveness of patches and selecting the most attractive patch.

To take advantage of incorporating movements into RSA yet circumvent the problems of using a discrete-time model, I embed a resource weighting function, which reflects the preferences of the animal, into a continuous-time movement model. For example, in Figure 1.3, there are five food patches in the landscape and I assume the animal decides which patch to move towards by comparing the patches' attractiveness, which is evaluated by a resource weighting function. After determining the destination, a continuous-time movement model is fitted to describe

(a)   (b)
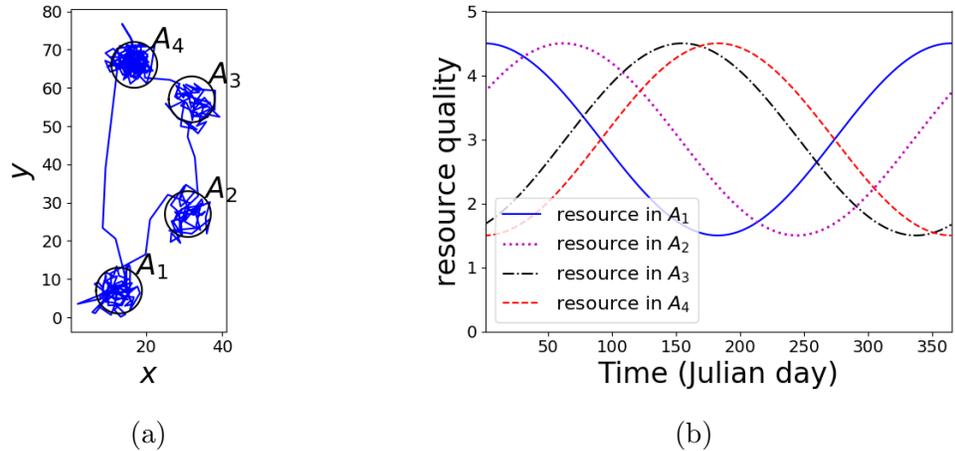
Figure 1.4: A simulated trajectory in a patchy landscape. (a) A simulation of migration generated on the assumption that the animal moves in response to the change of resource qualities in food patches. The animal was attracted to patch A_1 in the beginning and subsequently moved to patches A_2, A_3 and A_4. At last, it travelled back to patch A_1 (b) Resource qualities in the patches in Figure 1.4a.

how the animal approaches its target place. A simulated trajectory generated on these assumptions is shown in Figure 1.4a along with the resource qualities in food patches, given in Figure 1.4b. This scenario forms the basis of my first modelling framework, where the animal moves relying on complete knowledge of its environment. This strategy is the first to be able to consider movements triggered by factors in remote areas rather than being limited to decision-making at the scale of observation, although the idea of taking places far away into consideration has been mentioned in Bastille-Rousseau et al. (2018).

The second modelling framework assumes that animals move in response to local clues instead of considering resources across the land. This represents the situation where the dependence on perception predominates over the reliance on memory (cf. Bracis and Mueller (2017)). In this case, an animal is assumed to move in the direction up local resource gradient, which is calculated by evaluating the resource quality of nearby areas using a resource weighting function. Figure 1.5 gives an example. This is similar to movements with a drift term described by a potential function (Brillinger, 2010) and my modelling framework also allows for redirection by reassessing local resource qualities at any time.

To parameterise the models, I use a Bayesian inference procedure, which takes
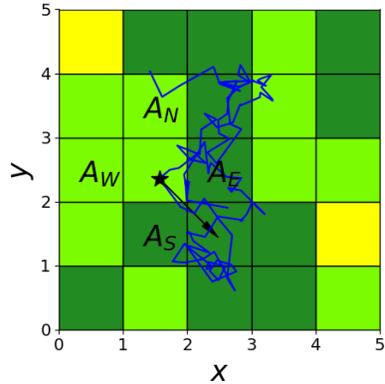
Figure 1.5: A simulation of movement following a local resource gradient. The patch colours, yellow, light green and dark green, represent low, medium and high resource qualities, assumed to be fixed in this example. The star is one of the simulated locations and the arrow points the direction up the resource gradient, determined by the resource qualities in patches A_N, A_S, A_E and A_W, where N, S, E, W stand for north, south, east and west. The animal is assumed to move in the direction up the resource gradient with some uncertainty.

into account the reality that movement decisions can be made at any time by augmenting the data with points where decision-making might occur, similar to the imputation of paths in Hanks et al. (2015). Moreover, my method is an advance in RSA as it is able to consider resource selection beyond the observation scale. The inference algorithm has been applied on both simulated and real data and the results show that the inference method is reliable at simultaneously parameterising a resource selection function and a movement process from movement data in a wide range of scenarios.

The simulated data examined consists of examples generated from the two modelling frameworks, one of which compares resource units across the landscape and the other only assesses local resources. For the former framework, three different situations were considered. In the first situation, I simulated movement in a landscape with several food patches. The resource qualities in the patches were assumed to change seasonally and be independent of the foraging activities of animals. This was used to simulate the scenario of migration and Figure 1.4 shows such an example. The second situation also used patchy landscapes but assumed the opposite for resource qualities. In this case, the resource qualities were depleted or renewed dependent on the time the animal spends in a resource patch. This can resemble movements of an animal in its home range. The third

situation was a raster-landscape version of the second. That is, simulated trajectories were generated in a rasterised landscape on the assumption that the resource in a cell was consumed as a result of the animal's presence in that cell and grew otherwise. Here, every food patch in a patchy landscape and every cell in a raster landscape is a resource unit, whose quality is assessed when deciding the target place of movement.

Simulated trajectories generated from the other modelling framework were in raster landscapes and the decision on movement direction only involved the four neighbouring cells in the north, south, east and west of the animal's current position (Figure 1.5). These simulated trajectories model movement relying on the perception of local clues, in contrast to the first modelling framework, which assumes a movement decision is made by assessing every resource unit, both far and near, in the landscape.

I applied the first modelling framework to analyse the spring migration data of 28 mule deer (*Odocoileus hemionus*) in the Greater Yellowstone Ecosystem to demonstrate the methodology and give a possible explanation for factors driving migration. The time and location data of individuals were collected every 2 hours from March to August 2016. The data points in the winter and summer ranges formed distinguishable clusters as the distance travelled from winter to summer ranges was about 75 kilometres on average. For some individuals, there were also small clusters of data points along the migration path. To represent resource qualities, I used the normalised difference vegetation index (NDVI) and instantaneous rate of green-up (IRG) over the area where the mule deer were observed. In addition, I also considered resource qualities defined by the integral of NDVI from a time of observation to the end of summer. Figure 1.6 shows the migration data of a mule deer along with resource qualities in recognised foraging patches.

1.2.1 PREVIOUS APPROACHES

The goal of the second part of the thesis is to develop a tool for RSA by integrating a RSF into a continuous-time movement model rather than a discrete-time
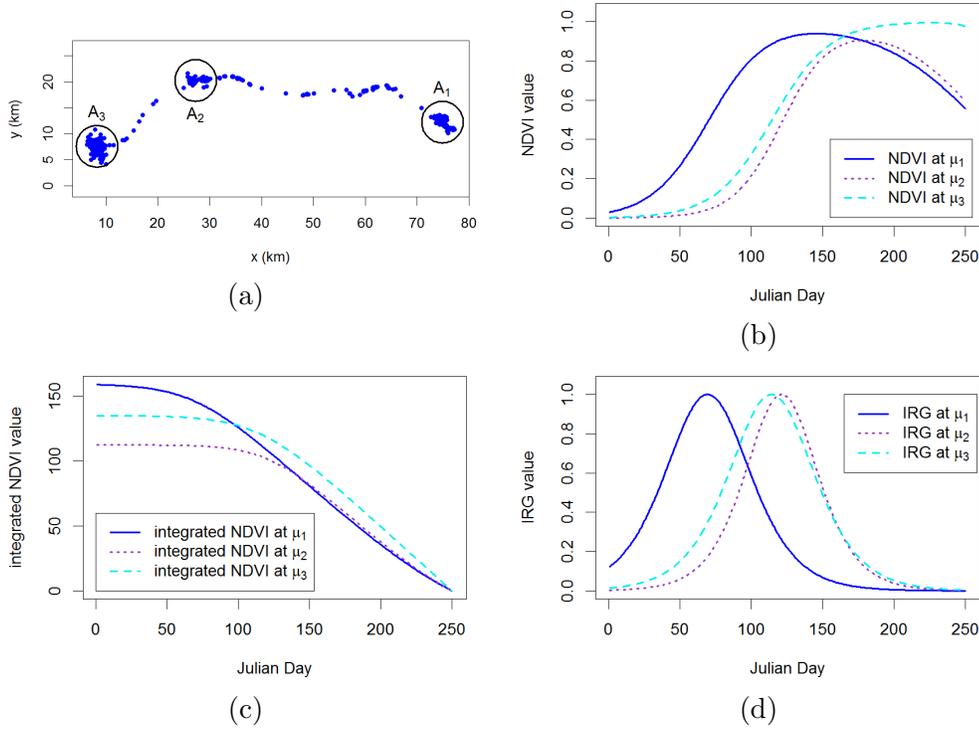
Figure 1.6: Migration data of a mule deer and relevant resource data. (a) The location data collected from March to August 2016 in the Greater Yellowstone Ecosystem. The circles A_1, A_2 and A_3 are identified patches with centres μ_1, μ_2 and μ_3 respectively. The deer migrated from patch A_1 to patch A_3 via patch A_2. Resource data were daily values from Julian day 1 to day 250 in 2016: (b) NDVI values at patch centres; (c) integrated NDVI at patch centres; (d) IRG values at patch centres.

movement model. In this review, I will focus on the modelling framework for considering resource selection across the whole space since this framework is the major achievement of the thesis. Here, I give a short review of a commonly used continuous-time movement model, the Ornstein-Uhlenbeck (OU) process, to explain why this is an ideal choice for my modelling framework. In addition, I briefly introduce some previous work on continuous-time movement models and how they relate to RSA to show gaps my work intends to close.

An OU process describes a biased random walk towards a centre of attraction and has been used for studying animal movement for a while. For instance, Dunn and Gipson (1977) studied the estimation of home range by using an OU process to represent movements attracted to a central place. Building on the idea

of Dunn and Gipson (1977), where individual movement is described by only a single OU process, Blackwell (1997) generalised the OU model by using different OU process for different situations. These different situations may stand for different locations of interest or different behavioural states such as resting and feeding. With a mixture of movement processes and assuming animals switch between these processes while moving, Blackwell (1997) has used such a model, termed a switching OU model, to describe more general space use patterns such as (i) movement with two centres of attraction, (ii) movement with large-scale excursions and (iii) movement with different behavioural states. These examples provide an initial step to model movement decisions made according to the selection of resources, which can be nearby or far away.

Parameterisation of the models is achieved by Bayesian inference strategies (Blackwell, 2003) and subsequent studies have extended the OU modelling framework to include heterogeneous spatial conditions (Harris and Blackwell, 2013; Blackwell et al., 2016). Since the OU models are ready to incorporate both heterogeneous landscapes and different behavioural states, it has great potential to strengthen the analysis of resource selection by inserting a RSF into an OU model. This was a direction for further research suggested by Schick et al. (2008) that the incorporation of a RSF in a movement model would provide a better understanding of the link between the environment, behavioural states and movement decisions. In particular, when using a switching OU model, a RSF can be used to identify the place where the animal intends to approach and thus can determine which OU process to follow at a specific moment. The incorporation of RSA into the OU framework in this way is the primary purpose of the second part of this thesis.

The use of an OU process in RSA has been proposed by Johnson et al. (2008). In Johnson et al. (2008), an OU process describes the probability of moving from one fix to the next to represent the density of available resources. However, they did not include possible decision changes between observations.

Apart from the switching OU model, there are several studies devoted to incorporating resource selection into a continuous-time movement model. An example is Horne et al. (2007), which introduced the Brownian bridge movement model (BBMM), which links pairs of two successive observations to estimate movement

paths. Estimation of movement paths can bring about the identification of places which are frequently used and then these areas can be related to environmental conditions to understand the selection of resources (Horne et al., 2007). However, as a Brownian bridge connecting two observations assumes a higher frequency of use in the extending area over the straight line between the two locations, that is, a step in SSA, it may also lead to misinterpretation of resource selection as SSA may do. For example, for cases other than crossing a road as in Horne et al. (2007), this approach may miss the situation where the animal aims at an area in the distance and have to pass a certain place which is in fact not of great interest. In addition, Horne et al. (2007) did not consider changes of movement decision along the path.

Hanks et al. (2015) proposed a model where the movement path is discretised and a resource selection function is embedded in the transition rate between neighbouring cells. However, it only focuses on transitions within a local area and implicitly assumes that the decision-making process is coincident with the scale of observation. It does not consider the situation where movement might be motivated by a long-term goal such as the migration between seasonal ranges.

Therefore, there is a lack of a model for simultaneously representing movement depending on resource selection at large scales in continuous time and considering decisions made between observations. On the other hand, the combination of a switching OU model and a RSF in this thesis provides a useful tool in this regard.

§1.3 Thesis outline

In Chapter 2, I investigate the accuracy of three partial differential equation (PDE) methods for estimating location redistributions by comparing approximations with actual distributions arising from movement kernels. I demonstrate the performance of the PDE methods on different assumptions for movement strategies and in different spatial conditions. Most of the contents in this chapter has been published in Wang and Potts (2017).

To understand how animals move in response to environmental conditions, Chapter 3 introduces a modelling framework for resource selection analysis using

continuous-time movement models. I use a resource selection function to identify the attraction centre of movement in space and a switching Ornstein-Uhlenbeck (OU) process to describe movements drifting to the most attractive place in the landscape.

Chapter 4 uses similar strategies as Chapter 3 but describes movements in a different situation, where only local information is considered. I describe movements using a random walk in the direction up the local resource gradient, which is determined by comparing the value of the resource selection function in the vicinity of the animal's position.

The movement modelling framework in Chapter 3 is applied to mule deer (*Odocoileus hemionus*) data in Chapter 5. I fit a migration model built on the framework in Chapter 3 to the migration data, using Bayesian inference techniques.

Some discussions and conclusions are given in Chapter 6.

Chapter 2

Partial differential equation techniques for analysing animal movement

Most contents of this chapter has been published in Wang and Potts (2017). Animal movement contributes to population distributions and the formation of space use patterns. By assuming all individuals move according to the same rule, the probability of an individual being at a certain position at a given time can represent the population density (Turchin, 1991). A movement rule can be represented by a movement kernel, which describes how an animal moves, for example, by explicitly giving the probability of the animal moving to a location in a fixed period of time given its current position. The strategy to convert individual movement kernels to patterns at population level dates back to Patlak (1953), which considers movement rules incorporating persistence in direction and bias caused by external factors. Patlak's work was later explained and popularised by Turchin (1991) and Turchin (1998). Recent examples include Rhodes et al. (2005), where animals' preference for a habitat is contained in a movement kernel to model habitat selection. It is also straightforward to incorporate other factors influencing movements such as interactions with the environment into a movement kernel (Potts, Mokross and Lewis, 2014).

To predict the population-level space use pattern from a movement kernel, partial differential equation (PDE) techniques provide an efficient tool to accomplish this task. This chapter gives a comparison between three such techniques to show under which conditions these techniques would provide more accurate approximations. The three techniques are the Hyperbolic Scaling method (Hillen and Painter, 2013), the Moment Closure method (Hillen and Painter, 2013) and Patlak's approach (Patlak, 1953).

§ 2.1   Movement kernel analysis

In this chapter, a movement kernel is denoted by $k_\tau(z|x)$, representing the probability of an animal arriving at position z in time τ given its current location x. A movement kernel can be related to a random walk by giving the location density after one time step. For example, the classical simple random walk model in 1-D assumes that in time τ, the walker can move a distance Δ either right or left with equal probability (Othmer et al., 1988; Codling et al., 2008). This simple random walk can be expressed by a movement kernel defined by

$$
k_\tau(z|x) = \begin{cases} \dfrac{1}{2} & \text{if } z = x \pm \Delta, \\[2mm] 0 & \text{otherwise.} \end{cases} \tag{2.1}
$$

By considering the position distribution after time $t = n\tau$ for a large $n \in N$ and taking the limit $\Delta, \tau \to 0$ such that $\delta^2/\tau = \text{constant}$, the simple ransom walk model brings about the density function, which is the solution to the diffusion equation

$$
\frac{\partial}{\partial t} u(x,t) = \frac{\Delta^2}{2\tau} \frac{\partial^2}{\partial x^2} u(x,t),
$$
$$
u(x,0) = \delta(x), \tag{2.2}
$$

where $u(x,t)$ is the population density at position x and time t and $\delta(x)$ is the Dirac distribution (Othmer et al., 1988; Codling et al., 2008). The diffusion equation (Equation 2.2) associated to the simple random walk can be derived by analysing the movement kernel in Equation 2.1 using the methods introduced below in Sections 2.1.1-2.1.3.

In this section, I will briefly introduce three methods, used in Potts et al. (2016) to transform a movement kernel into long-term space use distribution. These methods rely on different assumptions. The first technique, the Hyperbolic Scaling method, assumes that the major component of movement is the drift rather than diffusion term (Othmer et al., 1988; Hillen and Painter, 2013). The second method, the Moment Closure method, assumes that it is sufficient to describe movements using only the first and second moment while higher moments are at equilibrium and sufficiently small to be neglected (Hillen and Painter, 2013). In general, it is also possible to include the third, fourth or higher moments in a moment closure method, but this is beyond the scope of this chapter and will not be discussed here. The last method is Patlak's approach, which is based on similar assumptions about higher moments, but also assumes a slow change of movement kernel across space (Patlak, 1953).

2.1.1   Hyperbolic Scaling method

Here, I assume an animal is moving in a one-dimensional space according to a movement kernel, $k_\tau(z|x)$, and denote the probability density function of an animal's location distribution at time t by $u_H(x, t)$ with the subscript "H" standing for "Hyperbolic Scaling". Then the PDE arising from the Hyperbolic Scaling method is (Hillen and Painter, 2013; Potts et al., 2016)

$$\frac{\partial}{\partial t}u_H(x,t) = \frac{\tau}{2}\frac{\partial^2}{\partial x^2}[D(x)u_H(x,t)] - \frac{\partial}{\partial x}[c(x)u_H(x,t)] + \frac{\tau}{2}\frac{\partial}{\partial x}\left[c(x)\frac{\mathrm{d}c(x)}{\mathrm{d}x}u_H(x,t)\right],$$

(2.3)

where

$$c(x) = \frac{1}{\tau}\int_{-\infty}^{\infty}(z-x)k_\tau(z|x)\mathrm{d}z,$$

(2.4)

and

$$D(x) = \frac{1}{\tau^2}\int_{-\infty}^{\infty}(z-x)^2 k_\tau(z|x)\mathrm{d}z - c(x)^2.$$

(2.5)

Here, $c(x)$ is the advection coefficients, representing the mean distance moved over time τ, that is, the mean drift velocity of the animal, and $D(x)$ is the diffusion coefficient, describing the variance of this velocity.

The long-term population distribution can be represented by the solution to the PDE in Equation (2.3) at steady state (Smouse et al., 2010). Setting the left-hand side of Equation (2.3) to 0, Equation (2.3) becomes an ordinary differential equation (ODE) as follows:

$$\frac{\tau}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}[D(x)u_H^*(x)] - \frac{\mathrm{d}}{\mathrm{d}x}[c(x)u_H^*(x)] + \frac{\tau}{2}\frac{\mathrm{d}}{\mathrm{d}x}\left[c(x)\frac{\mathrm{d}c(x)}{\mathrm{d}x}u_H^*(x)\right] = 0 \qquad (2.6)$$

where $u_H^*(x)$ is the steady-state distribution. Imposing a zero-flux assumption at the steady state that

$$\frac{\tau}{2}\frac{\mathrm{d}}{\mathrm{d}x}[D(x)u_H^*(x)] - [c(x)u_H^*(x)] + \frac{\tau}{2}\left[c(x)\frac{\mathrm{d}c(x)}{\mathrm{d}x}u_H^*(x)\right] = 0, \qquad (2.7)$$

Equation (2.6) can be solved to give

$$u_H^*(x) = \frac{C_H}{D(x)}\exp\left(\frac{1}{\tau}\int_0^x \frac{2c(s) - \tau\frac{\mathrm{d}c}{\mathrm{d}s}c(s)}{D(s)}\mathrm{d}s\right), \qquad (2.8)$$

where

$$C_H = \left[\int_\Omega \frac{1}{D(x)}\exp\left(\frac{1}{\tau}\int_0^x \frac{2c(s) - \tau\frac{\mathrm{d}c}{\mathrm{d}s}c(s)}{D(s)}\mathrm{d}s\right)\mathrm{d}x\right]^{-1}$$

is a normalising constant and Ω is the domain where $u_H^*(x)$ is defined. Note that this expression relies on the assumption that Ω is connected in R and contains 0. This normalising constant is essential, since $u_H^*(x)$ is a probability density function, so must integrate to 1 across Ω.

2.1.2 MOMENT CLOSURE METHOD

Similar to the Hyperbolic Scaling method, I begin with a PDE in 1D when applying the Moment Closure method to estimate the steady-state distribution from a movement kernel. The PDE derived using the Moment Closure method is (Hillen and Painter, 2013; Potts et al., 2016)

$$\frac{\partial}{\partial t}u_M(x,t) = \frac{\tau}{2}\frac{\partial^2}{\partial x^2}[D(x)u_M(x,t)] - \frac{\partial}{\partial x}[c(x)u_M(x,t)], \qquad (2.9)$$

where $u_M(x, t)$ is the probability of location distribution with the subscript "M" referring to "Moment Closure". Here, $c(x)$ and $D(x)$ are defined by Equations (2.4) and (2.5) in the same way as using the Hyperbolic Scaling method. To solve the PDE in Equation (2.9) at steady state, I denote the steady-state distribution by $u_M^*(x)$ and assume the zero-flux condition

$$\frac{\tau}{2}\frac{\mathrm{d}}{\mathrm{d}x}[D(x)u_M^*(x)] - [c(x)u_M^*(x)] = 0. \tag{2.10}$$

Based on this assumption, the steady-state distribution $u_M^*(x)$ is obtained by solving the following ODE,

$$\frac{\tau}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}[D(x)u_M^*(x)] - \frac{\mathrm{d}}{\mathrm{d}x}[c(x)u_M^*(x)] = 0. \tag{2.11}$$

The solution to the ODE in Equation (2.11) is

$$u_M^*(x) = \frac{C_M}{D(x)}\exp\left(\frac{2}{\tau}\int_0^x \frac{c(s)}{D(s)}\mathrm{d}s\right), \tag{2.12}$$

where

$$C_M = \left[\int_\Omega \frac{1}{D(x)}\exp\left(\frac{2}{\tau}\int_0^x \frac{c(s)}{D(s)}\mathrm{d}s\right)\mathrm{d}x\right]^{-1}$$

is a normalising constant and Ω is the domain of definition of the distribution, $u_M^*(x)$.

2.1.3 Patlaks approach

The PDE considered when using Patlak's approach in a 1D space is given by (Patlak, 1953; Potts et al., 2016)

$$\frac{\partial}{\partial t}u_P(x, t) = \frac{\partial^2}{\partial x^2}\left[\frac{M_2(x)}{2\tau}u_P(x, t)\right] - \frac{\partial}{\partial x}\left[\frac{M_1(x)}{\tau}u_P(x, t)\right] \tag{2.13}$$

with

$$M_1(x) = \int_{-\infty}^{\infty}(z - x)k_\tau(z|x)\mathrm{d}z, \tag{2.14}$$

and

$$M_2(x) = \int_{-\infty}^{\infty} (z - x)^2 k_\tau(z|x) \mathrm{d}z, \tag{2.15}$$

where the subscript "P" refers to Patlak's approach, $M_1(x)$ and $M_2(x)$ are the first and second moments of the distance moved in time τ respectively. Unlike the Hyperbolic Scaling and Moment Closure methods, where the diffusion coefficient, defined by Equation (2.5), is in proportion to the variance of the distance moved in time τ, the diffusion coefficient, $M_2(x)$, in Equation (2.15) here is proportional to the second moment of the mean displacement in time τ.

To solve the PDE in Equation (2.13) at steady state, the term $\frac{\partial}{\partial t} u_P(x, t)$ is set to 0 to obtain the following ODE

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} \left[\frac{M_2(x)}{2\tau} u_P^*(x) \right] - \frac{\mathrm{d}}{\mathrm{d}x} \left[\frac{M_1(x)}{\tau} u_P^*(x) \right] = 0, \tag{2.16}$$

where $u_P^*(x)$ is the steady-state distribution. Assuming the ODE in Equation (2.16) satisfies the condition

$$\frac{\mathrm{d}}{\mathrm{d}x} \left[\frac{M_2(x)}{2\tau} u_P^*(x) \right] - \left[\frac{M_1(x)}{\tau} u_P^*(x) \right] = 0 \tag{2.17}$$

and solving Equation (2.16) give the solution to the original PDE in Equation (2.13) at steady state as follows:

$$u_P^*(x) = \frac{C_P}{M_2(x)} \exp \left(\int_0^x \frac{2M_1(s)}{M_2(s)} \mathrm{d}s \right) \tag{2.18}$$

with

$$C_P = \left[\int_\Omega \frac{1}{M_2(x)} \exp \left(\int_0^x \frac{2M_1(s)}{M_2(s)} \mathrm{d}s \right) \mathrm{d}x \right]^{-1}$$

a normalising constant and Ω the domain of definition.

§ 2.2   Comparison between the PDE techniques

Section 2.1 introduces the formulae for the PDEs and their steady-state solutions for scaling up individual movement processes to population distributions. This

section gives two approaches for assessing the capacity of the PDE methods of giving accurate approximation to real distributions arising from the underlying movement kernel. One approach is to inspect short-term distributions arising from the movement kernel, while the other considers long-term distributions.

2.2.1 EVALUATION OF SHORT-TERM ESTIMATION

For some particular movement kernels, it is possible to solve the PDEs in Equations (2.3),(2.9),(2.13) at a transient state. In this case, the transient distributions derived from the PDEs can be used to approximate the location distribution and be compared to the movement kernel directly. As a movement kernel describes the underlying process by which an animal moves, a good approximation should feature little discrepancy when comparing to the kernel.

Here I suppose a movement kernel, $k_\tau(z|x)$, is considered and the PDE used in the approximation procedure (Equations 2.3,2.9,2.13) can be solved to give $u(x,t)$, a distribution approximating the animal's location distribution at time t. I also assume no long-term correlation in the model, that is, the probability of moving to a location in time τ solely depends on the present condition. The notation $k_\tau(z|x)$ indicates the probability of an animal moving from its current position, x, to position z in a small period of time τ. Meanwhile, $u(x,t)$ represents the probability of the animal being at some location, x, at time t. Therefore, assuming the animal is currently at position x_0, meaning the initial condition is given by $u(x,0) = \delta(x_0)$, where $\delta(\cdot)$ is the Dirac delta function, the probability of observing the animal at position x in time τ, $u(x,\tau)$, should equal $k_\tau(x|x_0)$. That is, a precise approximation of $k_\tau(x|x_0)$ given by $u(x,\tau)$ should have

$$u(x,\tau) = k_\tau(x|x_0). \tag{2.19}$$

On the other hand, if an approximation cannot accurately capture the movement kernel, which controls the short-term movements, then the approximation method is more likely to provide misleading long-term estimations. That is, the cumulation of differences in each short step may result in large errors over a long period of time.

2.2.2 Evaluation of long-term estimation

Evaluation of the long-term approximation is less straightforward. The first step is to construct the "actual" space use distribution arising from a given movement kernel, $k_\tau(z|x)$. A technique that can be used to achieve this is the Master Equation. In general, the Master Equation describes the time evolution of the population density $u(x,t)$, given that $T(x,y)$ represents the probability of a jump from location y to x and a transition rate λ, as follows:

$$\frac{\partial}{\partial t}u(x,t) = -\lambda u + \lambda \int_\Omega T(x,y)u(y,t)dy, \tag{2.20}$$

where the first term on the right hand side of Equation 2.20 describe the rate of leaving position x and the second term gives the rate of arriving at x from all other locations in the domain Ω (Othmer et al., 1988). Equation 2.20 also applies to models of velocity jump processes, where transition between velocities is considered (Othmer et al., 1988).

Here, I will employ an alternative form for the Master Equation, which is also commonly used (Moorcroft and Barnett, 2008; Potts, Bastille-Rousseau, Murray, Schaefer and Lewis, 2014; Merkle et al., 2018):

$$u_I(x,t+\tau) = \int_{-\infty}^{\infty} k_\tau(x|y)u_I(y,t)dy, \tag{2.21}$$

where $u_I(x,t)$ is the probability density of the animal's position at time t and the subscript "I" stands for "Integral". As time t increases to infinity, Equation (2.21) becomes

$$u_I^*(x) = \int_{-\infty}^{\infty} k_\tau(x|y)u_I^*(y)dy, \tag{2.22}$$

where $u_I^*(x) = \lim_{t\to\infty} u_I(x,t)$. This distribution, $u_I^*(x)$, is the long-term population distribution propagated from the movement kernel. Nonetheless, it is often impossible to solve Equation (2.22) analytically, so it is necessary to iterate Equation (2.21) numerically to obtain the long-term distribution. (For an exception, see Barnett and Moorcroft (2008).) Therefore, in this chapter, I will numerically integrate Equation (2.21) to derive the desired real long-term distribution. To decide when to stop the iteration, I measure the difference be-

tween distributions by Kullback-Leibler divergence (KL-divergence; Kullback and
Leibler (1951)), which is commonly used to evaluate information loss when approximating a model (Horne and Garton, 2006). Applying the KL-divergence,
however, requires caution for situations where the probability density is 0. Once
the KL-divergence from $u_I(x, t + n\tau)$ to $u_I(x, t + (n-1)\tau)$ is less than 10^{-6}, I
set $u_I^*(x) = u_I(x, t + n\tau)$.

Having derived the real long-term distribution, $u_I^*(x)$, the next step is to evaluate the distance between $u_I^*(x)$ and the estimated distributions given by the
three PDE methods introduced in Section 2.1 (Equations 2.8, 2.12, 2.18). The
approximate distribution with the smallest KL-divergence to $u_I^*(x)$ is regarded
as the best approximation of the long-term distribution emerging from the movement kernel (Horne and Garton, 2006). For an alternative measurement, I use
Euclidean distance, which leads to very similar conclusions of the performance of
the PDE methods (see Appendix A). This suggests that the comparison between
methods is not sensitive to the metric used.

§2.3   An analytic example: a biased random walk

After introducing three PDE techniques for estimating population distribution
and approaches to comparing them (Sections 2.1 and 2.2), here I start with a
simple example to illustrate the comparison of the three PDE methods. This
example is a movement kernel for which the PDEs in Equations (2.3),(2.9),(2.13)
have analytical solutions at transient states, so the solutions can be compared
to the movement kernel as explained in Section 2.2.1. The results reveal that
Patlak's approach fails to provide a distribution corresponding to the movement
kernel even for this simple case, whereas the Hyperbolic Scaling and Moment
Closure methods are proven to be successful.

Assuming an animal is moving on an infinite 1D line, I consider a movement
kernel that describes a biased random walk using the normal distribution with
mean μ and variance σ^2, given by

$$k_\tau(z|x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(z-x-\mu)^2}{2\sigma^2}\right). \tag{2.23}$$

When using the Hyperbolic Scaling and Moment Closure methods, I calculate the mean velocity in Equation (2.4) and variance in Equation (2.5) with the movement kernel in Equation (2.23) to give

$$c(x) = \frac{\mu}{\tau}, \tag{2.24}$$

$$D(x) = \frac{\sigma^2}{\tau^2}. \tag{2.25}$$

Since the mean velocity function, $c(x)$, in Equation (2.24) is a constant, the derivative of $c(x)$ is 0 and the last term of the PDE in Equation (2.3) vanishes. Hence the Hyperbolic Scaling and Moment Closure methods give rise to the same PDE as follows:

$$\frac{\partial}{\partial t} u_M(x,t) = \frac{\sigma^2}{2\tau} \frac{\partial^2}{\partial x^2} u_M(x,t) - \frac{\mu}{\tau} \frac{\partial}{\partial x} u_M(x,t), \tag{2.26}$$

which is an advection-diffusion equation with constant coefficients. Given the initial condition $u(x,0) = \delta(x_0)$, where $\delta(\cdot)$ is the Dirac delta function, the solution to this PDE at time $t = \tau$ is (Codling et al., 2008)

$$u_M(x,\tau) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-x_0-\mu)^2}{2\sigma^2}\right). \tag{2.27}$$

Note that in this case, $u_H(x,\tau) = u_M(x,\tau)$.

Meanwhile, the first and second moments of the distance moved in time τ in Equations (2.14) and (2.15) are calculated for Patlak's approach:

$$M_1(x) = \mu, \tag{2.28}$$

$$M_2(x) = \sigma^2 + \mu^2. \tag{2.29}$$

Then inserting Equations (2.28) and (2.29) into Equation (2.13) leads to

$$\frac{\partial}{\partial t} u_P(x,t) = \frac{\partial^2}{\partial x^2} \left[\frac{\sigma^2 + \mu^2}{2\tau} u_P(x,t) \right] - \frac{\partial}{\partial x} \left[\frac{\mu}{\tau} u_P(x,t) \right]. \tag{2.30}$$

With the initial condition $u(x,0) = \delta(x_0)$, solving this PDE and calculating the probability density at time τ result in

$$u_P(x,\tau) = \frac{1}{\sqrt{2\pi(\sigma^2 + \mu^2)}} \exp\left(\frac{-(x - x_0 - \mu)^2}{2(\sigma^2 + \mu^2)} \right). \tag{2.31}$$

Comparing Equations (2.27) and (2.31) to the movement kernel $k_\tau(x|x_0)$ shows that $u_H(x,\tau) = u_M(x,\tau) = k_\tau(x|x_0)$ but $u_P(x,\tau) \neq k_\tau(x|x_0)$. Thus the Hyperbolic Scaling and Moment Closure methods successfully represent the movement kernel, whereas Patlak's approach fails even in this simple example.

The contrast between Patlak's approach and the other two methods stems from the different diffusion terms in Equations (2.26) and (2.30). Patlak's approach brings about a diffusion coefficient proportional to the second moment of velocity, while the diffusion coefficient derived by the other methods is proportional to the variance. Consequently, Patlak's approach would overestimate the variance in a transient probability distribution and the error becomes greater when the drift term rises (see Figure 2.1).

Note that for cases such as the example given in Equation 2.23 that the relocation is determined by a normal distribution, all PDE methods lead to solutions (Equations 2.27 and 2.31) allowing for infinite propagation, which is not realistic in a biological context.

In general, the Hyperbolic Scaling and Moment Closure methods would provide better predictions of population distribution than Patlak's approach does. This is because in an advection-diffusion equation with constant coefficients, the diffusion term represents the variance over time rather than the second moment. As revealed in this example, Patlak's formulation could be highly inaccurate except when the drift term is very small compared to the diffusion term.
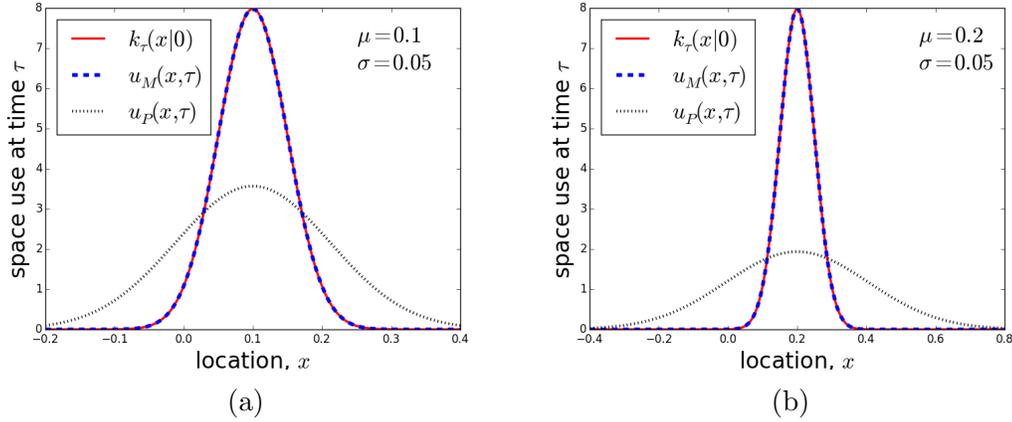
Figure 2.1: Errors arising from Patlak's approximation are corrected by the (more recent) Moment Closure approach. Here, we show the movement kernel from Equation (2.23) with values of mean, μ, and standard deviation, σ, as given in the panels, together with solutions of the PDEs for Patlak's approximation ($u_P(x,\tau)$; Equation 2.31) and the Moment Closure method ($u_M(x,\tau)$; Equation 2.27), given at time τ. Progressing from the left panel to the right, we see that a higher μ leads to a greater difference between the two methods, but the Moment Closure method always gives the correct result.

§2.4   A central-place foraging model with discontinuous mean velocity

Section 2.3 shows that Patlak's approach can give an inaccurate illustration of transient dynamics; hereafter, I investigate the performance of the three PDE techniques for estimating long-term population distributions. The steady-state solutions to the PDEs derived in the following examples do not require numerical integration, so they are helpful for the studies of steady state. The examples examined here are three central-place foraging models, differing by the smoothness of their mean velocity functions, and four models describing movements in a heterogeneous landscape.

The first central-place foraging model originates from the classical Holgate-Okubo model (Holgate, 1971; Okubo, 1980), which is the simplest model for home-ranging behaviour (Börger et al., 2008). Here, an animal is assumed to be moving on the real line with constant average velocity towards a central place, positioned at the origin $x = 0$ for convenience. This central place, for example, may be a place abundant in food or a den with young animals. The movement procedure
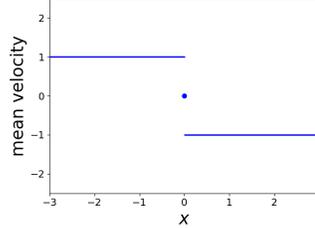
Figure 2.2: An example of a discontinuous mean velocity function, defined by Equation (2.33), derived from a central-place foraging model with the central place being located at 0 (Equation 2.32).

is defined by a movement kernel built on normal distributions as given below:

$$
k_\tau^1(z|x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z-x-\mu)^2}{2\sigma^2}\right) & \text{if } x < 0, \\[2ex] \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z-x+\mu)^2}{2\sigma^2}\right) & \text{if } x > 0, \\[2ex] \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z-x)^2}{2\sigma^2}\right) & \text{if } x = 0, \end{cases} \tag{2.32}
$$

where the superscript "1" is used to refer to this model, $\mu > 0$ and σ^2 are the mean and variance of the distance the animal moves over a time τ.

2.4.1 Analysis of movement kernel $k_\tau^1(z|x)$ by the PDE techniques

To apply the Hyperbolic Scaling and Moment Closure methods, the drift and diffusion terms in Equations (2.4) and (2.5) are calculated by inserting the movement kernel in Equation (2.32) to give

$$
c_1(x) = \begin{cases} \dfrac{\mu}{\tau} & \text{if } x < 0, \\[2ex] -\dfrac{\mu}{\tau} & \text{if } x > 0, \\[2ex] 0 & \text{if } x = 0, \end{cases} \tag{2.33}
$$

and

$$
D_1(x) = \frac{\sigma^2}{\tau^2}. \tag{2.34}
$$

Since $c_1(x)$ is piecewise constant (Figure 2.2), the derivative of $c_1(x)$ is 0 for

$x \neq 0$. By comparing Equations (2.8) and (2.12), it is obvious that in this situation the steady state distribution obtained using the Hyperbolic Scaling method, denoted by $u_H^1(x)$, reduces to that using the Moment Closure method, denoted by $u_M^1(x)$. That is, $u_H^1(x) = u_M^1(x)$, so only the latter is calculated. Note that because the mean velocity function, $c_1(x)$, is piecewise defined, the PDEs involved and the ODEs at steady state are thus piecewise defined. Therefore, I solve the ODE in Equation (2.11) for $x > 0$ and $x < 0$ separately and assume that the solution is continuous. Now placing expressions (2.33) and (2.34) into Equation (2.12) gives

$$
u_M^1(x) = \begin{cases} C_{M1}^1 \dfrac{\tau^2}{\sigma^2} \exp\left(\dfrac{2\mu}{\sigma^2}x\right) & \text{if } x < 0, \\[4mm] C_{M2}^1 \dfrac{\tau^2}{\sigma^2} \exp\left(-\dfrac{2\mu}{\sigma^2}x\right) & \text{if } x > 0, \end{cases} \tag{2.35}
$$

where C_{M1}^1 and C_{M2}^1 are arbitrary constants. The continuity assumption implies that $C_{M1}^1 = C_{M2}^1$ must hold, because $\lim_{x \to 0^+} u_M^1(x) = \lim_{x \to 0^-} u_M^1(x)$. To ensure $u_M^1(x)$ integrates to 1, C_{M1}^1 is given by

$$
C_{M1}^1 = \left[\int_{-\infty}^0 \frac{\tau^2}{\sigma^2} \exp\left(\frac{2\mu}{\sigma^2}x\right) dx + \int_0^\infty \frac{\tau^2}{\sigma^2} \exp\left(-\frac{2\mu}{\sigma^2}x\right) dx \right]^{-1} = \frac{\mu}{\tau^2}. \tag{2.36}
$$

Placing Equation (2.36) into Equation (2.35) and setting $u_M^1(0) = \lim_{x \to 0} u_M^1(x) = \mu/\sigma^2$ result in

$$
u_M^1(x) = \begin{cases} \dfrac{\mu}{\sigma^2} \exp\left(\dfrac{2\mu}{\sigma^2}x\right) & \text{if } x < 0, \\[4mm] \dfrac{\mu}{\sigma^2} \exp\left(-\dfrac{2\mu}{\sigma^2}x\right) & \text{if } x \geq 0. \end{cases} \tag{2.37}
$$

This is the probability density function of a Laplace distribution with mean 0 and variance $\sigma^4/2\mu^2$.

When using Patlak's approach to derive the steady state distribution, the first step involves the computation of the first and second moments of distance moved over time τ. Inserting the movement kernel in Equation (2.32) into Equations (2.14) and (2.15) gives

$$M_1^1(x) = \begin{cases} \mu & \text{if } x < 0, \\ -\mu & \text{if } x > 0, \\ 0 & \text{if } x = 0, \end{cases} \qquad (2.38)$$

and

$$M_2^1(x) = \sigma^2 + \mu^2. \qquad (2.39)$$

Equation (2.38) is also piecewise defined, so the steady state distribution, $u_P^1(x)$, is calculated for $x > 0$ and $x < 0$ with the continuity assumption $u_P^1(0) = \lim_{x \to 0^+} u_P^1(x) = \lim_{x \to 0^-} u_P^1(x)$. Substituting Equations (2.38) and (2.39) into Equation (2.18) with the condition that $u_P^1(x)$ integrates to 1 yields

$$u_P^1(x) = \begin{cases} \dfrac{\mu}{\sigma^2 + \mu^2} \exp\left(\dfrac{2\mu}{\sigma^2 + \mu^2} x\right) & \text{if } x < 0, \\[3mm] \dfrac{\mu}{\sigma^2 + \mu^2} \exp\left(-\dfrac{2\mu}{\sigma^2 + \mu^2} x\right) & \text{if } x \geq 0, \end{cases} \qquad (2.40)$$

which is a Laplace distribution with mean 0 and variance $(\sigma^2 + \mu^2)^2/2\mu^2$. The variances of Equations 2.40 is $(\sigma^2 + \mu^2)^2/\sigma^4$ times larger than the variance of Equation 2.37 when $\mu > 0$.

2.4.2   NUMERICAL ANALYSIS OF MOVEMENT KERNEL $k_\tau^1(z|x)$

Having obtained steady-state distributions by the three PDE techniques (Equations 2.37 and 2.40), I now compare these distributions to the actual long-term population distribution, given by Equation (2.22), to decide which PDE method performs better. As described in Section 2.2.2, the actual long-term population distribution, $u_I^1(x)$, is represented by the integral of the ME in Equation (2.21), numerically computed given the moment kernel in Equation (2.32).

The movement kernel in Equation (2.32) contains two parameters, μ and σ, hence I explore the parameter space to understand the impact of these parameters on the accuracy of approximations. The plots of contour lines of the KL-divergence on the $\mu - \sigma$ plane reveal that both the KL-divergence of $u_I^1(x)$ from $u_M^1(x)$ and the KL-divergence of $u_I^1(x)$ from $u_P^1(x)$ grow with increasing μ/σ (Figures

(a) Moment Closure

(b) Patlak's method

(c) $0.05 \leq \mu \leq 0.2$, $\sigma = 0.05$

(d) $\mu = 0.05$, $0.05 \leq \sigma \leq 0.2$

(e) $\mu = 0.01$, $\sigma = 0.05$

(f) $\mu = 0.1$, $\sigma = 0.05$
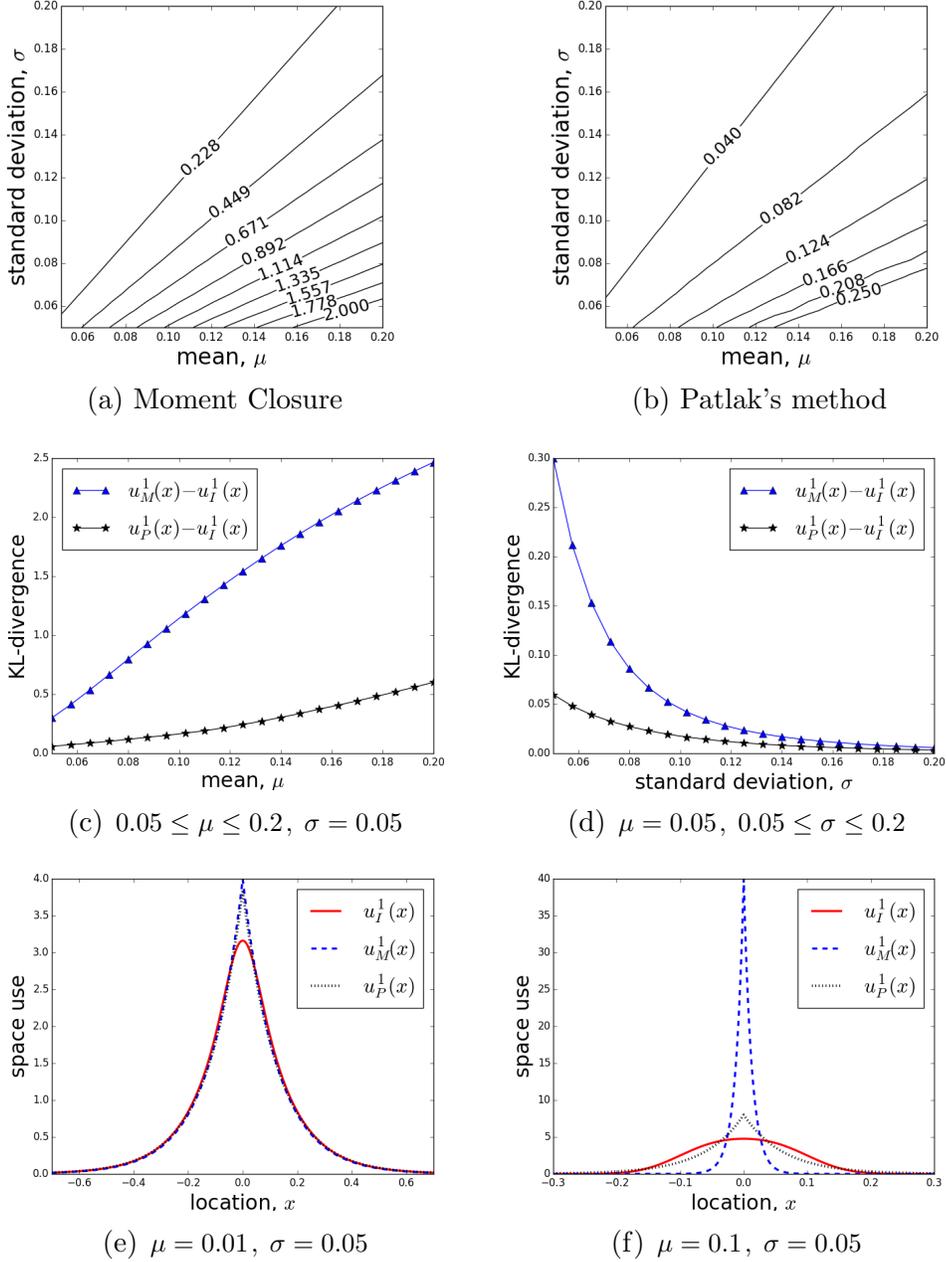
Figure 2.3: Discontinuous mean velocity movement kernel $k_\tau^1(z|x)$ with μ the mean move length in one step and σ the standard deviation of move length: The contours of the KL-divergence of the numerical solution, $u_I^1(x)$, (a) from the analytic solution, $u_M^1(x)$ (Equation 2.37), derived using a moment closure technique, $\mu, \sigma \in [0.05, 0.2]$. (b) from the analytic solution, $u_P^1(x)$ (Equation 2.40), derived using Patlak's method, $\mu, \sigma \in [0.05, 0.2]$. (c) KL-divergence between $u_M^1(x)$ and $u_I^1(x)$ (▲), and $u_P^1(x)$ and $u_I^1(x)$ (⋆) with $0.05 \leq \mu \leq 0.2$ and $\sigma = 0.05$. (d) KL-divergence between $u_M^1(x)$ and $u_I^1(x)$ (▲), and $u_P^1(x)$ and $u_I^1(x)$ (⋆) with $0.05 \leq \sigma \leq 0.2$ and $\mu = 0.05$. (e) steady-state distributions with $\mu = 0.01$ and $\sigma = 0.05$. (f) steady-state distributions with $\mu = 0.1$ and $\sigma = 0.05$.

2.3a,b). Note that steady-state distribution $u_H^1(x)$ equals to $u_M^1(x)$ in this case (see Section 2.4.1).

Figures 2.3c,d show that the KL-divergence of $u_I^1(x)$ from $u_M^1(x)$ is greater than the KL-divergence of $u_I^1(x)$ from $u_P^1(x)$, meaning $u_P^1(x)$ is a better approximation of $u_I^1(x)$. This implication is in opposition to the analytic analysis in Section 2.3, which induces the conjecture that $u_M^1(x)$ might be closer to $u_I^1(x)$ than $u_P^1(x)$. Nevertheless, neither of the two PDE methods captures the dynamics of the movement kernel properly. Both $u_M^1(x)$ and $u_P^1(x)$ have sharp peaks at $x = 0$, contrasting with the relatively smooth shape of $u_I^1(x)$, as shown in Figures 2.3e,f. In addition, since $\mu/(\sigma^2 + \mu^2) < \mu/\sigma^2$ for $\mu > 0$, comparing Equations (2.37) and (2.40) reveals that $u_P^1(0) < u_M^1(0)$ and the variance of $u_P^1(x)$ is always greater than that of $u_M^1(x)$ for $\mu > 0$. This contributes to the smaller KL-divergence from $u_I^1(x)$ to $u_P^1(x)$. That is, while all PDE techniques, including Patlak's approach, overestimate the probability density near the central place, Patlak's approach also overestimate the variance, resulting in a flatter distribution closer to the real long-term distribution. However, note that the apparently greater variance of $u_P^1(x)$ observed in Figure 2.3f agrees with the analytic observations of Section 2.3.

§ 2.5   A central-place foraging model with continuous mean velocity

The movement kernel in Equation (2.32) in Section 2.4 describes movement towards an attraction centre with a fixed mean velocity, which is discontinuous at the central place, $x = 0$, as the direction changes. Results of analysing this movement kernel by the PDE methods show that there is a sharp spike in the steady-state distribution at the point where the mean velocity function is discontinuous (Figures 2.3e,f). This sharp spike contrasts with the smooth shape of the actual long-term distribution derived from the ME in Equation (2.21). This means the distributions given by the PDE methods are very poor approximations to the actual long-term distribution. To understand how continuity of the mean velocity function affects the approximations, I consider another central-place foraging model, where the mean velocity function is continuous over the real line.

Setting the central place at $x = 0$, the movement kernel studied here is

$$k_\tau^2(z|x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z-x-\mu)^2}{2\sigma^2}\right) & \text{if } x < -\mu, \\[3mm] \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(\dfrac{-z^2}{2\sigma^2}\right) & \text{if } -\mu \leq x \leq \mu, \\[3mm] \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z-x+\mu)^2}{2\sigma^2}\right) & \text{if } x > \mu. \end{cases} \qquad (2.41)$$

Here, the mean displacement over time τ is constant when the distance between the animal and the central place is greater than μ (*i.e.* $|x - 0| > \mu$), but equal to $|x|$ when the animal is located in the interval $[-\mu, \mu]$, the vicinity of the central place.

2.5.1 Analysis of movement kernel $k_\tau^2(z|x)$ by the PDE techniques

When using the Hyperbolic Scaling and Moment Closure methods, the drift and diffusion coefficients are calculated by inserting the movement kernel in Equation (2.41) into Equations (2.4) and (2.5) to give

$$c_2(x) = \begin{cases} \dfrac{\mu}{\tau} & \text{if } x < -\mu, \\[3mm] -\dfrac{x}{\tau} & \text{if } -\mu \leq x \leq \mu, \\[3mm] -\dfrac{\mu}{\tau} & \text{if } x > \mu, \end{cases} \qquad (2.42)$$

and

$$D_2(x) = \frac{\sigma^2}{\tau^2}. \qquad (2.43)$$

Here, the mean velocity function, $c_2(x)$, is continuous and the speed decreases as the animal approaches the central place when the distance to it is less than μ (Figure 2.4).

Although $c_2(x)$ is continuous, the derivative of $c_2(x)$ can only be defined piecewise because $c_2(x)$ is not differentiable at points $x = \pm\mu$. Therefore, I derive the steady-state distribution by substituting Equations (2.42) and (2.43) into Equation (2.8) in three cases for $x < -\mu$, $-\mu < x < \mu$ and $x > \mu$. Then assuming
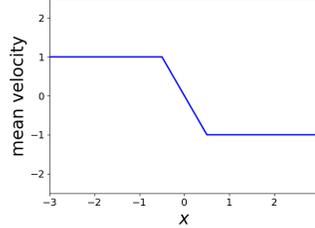
Figure 2.4: An example of a continuous mean velocity function, defined by Equation (2.42), derived from a central-place foraging model with the central place being located at 0 (Equation 2.41).

continuity of the solution gives

$$
u_H^2(x) =
\begin{cases}
C_H^2 \exp\left(\dfrac{2\mu}{\sigma^2}x + \dfrac{\mu^2}{2\sigma^2}\right) & \text{if } x < -\mu, \\[2mm]
C_H^2 \exp\left(-\dfrac{3}{2\sigma^2}x^2\right) & \text{if } -\mu \leq x \leq \mu, \\[2mm]
C_H^2 \exp\left(-\dfrac{2\mu}{\sigma^2}x + \dfrac{\mu^2}{2\sigma^2}\right) & \text{if } x > \mu,
\end{cases}
\tag{2.44}
$$

where

$$
C_H^2 = \left[\frac{\sigma^2}{\mu}\exp\left(-\frac{3\mu^2}{2\sigma^2}\right) + \frac{\sqrt{2\pi}\sigma}{\sqrt{3}}\mathrm{erf}\left(\frac{\sqrt{3}\mu}{\sqrt{2}\sigma}\right)\right]^{-1}
\tag{2.45}
$$

is a normalising constant ensuring that the distribution integrates to 1.

Applying the Moment Closure method to analyse the movement kernel in Equation (2.41) is carried out in the same way. Placing Equations (2.42) and (2.43) into Equation (2.12) for $x < -\mu$, $-\mu < x < \mu$ and $x > \mu$ and assuming continuity lead to the following steady-state distribution

$$
u_M^2(x) =
\begin{cases}
C_M^2 \exp\left(\dfrac{2\mu}{\sigma^2}x + \dfrac{\mu^2}{\sigma^2}\right) & \text{if } x < -\mu, \\[2mm]
C_M^2 \exp\left(-\dfrac{x^2}{\sigma^2}\right) & \text{if } -\mu \leq x \leq \mu, \\[2mm]
C_M^2 \exp\left(-\dfrac{2\mu}{\sigma^2}x + \dfrac{\mu^2}{\sigma^2}\right) & \text{if } x > \mu,
\end{cases}
\tag{2.46}
$$

where

$$C_M^2 = \left[\frac{\sigma^2}{\mu} \exp\left(-\frac{\mu^2}{\sigma^2}\right) + \sqrt{\pi}\sigma \operatorname{erf}\left(\frac{\mu}{\sigma}\right) \right]^{-1} \tag{2.47}$$

is a normalising constant. Here, since $c_2(x)$ (Equation 2.42) is non-constant for $-\mu \leq x \leq \mu$ and hence the derivative of $c_2(x)$ is non-zero on this interval, the integral in the steady-state solutions in Equations 2.8 and 2.12 are different on this interval. Consequently, the expressions for the solutions for $-\mu \leq x \leq \mu$ in Equations 2.44 and 2.46 are not the same and lead to different normalising terms (Equations 2.45 and 2.47).

To use Patlak's approach, I calculate the first and second moments of displacement by placing the movement kernel in Equation (2.41) into Equations (2.14) and (2.15) to obtain

$$M_1^2(x) = \begin{cases} \mu & \text{if } x < -\mu, \\ -x & \text{if } -\mu \leq x \leq \mu, \\ -\mu & \text{if } x > \mu, \end{cases} \tag{2.48}$$

and

$$M_2^2(x) = \begin{cases} \sigma^2 + \mu^2 & \text{if } x < -\mu \text{ or } x > \mu, \\ \sigma^2 + x^2 & \text{if } -\mu \leq x \leq \mu. \end{cases} \tag{2.49}$$

Substitute Equations (2.48) and (2.49) into Equation (2.18) and assuming the distribution is continuous give the steady-state distribution derived by Patlak's approach as follows:

$$u_P^2(x) = \begin{cases} \dfrac{C_P^2}{(\sigma^2+\mu^2)^2} \exp\left(\dfrac{2\mu}{\sigma^2+\mu^2}x + \dfrac{2\mu^2}{\sigma^2+\mu^2}\right) & \text{if } x < -\mu, \\[3mm] \dfrac{C_P^2}{(\sigma^2+x^2)^2} & \text{if } -\mu \leq x \leq \mu, \\[3mm] \dfrac{C_P^2}{(\sigma^2+\mu^2)^2} \exp\left(\dfrac{-2\mu}{\sigma^2+\mu^2}x + \dfrac{2\mu^2}{\sigma^2+\mu^2}\right) & \text{if } x > \mu, \end{cases} \tag{2.50}$$

where

$$C_P^2 = \left[\frac{1}{\mu(\sigma^2+\mu^2)} + \frac{\arctan(\mu/\sigma)}{\sigma^3} + \frac{\mu}{\sigma^2(\sigma^2+\mu^2)} \right]^{-1} \tag{2.51}$$

is a normalising constant.

2.5.2   NUMERICAL ANALYSIS OF MOVEMENT KERNEL $k_\tau^2(z|x)$

The plots of contour lines of KL-divergence show similar patterns to those observed in the analysis of the first central-place foraging model (Figures 2.5a-c). However, unlike the case with movement kernel $k_\tau^1(z|x)$, the $\mu - \sigma$ plane here consists of two regions, one where the KL-divergence of $u_I^2(x)$ from $u_P^2(x)$ is smaller and another where $u_M^2(x)$ is closer to $u_I^2(x)$ (Figure 2.5d). The steady-state distribution arising from the Moment Closure method, $u_M^2(x)$, outperforms others when μ/σ is larger.

In the region where the KL-divergence of $u_I^2(x)$ from $u_P^2(x)$ is the smallest, the KL-divergence of $u_I^2(x)$ from $u_M^2(x)$ is in fact only slightly higher (Figures 2.6a,b). That is, the approximations using Patlak's approach and the Moment Closure method reach similar accuracy (e.g. Figure 2.6c). Although the Moment Closure method appears to give relatively accurate approximations for larger μ, $u_M^2(x)$ still tends to overestimate the probability near the central place considerably as the other two methods do (Figure 2.6d). Note that $u_P^2(x)$ has a higher variance, corresponding to the implication made by the analytic example in Section 2.3. Overall, the steady-state distributions are more smooth near the central place, compared to those derived when analysing the movement kernel in Equation (2.32) but still overestimate the probabilities around the central place remarkably.

§ 2.6   A central-place foraging model with differentiable mean velocity

The results in Section 2.5 show that the PDE methods provide slightly better approximations to the actual long-term distribution than in Section 2.4 since the movement kernel in Equation (2.41) has a continuous mean velocity function. However, the overestimation of probabilities near the central place remains notably large. Therefore, to investigate whether a more smooth kernel can improve the approximations further, I consider an example of central-place foraging model where the mean velocity function is continuously differentiable. Here, the example of central-place foraging in a 1D space is described by a movement kernel where the average distance moved over time τ decreases as the animal approaches the
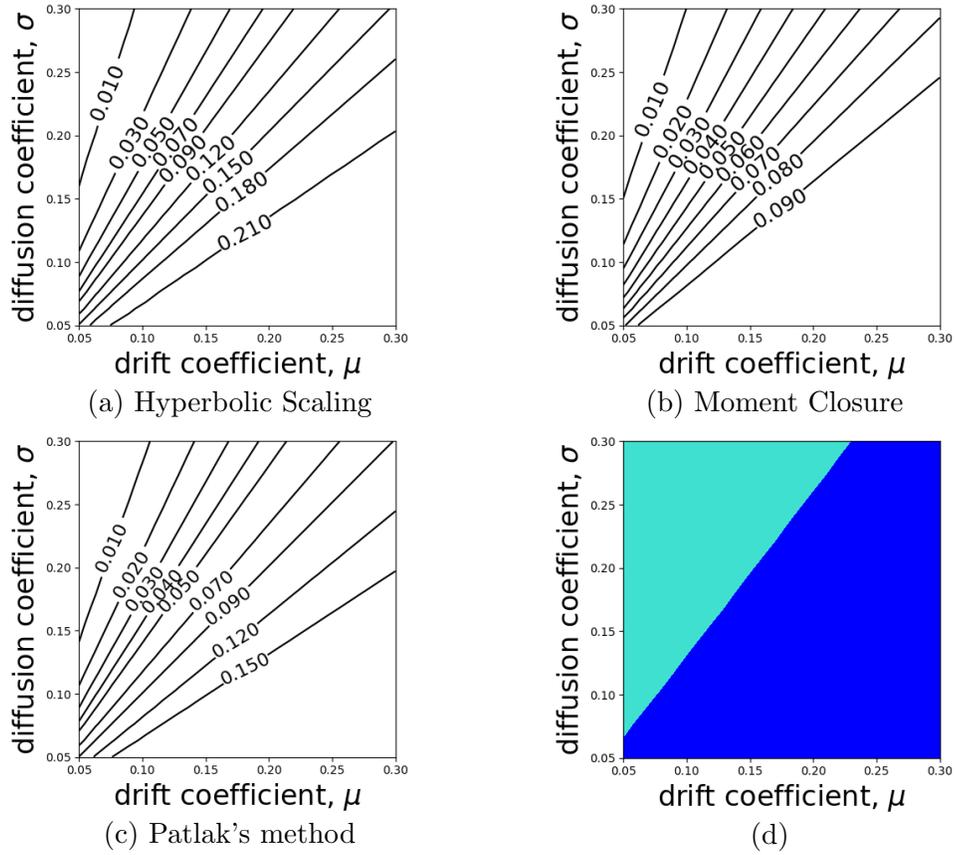
Figure 2.5: Continuous mean velocity movement kernel $k_\tau^2(z|x)$ with μ (resp. $|x|$) the mean move length in one step for $|x| > \mu$ (resp. $|x| \leq \mu$) and σ the standard deviation of move length: The contours of the KL-divergence of the numerical solution, $u_I^2(x)$, (a) from the analytic solution, $u_H^2(x)$ (Equation 2.44), derived from a Hyperbolic Scaling method; (b) from $u_M^2(x)$ (Equation 2.46), derived from a moment closure technique; (c) from $u_P^2(x)$ (Equation 2.50), derived from Patlak's method. (d) Turquoise region: the KL-divergence of $u_I^2(x)$ from $u_P^2(x)$ is smaller than from $u_M^2(x)$ or $u_H^2(x)$. Blue region: the KL-divergence of $u_I^2(x)$ from $u_M^2(x)$ is the smallest.
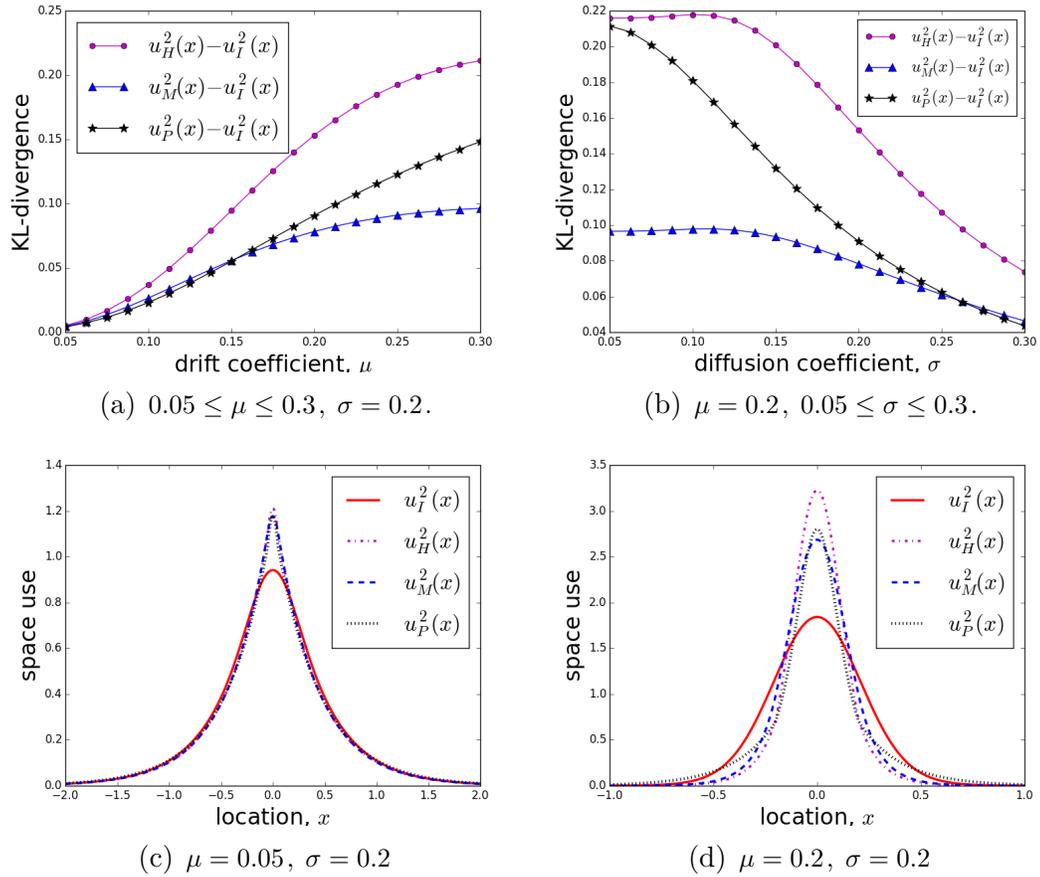
Figure 2.6: Continuous mean velocity movement kernel $k_\tau^2(z|x)$ with μ (resp. $|x|$) the mean
move length in one step for $|x| > \mu$ (resp. $|x| \le \mu$) and σ the standard deviation of move
length: (a) KL-divergence between $u_H^2(x)$ and $u_I^2(x)$ ($\bullet$), $u_M^2(x)$ and $u_I^2(x)$ ($\blacktriangle$), and $u_P^2(x)$
and $u_I^2(x)$ ($\star$) with $0.05 \le \mu \le 0.3$ and $\sigma = 0.2$. (b) KL-divergence between $u_H^2(x)$ and
$u_I^2(x)$ ($\bullet$), $u_M^2(x)$ and $u_I^2(x)$ ($\blacktriangle$), and $u_P^2(x)$ and $u_I^2(x)$ ($\star$) for $\mu = 0.2$, $0.05 \le \sigma \le 0.3$.
(c) steady-state distributions with $\mu = 0.05$ and $\sigma = 0.2$. (d) steady-state distributions with
$\mu = 0.2$ and $\sigma = 0.2$.
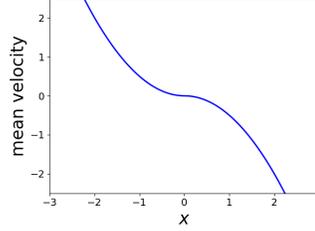
Figure 2.7: An example of a differentiable mean velocity function, defined by Equation (2.53), derived from a central-place foraging model with the central place being located at 0 (Equation 2.52).

central place. In particular, the mean velocity function of this model is continuously differentiable. The movement kernel is given by

$$
k_\tau^3(z|x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z - x - \mu x^2)^2}{2\sigma^2} \right) & \text{if } x < 0, \\[3mm] \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z - x + \mu x^2)^2}{2\sigma^2} \right) & \text{if } x \geq 0, \end{cases} \tag{2.52}
$$

where μ is the drift coefficient and σ is the diffusion coefficient. Note that this movement kernel (Equation 2.52) implies that the mean velocity could be arbitrarily large for large x, which is unrealistic for animal movement.

2.6.1 Analysis of movement kernel $k_\tau^3(z|x)$ by the PDE techniques

The drift and diffusion terms for the PDEs using the Hyperbolic Scaling and Moment Closure methods are computed by placing the movement kernel in Equation (2.52) into Equations (2.4) and (2.5), so that

$$
c_3(x) = \begin{cases} \dfrac{\mu x^2}{\tau} & \text{if } x < 0, \\[3mm] -\dfrac{\mu x^2}{\tau} & \text{if } x \geq 0, \end{cases} \tag{2.53}
$$

and

$$
D_3(x) = \frac{\sigma^2}{\tau^2}. \tag{2.54}
$$

Here, the mean velocity function (Equation 2.53) is differentiable everywhere (Figure 2.7). When using the Hyperbolic Scaling method, I insert these expres-

sions in Equations (2.53) and (2.54) into Equation (2.8) to give the steady-state
distribution as follows:

$$u_H^3(x) = \begin{cases} C_H^3 \exp\left(\dfrac{2\mu}{3\sigma^2}x^3 - \dfrac{\mu^2}{2\sigma^2}x^4\right) & \text{if } x < 0, \\[3mm] C_H^3 \exp\left(-\dfrac{2\mu}{3\sigma^2}x^3 - \dfrac{\mu^2}{2\sigma^2}x^4\right) & \text{if } x \geq 0, \end{cases} \tag{2.55}$$

where

$$C_H^3 = \left[\int_{-\infty}^0 \exp\left(\frac{2\mu}{3\sigma^2}x^3 - \frac{\mu^2}{2\sigma^2}x^4\right) dx + \int_0^\infty \exp\left(-\frac{2\mu}{3\sigma^2}x^3 - \frac{\mu^2}{2\sigma^2}x^4\right) dx\right]^{-1} \tag{2.56}$$

is a normalising constant. As in Sections 2.4 and 2.5, here I also calculate the
solution for $x > 0$ and $x < 0$ separately and assume continuity because $c_3(x)$ is
piecewise defined.

Likewise, substituting Equations (2.53) and (2.54) into Equation (2.12) for $x > 0$
and $x < 0$ and assuming continuity result in

$$u_M^3(x) = \begin{cases} C_M^3 \exp\left(\dfrac{2\mu}{3\sigma^2}x^3\right) & \text{if } x < 0, \\[3mm] C_M^3 \exp\left(-\dfrac{2\mu}{3\sigma^2}x^3\right) & \text{if } x \geq 0, \end{cases} \tag{2.57}$$

where $u_M^3(x)$ is the steady-state distribution derived by the Moment Closure
method and

$$C_M^3 = \left[\int_{-\infty}^0 \exp\left(\frac{2\mu}{3\sigma^2}x^3\right) dx + \int_0^\infty \exp\left(-\frac{2\mu}{3\sigma^2}x^3\right) dx\right]^{-1} \tag{2.58}$$

is a normalising constant.

Applying Patlak's approach involves the first and second moments of displace-
ment, which are given by placing the movement kernel in Equation (2.52) into
Equations (2.14) and (2.15) to give

$$M_1^3(x) = \begin{cases} \mu x^2 & \text{if } x < 0, \\ -\mu x^2 & \text{if } x \geq 0, \end{cases} \tag{2.59}$$

and

$$M_2^3(x) = \sigma^2 + \mu^2 x^4. \tag{2.60}$$

Since $M_1^3(x)$ is piecewise defined, the steady-state solution to the PDE is calculated for $x > 0$ and $x < 0$ with the assumption of continuity. Under these conditions, inserting Equations (2.59) and (2.60) into Equation (2.18) gives the steady-state distribution obtained by Patlak's approach as follows:

$$u_P^3(x) = \begin{cases} \dfrac{C_P^3}{\sigma^2 + \mu^2 x^4} \exp\left(-\sqrt{\dfrac{1}{\mu\sigma}}\left[2^{-\frac{3}{2}}\ln\left(\dfrac{|\frac{\mu}{\sigma}x^2 + \sqrt{\frac{2\mu}{\sigma}}x + 1|}{|\frac{\mu}{\sigma}x^2 - \sqrt{\frac{2\mu}{\sigma}}x + 1|}\right)\right.\right. \\ \qquad\left.\left. + \dfrac{1}{\sqrt{2}}\arctan\left(-\sqrt{\dfrac{2\mu}{\sigma}}x + 1\right) + \dfrac{1}{\sqrt{2}}\arctan\left(-\sqrt{\dfrac{2\mu}{\sigma}}x - 1\right)\right]\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } x < 0, \\[2ex] \dfrac{C_P^3}{\sigma^2 + \mu^2 x^4} \exp\left(-\sqrt{\dfrac{1}{\mu\sigma}}\left[2^{-\frac{3}{2}}\ln\left(\dfrac{|\frac{\mu}{\sigma}x^2 - \sqrt{\frac{2\mu}{\sigma}}x + 1|}{|\frac{\mu}{\sigma}x^2 + \sqrt{\frac{2\mu}{\sigma}}x + 1|}\right)\right.\right. \\ \qquad\left.\left. + \dfrac{1}{\sqrt{2}}\arctan\left(\sqrt{\dfrac{2\mu}{\sigma}}x + 1\right) + \dfrac{1}{\sqrt{2}}\arctan\left(\sqrt{\dfrac{2\mu}{\sigma}}x - 1\right)\right]\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } x \geq 0, \end{cases} \tag{2.61}$$

where C_P^3 is the usual normalising constant.

2.6.2 Numerical analysis of movement kernel $k_\tau^3(z|x)$

The plots of contour lines of KL-divergence show substantially different patterns to the observations of Sections 2.4 and 2.5 (Figures 2.8a-c). The KL-divergence of the real long-term pattern, $u_I^3(x)$ from the distributions given by the PDE methods increases as parameters μ or σ become larger (Figure 2.8d). When both μ and σ are very small, all three PDE methods provide estimates quite close to the real long-term distribution (2.9a). The inaccuracy of the approximate distributions increases as μ and σ rise (e.g. 2.9b), but the Moment Closure method performs best among the three PDE techniques.

(a) Hyperbolic Scaling

(b) Moment Closure

(c) Patlak's method

(d) $0.05 \leq \mu \leq 0.5$, $\sigma = 0.1$
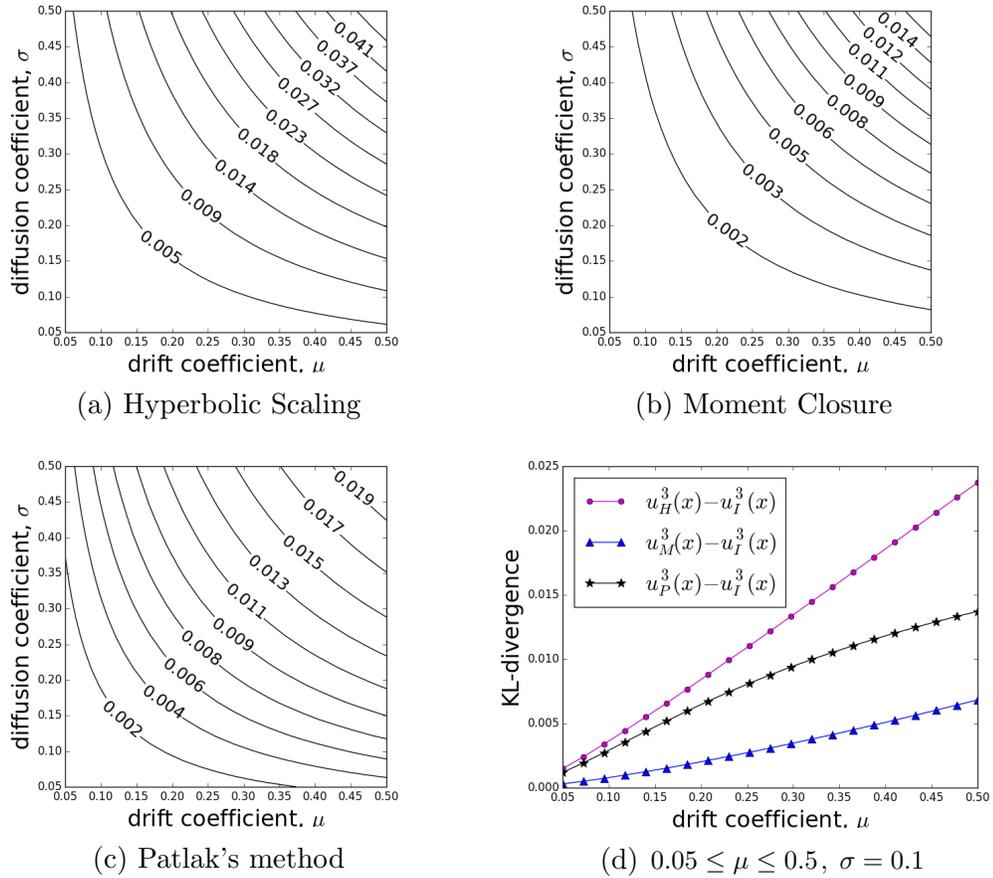
Figure 2.8: Differentiable mean velocity movement kernel $k_\tau^3(z|x)$ with μx^2 the mean move length in one step and σ the standard deviation of the move length: The contours of the KL-divergence of the numerical solution, $u_I^3(x)$, (a) from the analytic approximation, $u_H^3(x)$ (Equation 2.55), obtained using a Hyperbolic Scaling method, $\mu, \sigma \in [0.05, 0.5]$. (b) from the analytic approximation, $u_M^3(x)$ (Equation 2.57), obtained using a moment closure technique, $\mu, \sigma \in [0.05, 0.5]$. (c) from the analytic approximation, $u_P^3(x)$ (Equation 2.61), obtained using Patlak's method, $\mu, \sigma \in [0.05, 0.5]$. (d) KL-divergence between $u_H^3(x)$ and $u_I^3(x)$ ($\bullet$), $u_M^3(x)$ and $u_I^3(x)$ ($\blacktriangle$), and $u_P^3(x)$ and $u_I^3(x)$ ($\star$) with $0.05 \leq \mu \leq 0.5$ and $\sigma = 0.1$.

(a) $\mu = 0.05$, $\sigma = 0.05$ (b) $\mu = 0.8$, $\sigma = 0.5$
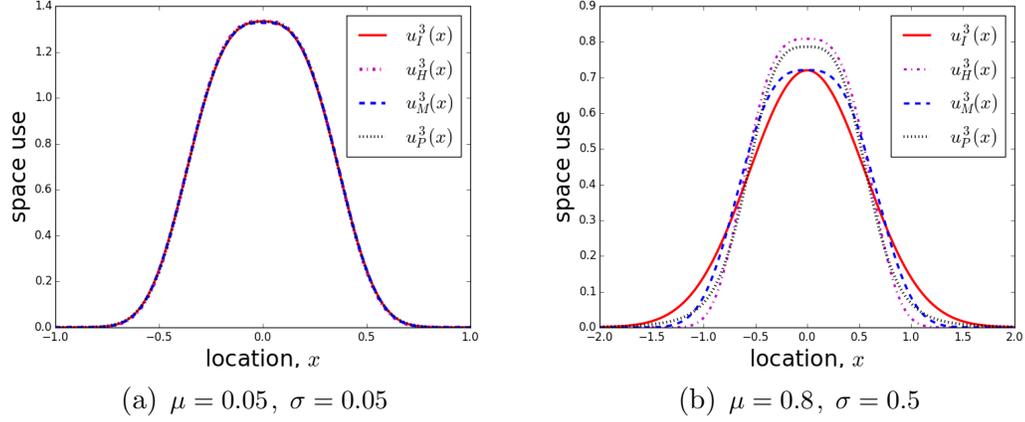
Figure 2.9: Differentiable mean velocity movement kernel $k_\tau^3(z|x)$ with μx^2 the mean move length in one step and σ the standard deviation of the move length: (a) steady-state distribution with $\mu = 0.05$ and $\sigma = 0.05$. (b) steady-state distribution with $\mu = 0.8$ and $\sigma = 0.5$.

The smooth mean velocity function in Equation (2.53) leads to very small changes in velocity near the central place, so the resulting approximations are much more smooth and flatter than the approximations derived in Sections 2.4.2 and 2.5.2. However, since the mean velocity function (Equations 2.53) are only twice differentiable, the PDE techniques might perform poorly when involving higher order moments and large μ.

§ 2.7   Movements on heterogeneous landscapes

The central-place foraging models in Sections 2.4-2.6 demonstrate analysis of movement in a 1D space with a bias towards an attraction centre. They are simple examples where the steady-state solutions to PDEs (Equations 2.8, 2.12, 2.18) can be calculated without numerical integration. Therefore, they are a good start to investigate to which extent these steady-state solutions can accurately approximate actual distributions emerging from the movement kernels. The central-place foraging models are also simple examples of movement in heterogeneous environments since the central place is the only one point affecting movement. In general, a heterogeneous space may contain more than one centres of attraction or repulsion and different types of habitats. In this section, I examine movement in heterogeneous environments in slightly more complex situations

than central-place foraging. Based on the framework of the step-selection func-
tion in Rhodes et al. (2005), the general form of the movement kernel considered
here is

$$k_\tau(z|x) = \frac{\varphi_\tau(z|x)w(z)}{\int_\Omega \varphi_\tau(y|x)w(y)\mathrm{d}y},$$ (2.62)

which describes the probability of an animal being at position z in time τ, given
its current position, x. The function $\varphi_\tau(z|x)$ itself is also a movement kernel,
representing the probability of moving to z from x in time τ in a homogeneous
environment. The function $w(z)$ is a weighting function, evaluating the influence
of the environmental conditions such as resources at position z on the movement.

To represent the movement process without considering environmental factors,
that is, the resource-independent movement kernel, here I use a Laplace distribu-
tion as an example in addition to a normal distribution. The purpose of consid-
ering an alternative distribution for the resource-independent movement kernel is
to show a possible way to construct a movement kernel using distributions other
than a normal distribution and how this kernel affects the performance of the
PDE methods. Choosing a Laplace distribution is because it is also symmetric
and simple. The resource-independent movement kernels are given by

$$\varphi_\tau^N(z|x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(z-x)^2}{2\sigma^2}\right),$$ (2.63)

$$\varphi_\tau^L(z|x) = \begin{cases} \dfrac{1}{2b} \exp\left(\dfrac{z-x}{b}\right) & \text{if } z < x, \\[2mm] \dfrac{1}{2b} \exp\left(\dfrac{x-z}{b}\right) & \text{if } z \geq x, \end{cases}$$ (2.64)

where σ^2 and $2b^2$ are the variance of distance moved over time τ, while the su-
perscripts "N" and "L" stand for normal and Laplace distributions respectively.

Here I assume the main environmental factor affecting movements is the resource
quality, which is uneven across the landscape and described by a weighting func-
tion. Two types of weighting functions are considered. One weighting function
consists of three distinct regions, while the other represent some resource quality

changing smoothly. The former is given by

$$w_T(z) = \begin{cases} 1 & \text{if } z \in [0, 1/3] \cup (2/3, 1], \\ 2 & \text{if } z \in (1/3, 2/3], \end{cases} \tag{2.65}$$

which features the shape of a top hat (Figure 2.10a). Therefore it is called a "top hat" function and the subscript "T" refers to it. This type of weighting function illustrates the environments where resources cluster in certain regions, forming disconnected food patches (e.g. Sawyer and Kauffman (2011); Merkle et al. (2014)).

The second weighting function describes the resource quality by a sine function, referred to by a subscript "S" as follows (Figure 2.10b):

$$w_S(z) = \sin(3\pi z) + 2. \tag{2.66}$$

This type of weighting function describes resources changing continuously over the space. Considering both discontinuous and smooth weighting functions is the same strategy as that in Sections 2.4-2.6 to understand how smoothness of movement kernels affects the approximations derived from the PDE techniques.

There are four possible ways to construct a movement kernel by inserting one of Equations (2.63) and (2.64) into the $\varphi_\tau(z|x)$ part in Equation (2.62) and one of Equations (2.65) and (2.66) into the $w(z)$ part in Equation (2.62). The first movement kernel uses a normal distribution for movement and the top-hat function for landscape resource, leading to

$$\begin{aligned} k_\tau^4(z|x) &= \frac{\varphi_\tau^N(z|x)w_T(z)}{\int_0^1 \varphi_\tau^N(y|x)w_T(y)\mathrm{d}y} \\ &= \begin{cases} \dfrac{1}{g_4(x)\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z-x)^2}{2\sigma^2}\right) & \text{if } z \in [0, 1/3] \cup (2/3, 1], \\[3mm] \dfrac{2}{g_4(x)\sqrt{2\pi}\sigma} \exp\left(\dfrac{-(z-x)^2}{2\sigma^2}\right) & \text{if } z \in (1/3, 2/3], \end{cases} \end{aligned} \tag{2.67}$$

where

$$
\begin{aligned}
g_4(x) &= \int_0^1 \varphi_\tau^N(y|x) w_T(y) \mathrm{d}y \\[2mm]
&= \frac{1}{2} \left[\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma}\right) + \mathrm{erf}\left(\frac{x - 1/3}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{x - 2/3}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{x - 1}{\sqrt{2}\sigma}\right) \right]
\end{aligned}
\tag{2.68}
$$

is a normalising function ensuring the distribution integrates to 1 over the domain.

The next model describes movements depending on a normal distribution in a landscape where the resource quality changes smoothly, given by

$$
\begin{aligned}
k_\tau^5(z|x) &= \frac{\varphi_\tau^N(z|x) w_S(z)}{\int_0^1 \varphi_\tau^N(y|x) w_S(y) \mathrm{d}y} \\[2mm]
&= \frac{1}{g_5(x)\sqrt{2\pi}\sigma} \exp\left(\frac{-(z-x)^2}{2\sigma^2}\right) (\sin(3\pi z) + 2),
\end{aligned}
\tag{2.69}
$$

where

$$
\begin{aligned}
g_5(x) &= \int_0^1 \varphi_\tau^N(y|x) w_S(y) \mathrm{d}y \\[2mm]
&= \int_0^1 \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(y-x)^2}{2\sigma^2}\right) (\sin(3\pi y) + 2) \mathrm{d}y.
\end{aligned}
\tag{2.70}
$$

The remaining two movement kernels use a Laplace distribution for the resource-independent movement kernel. With the resource quality in space being described by the top hat function in Equation (2.65), the movement kernel is

$$k_\tau^6(z|x) = \frac{\varphi_\tau^L(z|x)w_T(z)}{\int_0^1 \varphi_\tau^L(y|x)w_T(y)\mathrm{d}y}$$

$$= \begin{cases} \dfrac{1}{2bg_{61}(x)} \exp\left(\dfrac{z-x}{b}\right) & \text{if } x \in [0,1/3] \text{ and } z \in [0,x], \\[3mm] \dfrac{1}{2bg_{61}(x)} \exp\left(\dfrac{x-z}{b}\right) & \text{if } x \in [0,1/3] \text{ and } z \in [x,1/3] \cup (2/3,1], \\[3mm] \dfrac{1}{bg_{61}(x)} \exp\left(\dfrac{x-z}{b}\right) & \text{if } x \in [0,1/3] \text{ and } z \in (1/3,2/3], \\[3mm] \dfrac{1}{2bg_{62}(x)} \exp\left(\dfrac{z-x}{b}\right) & \text{if } x \in (1/3,2/3] \text{ and } z \in [0,1/3], \\[3mm] \dfrac{1}{bg_{62}(x)} \exp\left(\dfrac{z-x}{b}\right) & \text{if } x \in (1/3,2/3] \text{ and } z \in (1/3,x], \\[3mm] \dfrac{1}{bg_{62}(x)} \exp\left(\dfrac{x-z}{b}\right) & \text{if } x \in (1/3,2/3] \text{ and } z \in (x,2/3], \\[3mm] \dfrac{1}{2bg_{62}(x)} \exp\left(\dfrac{x-z}{b}\right) & \text{if } x \in (1/3,2/3] \text{ and } z \in (2/3,1], \\[3mm] \dfrac{1}{2bg_{63}(x)} \exp\left(\dfrac{z-x}{b}\right) & \text{if } x \in (2/3,1] \text{ and } z \in [0,1/3] \cup (2/3,x], \\[3mm] \dfrac{1}{bg_{63}(x)} \exp\left(\dfrac{z-x}{b}\right) & \text{if } x \in (2/3,1] \text{ and } z \in (1/3,2/3], \\[3mm] \dfrac{1}{2bg_{63}(x)} \exp\left(\dfrac{x-z}{b}\right) & \text{if } x \in (2/3,1] \text{ and } z \in (x,1], \end{cases}$$

(2.71)

where

$$g_{61}(x) = 1 - \frac{1}{2}\left[\exp\left(\frac{-x}{b}\right) - \exp\left(\frac{x-1/3}{b}\right) + \exp\left(\frac{x-2/3}{b}\right) + \exp\left(\frac{x-1}{b}\right)\right],$$

(2.72)

$$g_{62}(x) = 2 - \frac{1}{2}\left[\exp\left(\frac{-x}{b}\right) + \exp\left(\frac{1/3-x}{b}\right) + \exp\left(\frac{x-2/3}{b}\right) + \exp\left(\frac{x-1}{b}\right)\right],$$

(2.73)

$$g_{63}(x) = 1 - \frac{1}{2}\left[\exp\left(\frac{-x}{b}\right) + \exp\left(\frac{1/3 - x}{b}\right) - \exp\left(\frac{2/3 - x}{b}\right) + \exp\left(\frac{x-1}{b}\right)\right].$$

(2.74)

are normalising functions for $x \in [0, 1/3]$, $x \in (1/3, 2/3]$ and $x \in (2/3, 1]$ respectively.

The last movement kernel combines the Laplace distribution in Equation (2.64) with the sine function in Equation (2.66) to give

$$
\begin{aligned}
k_\tau^7(z|x) &= \frac{\varphi_\tau^L(z|x)w_S(z)}{\int_0^1 \varphi_\tau^L(y|x)w_S(y)\mathrm{d}y} \\[2mm]
&= \begin{cases} \dfrac{1}{2bg_7(x)}\exp\left(\dfrac{z-x}{b}\right)(\sin 3\pi z + 2) & \text{if } z < x, \\[4mm] \dfrac{1}{2bg_7(x)}\exp\left(\dfrac{x-z}{b}\right)(\sin 3\pi z + 2) & \text{if } z \geq x, \end{cases}
\end{aligned}
$$

(2.75)

where

$$
\begin{aligned}
g_7(x) &= \int_0^1 \varphi_\tau^L(y|x)w_S(y)\mathrm{d}y \\[2mm]
&= 2 - \frac{4}{(18\pi^2 b^2)^2 - 4}\sin(3\pi x) - \frac{108\pi^3 b^3}{(18\pi^2 b^2)^2 - 4}\cos(3\pi x) \\[2mm]
&\quad + \left(\frac{3\pi b}{18\pi^2 b^2 + 2} - 1\right)\exp\left(\frac{-x}{b}\right) - \left(\frac{3\pi b}{18\pi^2 b^2 - 2} + 1\right)\exp\left(\frac{x-1}{b}\right)
\end{aligned}
$$

(2.76)

is a normalising function. The expressions in Equations 2.68, 2.72, 2.73, 2.74 and 2.76 are obtained by using Integral Calculator *Integral Calculator* (2019).

Having constructed movement kernels for situations where an animal is moving in a 1D interval with heterogeneous resource quality, now I employ the PDE techniques introduced in Section 2.1 to estimate the long-term population distributions. The approximate distributions are represented by the steady-state solutions to relevant PDEs used in the Hyperbolic Scaling method (Equation 2.8), the Moment Closure method (Equation 2.12) and Patlak's approach (Equation 2.18), denoted by subscripts "H", "M" and "P" respectively. The real long-term distributions arising from the movement kernels are described by the

(a) $w_t(x) = \begin{array}{ll} 1 & \text{if } x \in [0, 1/3] \cup (2/3, 1] \\ 2 & \text{if } x \in (1/3, 2/3] \end{array}$

(b) $w_s(x) = \sin(3\pi x) + 2$

(c)

(d)

(e)

(f)
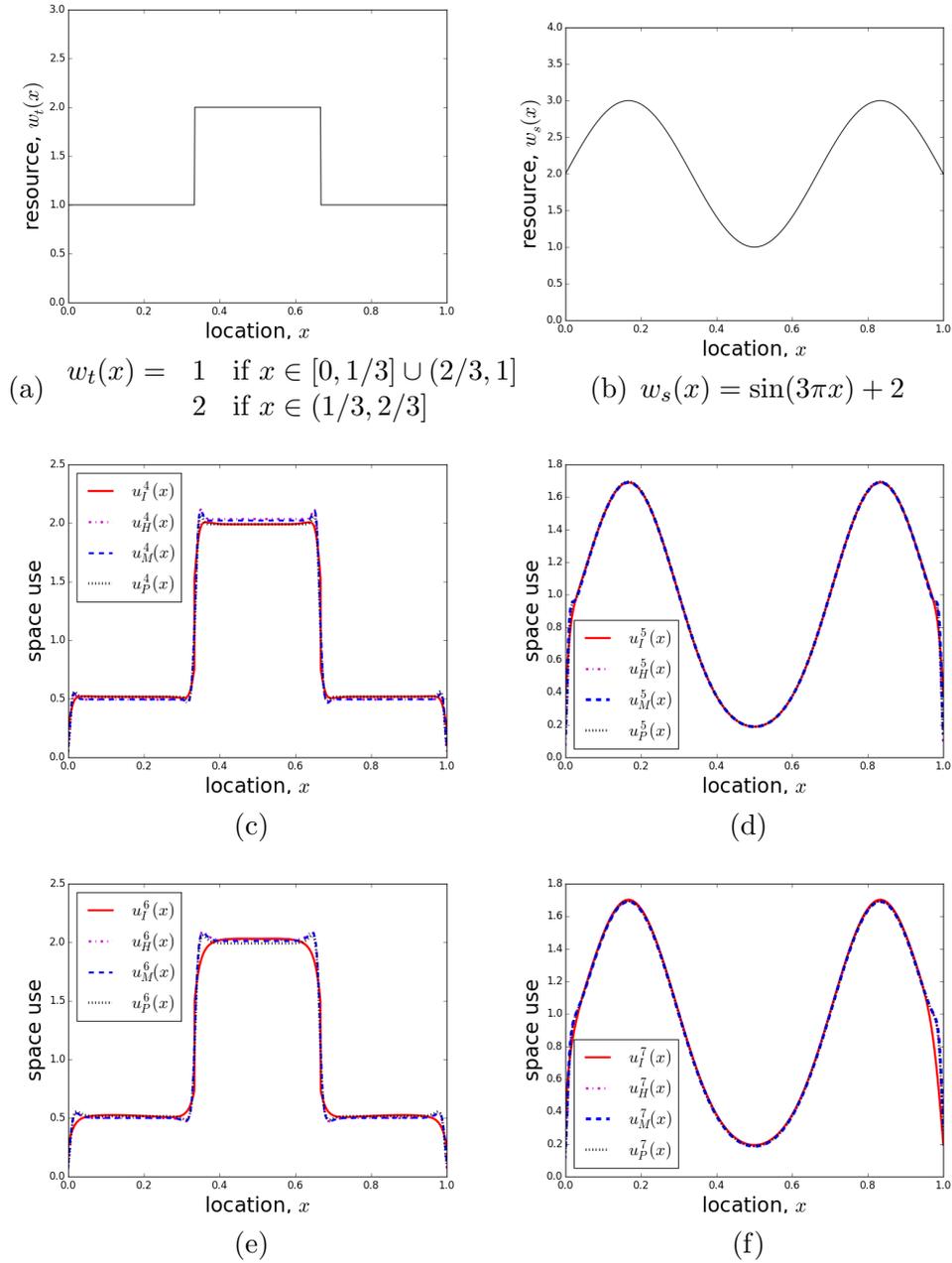
Figure 2.10: Steady-state distributions emerging from movement on heterogeneous landscapes. (a) The weighting function $w_t(x)$ (Equation 2.65). (b) The weighting function $w_s(x)$ (Equation 2.66). (c) Movement according to kernel $k_\tau^4(z|x)$ (Equation 2.67) based on a Normal distribution with $w_t(x)$ as the weighting function. (d) Movement according to kernel $k_\tau^5(z|x)$ (Equation 2.69) based on a Normal distribution with $w_s(x)$ as the weighting function. (e) Movement according to kernel $k_\tau^6(z|x)$ (Equation 2.71) based on a Laplace distribution with $w_t(x)$ as the weighting function. (f) Movement according to kernel $k_\tau^7(z|x)$ (Equation 2.75) based on a Laplace distribution with $w_s(x)$ as the weighting function.

(a)   (b)
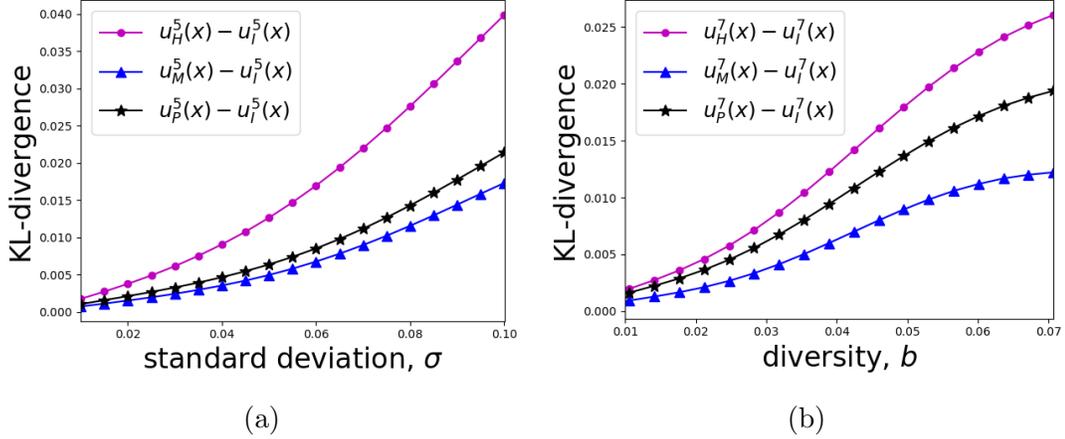
Figure 2.11: Movements in a landscape with smooth resource change, defined by movement kernels $k_\tau^5(z|x)$ (Equation 2.69) and $k_\tau^7(z|x)$ (Equation 2.75) with variance σ^2 and $2b^2$. (a) KL-divergence between $u_H^5(x)$ and $u_I^5(x)$ ($\bullet$), $u_M^5(x)$ and $u_I^5(x)$ ($\blacktriangle$), and $u_P^5(x)$ and $u_I^5(x)$ ($\star$) with $0.01 \leq \sigma \leq 0.1$. (b) KL-divergence between $u_H^7(x)$ and $u_I^7(x)$ ($\bullet$), $u_M^7(x)$ and $u_I^7(x)$ ($\blacktriangle$), and $u_P^7(x)$ and $u_I^7(x)$ ($\star$) with $0.01 \leq b \leq 0.07$.

integration of the Master Equation (2.21) and referred to using a subscript "I".

Unlike the cases examined in Sections 2.4-2.6, here, the integrals in the steady-state distributions (Equations 2.8,2.12,2.18) can only be calculated numerically. Figures 2.10c-f illustrate an example for models $k_\tau^4(z|x)$, $k_\tau^5(z|x)$, $k_\tau^6(z|x)$ and $k_\tau^7(z|x)$ when fixing the variance of the functions $\varphi_\tau^N(z|x)$ and $\varphi_\tau^L(z|x)$ in Equations (2.63) and (2.64) at $\sigma^2 = 2b^2 = 10^{-4}$. In all four cases of moving in a heterogeneous landscape, the steady-state distributions obtained by the three PDE techniques are quite similar to each other. All approximation methods give poor estimations of the real long-term patterns at discontinuous points (Figures 2.10c,e).

For models built on movement kernels in Equations (2.69) and (2.75), the resource qualities are smoothly distributed and the Moment Closure method always gives the best approximation of the real long-term distributions (Figure 2.11). This coincides with the observations in Section 2.6, where the mean velocity function of movement in a central-place foraging model is smooth. In both cases, the distance between steady-state distributions derived by PDE methods and the real long-term distributions rises as the standard deviation (σ, in Equation 2.63 or diversity, b, in Equation 2.64) increases (Figure 2.11).

§ 2.8   Summary

I have briefly introduced three PDE techniques for scaling up individual movement kernels to population distributions. These three methods are Patlak's approach (Patlak, 1953) and two recently reviewed methods, the Hyperbolic Scaling and Moment Closure methods (Hillen and Painter, 2013). There are two approaches to evaluate the accuracy of using the PDE methods to approximate patterns arising from the movement kernel. One is to compare the transient distributions derived from the PDEs in Equations (2.3), (2.9), (2.13) to the movement kernel, while the other is to compare the steady-state distributions in Equations (2.8), (2.12) and (2.18) to the long-term distribution obtained by the Master Equation (Equation 2.21).

By analysing a simple example of a biased random walk, I have demonstrated a comparison between the movement kernel and transient distributions derived by the PDE methods. Applying Patlak's approach to analyse this example results in a very poor approximation with an overestimated variance, while the Hyperbolic Scaling and Moment Closure methods correctly describe the movement kernel (Figure 2.1).

To understand how PDE techniques performs when estimating the long-term distribution, I have examined three central-place foraging models, characterised by mean velocity functions with different levels of smoothness. When analysing a non-smooth movement kernel, all PDE methods give poor approximations to the long-term distribution (Figures 2.3, 2.6). On the other hand, when considering the model with a differentiable mean velocity function, the accuracy of approximations improves and the Moment Closure methods performs better than the others in the range studied (Figures 2.8d, 2.9).

In addition, I have investigated some simple examples of movement in heterogeneous environments and shown similar results to those observed in central-place foraging models. That is, the PDE methods provide poor approximations if the movement kernel is non-smooth and perform well for smooth models (Figure 2.10). For smooth movement kernels, the Moment Closure method outperforms others (Figure 2.11).

Chapter 3

Resource selection analysis by continuous-time movement models

Resource selection analysis (RSA) is a fundamental tool for understanding mechanisms behind abundance and distributions. It has been strengthened by methods such as step selection analysis (SSA) and integrated step selection analysis (iSSA) because the incorporation of movements makes it straightforward to define the availability of a resource unit by mobility. However, both SSA and iSSA rely on a discrete-time movement model, subject to fixed observation intervals. Therefore, RSA can be improved even further by integrating resource selection into a continuous-time movement modelling framework rather than a discrete-time framework. Moreover, SSA and iSSA compare a used 'step', defined by two successive observations, to some available steps starting with the same source point as the used step. This means SSA and iSSA consider selection at the scale of steps and hence depend on the assumption of a correspondence between decision making and observation scales. By considering movements in continuous time, it is straightforward to deal with irregularly collected data and allow changes in movement decisions. This chapter introduces such a modelling framework on the assumption that animals would be attracted to the place with the best resource quality.

§ 3.1   Modelling framework

Using the OU process as a building block as in Blackwell et al. (2016), I construct a switching OU process, composed of a set of OU processes, each of which represents a random walk towards a target place, or an attraction centre, different to others. That is, as the destination of movement switches because of the change of resources, the OU process used is switched to another OU process which models movement attracted to the new target place. In this way, I model movements in response to the change of resources over time.

Assuming an animal is moving in a 2-dimensional space and its location at time t is $\mathbf{x}(t)$, the OU process gives the probability of the animal's location in time τ as follows:

$$\mathbf{x}(t+\tau)|\mathbf{x}(t) \sim MVN(\boldsymbol{\mu}(t) + e^{B\tau}(\mathbf{x}(t) - \boldsymbol{\mu}(t)), \Lambda - e^{B\tau}\Lambda e^{B'\tau}), \qquad (3.1)$$

where MVN refers to "multi-variate normal" distribution and here the 2-dimensional version is considered. The attraction centre at time t is $\boldsymbol{\mu}(t)$. The tendency towards the attraction centre and uncertainty of the movement are controlled by a 2×2 matrix B and the 2×2 covariance matrix Λ respectively. Here, I assume $B = -bI$ and $\Lambda = vI$ with $b, v > 0$ and I the 2×2 identity matrix. The 0s off the diagonal of Λ indicates no correlation between the two coordinates. The parameter b governs the strength of the drift towards the central point, while the parameter v determines the range of strolling around the centre (Blackwell, 1997). Equation (3.1) is a continuous-time analogue of a movement kernel such as those in Chapter 2 but time τ here can be any value rather than being fixed.

Here, I assume that the attraction centre $\boldsymbol{\mu}(t)$ in Equation (3.1) is decided by using a weighting function to assess the attractiveness of locations in space. A resource selection function (RSF) is commonly used for this purpose as it reflects the probability of an animal using a resource unit in space. It is usually formulated by an exponential function such as (Boyce et al., 2002)

$$w(\mathbf{z}(\mathbf{x})) = \exp(\beta_1 z_1(\mathbf{x}) + \beta_2 z_2(\mathbf{x}) + \cdots + \beta_k z_k(\mathbf{x})), \qquad (3.2)$$

where $\mathbf{x}$ is a resource unit, which can be an area in space. The factors influencing movement decisions are incorporated in the vector of predictor covariates, $\mathbf{z}(\mathbf{x}) = (z_1(\mathbf{x}), \cdots, z_k(\mathbf{x}))$, with coefficients $\beta_1, \cdots, \beta_k$ representing the extent to which these factors affect movements. Possible drivers of movement can be categorical such as vegetation types, or continuous such as distance to human constructions, etc. (Manly et al., 2002).

I also assume that the animal has complete knowledge of the environment and determines its movement centre $\boldsymbol{\mu}(t)$ at time t by selecting the most attractive place in space. That is, the target place is decided by comparing the RSF in Equation (3.2) at all potential destinations and given by (cf. Avgar et al. (2017), Bastille-Rousseau et al. (2017))

$$\boldsymbol{\mu}(t) = \boldsymbol{\mu}_i \qquad \text{where} \qquad w(\mathbf{z}(\boldsymbol{\mu}_i)) = \max_{j \in \Omega} w(\mathbf{z}(\boldsymbol{\mu}_j)). \qquad (3.3)$$

Here, the attraction centre, $\boldsymbol{\mu}_i$, is the centre of a resource unit, which may be a food patch of any shape. The notation Ω stands for the finite collection of all resource units considered. If the attraction centre determined by Equation (3.3) is not unique because of the equal attractiveness of more than one resource units, then further steps are required to make a decision. For example, randomly select one of the most attractive places. Alternatively, it may be feasible to exclude some minor factors from the RSF rather than considering all possible factors. However, in real life, $\boldsymbol{\mu}_i$ will almost always be unique if the RSF involves continuous covariates.

Note that in this modelling framework, reassessment of the movement centre can occur at any given time. Therefore, to include points where the reassessment might happen in the inference procedure, a Poisson process is used to simulate such points between observations (Blackwell et al., 2016). This is similar to a velocity jump process, which is also a continuous-time movement model and assumes the change of velocity has a Poisson distribution in time (Othmer et al., 1988). While the change of velocity in a velocity jump process relies on a turning kernel, the change of attraction centre in a switching OU process is determined by a RSF. Although it is possible to develop a velocity jump process which describe movement biased towards an attraction centre, it is straightforward to use OU

processes as they intrinsically describe this bias.

§ 3.2   Inference by Markov chain Monte Carlo

To parameterise the model given in Section 3.1 from movement data, I develop a Markov chain Monte Carlo (MCMC) algorithm based on Blackwell et al. (2016). A key feature of the inference procedure introduced in Blackwell et al. (2016) is that it takes into account the fact that an animal can make a movement decision at any given time. This has successfully resolved one of the major problems of discrete-time movement models, which assume that behavioural changes would not occur between observations.

The major difference between Blackwell et al. (2016) and this work is the determination of the attraction centre in the OU process in Equation (3.1). Rather than determining the attraction centre by transition rates defined by habitat and time as in Blackwell et al. (2016), here, the attraction centre is decided by the comparison of the values of the RSF at potential destinations.

The inference procedure generates a chain of samples for each parameter by iteration and is composed of two major stages, that is, the augmentation of observed data and the update on parameters. Both of these two stages follow a typical technique for applying Bayesian inference, the Metropolis-Hastings algorithm, which accepts a proposed sample state according to a Hastings ratio determined by the likelihoods of the samples (Chib and Greenberg, 1995). Being based on likelihoods makes the Metropolis-Hastings algorithm flexible enough to adjust for many situations, including updating trajectories here. However, it can take hours or days to generate a sufficient large amount of samples to derive a better estimation of parameters.

In the first stage of the inference procedure here, a subset of the observed data is augmented with potential points where the switch of the attraction centre might occur. That is, given the observed data, $\mathbf{x}(t_0), \cdots, \mathbf{x}(t_n)$, where $\mathbf{x}(t_i)$ is the animal's location at time t_i, for each iteration of the inference procedure, I start with the selection of a subset of data points, $\mathbf{x}(t_k), \mathbf{x}(t_{k+1}), \cdots, \mathbf{x}(t_l)$, where $t_0 \leq t_k < t_l \leq t_n$. Once a time interval t_k, t_l has been selected, I generate a series

of time points denoted by $\{t'_{pq}, p = k, \cdots, l-1, q = 1, \cdots, M'_p\}$ with M'_p the number of switching points between t_p and t_{p+1} on this interval, assuming the switch of destination is a Poisson process with mean $\kappa(t_l - t_k)$, as in Blackwell et al. (2016) (e.g. Figure 3.1a).

After generating time points between times t_k and t_l, at each time point, the location and attraction centre are determined. If a time point is an observed data point, then the animal's location at that time is fixed and the attraction centre can be determined by Equation (3.3). When considering a proposed time point at which the location is unknown, a simulated position is given by the OU process in Equation (3.1) depending on the location and attraction centre at the previous time. Once a simulated location is selected, the attraction centre at that point is then decided using Equation (3.3) and then the next location can be generated if necessary (e.g. Figure 3.1b).

Having decided the locations and attraction centres along the proposed path, I calculate the likelihood of this path and the likelihood of the existing data between times t_k and t_l to give the Hastings ratio as follows (Blackwell et al., 2016):

$$\prod_{i=k}^{l-1} \frac{f(\mathbf{x}(t_{i+1})|\mathbf{x}(t'_{i,M'_i}), \boldsymbol{\mu}(t'_{i,M'_i}))}{f(\mathbf{x}(t_{i+1})|\mathbf{x}(t_{i,M_i}), \boldsymbol{\mu}(t_{i,M_i}))}. \tag{3.4}$$

Here, the denominator is the likelihood of moving to observed points $\mathbf{x}(t_{i+1}), i \in \{k, k+1, \cdots, l-1\}$ from their existing predecessors $\mathbf{x}(t_{i,M_i}), i \in \{k, k+1, \cdots, l-1\}$, where the centres of attraction are $\boldsymbol{\mu}(t_{i,M_i}), i \in \{k, k+1, \cdots, l-1\}$. In the numerator, $\mathbf{x}(t'_{i,M'_i}), i \in \{k, k+1, \cdots, l-1\}$ are candidate predecessors of the observed points from the proposed path and $\boldsymbol{\mu}(t'_{i,M'_i}), i \in \{k, k+1, \cdots, l-1\}$ are the attraction centres associated to these candidates. The proposed path is accepted with a probability of

$$\text{acceptance rate} = \min\{1, \text{Hastings ratio}\}. \tag{3.5}$$

In practice, a random number u is generated from the uniform distribution on the interval $[0, 1]$ and compared to the acceptance rate. If $u \leq$ acceptance rate, then the proposed path is accepted and the existing switching points between
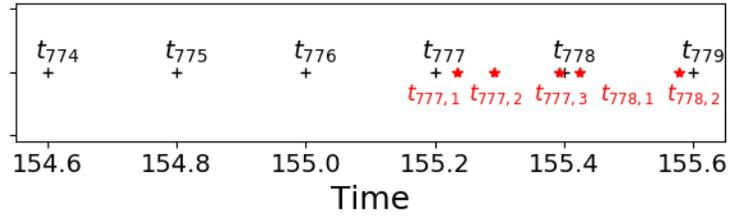
(a)
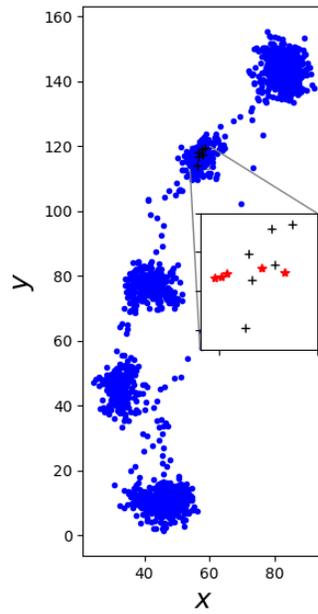
(b)

Figure 3.1: Augmentation of a subset of observed data. (a) A selected time interval containing observation points from t_{774} to t_{779}, marked by black crosses. Time points when a switch of destination might occur on this interval are generated and marked by red stars. (b) After proposed time points are generated, locations at these times are decided by following an OU process according to the states at predecessors. Blue dots show the original observed locations.

observations on the selected interval are replaced by the newly generated points.
If u is greater than the acceptance rate, then the proposed path is rejected and
the existing points where decisions might change on this interval are kept.

The augmentation of observed data enables the detection of decision changes
occurring between observations. This is especially crucial for analysing data with
coarse time resolution if the observation scale is much larger than the decision-
making scale. However, the incorporation of additional points will reduce the
efficiency of the algorithm in terms of computer time as more points need to
be processed. Therefore, it may require some prior knowledge of the behaviours
of the animal or pilot studies to determine an appropriate value for κ, which
partially decides the average number of points added in a unit time interval.

In the second stage, the parameters for the model consisting of Equations (3.1)
and (3.2) are updated according to the accepted trajectory determined in the
first stage. The parameters to be updated are the drift and diffusion coefficients
b and v in Equation (3.1) and the resource coefficients $\beta_1, \cdots, \beta_k$ in Equation
(3.2). Here I will use $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_k)$ to ease the notation. For each parameter,
a candidate is generated from a chosen probability distribution, termed a proposal
distribution. Here, I use a normal distribution for a proposal distribution. For
example, if the current value of parameter b in this MCMC sampling chain is
$b^{(0)}$, then a candidate b' for parameter b is selected from a normal distribution
with a probability density function

$$g_b(b'|b^{(0)}) = \frac{1}{\sqrt{2\pi}\sigma_b} \exp\left(\frac{-(b' - b^{(0)})^2}{2\sigma_b^2}\right), \tag{3.6}$$

where σ_b^2 is the variance of this distribution.

Next, the candidate parameter, $\boldsymbol{\beta}'$, for the RSF in Equation (3.2) is used to
update the attraction centre at each point of the accepted trajectory. With the
updated attraction centres, I calculate the likelihood of the trajectory using the
candidate parameters, b' and v', for the OU process in Equation (3.1). Then this
likelihood is compared with the likelihood of the trajectory computed using the
existing parameter values, $b^{(0)}$, $v^{(0)}$ and $\boldsymbol{\beta}^{(0)}$, to calculate the Hastings ratio in
the following:

$$\prod_{i=0}^{n} \left(\frac{f(\mathbf{x}(t_{i,1})|\mathbf{x}(t_i), b', v', \boldsymbol{\beta}')}{f(\mathbf{x}(t_{i,1})|\mathbf{x}(t_i), b^{(0)}, v^{(0)}, \boldsymbol{\beta}^{(0)})} \left(\prod_{j=1}^{M_i-1} \frac{f(\mathbf{x}(t_{i,j+1})|\mathbf{x}(t_{i,j}), b', v', \boldsymbol{\beta}')}{f(\mathbf{x}(t_{i,j+1})|\mathbf{x}(t_{i,j}), b^{(0)}, v^{(0)}, \boldsymbol{\beta}^{(0)})} \right) \frac{f(\mathbf{x}(t_{i+1})|\mathbf{x}(t_{i,M_i}), b', v', \boldsymbol{\beta}')}{f(\mathbf{x}(t_{i+1})|\mathbf{x}(t_{i,M_i}), b^{(0)}, v^{(0)}, \boldsymbol{\beta}^{(0)})} \right).$$

$$(3.7)$$

Subsequently, the parameter samples $b^{(0)}$, $v^{(0)}$ and $\boldsymbol{\beta}^{(0)}$ are replaced by the candidates with the probability of the Hastings ratio or 1, whichever is smaller. For details of the inference method, see Blackwell et al. (2016). The inference algorithm is carried out using R (R Core Team, 2017).

In general, in terms of computer time, since the approach introduced here is based on the Metropolis-Hastings algorithm, allowing simultaneous update on multiple variables, it should be more efficient than using another commonly used MCMC technique, the Gibbs sampling method, which updates one parameter at a time Albert (2007). In practice, some initial sample points of the MCMC chains are discarded so the remaining samples are in the high probability region of the posterior distribution. That is, the chains converges after the discarded portion of iterations. The number of iterations before converging can depend on initial values and the variance of proposal distributions. However, the chains should converge faster when starting with values closer to the real parameter values.

Before applying this inference method on real movement data, I test it on simulated trajectories to examine its capacity to estimate the parameters used in the simulations. The simulation models investigated in this chapter include migration models and models for movements depending on the depletion and renewal of resources.

§ 3.3   Migration models

The first example of the modelling framework introduced in Section 3.1 is a model of migration. As a simple example, here I assume migration decisions are made solely dependent on the trade-off between the resource quality of a potential target patch and the travel distance to the patch. Note that migration is triggered by resource qualities changing seasonally and the distance between the animal and a patch is also dependent on time. Thus the RSF in Equation (3.2) is modified

to incorporate time into the function as a variable to give

$$w(\mathbf{z}(\boldsymbol{\mu}, t)) = \exp(\beta_1 R(\boldsymbol{\mu}, t) + \beta_2 |\boldsymbol{\mu} - \mathbf{x}(t)|), \tag{3.8}$$

where $\mathbf{z}(\boldsymbol{\mu}, t) = (R(\boldsymbol{\mu}, t), |\boldsymbol{\mu} - \mathbf{x}(t)|)$ is the vector of predictor covariates, containing $R(\boldsymbol{\mu}, t)$, the resource quality in a food patch with centre $\boldsymbol{\mu}$ at time t, and $|\boldsymbol{\mu} - \mathbf{x}(t)|$, the distance between $\boldsymbol{\mu}$ and $\mathbf{x}(t)$, the animal's position at time t. Here I assume the animal favours resources of higher quality and is averse to long-distance travel, so that $\beta_1 > 0$ and $\beta_2 < 0$.

In the inference process described in Section 3.2, the values of the RSF at all potential destinations are compared to decide the attraction centre at a given time. In other words, the aim is to order the places by the RSF values rather than considering the actual values of the RSF. Therefore, in practice, the results of the comparison will not change when the exponent of the RSF is multiplied by a constant. Consequently, one of the coefficients in Equation (3.8) can be factored out and it is sufficient to consider an alternative RSF given by

$$w(\mathbf{z}(\boldsymbol{\mu}, t)) = \exp(\beta R(\boldsymbol{\mu}, t) - |\boldsymbol{\mu} - \mathbf{x}(t)|), \tag{3.9}$$

where $\beta = -\beta_1/\beta_2$. The parameters to be inferred are then β in Equation (3.9) and the drift and diffusion coefficients, b and v, in the OU process in Equation (3.1).

3.3.1   SIMULATIONS

Here, I assume an animal moves in a two-dimensional landscape comprising N food patches, which do not overlap each other. These patches are denoted by A_i, $i \in \{1, \cdots, N\}$ and ordered by latitude. The centre of patch A_i is denoted by $\boldsymbol{\mu}_i = (x_i, y_i)$, where $y_1 \leq y_2 \cdots \leq y_N$ are latitudinal coordinates. The resource quality in patch A_i changes with period one year. The resource quality in a patch is defined by a cosine function with a 365-day period as follows:

$$R(A_i, t) = a \cos\left(\frac{2\pi}{365} t - \frac{y_i - y_1}{y_N - y_1} \pi\right) + m, \tag{3.10}$$

where a is the amplitude and m is the mean of the resource quality. Equation 3.10 is an example of resource qualities vary periodically such as the bulk of the growth of plants. For simulations, a and m are chosen arbitrarily as long as $m \geq a$ so that the resource quality $R(A_i, t)$ in Equation 3.10 is always non-negative.

Assuming that the resource quality in a patch is given by Equation (3.10) with $a = 50$ and $m = 150$, I generated migration trajectories with $0.1 \leq b \leq 0.8$, $2 \leq v \leq 30$ for the OU process (Equation 3.1) and $1.5 \leq \beta \leq 8.5$ for the RSF (Equation 3.9) in a 90×160 unit^2 landscape with 10 non-overlapping food patches. These parameters were chosen in these ranges so that migration from patches in the south to those in the north and back to the south late in the year would occur in the landscape.
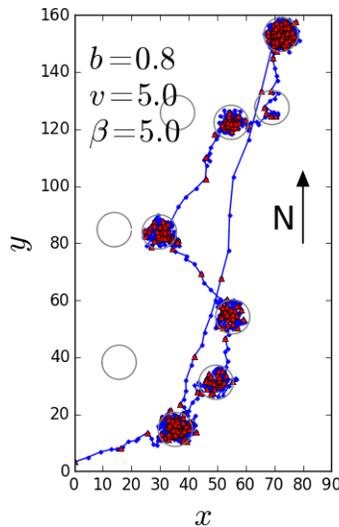
Figure 3.2: A simulation of migration corresponding to the model described in Section 3.3.1. The blue dots and line segments show the whole set of data points, and the red triangles are those used in the MCMC algorithm. The animal moves towards the north from patches in the south and comes back to the south.

Three sets of simulations were generated by varying one of the parameters, b, v and β, with the other two being fixed. For the first set of simulations, I used different values for b and fixed v and β at 10 and 5 respectively. The second set of simulations was produced using various v with $b = 0.5$ and $\beta = 5$. With $b = 0.5$

and $v = 15$, I created the last set of simulations by changing β. The values for
the fixed parameters were selected arbitrarily from the parameter ranges stated
in the last paragraph. The parameter ranges were determined in such a way that
the change of attraction centres may occur. Figure 3.2 shows a simulation of
migration from the winter range in the south to the summer range in the north
via five stopovers and back to the south over a year. The red triangles in Figure
3.2 are the points selected to test the effectiveness of the inference method in
dealing with missing data.

3.3.2   Inference from simulations
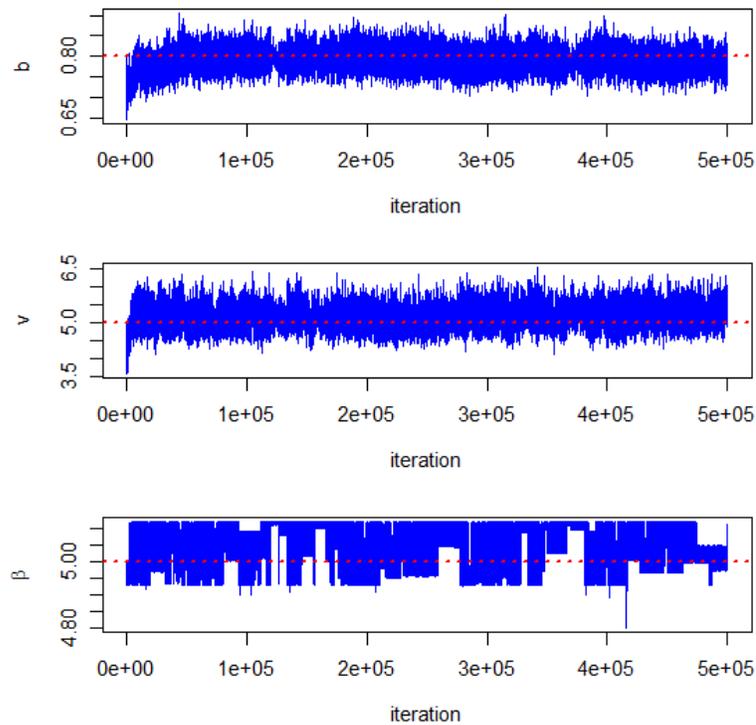
Figure 3.3: Trace plots of MCMC chains when inferring parameters from the simulation in
Figure 3.2. The sampling chains of the drift and diffusion coefficients b and v in Equation (3.1),
resource coefficient β in Equation (3.9). After around 50,000 iterations, all chains converge.

Having generated the simulated trajectories, the inference procedure in Section
3.2 was applied to infer the model parameters, namely b and v in the OU process
(Equation 3.1) and the coefficient β in the RSF (Equation 3.9). Figure 3.3 gives
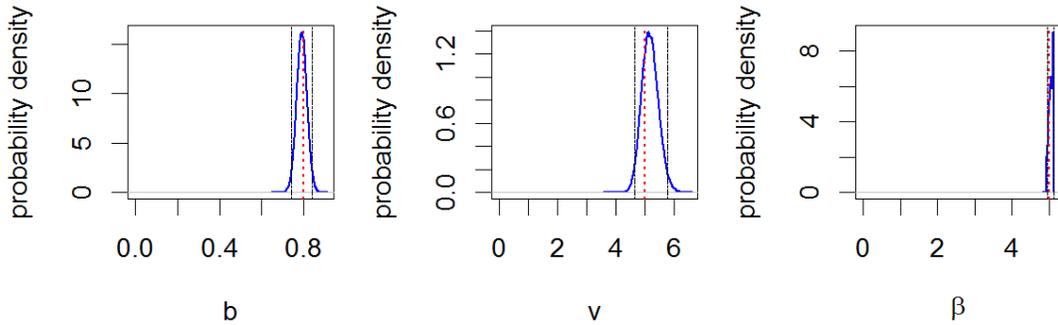
Figure 3.4: Posterior distributions derived when analysing a simulation of migration. The simulated trajectory analysed is given in Figure 3.2. Parameters inferred using the MCMC algorithm include the movement coefficients b and v in the OU process in Equation (3.1) and the resource coefficient β in the RSF in Equation (3.9). Red dotted lines indicate real values used in the simulation and black dot-dashed lines shows 95% quantile intervals.

the trace plots of MCMC chains and the convergence of the chains was decided by applying the burnin function in R package LaplacesDemon (Statisticat and LLC., 2018). This burnin function is easy to use and some other tools available for MCMC convergence diagnostics include R packages coda and boa (Plummer et al., 2006; Smith, 2007).

Figure 3.4 illustrates the posterior distributions of parameters b, v and β obtained by analysing the simulation in Figure 3.2 using the MCMC inference method. The posterior distributions successfully captured the real values of the parameters, indicated by red dotted lines, with 95% central posterior intervals, shown by black dashed lines in Figure 3.4.

I tested the inference method on the three sets of simulations using different parameters, mentioned in Section 3.3.1, to investigate the effect of parameter values on the efficiency of the inference method. Here I fixed the initial values at real values used in simulation and will examine the impact of initial values on the efficiency of the inference procedure later in this section. When applying this model approach to real data, where real parameter values are unknown, a common practice is to carry out the MCMC algorithm for several times with different initial values. Figure 3.5 shows the number of iterations needed for the MCMC sampling chains to converge when applying the inference algorithm on the three sets of simulations. For example, Figure 3.5a gives the iterations before
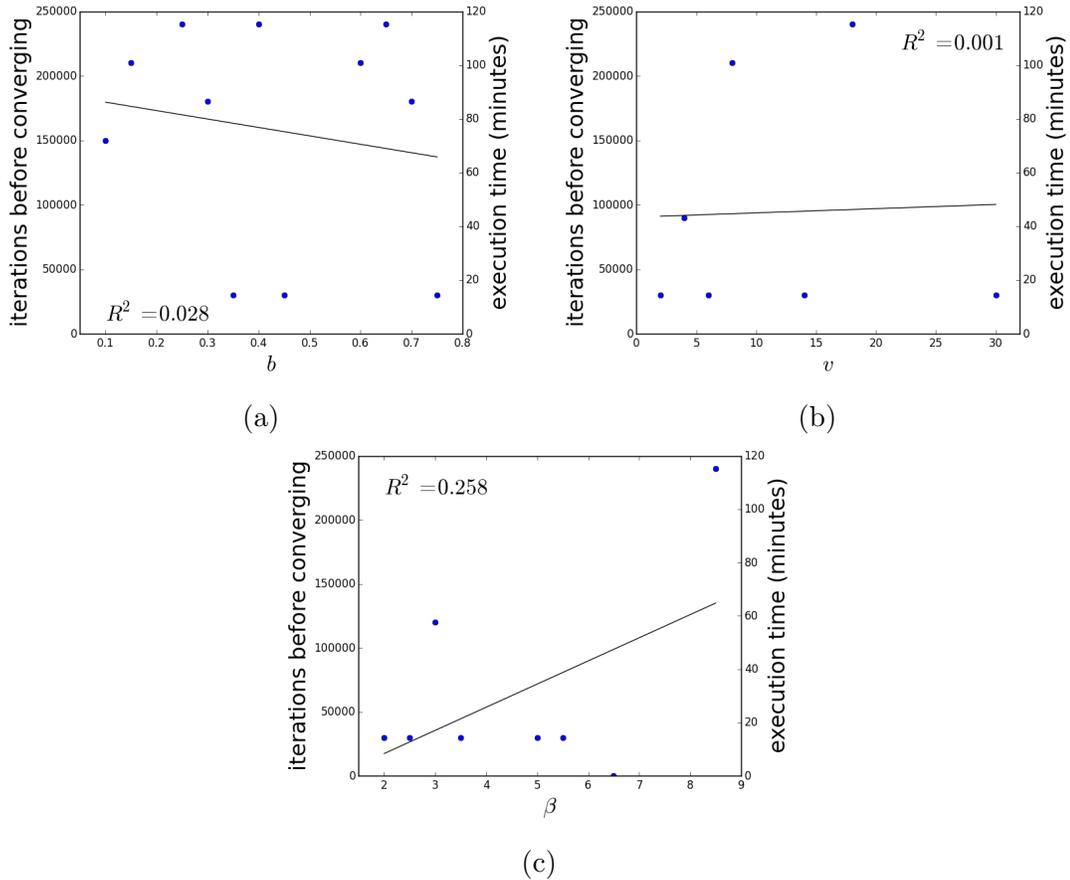
Figure 3.5: The relationship between iterations before converging and parameters in simulations of migration. Coefficients b and v are the drift and diffusion terms in the OU process (Equation 3.1) and β is the resource coefficient in the resource selection function (Equation 3.9). The initial values of the chains for b, v and β are fixed at the real values used in the simulations. The corresponding approximate execution time in minutes is shown in right y-axis. Applying MCMC inference on simulations generated (a) using different b but the same v and β. (b) using different v but the same b and β. (c) using different β but the same b and v.

converging when analysing the first set of simulations, generated using different values for b and fixed v and β. In general, the parameters used had little impact on the number of iterations needed for the chains to converge (Figure 3.5).

With regard to the accuracy of the estimation by the inference, the 95% central intervals of posterior distributions were calculated to show if they contained the real values of parameters inferred. In Figure 3.6, the dots represent the log ratios between sample means (the means of posterior distributions) and real values and the vertical intervals illustrate the corresponding 95% central intervals. The grey dashed lines mark the log ratios equal to 0, standing for the positions of real values in the plots. The blue intervals contain 0, meaning that the original 95% central posterior intervals contain the real values, whereas the red intervals represent cases where 95% central posterior intervals fail to cover the real values. Figures 3.6a,c reveal that the posterior distributions tend to underestimate the drift coefficient b and overestimate the resource coefficient β for larger values of b used in simulations. On the other hand, the real values of v and β had no significant effect on the accuracy of estimation (Figures 3.6d-i).

Overall, 95% central posterior intervals of posterior distributions captured the real values of b for about 2/3 of the simulations but underestimated them for the remaining cases (Figures 3.6a,d,g, Table 3.1). The real values of v were well estimated by 95% central posterior intervals for more than 2/3 of the simulations and were overestimated for all but one other cases (Figures 3.6b,e,h, Table 3.1). The estimation of the resource coefficient β was less accurate, with real values lying in 95% central posterior intervals for only above 2/5 of the cases and others being overestimated (Figures 3.6c,f,i, Table 3.1). Note that in Equations (3.1) and (3.9), b represents the strength of tendency towards the attraction centre, while β describes the extent to which resource qualities influence an animal's selection decisions. As a consequence, both parameters b and β would bring about more direct movement towards the attraction centre. That is, b and β are related to each other. Therefore, the inference algorithm may fail to separate their impacts on movement, resulting in the underestimate of b and overestimate of β at the same time (Figure 3.6).

When applying the MCMC algorithm, a parameter κ is employed to determine
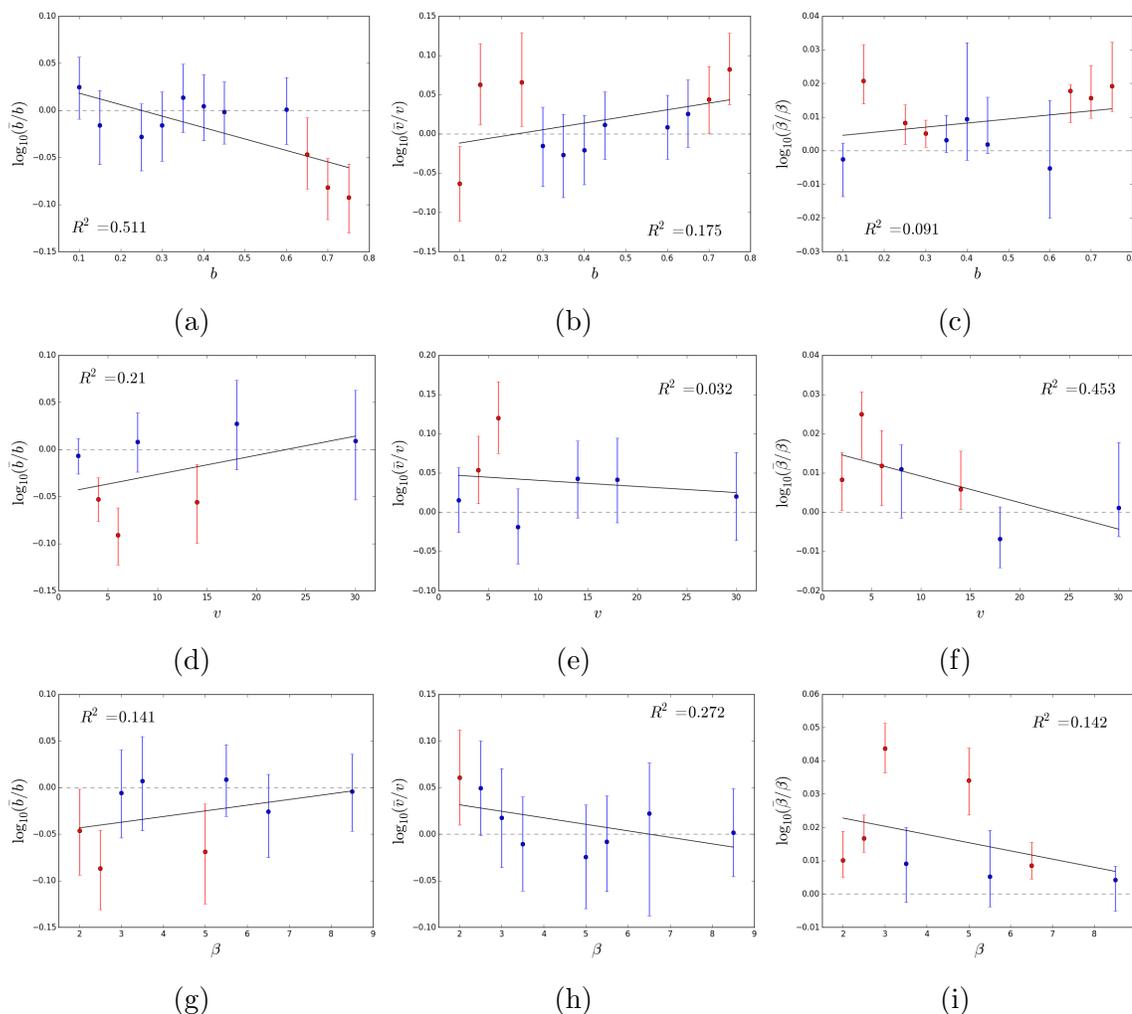
Figure 3.6: The log ratios between sample means $\bar{b}$, $\bar{v}$ and $\bar{\beta}$ and real values of b, v and β with 95% central posterior intervals when applying MCMC inference on simulations of migration. Red intervals do not contain 0. Applying MCMC on simulations generated (a)(b)(c) using different b but the same v and β. (d)(e)(f) using different v but the same b and β. (g)(h)(i) using different β but the same b and v.
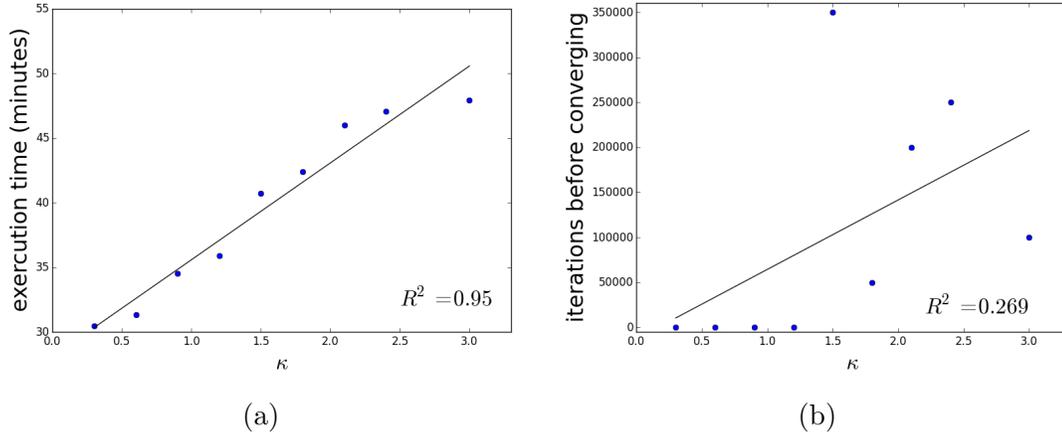
Figure 3.7: The relationship between the efficiency of the MCMC algorithm and κ, the mean of number of proposed switching points during a unit time interval. (a) Approximate execution time for $100,000$ iterations when applying the MCMC algorithm with different κ. (b) Iterations before converging when applying the MCMC algorithm on the same simulation of migration using different κ. The corresponding execution time increases as κ rises (see Figure 3.7a).

the average number of possible switching points inserting between two observations, as described in Section 3.2. Analysing the simulation in Figure 3.2 showed that larger κ led to a larger amount of data considered in the algorithm and therefore longer execution time (Figure 3.7a). Furthermore, as κ grew, the number of iterations needed for the MCMC chains to converge increased (Figure 3.7b). These imply that a limited number of augmented points is allowed to ensure the efficiency of the inference procedure. Nevertheless, Figure 3.8 shows that when analysing the trajectory in Figure 3.2 using different κ, the real values of parameters were contained in the 95% central posterior intervals for each κ. That is, the accuracy of estimation was not affected by the value of κ.

In practice, choosing a value for κ depends on the frequency of data collection and the decision-making process underlying movement. One can select a κ according to some background knowledge about the species studied. If the change of attraction centre does not happen very often, as in the migration model here, and the rate of observation is relatively high, it is unnecessary to use a large κ. It is convenient to start with $\kappa = 1/(\text{time between observations})$, that is, to add a point to a unit time interval on average.

The simulation in Figure 3.2 was parameterised using different initial values for
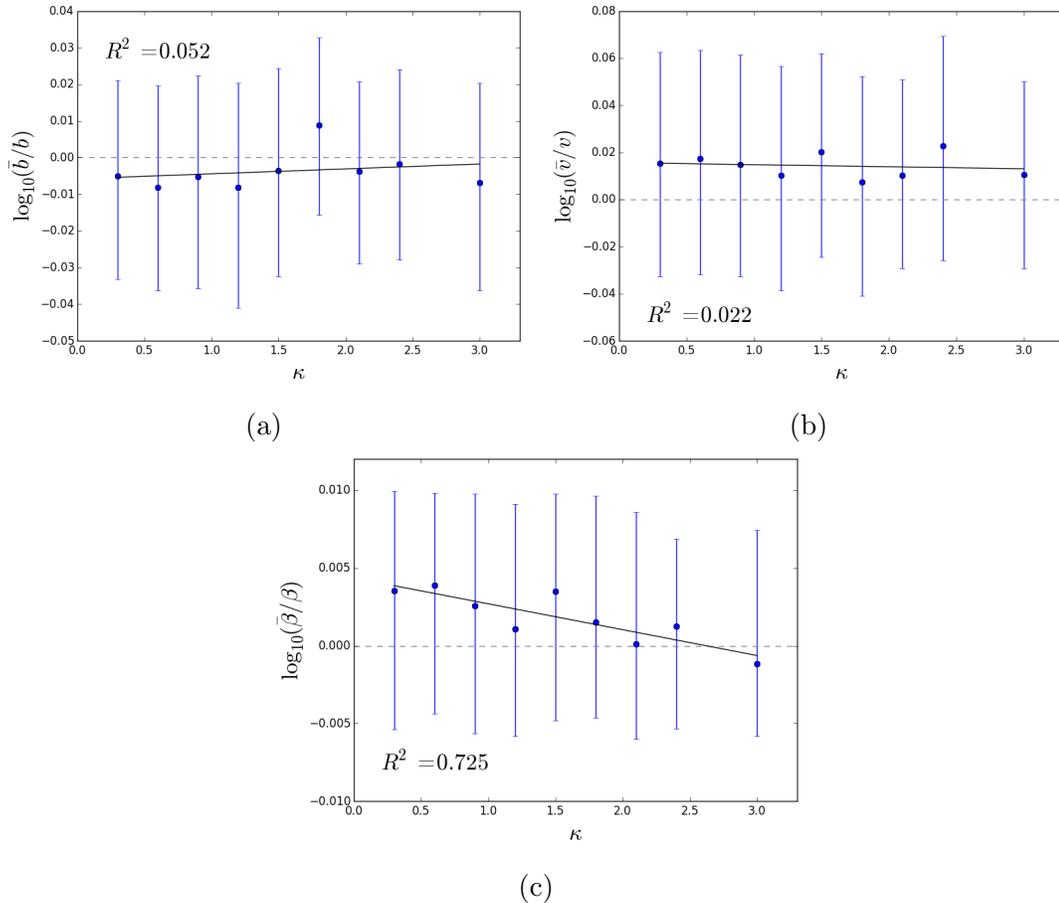
Figure 3.8: The relationship between the accuracy of the MCMC algorithm and κ, the mean of number of proposed switching points during a unit time interval. When using different κ, the log ratio between (a) sample means $\bar{b}$ and the real values of b. (b) sample means $\bar{v}$ and the real values of v. (c) sample means $\bar{\beta}$ and the real values of β.

each parameter to examine the relationship between the initial values and iterations needed before converging. The red stars in Figure 3.9 indicate the situation where the MCMC algorithm was carried out with initial values of every parameter being set at the real values. The sample chains of the drift and diffusion coefficients b and v converged faster when the initial values were small but the speed of convergence was not affected by larger initial b or v (Figures 3.9a-d). Little difference of time for convergence was observed for various initial values of β (Figures 3.9e,f).

When running the inference algorithm on different simulations of the migration
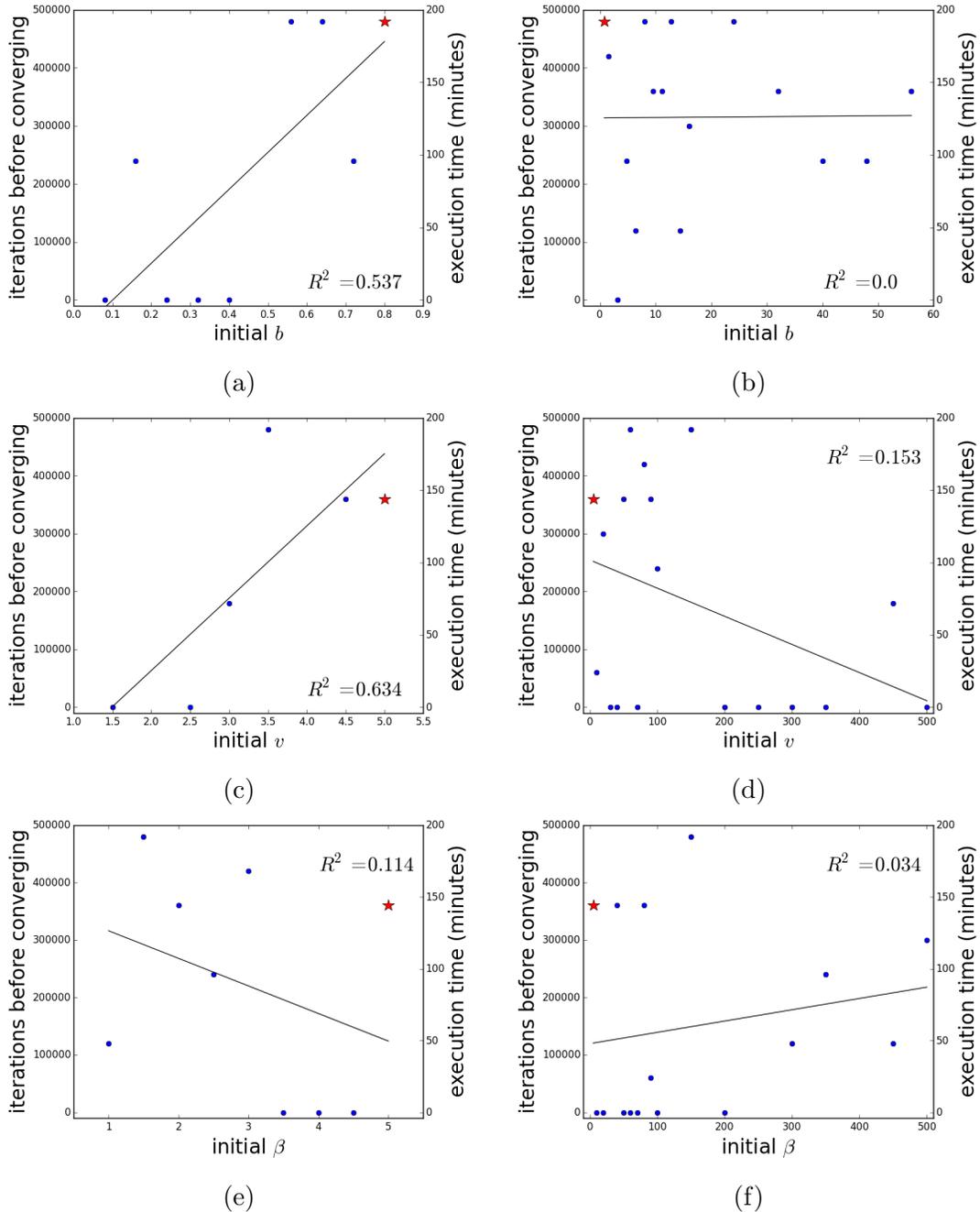
Figure 3.9: The number of iterations before converging when using different initial values in MCMC inference on a simulation of migration. The red stars indicate the case when applying MCMC algorithm with initial values of all parameters being fixed at real values. The corresponding approximate execution time in minutes is shown in right y-axis. (a)(b) Applying MCMC with various initial values of b. The initial values of v and β are fixed at real values. (c)(d) Applying MCMC with various initial values of v. The initial values of b and β are fixed at real values. (e)(f) applying MCMC with various initial values of β. The initial values of b and v are fixed at real values.

model or with different initial values as mentioned earlier in this section, the algorithm did not converge for nearly 10% of the cases. The algorithm may converge eventually when increasing the number of iterations. However, if the data is too sparse, the algorithm may fail to converge due to insufficient information. The structure of the model may also hinder the convergence of the algorithm and it may be helpful to consider an alternative model, for example, a simpler version.

§ 3.4   Resource depletion-renewal models in a patchy landscape

In the migration models described in Section 3.3, the change of resource quality is merely dependent on the time of the year but not animals' foraging behaviours. In this section, I simulate the situation where the fluctuation in resource quality relies on the residential time of an animal in a food patch. To put it in another way, the resource in a patch is consumed when an animal is foraging in the patch and renews when no animal is present there.

3.4.1   Simulations

Based on the modelling framework described in Section 3.1, here I also use a switching OU process (Equation 3.1) to model an animal's movements and the RSF in Equation (3.9) to decide the attraction centre at a given time. In addition, the animal is assumed to be moving in a landscape with N circular food patches, A_i, $i \in \{1, \cdots, N\}$. As with the migration models in Section 3.3, the movement decisions are assumed to be determined by the resource qualities and the distance to the resource units. Unlike the migration model, here I assume that the resource quality in a food patch declines exponentially when the animal is foraging in it and grows logistically otherwise. That is, if the animal is foraging in food patch A_i at time t, then the resource quality in each food patch in time τ is given by (Van Moorter et al., 2009; Ford, 1983)

$$R(A_j, t + \tau) = \begin{cases} R(A_j, t)e^{-d_j \tau} & \text{if } j = i, \\ \dfrac{K_j R(A_j, t)e^{r_j \tau}}{K_j + R(A_j, t)(e^{r_j \tau} - 1)} & \text{for } j \neq i, \end{cases} \quad (3.11)$$
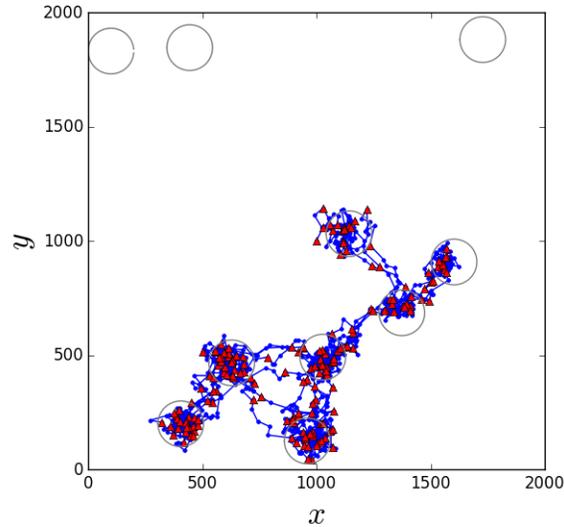
Figure 3.10: A simulated trajectory in a patchy landscape with the resource depletion-renewal model. The resource quality changes according to Equation (3.11) and the movement process is decided by the OU process in Equation (3.1) and RSF in Equation (3.9). The blue dots and line segments show the whole set of data points, and the red triangles are those used in the MCMC algorithm.

where d_j, r_j and K_j are the depletion rate, growth rate and carrying capacity in patch A_j respectively. Hereafter, I assume all patches are equal in depletion and growth rates and carrying capacity, which are then denoted by d, r and K respectively.

I generated simulated trajectories in a 2000×2000 unit^2 landscape with ten non-overlapping circular food patches of radius 100, similar to the spatial scale in Van Moorter et al. (2009). Figure 3.10 shows an example of simulations of movements in such a landscape in response to resource changes defined by Equation (3.11). This landscape layout was chosen only for demonstrating the methodology. Nevertheless, a patchy landscape is widely used for theoretical models (Ford, 1983; Van Moorter et al., 2009; Harris and Blackwell, 2013; Riotte-Lambert et al., 2015) as well as for real situations where it is possible to clearly delineate resource patches (Zweifel-Schielly et al., 2009; Sawyer and Kauffman, 2011; Merkle et al., 2014; Shariatinajafabadi et al., 2014). If the resources are distributed in the environment continuously, then it may be better to partition the landscape using a grid other than non-overlapping patches as Section 3.5 will explain later.
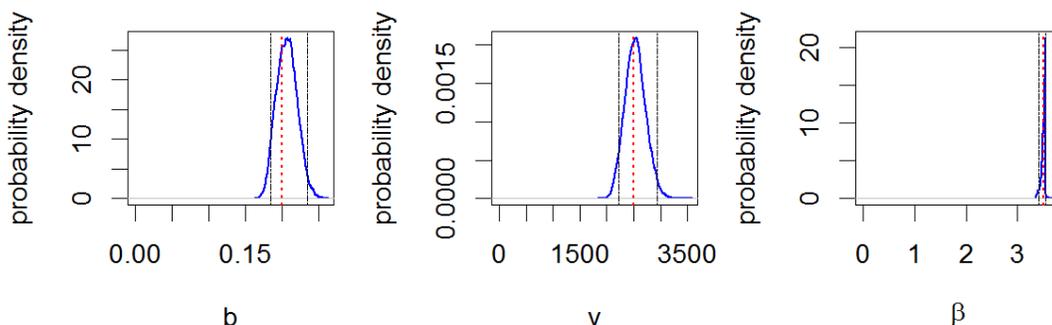
Figure 3.11: Posterior distributions derived when analysing a simulation of movement in a patchy landscape with the resource depletion-renewal model. The simulated trajectory analysed is given in Figure 3.10. Parameters inferred include the movement coefficients b and v in the OU process in Equation (3.1) and the resource coefficient β in the RSF in Equation (3.9). Red dotted lines indicate real values used in simulations and black dashed lines shows 95% central posterior intervals.

3.4.2   INFERENCE FROM SIMULATIONS

Applying the inference method in Section 3.2 on the example shown in Figure 3.10 resulted in posterior distributions capturing the real values with 95% central posterior intervals (Figure 3.11). This shows the ability of the inference procedure to estimate the parameters to a good degree of accuracy in such cases.

The real values of the drift coefficient b in Equation (3.1) and the resource coefficient β in Equation (3.9) had little impact on convergence time (Figures 3.12a,c). Meanwhile, it took longer for the sampling chains to converge when the real value of the diffusion coefficient v in Equation (3.1) used in simulations was larger (Figure 3.12b).

The relationship between the accuracy of inference and the parameters used in simulations is illustrated in Figure 3.13. The inference procedure tended to overestimate b when the real value of b was smaller and overestimate v for larger b (Figures 3.13a,b). The resource coefficient β (Equation 3.9) was often overestimated except when using smaller b to generate simulated trajectories (Figures 3.13c,f,i). Otherwise, the accuracy was generally good.

When inferring from the example given in Figure 3.10 with different initial values for parameters, no apparent trends were observed for various initial values of b

and v or smaller initial values of β (Figures 3.14a-e). On the other hand, the sampling chains converged faster for larger initial values of β (Figures 3.14f).

(a)

(b)   (c)
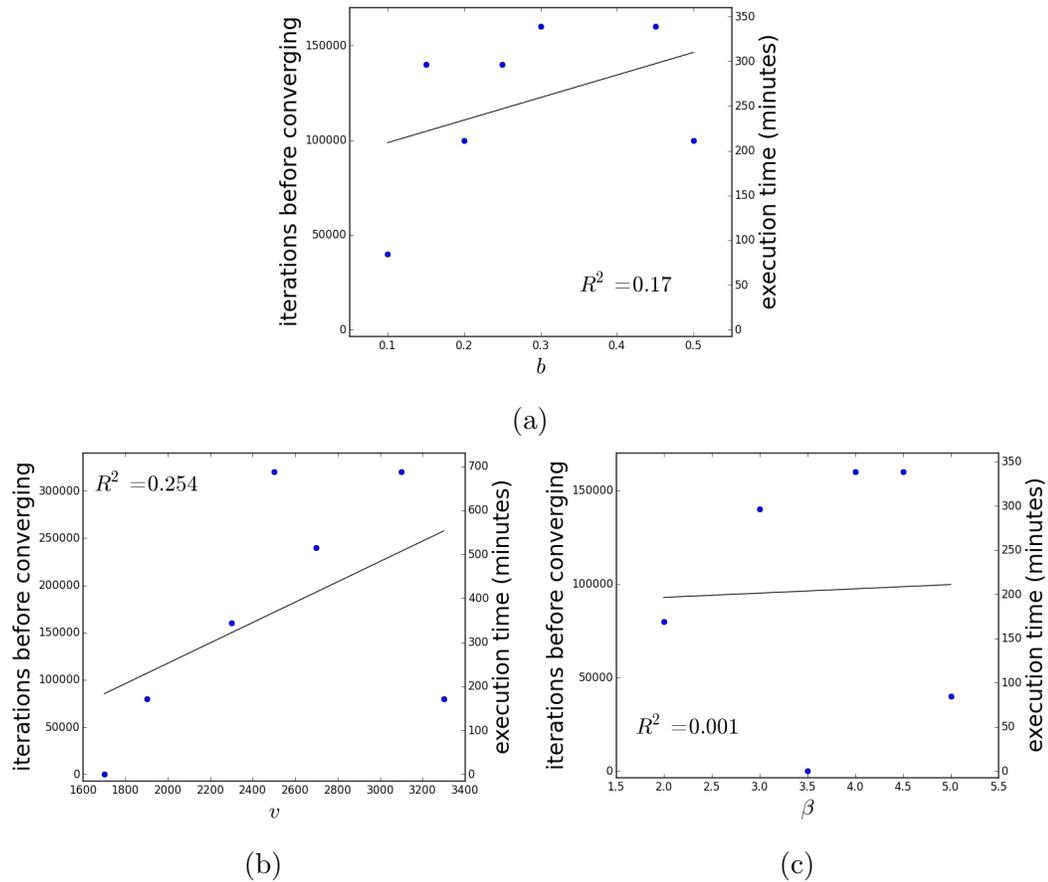
Figure 3.12: The relationship between iterations before converging and parameters in simulations of movements depending on resource depletion or renewal in a patchy landscape. Coefficients b and v are the drift and diffusion terms in the OU process (Equation 3.1) and β is the resource coefficient in the resource selection function (Equation 3.9). The initial values of the chains for b, v and β are fixed at the real values used in the simulations. The corresponding approximate execution time in minutes is shown in right y-axis. (a) Applying MCMC inference on simulations generated using different b but the same v and β. (b) Applying MCMC inference on simulations generated using different v but the same b and β. (c) Applying MCMC inference on simulations generated using different β but the same b and v.

Figure 3.13: The log ratios between sample means $\bar{b}$, $\bar{v}$ and $\bar{\beta}$ and real values of b, v and β with 95% central posterior intervals when applying MCMC inference on simulations of movements depending on resource depletion or renewal in a patchy landscape. Red intervals do not contain 0. (a)(b)(c) Applying MCMC on simulations generated using different b but the same v and β. (d)(e)(f) Applying MCMC on simulations generated using different v but the same b and β. (g)(h)(i) Applying MCMC on simulations generated using different β but the same b and v.
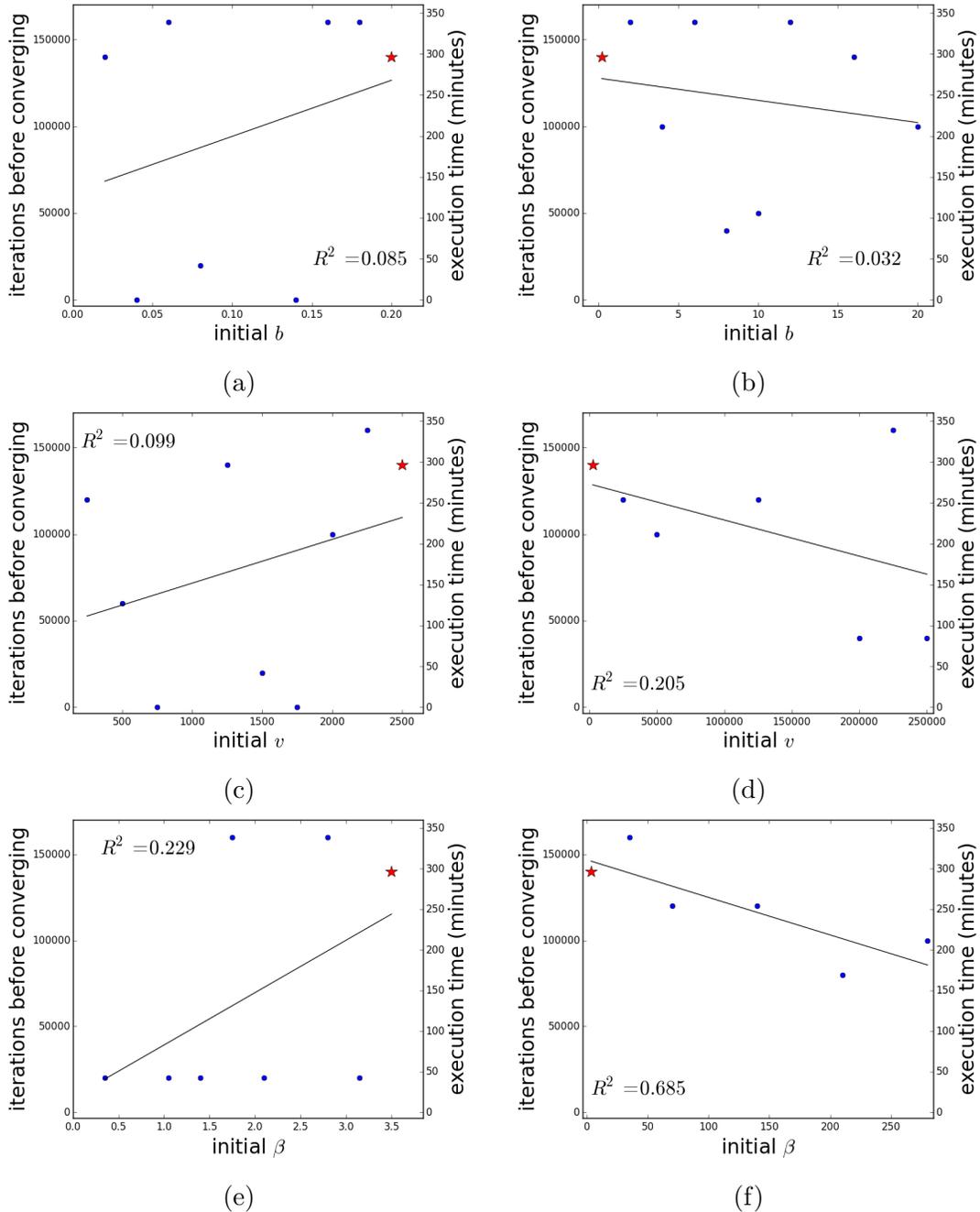
Figure 3.14: The number of iterations before converging when using different initial values in MCMC inference on a simulation of movements dependent on resource depletion or renewal in a patchy landscape. The red stars indicate the case when applying MCMC algorithm with initial values of all parameters being fixed at real values. The corresponding approximate execution time in minutes is shown in right y-axis. (a)(b) Applying MCMC with various initial values of b. The initial values of v and β are fixed at real values. (c)(d) Applying MCMC with various initial values of v. The initial values of b and β are fixed at real values. (e)(f) applying MCMC with various initial values of β. The initial values of b and v are fixed at real values.
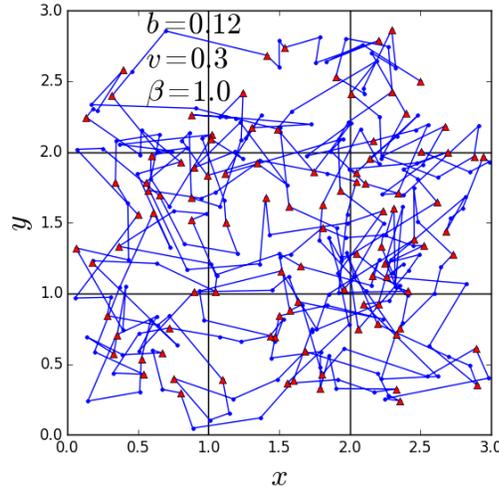
Figure 3.15: A simulated trajectory in a raster landscape with the resource depletion-renewal model. The resource quality changes according to Equation (3.11) and the movement process is decided by the OU process in Equation (3.1) and RSF in Equation (3.9). The blue dots and line segments show the whole set of data points, and the red triangles are those used in the MCMC algorithm.

§ 3.5 Resource depletion-renewal models in a raster landscape

Landscapes in Sections 3.3 and 3.4 contain non-overlapping resource patches and these might represent some real situations well (Sawyer and Kauffman, 2011; Merkle et al., 2014). However, for some other situations, resource patches are not clearly disjoint as resource qualities change continuously over the landscape and it is better to use a rasterised grid for this case, e.g. Potts, Bastille-Rousseau, Murray, Schaefer and Lewis (2014).

3.5.1 SIMULATIONS

For simplicity, I assumed a landscape of a 3×3 grid consisting of unit squares where the depletion rate, growth rate and carrying capacity of resources are equal across the land. At a given time, the animal was assumed to move towards a central point, which is the centre of a cell. That is, every cell is regarded as a foraging patch and the centre of a cell is a potential attraction centre of movement. I generated simulated trajectories following Equation (3.1) with resource qualities

controlled by Equation (3.11) in a raster landscape with $0.05 \leq b \leq 0.14$, $0.05 \leq v \leq 0.5$ and $0.2 \leq \beta \leq 2$. These ranges for parameter values were chosen to ensure that the distance moved between two consecutive simulated locations was small relative to the spatial scale of the landscape. Therefore, the animal would stay in a cell for a while before moving to another cell. If a simulated location falls outside the landscape, then the point is discarded and replaced by a new position. This is to ensure the animal always stays inside the given landscape. Figure 3.15 gives such a simulation, where every 3rd data point, marked by red triangles in the plot, were used to parameterise the model.

3.5.2 INFERENCE FROM SIMULATIONS

The inference from the example shown in Figure 3.15 was successful as the 95% central posterior intervals of posterior distributions contain the real parameter values (Figure 3.16).
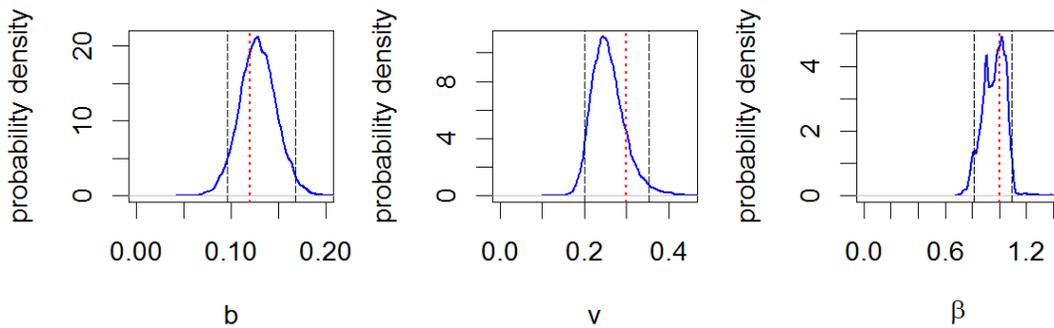
Figure 3.16: Posterior distributions derived when analysing a simulation of movement in a raster landscape with the resource depletion-renewal model. The simulated trajectory analysed is given in Figure 3.15. Parameters inferred include the movement coefficients b and v in the OU process in Equation (3.1) and the resource coefficient β in the RSF in Equation (3.9). Red dotted lines indicate real values used in simulations and black dashed lines shows 95% central posterior intervals.

Applying the inference procedure on simulations generated using different parameter values showed that the number of iterations needed for the chains to converge was not affected by the parameter values used (Figure 3.17).

As the identification of attraction centres in a raster landscape is less obvious, the accuracy of the inference results might be more sensitive to the density of
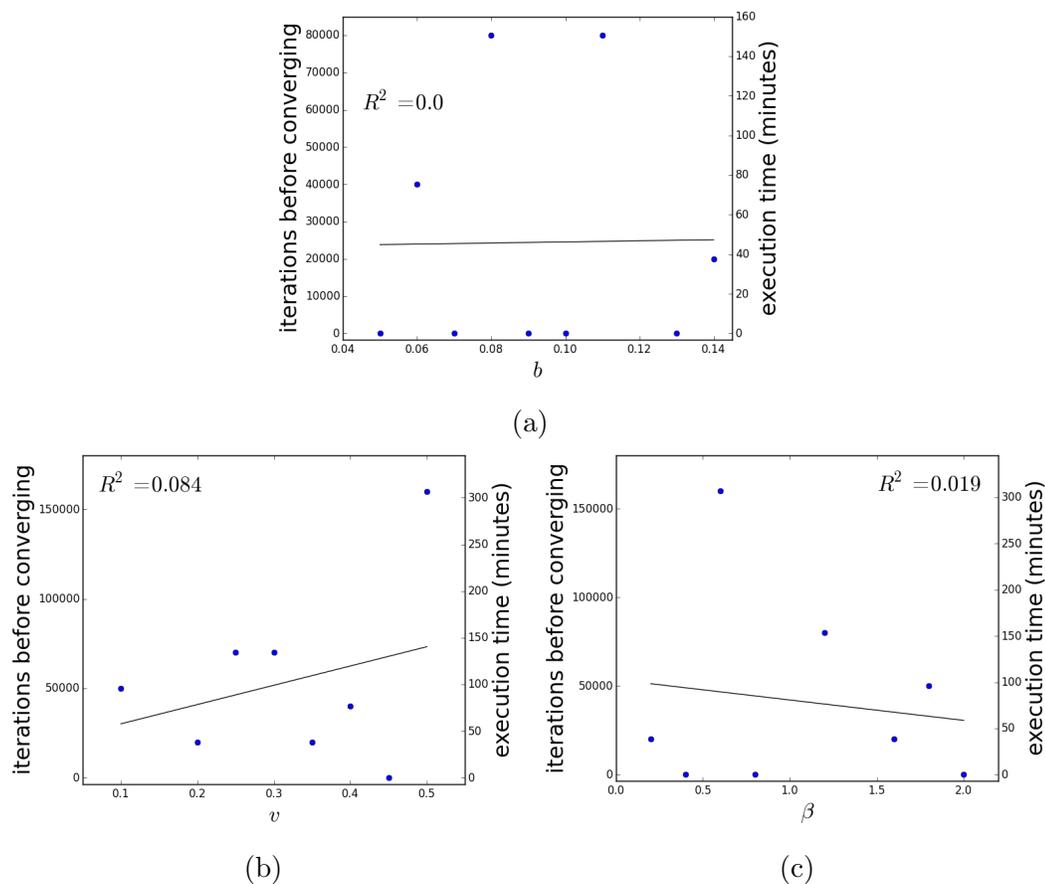
Figure 3.17: The relationship between iterations before converging and parameters in simulations of movements depending on resource depletion or renewal in a raster landscape. Coefficients b and v are the drift and diffusion terms in the OU process (Equation 3.1) and β is the resource coefficient in the resource selection function (Equation 3.9). The initial values of the chains for b, v and β are fixed at the real values used in the simulations. The corresponding approximate execution time in minutes is shown in right y-axis. Applying MCMC inference on simulations generated (a) using different b but the same v and β. (b) using different v but the same b and β. (c) using different β but the same b and v.

data points. To investigate how the frequency of data collection influences the accuracy of inference, I tested the inference procedure on the same simulations using different subsets of data points. The log ratios between the real values of b, v and β and posterior means, $\bar{b}$, $\bar{v}$ and $\bar{\beta}$, are illustrated in Figures 3.18-3.20. The results of inference using every 3rd and every 5th points were similarly good when estimating the parameters using 95% central posterior intervals. However, the difference between the posterior means of β and the real values was reduced for most simulations when more data points were included in the inference procedure (Figures 3.18e,f, 3.19e,f, 3.20e,f).

I tested the inference method on the simulated trajectory shown in Figure 3.15 by varying the initial value of one of the parameters b, v and β and fixing the initial values of the other two at real values. Figure 3.21 shows that in general, carrying out the MCMC algorithm with initial values nearer to real values would require fewer iterations before the sampling chains converged.

Model	parameter	real values lie within 95% CPI (%)	real values lie within 99% CPI (%)
	b	65.38	73.07
Migration	v	69.23	88.46
	β	42.30	50
	b	90	90
Resource depletion/renewal - patch	v	90	90
	β	50	60
	b	100	100
Resource depletion/renewal - raster	v	83.33	100
	β	78.57	90.48
(Overall average)		76.52	85.23

Table 3.1: The overall performance of the inference method when analysing simulations in Chapter 3. The figures represent the percentage of cases where central posterior intervals (CPIs) of posterior distributions contain the real values of parameters used in simulations.
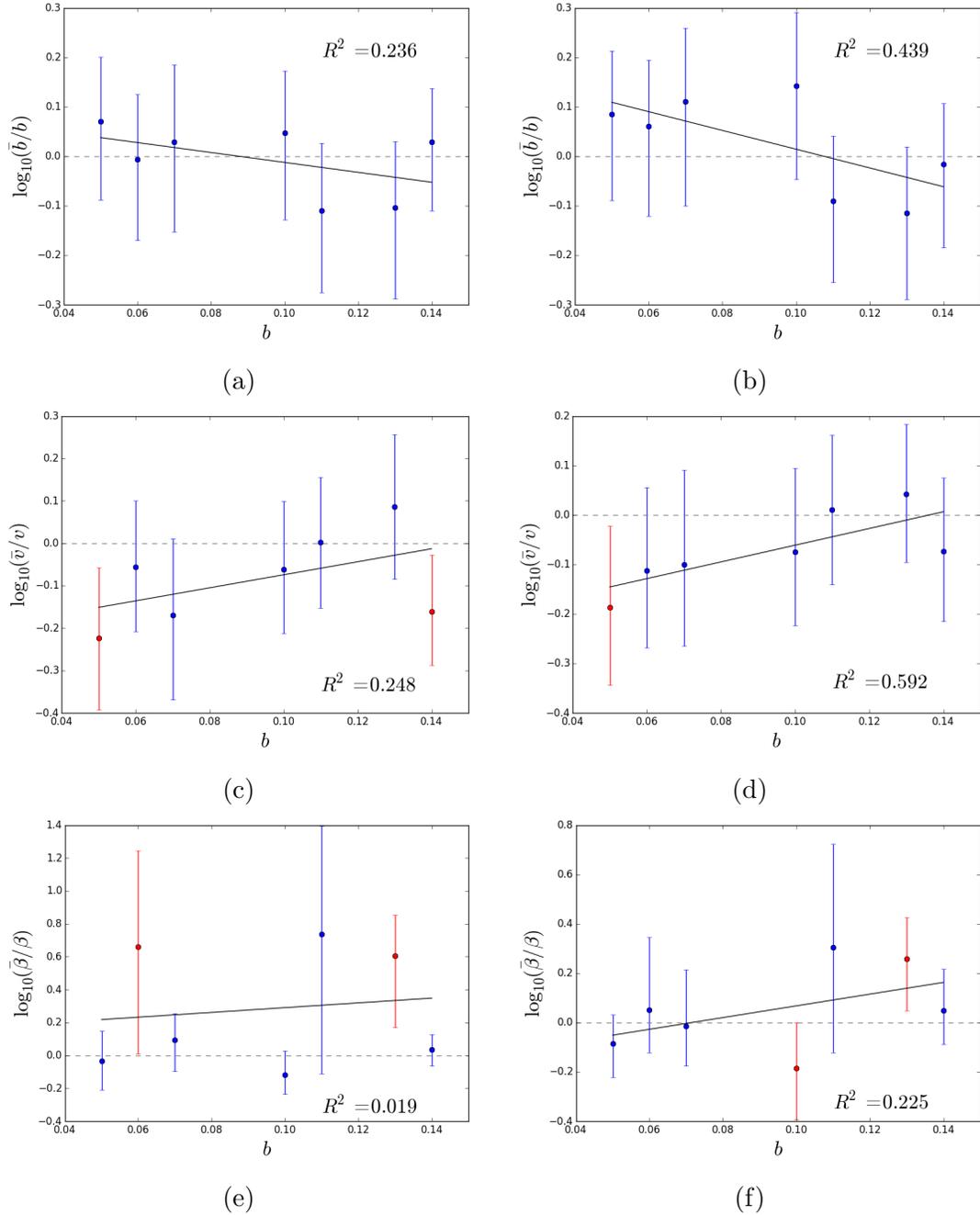
Figure 3.18: The log ratios between sample means $\bar{b}$, $\bar{v}$ and $\bar{\beta}$ and real values of b, v and β when applying MCMC inference on simulations of resource depletion-renewal models in a raster landscape. The simulations are generated using different b but the same v and β. (a)(c)(e) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every 5th data points. (b)(d)(f) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every 3rd data points.
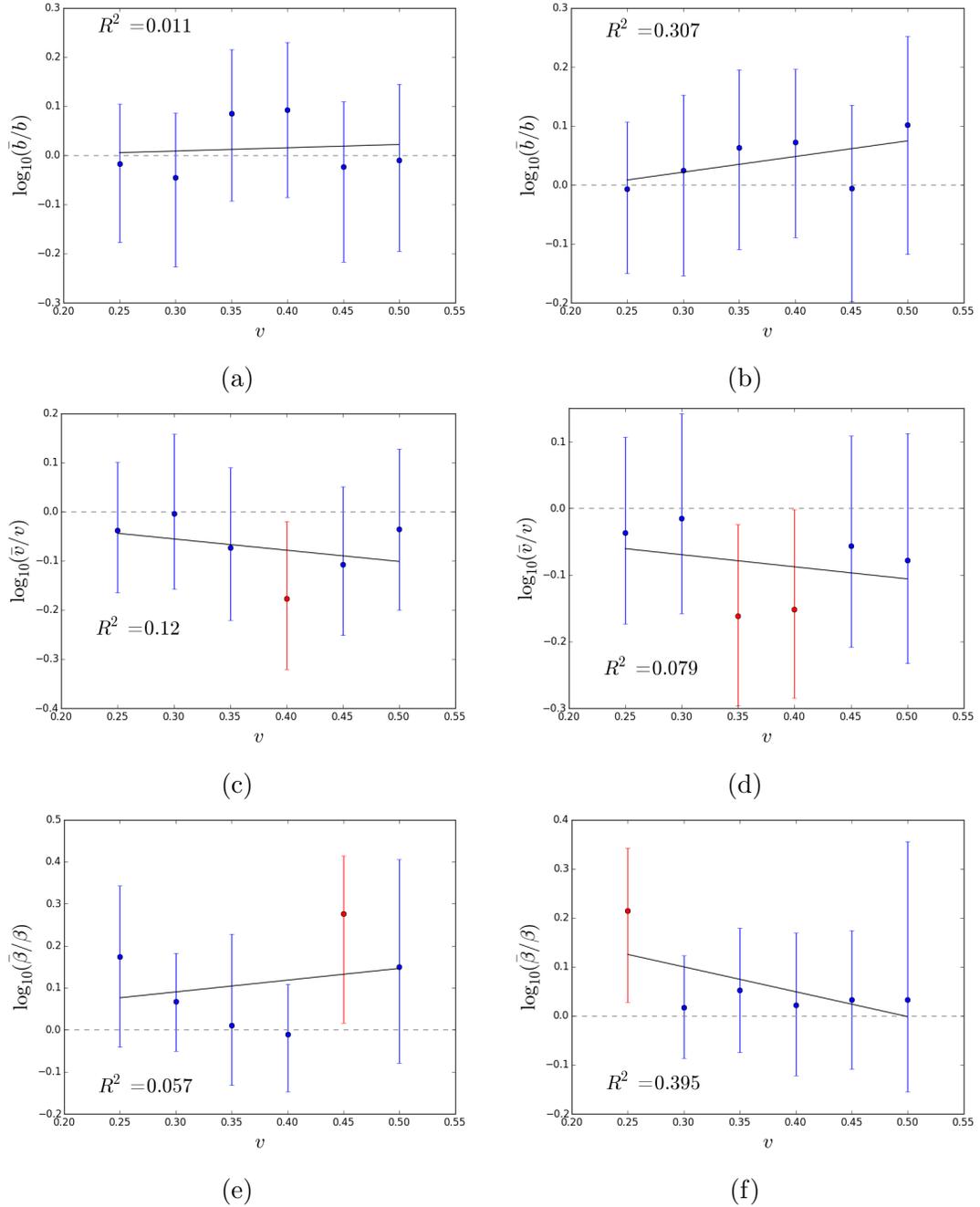
Figure 3.19: The log ratios between sample means $\bar{b}$, $\bar{v}$ and $\bar{\beta}$ and real values of b, v and β when applying MCMC inference on simulations of resource depletion-renewal models in a raster landscape. The simulations are generated using different v but the same b and β. (a)(c)(e) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every 5th data points. (b)(d)(f) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every 3rd data points.
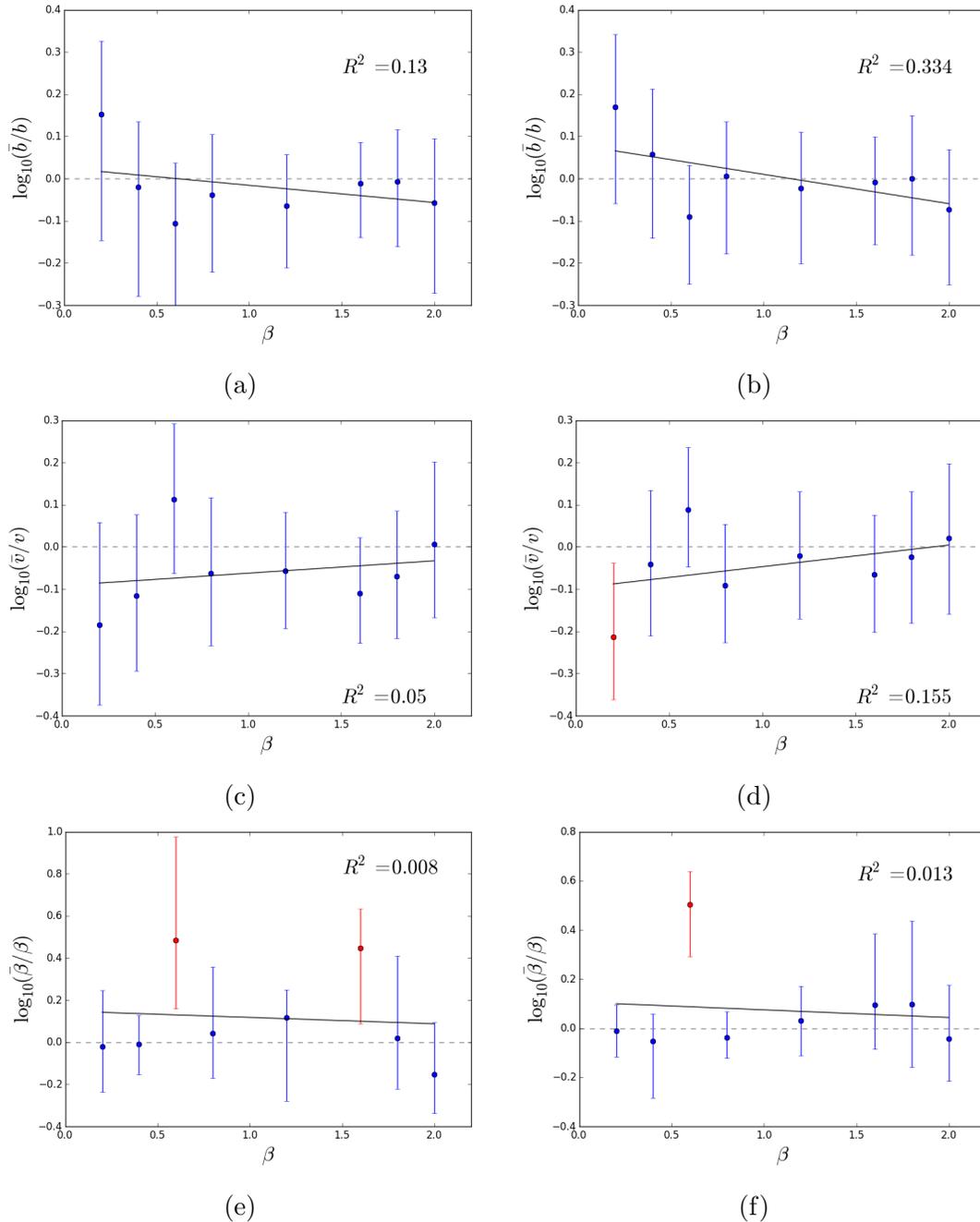
Figure 3.20: The log ratios between sample means $\bar{b}$, $\bar{v}$ and $\bar{\beta}$ and real values of b, v and β when applying MCMC inference on simulations of resource depletion-renewal models in a raster landscape. The simulations are generated using different β but the same b and v. (a)(c)(e) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every 5th data points. (b)(d)(f) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every 3rd data points.
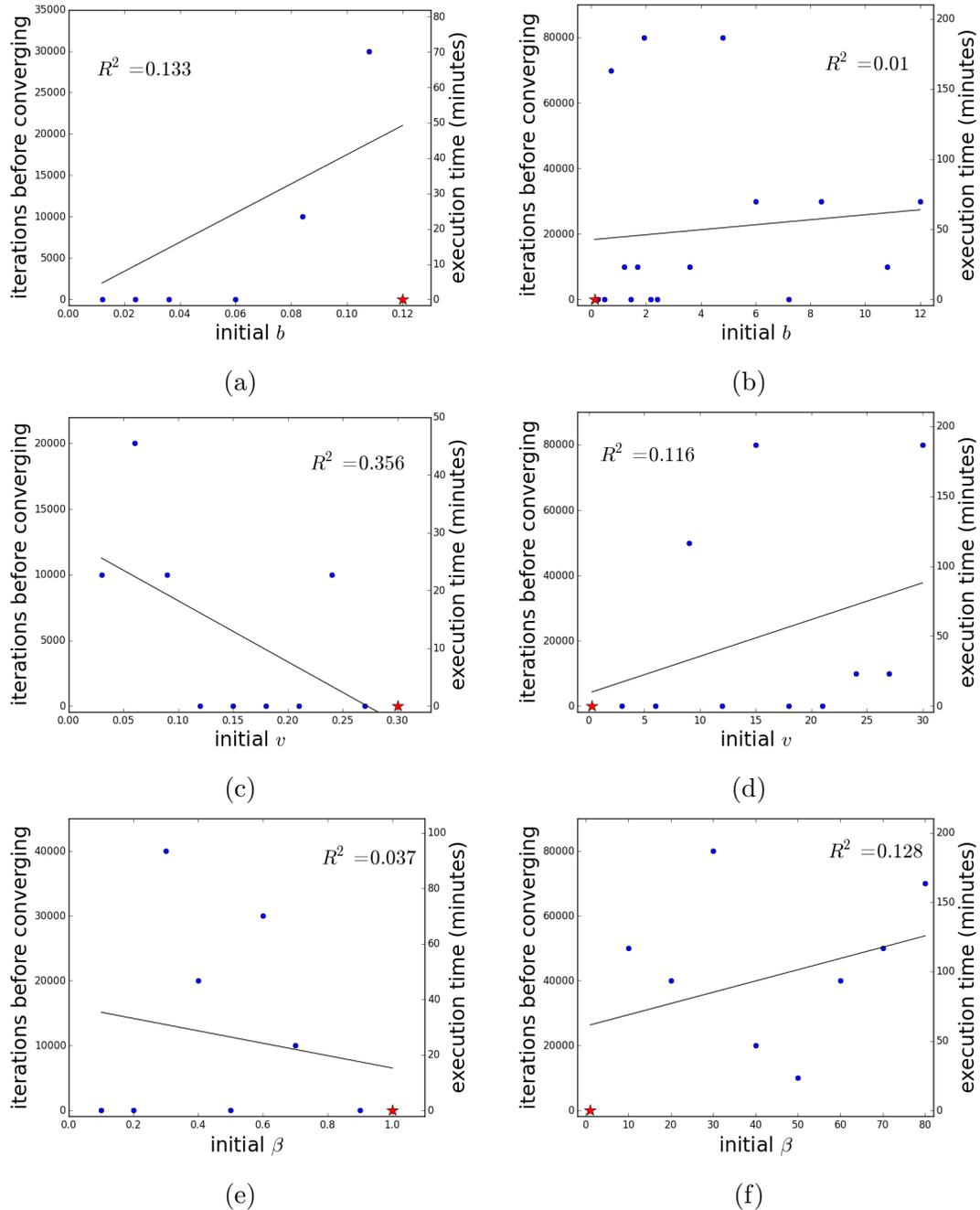
Figure 3.21: The number of iterations before converging when using different initial values in MCMC inference on a simulation of movements dependent on resource depletion or renewal in a raster landscape. The red stars indicate the case when applying MCMC algorithm with initial values of all parameters being fixed at real values. The corresponding approximate execution time in minutes is shown in right y-axis. (a)(b) Applying MCMC with various initial values of b. The initial values of v and β are fixed at real values. (c)(d) Applying MCMC with various initial values of v. The initial values of b and β are fixed at real values. (e)(f) applying MCMC with various initial values of β. The initial values of b and v are fixed at real values.

§ 3.6   Discussion

In this chapter, I have incorporated a resource weighting function in a continuous-time movement model to infer the selection of resources by animals from movement data. The modelling framework relies on the assumption that animals have complete knowledge of the whole landscape, so they would select the most attractive place and move towards it, Based on this modelling framework, I have considered three types of models, namely a migration model, a resource depletion-renewal model in patchy landscapes and a resource depletion-renewal model in raster landscapes. The migration model differs from the resource depletion-renewal models by the way in which resource qualities change. In the migration model, the change of resource qualities is independent of the foraging activities of animals, while in the other two models, the foraging time in a patch explains the decrease of resources.

In general, the parameter values used in simulations or initial values for the MCMC algorithm did not have much impact on the efficiency of the inference procedure. Meanwhile, the efficiency of the algorithm might be mainly affected by the parameter, κ, used to determine the average number of points added to a unit time interval in the augmentation of data. If κ is larger, there are more simulated points between observations to be included and therefore, it may take longer for the procedure of updating the trajectory to converge.

The accuracy of estimating the drift and diffusion coefficients b and v (Equation 3.1) was generally good. However, the resource coefficient β (Equation 3.9) was more likely to be overestimated. For the resource depletion-renewal model in a raster landscape, the accuracy of parameterising the model depends more heavily on the time resolution of the data.

In the models considered in Sections 3.3-3.5, I assume that the selection of resources is determined by the trade-off between resource qualities and the distance to a target place. These two factors are the two covariates in the resource weighting function in Equation 3.9. In Equation 3.9, the coefficient of resource quality, β, is assumed to be positive, while the coefficient of the distance to an attraction centre is fixed at -1. For a general situation, it may be more appropriate to

relax the assumption that a coefficient is positive or negative because the effect of a covariate on movement may be unclear. For example, some animals may avoid roads while some others may tend to use them. Nevertheless, it is still necessary to fix one of the coefficients such as the distance in the models here to enable the inference of other coefficients as mentioned in Section 3.3. Otherwise, the algorithm will not converge because multiplying all coefficients by a same number will lead to the same selection results.

Employing a similar strategy as SSA and iSSA, the modelling framework introduced in this chapter also uses movement as a constraint on the availability of resources. However, unlike SSA or iSSA, where resource selection is restricted to the selection from resources available in one step, my method attempt to identify the place of actual interest beyond the scale of steps. In Merkle et al. (2014), SSA has developed into a modelling framework for patch selection by comparing a used patch to other potential target patches instead of comparing an observed step to available steps. That is, the approach proposed by Merkle et al. (2014) focuses on animals' interest in foraging patches and omits detailed steps between patches. Therefore, the method in Merkle et al. (2014) might provide a better tool for understanding the preference of animals for important habitats. Although both modelling frameworks in Merkle et al. (2014) and in this chapter consider the selection of patches, these two approaches are based on different assumptions. Merkle et al. (2014) implicitly assumes that the arrival of animals in a patch means the patch is preferable. However, this assumption may not necessarily be the case since an animal may visit a patch simply in order to reach another patch, which is its actual destination. On the other hand, rather than interpreting presence as preference, my approach assumes animals are attracted to the best patch even before reaching it and hence may be more capable of uncovering the actual preferences of animals for resources.

In addition, the modelling framework introduced here considers conditions in both the patch an animal is leaving and potential target places as some other methods do (e.g. Avgar et al. (2017)). Comparing resources in the source and target patches should better reveal how animals select a patch according to resources. Therefore, the modelling framework in this chapter would have an advantage over

Breed et al. (2017), which also applies the concept of a switching OU model but only takes resources in the source patch into account.

Chapter 4

Analysis of movements following a resource gradient

The switching OU process model introduced in Chapter 3 assumes animals have complete knowledge of the environment when making movement decisions. Rather than making this assumption, in this chapter, I consider the other extreme, where animals only take account of environmental conditions in their immediate vicinity. Animals are assumed to follow the resource gradient in the neighbouring area predominantly by perception instead of memory (cf. Bracis and Mueller (2017)). That is, movement decisions are made solely according to the information perceived in the surrounding area and not depending on previous experiences.

§ 4.1 Modelling framework

Similar to Chapter 3, movements following local resource gradient can be described by a biased random walk. However, unlike in Chapter 3, where the direction points to a fixed centre in space, in this chapter, the direction is determined by local conditions. A continuous-time model of this type is a process $X(t)$ given by the solution to a stochastic differential equation where the drift term is constant as follows (Preisler et al., 2004):

$$\mathrm{d}X(t) = \boldsymbol{\alpha}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}W(t). \tag{4.1}$$

Here, $\boldsymbol{\alpha}$ is the drift term, an n-dimensional constant vector representing the
direction of drift, $\boldsymbol{\sigma}$ is the diffusion term, an $n \times n$ matrix, and $W(t)$ is Brownian
motion in an n-dimensional space. In a two-dimensional space, the conditional
distribution of this process is given by (Platen and Bruti-Liberati, 2010)

$$\mathbf{x}(t + \tau)|\mathbf{x}(t) \sim MVN(\mathbf{x}(t) + \alpha\boldsymbol{\rho}(t)\tau, \boldsymbol{\Sigma}\tau), \qquad (4.2)$$

where $\mathbf{x}(t)$ is the animal's location at time t and τ is a short period of time.
MVN stands for "multi-variate normal" and a two-dimensional normal distribu-
tion is considered here. The drift term here is split into two elements, namely the
average speed, α, which is a constant, and direction, $\boldsymbol{\rho}$, which is usually a 2D
vector of length 1. The vector $\boldsymbol{\rho}$ becomes a zero vector If the resource quality is
even in the surrounding area of the animal's position, meaning there is no drift
in the movement. The covariance matrix is given by $\boldsymbol{\Sigma} = \sigma^2 I$, where σ is a con-
stant and I is the 2×2 identity matrix. This is similar to the movement kernels
describing central-place foraging behaviours in Chapter 2, but here the time τ is
a continuous variable rather than a fixed constant. Therefore, it is straightfor-
ward to fit this model to data irregular in time. In theory, animals in this model
can move to any position, but in practice, I will only accept trajectories entirely
restricted inside the landscape in simulations, described in more detail below in
Section 4.3.

Now I assume an animal moves following the process given in Equation (4.2) in a
rasterised landscape. It can be subdivided into a square or rectangular lattice, but
only square grids will be used for simulations in this chapter. I assume the animal
decides its direction, $\boldsymbol{\rho}(t)$ in Equation (4.2), by assessing the values of a resource
weighting function in the four adjoining squares (east, west, north, and south)
next to the one where it is present. Considering these four cells is sufficient when
the animal is not close to the corners of a cell. Therefore, I will not take account
of the four cells diagonally adjacent to the animal for simplicity, meaning that the
influence of diagonally adjacent cells on movement is neglected even if the animal
is located near a corner. For a more realistic model, one may incorporate the
diagonally adjacent cells or use a landscape composed of hexagonal cells. Here,
the direction is assumed to be up the resource gradient, which is given by the
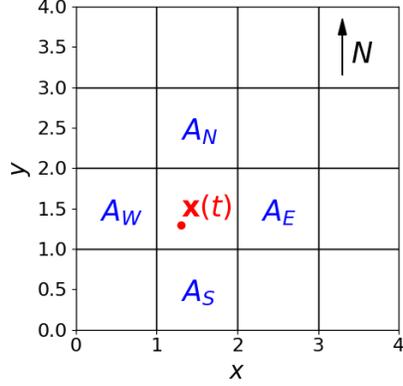
Figure 4.1: Neighbouring patches used to determine resource gradient in a rasterised landscape. The animal is located at $\mathbf{x}(t)$ at time t. Patches A_N, A_W, A_E and A_S are the adjacent squares to the patch where the animal is located. N, W, E, S stand for the north, west, east and south respectively. When calculating the resource gradient in the nearby area of $\mathbf{x}(t)$, the resource qualities in the four adjacent patches are considered, which means the animal only assesses resource qualities in neighbouring areas to determine its moving direction.

ratio of the difference between the values of the resource weighting function in these neighbouring squares as follows (cf. Preisler et al. (2013)):

$$\nabla w(\mathbf{x}, t) := \left(\frac{w(A_E, t) - w(A_W, t)}{\Delta x}, \frac{w(A_N, t) - w(A_S, t)}{\Delta y} \right), \qquad (4.3)$$

where $\mathbf{x}$ is the animal's location at time t; $w(A_E, t)$, $w(A_W, t)$, $w(A_N, t)$ and $w(A_S, t)$ are the resource selection weightings for the adjoining squares A_E, A_W, A_N and A_S in the east, west, north, and south respectively at time t (Figure 4.1). The notations Δx and Δy represent the distance between the centres of squares A_E and A_W and the distance between the centres of patches A_N and A_S respectively. Subsequently, the resource gradient defined by Equation (4.3) is normalised to give the unit vector $\boldsymbol{\rho}(t)$ in Equation (4.2):

$$\boldsymbol{\rho}(t) = \frac{\nabla w(\mathbf{x}, t)}{|\nabla w(\mathbf{x}, t)|}. \qquad (4.4)$$

§ 4.2 Inference by Markov chain Monte Carlo

As in Chapter 3, I use a Markov chain Monte Carlo (MCMC) algorithm, based
on Blackwell et al. (2016), to parameterise the model introduced in Section 4.1.
The algorithm contains two fundamental parts, namely the update of trajectory
and the update of parameters. For details, see Section 3.2 and Blackwell et al.
(2016).

The first part of the algorithm, the update of trajectory, is accomplished by
adding points where the animal might have changed its direction. The change
of direction may happen at any instance, though I assume the animal does not
reassess its direction continuously. The reassessment of direction is given by
Equations (4.3) and (4.4) and assumed to be a Poisson process in time, as in
Blackwell et al. (2016).

In every iteration, a proposed trajectory is generated by selecting a subset of
observed data and adding proposed switching points to this subset. At each point,
Equations (4.3) and (4.4) are computed to determine the direction of movement,
$\boldsymbol{\rho}(t)$, in Equation (4.2). Given the model in Equation (4.2), I calculate the
Hastings ratio by comparing the likelihood of this proposal trajectory to the
likelihood of the existing trajectory within the selected time frame. The proposed
switching points are accepted with a probability of

$$\text{acceptance rate} = \min\{1, \text{Hastings ratio}\}, \tag{4.5}$$

where the Hastings ratio is given by

$$\prod_{i=k}^{l-1} \frac{f(\mathbf{x}(t_{i+1})|\mathbf{x}(t'_{i,M'_i}), \boldsymbol{\mu}(t'_{i,M'_i}))}{f(\mathbf{x}(t_{i+1})|\mathbf{x}(t_{i,M_i}), \boldsymbol{\mu}(t_{i,M_i}))} \tag{4.6}$$

with $\mathbf{x}(t'_{i,M'_i})$ and $\mathbf{x}(t_{i,M_i})$ the proposed and existing predecessor of the observed
point $\mathbf{x}(t_{i+1})$ respectively and $i = \{k, k+1, \cdots, l\}$ the indexes of the selected
interval.

Secondly, to update the parameters α and σ in Equation (4.2), a set of candidate
parameters is generated. After generating the candidate parameters, I use them
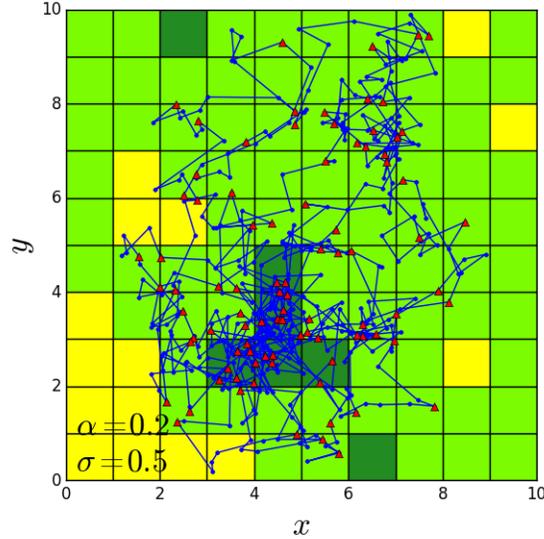
Figure 4.2: A simulation of movement following a resource gradient, according to Equations (4.2-4.4). The different colours in the landscape represent different resource qualities. Dark green, light green and yellow stand for high, medium and low resource quality respectively. The blue dots and line segments show the whole set of data points, and the red triangles are those used in the MCMC algorithm.

and the existing parameters to calculate the likelihood of the accepted trajectory to give the Hastings ratio for updating parameters. Once the Hastings ratio is obtained and if the candidate parameters are accepted, the existing parameters are replaced with the candidate parameters.

§ 4.3   Simulations

To start with a simple situation, I assumed that the weighting function in Equation (4.3) was the resource quality, that is, $w(A, t) = R(A, t)$, the resource quality of a cell A in a rasterised landscape at time t. To simplify this even further, I assumed the resource quality was static, so that $R(A, t) = R(A)$. I constructed simulated landscapes, each of which consists of a 10×10 grid, composed of 3 different types of cells, featuring different resource qualities. Every such simulated landscape was simplified from a Gaussian random field generated using R (R Core Team, 2017) by classifying the origin values for cells into three levels. The grid in Figure 4.2 shows an example of such a landscape, where yellow, light

green and dark green cells possess low, medium and high resource qualities, fixed
at 5, 10, and 15 respectively. These three different positive values were chosen
arbitrarily just to show different levels of resource qualities. The resource quality
outside the landscape was assumed to be 0, also selected arbitrarily as long as it is
smaller than the resource qualities inside the landscape. Therefore, the direction
of the resource gradient along the boundaries always pointed inside the landscape
and animals located near the boundaries would tend to move inwards.

In each simulated landscape, I generated a simulated trajectory by Equation
4.2. I assumed that animals never cross the boundaries of the landscape. As
in Section 3.5.1, I rejected a simulated position if it was outside the boundaries
and generated another position until a point inside the boundaries was obtained.
This assumption about movement near the boundaries should only have minor
impact on the model results. This is because the probability of remaining inside
the landscape is always larger than crossing the boundaries according to the given
resource qualities, described in the last paragraph. In this way, two series of paths
in different landscapes were generated. For the first series, the drift coefficient, α,
in Equation (4.2) were varied between 0.1 and 1, while the diffusion coefficient,
σ, was fixed at 0.5. The second series of simulations used a fixed $\alpha = 0.2$ and
different values between 0.1 and 1 for σ. Figure 4.2 gives a simulated trajectory
of movement following a resource gradient in a raster landscape.

§4.4   Inference from simulated data

Figure 4.3 illustrates the posterior distributions derived from applying the infer-
ence algorithm on the simulation shown in Figure 4.2. The real values of the
drift coefficient, α, and diffusion coefficient, σ, in Equation (4.2) are indicated
by the red dotted lines in the figure. The real values were well estimated by
the posterior distributions, as they lie within the 95% central posterior intervals,
which are shown by black dashed lines.

Figures 4.4 and 4.5 show the results of applying the inference procedure on the
data shown in Figure 4.2 with different κ, which controls the average number of
proposed switching points over a unit time interval. Larger κ tended to increase
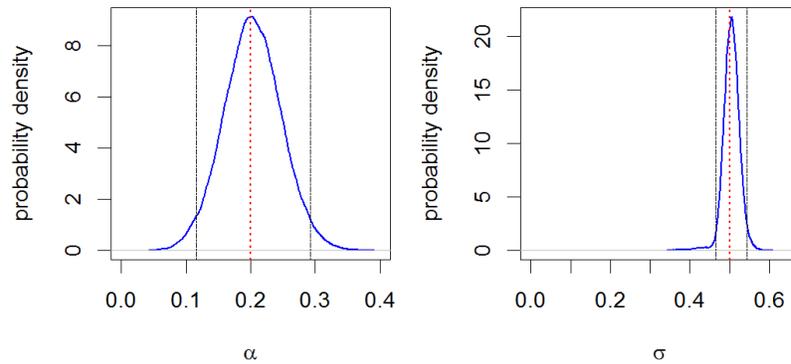
Figure 4.3: Posterior distributions derived when analysing a simulation of movement following a resource gradient in Figure 4.2. Parameters inferred using the MCMC algorithm include the drift and diffusion coefficients α and σ in Equation (4.2). Red dotted lines indicate real values used in the simulation and black dashed lines shows 95% quantile intervals.

(a)   (b)
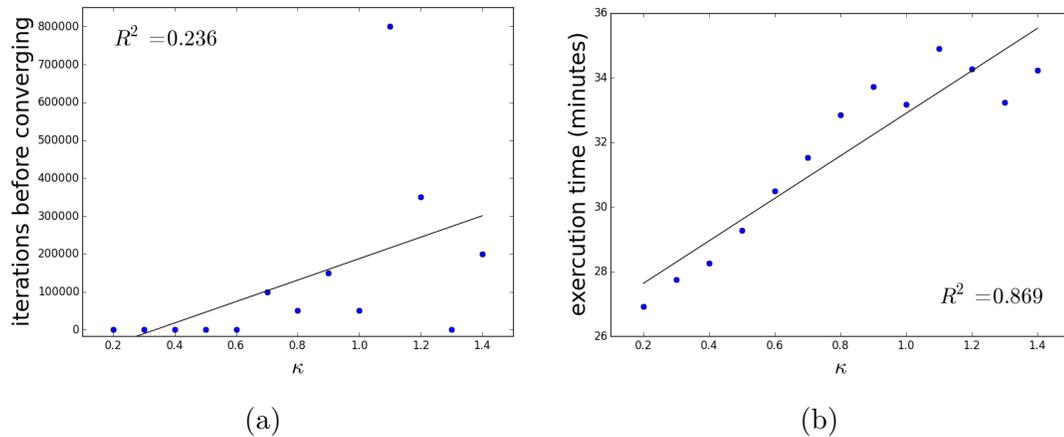
Figure 4.4: The relationship between the efficiency of the MCMC algorithm and κ, the mean of number of proposed switching points during a unit time interval. (a) Iterations before converging when applying the MCMC algorithm on the same simulation of movements following resource gradient using different κ. The corresponding execution time increases as κ rises (see Figure 4.4b). (b) Approximate execution time for $100,000$ iterations when applying the MCMC algorithm with different κ.
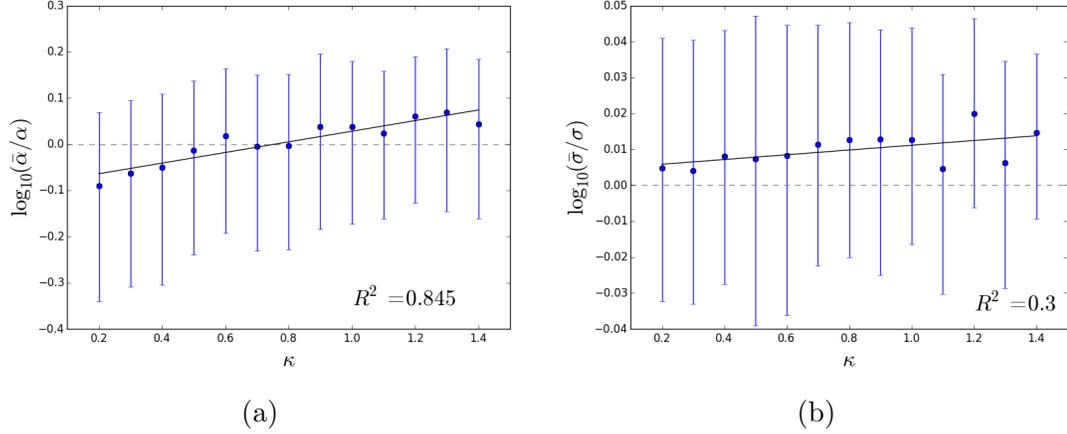
Figure 4.5: The relationship between the accuracy of the MCMC algorithm and κ, the mean of number of proposed switching points during a unit time interval. The log ratio between (a) sample means of $\bar{\alpha}$ and the real value of α (b) sample means of $\bar{\sigma}$ and the real value of σ when using different κ along with 95% confidence intervals.

the number of iterations needed before the algorithm converged (Figure 4.4a). In addition, the execution time was in proportion to the value of κ as larger κ brought about more proposed points to be processed. (Figure 4.4b). As for accuracy, it was not affected by the value of κ, as Figure 4.5 illustrates that the 95% central posterior intervals contain the real values of α and σ for all κ used. However, the posterior mean of α increased when κ became larger (Figures 4.5a). When analysing the case in Figure 4.2 with different initial values for the parameters inferred, initial values had little influence on the convergence time (Figure 4.6). Initial values far away from the real values did not necessarily lead to slower convergence.

Applying the inference procedure on simulations generated using different values for the drift coefficient, α (Equation 4.2), reveals that the algorithm usually converged fast but slower for some larger α (Figure 4.7a). On the other hand, the relationship between the value of σ and convergence time was insignificant (Figure 4.7b).

To test the accuracy of the algorithm, I parameterised the model introduced in Section 4.1 from simulations generated using various values for the drift and diffusion coefficients (Section 4.3). In addition, for each simulated trajectory, I compared the difference between the inference from every 3rd data point and
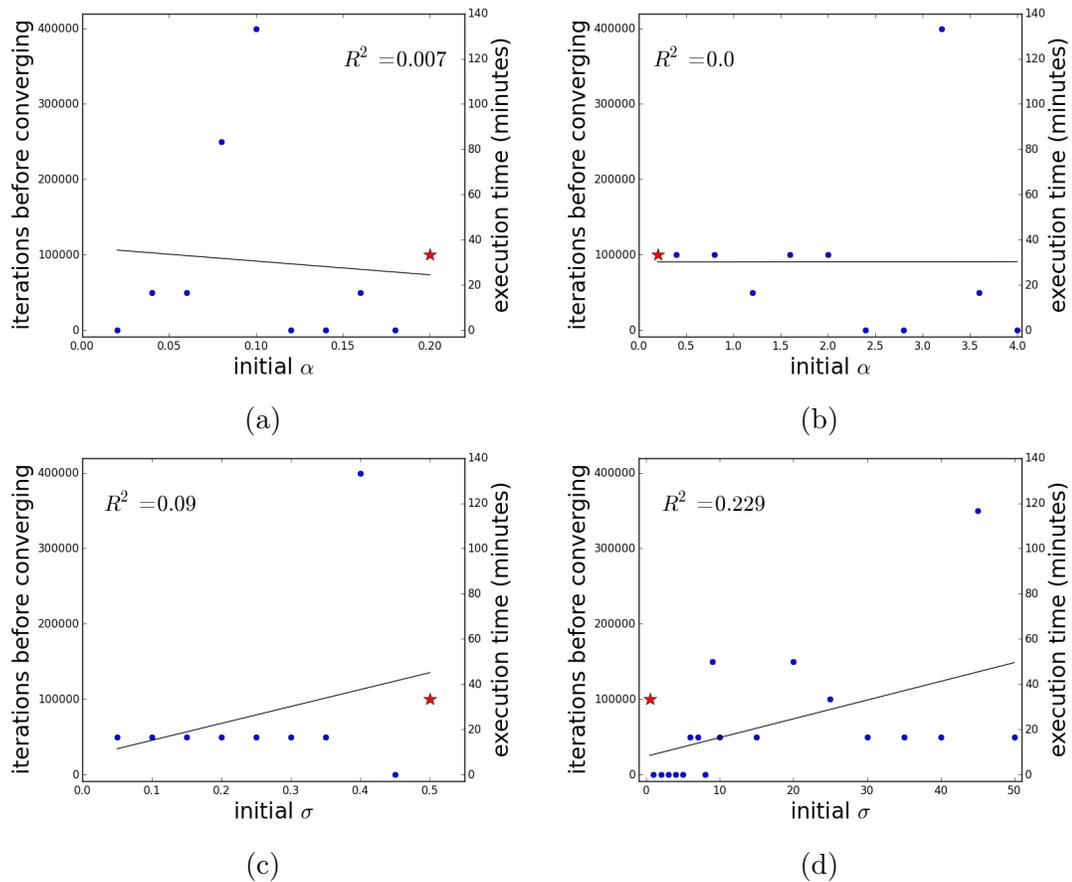
Figure 4.6: The number of iterations before converging when using different initial values in MCMC inference on a simulation of movements following local resource gradient. The red stars indicate the case when applying MCMC algorithm with initial values of all parameters being fixed at real values. The real values are $\alpha = 0.2, \sigma = 0.5$ The corresponding approximate execution time in minutes is shown in right y-axis. (a)(b) Applying MCMC with various initial values of α. The initial values of σ are fixed at real values.(c)(d) Applying MCMC with various initial values of σ. The initial values of α are fixed at real values.

(a)   (b)
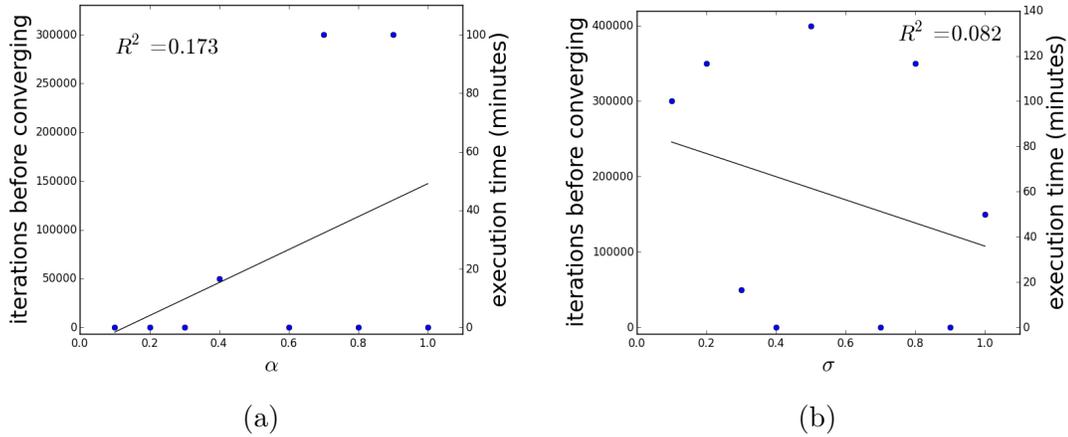
Figure 4.7: The relationship between iterations before converging and parameters in simulations of movements following local resource gradient. Coefficients α and σ are the drift and diffusion terms in the gradient following model respectively (Equation 4.2). The initial values of the chains for α and σ are fixed at the real values used in the simulations. The corresponding approximate execution time in minutes is shown in right y-axis. (a) Applying MCMC inference on simulations generated using different α but the same σ. (b) Applying MCMC inference on simulations generated using different σ but the same α.

from every data point to examine the impact of the density of data on the accuracy of the inference (Figures 4.8,4.9). When applying the MCMC algorithm on simulations generated by various values of α and using a third of data points, the 95% central posterior intervals contained the real values of α and σ only for cases generated by smaller α (Figures 4.8a,c). As for analysing trajectories generated by different values of σ, α was well-estimated for cases where medium values for σ were used, while the real values of σ were underestimated when larger σ was used (Figures 4.9a,c).

Comparing Figures 4.8a,4.9a to Figures 4.8b,4.9b reveals that the accuracy of the estimation of the drift coefficient, α, did not change notably when all data points were used in the inference algorithm. Meanwhile, the estimation of the diffusion coefficient, σ, was significantly improved as Figures 4.8c,d 4.9c,d show. When carrying out the algorithm with all data points, 95% central posterior intervals captured the real values of α and σ for about 67% and 87% simulations respectively.
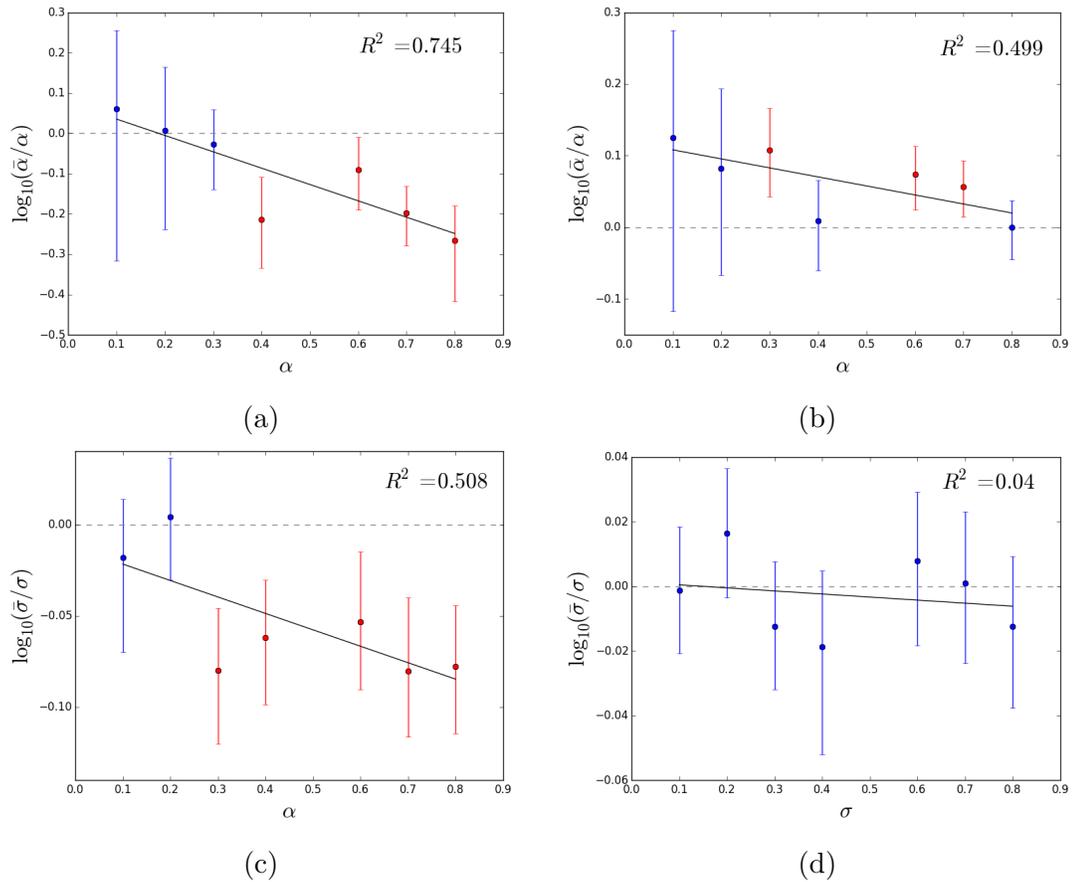
Figure 4.8: The log ratios between sample means $\bar{\alpha}$ and $\bar{\sigma}$ and real values of α and σ with 95% central posterior intervals when applying MCMC inference on simulations of gradient-following movements. The simulations were generated using different α but the same σ. Red intervals do not contain 0. (a)(c) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every 3rd data point. (b)(d) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every data point.
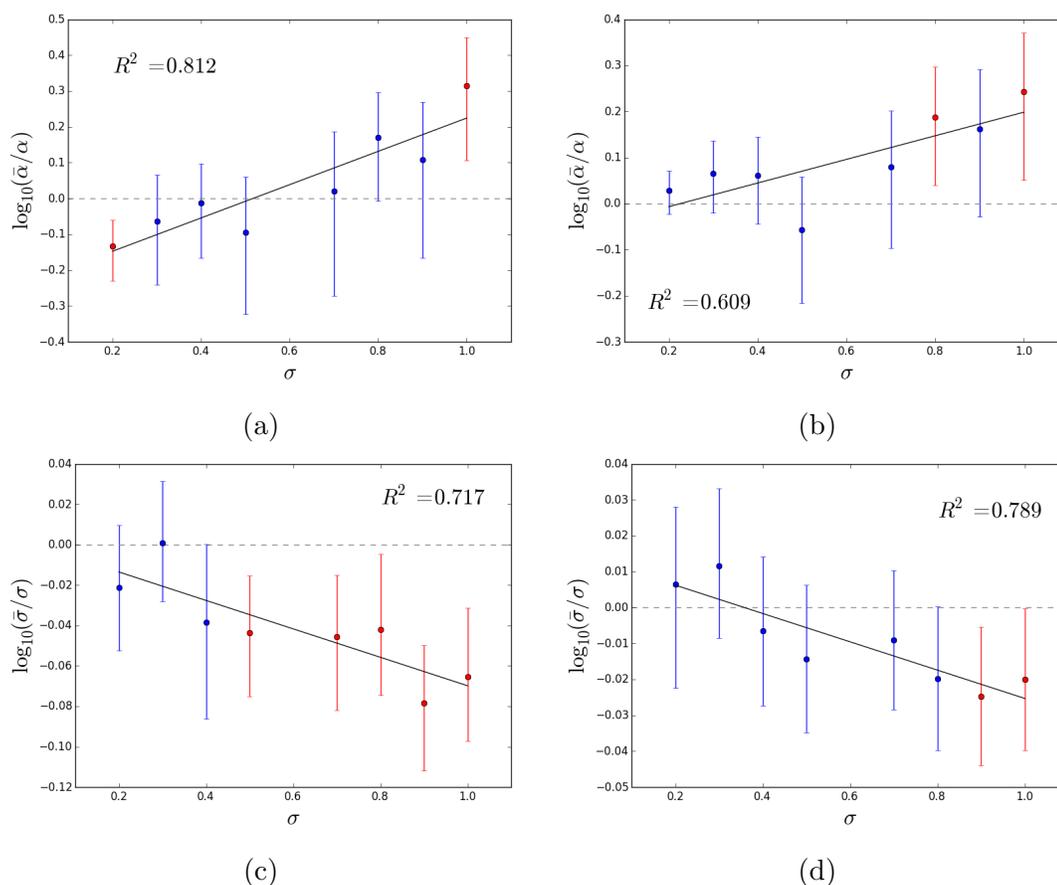
Figure 4.9: The log ratios between sample means $\bar{\alpha}$ and $\bar{\sigma}$ and real values of α and σ with 95% central posterior intervals when applying MCMC inference on simulations of gradient-following movements. The simulations were generated using different σ but the same α. Red intervals do not contain 0. (a)(c) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every 3rd data point. (b)(d) Applying the MCMC algorithm on the subsets of the simulated trajectories which contain every data point.

§ 4.5   Discussion

I have introduced a modelling framework, describing movement drifting in the direction up the local resource gradient. That is, animals are assumed to determine the direction of movement according to the perception of cues in the surrounding area rather than conditions over the whole landscape, in contrast to Chapter 3. The scenario described in this chapter can be applied to movement in oriented environments. For example, bacteria or cells move along the gradient of a chemical substance Alt (1980), and butterflies fly up slopes as a strategy for optimising mating success (Painter, 2014). I have applied an MCMC algorithm to infer the drift and diffusion terms of the model (Equation 4.2), where the drift term represents the strength of resource selection and the diffusion term accounts for the uncertainty.

When analysing a simulated trajectory, the parameter determining the average number of points inserted into a unit time interval, κ, did not have a remarkable effect on the accuracy of estimating the drift and diffusion terms α and σ in Equation 4.2. Nonetheless, the efficiency of the algorithm declined as κ increased (Figure 4.5). The benefit of inserting points between observations should be more clear when the data is much coarser than the scale of decision making.

Initial values for the MCMC algorithm had unimportant impact on the number of iterations to discard before converging (Figure 4.6). It appeared that κ dominated the efficiency of the algorithm. This might be because for a larger κ, more points were proposed to augment a selected subset of the observed data, and such proposed trajectory is more likely to be rejected. As a consequence, it would take longer for the algorithm for the update of trajectory to converge and this decelerates the overall inference procedure. Blackwell et al. (2016) has proposed an additional strategy to accelerate the update of trajectory by updating a selected point in each iteration. However, this did not improve the inference of models in this thesis so I did not include it here.

In general, the real values of α and σ (Equation 4.2) had little impact on the efficiency of the inference algorithm (Figure 4.7). However, larger α or σ were often poorly estimated (Figures 4.8,4.9). The estimation of parameters would be

more accurate if the time resolution of the data is higher.

The model described here considers movement in a rasterised landscape, similar to Hanks et al. (2015), where continuous-time movement models are also incorporated in the inference of resource selection. However, Hanks et al. (2015) only infers the selection of resources, while the inference based on the model in this chapter also infers the stochastic component of the movement process, namely the diffusion term. See Appendix B for more details.

Chapter 5

A case study of mule deer data in the Greater Yellowstone Ecosystem

Having constructed modelling frameworks for inferring movement responses to resource changes in the environment, I apply the inference algorithm on some mule deer (*Odocoileus hemionus*) data to demonstrate the methodology. Mule deer is one of the only two deer species living in the United States (Heffelfinger, 2018). They migrate from winter ranges to high-elevation summer ranges during spring and travel back to low-elevation ranges during autumn. The migration distance can range from around 20 km to more than 150 km (e.g. Sawyer et al. (2005); Sawyer and Kauffman (2011)). Understanding their migration behaviour is urgent because anthropogenic disturbances including housing and energy development have affected their ranges and migration routes (Sawyer et al., 2005, 2006, 2009; Lendrum et al., 2013).

The mule deer I study in this chapter belong to the Greater Yellowstone Ecosystem, famous for being one of the largest intact ecosystems in the world. This area is rich in not only wildlife but also natural resources such as mineral and petroleum. As a consequence, the management of the Greater Yellowstone Ecosystem remains challenging because of the trade-off between the preservation of nat-

ural processes and the exploitation of resources (Brussard, 1992).

§ 5.1   The data and models

5.1.1   THE MOVEMENT AND RESOURCE DATA

GPS collar data were collected from 28 adult female mule deer, each more than
1.5 years old. They were captured using a netgun fired from a helicopter near
Cody, Wyoming (USA), guided by protocols in agreement with the University
of Wyoming standards. Data used in this thesis was collected every two hours
from March to August in 2016 using collars (ATS, Iridium, Isanti, Minnesota,
USA). These data was part of Eastern Greater Yellowstone Mule Deer Project
and collected by Matt Kauffman and colleagues (*Eastern Greater Yellowstone
Mule Deer Project*, n.d.).

Resource data used in this chapter consisted of the normalised difference vegeta-
tion index (NDVI) and instantaneous rate of green-up (IRG), which were com-
piled from the MODIS satellite based on the methods of Bischof et al. (2012) and
Merkle et al. (2016).

5.1.2   THREE MODELS FOR RESOURCE SELECTION

As in Chapter 3, I assumed the deer moved following a switching OU process,
given by Equation (3.1). The attraction centre in the OU process was decided
by comparing the value of the RSF in Equation (3.9) for every patch centre. The
RSF in Equation (3.9) involves the resource quality in a target patch and the
distance between the animal and the patch. Here, I used three definitions for the
resource quality in Equation (3.9) to build three different migration models. The
first model, Model 1, used the NDVI values to stand for the resource qualities.
Model 2 used the integral of the NDVI values from current time t to the end of
the available resource data (Julian day 250, $t_{\max}$), given by

$$R(\boldsymbol{\mu}, t) = \int_t^{t_{\max}} \mathrm{NDVI}(\boldsymbol{\mu}, t)\mathrm{d}t, \qquad (5.1)$$
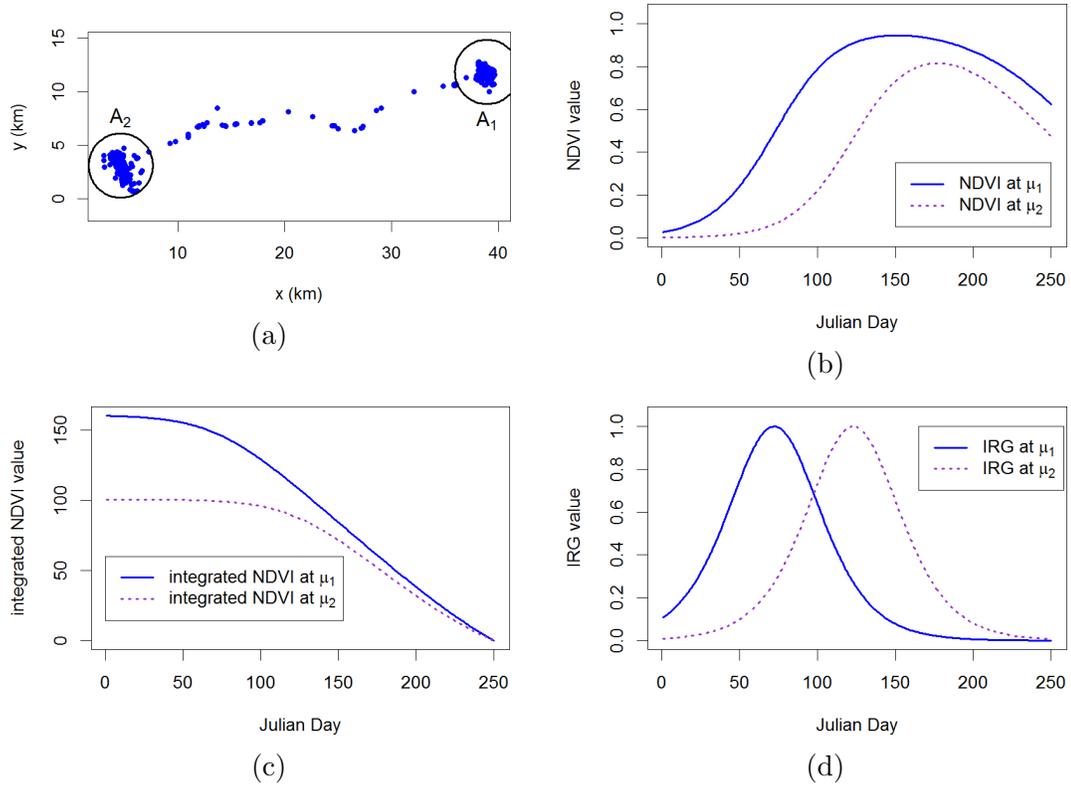
Figure 5.1: A case study of mule deer data. (a) The migration trajectory of mule deer No.4. The blue dots are observed locations collected between March and August 2016. The circles A_1 and A_2 are foraging patches. (b)(c)(d) The NDVI values, integrated NDVI and IRG at $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, the centres of patches A_1 and A_2 illustrated in Figure 5.1a.

where $\mathrm{NDVI}(\boldsymbol{\mu}, t)$ is the NDVI value at a potential attraction centre $\boldsymbol{\mu}$ at time t. I considered Model 2 to examine the influence of the expected resource in the rest of the season on the migration decisions. The last model, Model 3, employed the IRG values for resource qualities on the assumption that migration was triggered by the instant change of resources. I applied the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) to decide the best-fit model.

§5.2 Inference from data

Before applying the inference procedure introduced in Chapter 3, I identified foraging patches by grouping data points containing data of more than three days and forming a cluster within a circle of radius 3 km (c.f. Zweifel-Schielly et al.

(a)

(b)
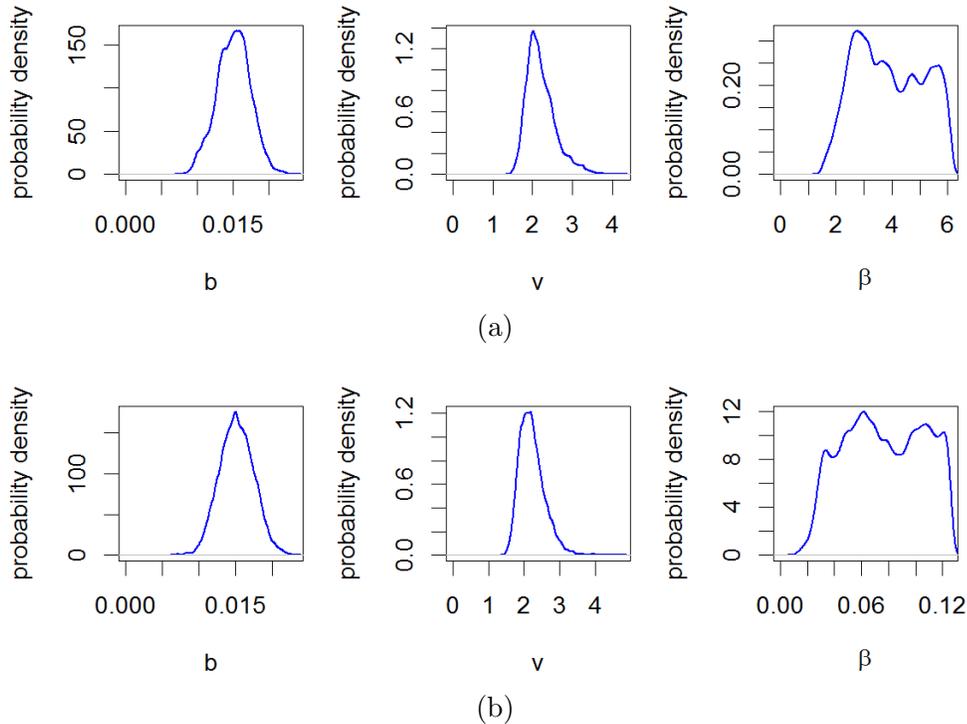
Figure 5.2: Inference from mule deer data using NDVI and integrated NDVI. The posterior distributions of movement coefficients b and v in Equation (3.1) and resource coefficient β in Equation (3.9) when applying the MCMC algorithm on the data of mule deer No.4 using (a) NDVI values (b) integrated NDVI values.

(2009)). In some cases, mule deer moved back and forth between neighbouring areas frequently, so larger patches were identified. For example, Figure 5.1a shows the trajectory of mule deer No.4 along with the identified circular patches.

After delineating foraging patches, the centre of a patch was defined by the average longitude and average latitude of points belonging to the patch. Subsequently, for each patch centre, I extracted the NDVI (Figure 5.1b) and IRG (Figure 5.1d) for Julian days 1 to 250 in 2016 from the correspondent pixels in the NDVI and IRG images. Besides, the integrated NDVI was calculated for Model 2 (Figure 5.1c).

The posterior distributions of fitting Models 1, 2 and 3 to the data of mule deer No.4 are shown in Figures 5.2a, 5.2b, and 5.3 respectively. In this case, Model 3, which uses the values of IRG for resource qualities, was best fitted to
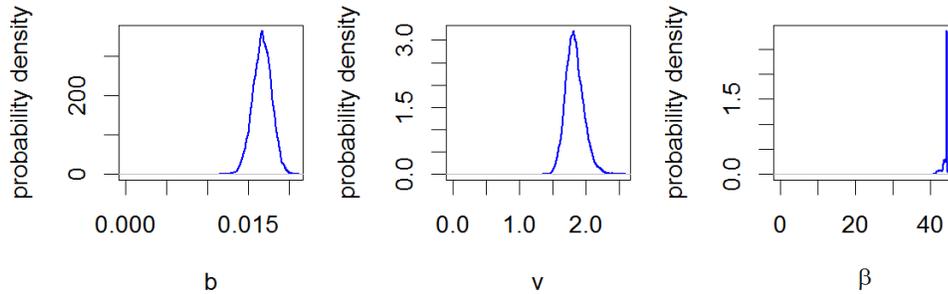
Figure 5.3: Inference from mule deer data using IRG. The posterior distributions of movement coefficients b and v in Equation (3.1) and resource coefficient β in Equation (3.9) when applying the MCMC algorithm on the data of mule deer No.4 using IRG values.
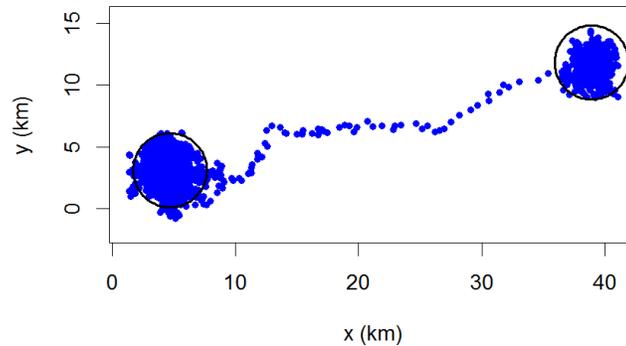
Figure 5.4: A simulated trajectory of mule deer migration. This simulated trajectory is generated using posterior means derived from analysing the data of deer No.4 with Model 3.

the data. Figure 5.4 illustrates a simulated trajectory generated by using the posterior means when fitting Model 3. It shows a similar migration pattern as observations.

The inference algorithm converged for all three models introduced in Section 5.1.2 when fitting them to the movement data of 17 out of 28 mule deer (Tables 5.1, 5.2). Among these seventeen individuals, Model 1 was best fitted to mule deer No.19 (5.2). Model 2 was best fitted to seven of them (Nos.6,7,8,9,10,20,24, Tables 5.1, 5.2), and Model 3 was the best for the other nine deer (Nos.1,4,11,12,13,16,18,23,25, Tables 5.1, 5.2).

For ten individuals, fitting Model 3 and one of Models 1 and 2 was successful, that is, the algorithm converged when using two of the three models. In these cases, only mule deer No.26 was best fitted by Model 1. For five of these individuals,

Model 2 was the best-fit model and the remaining four were best fitted by Model 3 (Table 5.3). For mule deer No.28, only Model 3 was fitted successfully (Table 5.3).

No.	Model	simulated departure date	observed departure date	ΔDIC
1	1	143,154		31.431
	2	143,153	132,150	29.326
	3	**100,151**		0
4	1	128		157.575
	2	128	125	154.359
	3	**126**		0
6	1	138		117.228
	2	**134**	133	0
	3	134		24.511
7	1	152		521.89
	2	**149**	148	0
	3	151		349.455
8	1	160		247.377
	2	**155**	154	0
	3	157		23.411
9	1	149		235.231
	2	**145**	144	0
	3	147		161.151
10	1	124,132		97.914
	2	**124,131**	123,128	0
	3	124,132		256.25

Table 5.1: The comparison of models for the mule deer data, where all 3 models can be fitted. For cases where the switch of movement centre occurred on two days or more, the numbers for the Julian dates are separated with a comma. Figures in bold indicate the model with the smallest DIC value on that individual.

In Models 1 to 3, I tested a simple situation where the selection of foraging patches was totally decided by the trade-off between resource qualities and travel distance. This implies if the resource quality in a target patch is never higher than that in the source patch, then the target patch will never become more attractive to animals being located in the source patch. Consequently, the inference algorithm may fail if an animal moves to a place which is regarded as less attractive by the

No.	Model	simulated departure date	observed departure date	ΔDIC
	1	127,152		188.879
11	2	135	125,149	194.51
	3	**126,151**		0
	1	148		254.247
12	2	141	140	170.427
	3	**146**		0
	1	148		67.66
13	2	148	140	30.176
	3	**142**		0
	1	114,137,149		29.094
16	2	119,147	112,136,147	52.185
	3	**101,138,148**		0
	1	130,158		85.453
18	2	124	123,157	2.164
	3	**124,158**		0
	1	**153**		0
19	2	153	124	74.003
	3	153		88.652
	1	135,146		128.348
20	2	**134**	133,145	0
	3	134,146		81.778
	1	137,137,137		109.592
23	2	134	133	64.982
	3	**134**		0
	1	151		38.284
24	2	**140**	139	0
	3	140		76.941
	1	142		471.274
25	2	142	140	468.167
	3	**141**		0

Table 5.2: The comparison of models for the mule deer data, where all 3 models can be fitted. For cases where the switch of movement centre occurred on two days or more, the numbers for the Julian dates are separated with a comma. Figures in bold indicate the model with the smallest DIC value on that individual.

No.	simulated departure date (Model 1)	simulated departure date (Model 2)	simulated departure date (Model 3)	observed departure date	ΔDIC
2	–	**148**	153	147	127.393
3	–	**100,156**	126,158	125,154	197.631
5	–	143	**143**	142	14.982
14	–	?	**140**	139	228.78
15	169	–	**149**	139	317.64
17	–	143	**142,143,143**	139	13.078
21	–	**148**	149,150,150	144	19.155
22	–	**147**	149	146	340.787
26	**126,142**	–	125,141	124,137	61.185
27	–	**147**	148	128	76.445
28	–	–	134	131	–

Table 5.3: The comparison of models for the mule deer data, where 1 or 2 models can be fitted. For cases where the switch of movement centre occurred on two days or more, the numbers for the Julian dates are separated with a comma. Figures in bold indicate the model with the smaller DIC value on that individual. The '?' in No.14 Model 2 indicates the algorithm converged but failed to estimate departure dates, meaning the model may not be a good fit.

model.

For each individual, the posterior mean of β (Equation 3.9) of the best-fit model was used to estimate departure dates of migration. Simulated departure dates were those dates when the centre of attraction switched according to the RSF in Equation (3.9). I compared these simulated departure dates to observed departure dates on which successive points leading away from a winter range were observed for the first time (Tables 5.1, 5.2 and 5.3). The correspondence between simulated and observed departure dates was generally good. This means the timing of migration can be explained by the trade-off between resources and travel distance. In addition, both integrated NDVI and IRG are proper measurements for resource qualities.

§ 5.3   Discussion

I have demonstrated the application of the modelling framework introduced in Chapter 3 to the migration data of 28 mule deer. The deer data were fitted by

three separate migration models, namely Models 1, 2 and 3, using NDVI, integrated NDVI and IRG values for resource qualities respectively. The results show a significant difference between zero and the resource coefficient β (Equation 3.9) with $p < 10^{-5}$, meaning resource qualities have an impact on the movement of mule deer. That is, mule deer prefer better resource qualities and would select patches with resource qualities high enough and worth travelling to. These findings also correspond to the green wave hypothesis (Drent et al., 1978; Bischof et al., 2012), which assumes migratory species follow better resources as spring progresses from places at low latitudes or altitudes. In addition, representing resource qualities by IRG, Models 3 supports the forage maturation hypothesis (Fryxell, 1991; Hebblewhite et al., 2008), which assumes the intermediate stage of maturation, corresponding to a high IRG, is more preferable since the nutrition value declines as plants become mature.

I have tested the inference procedure with models containing only two covariates, resource quality and distance, as a simple example. However, more factors influencing mule deer migration have been considered in previous work. Some factors studied include energy development, weather, canopy cover, distance to road, distance to hiding cover, slope, nutritional condition, reproductive status, age, and are shown to influence the movement of mule deer (Ager et al., 2003; Monteith et al., 2011; Lendrum et al., 2013). Therefore, it would be worth integrating these factors into the RSF in Equation 3.2 as covariates to examine how they impact resource selection and compare the results to previous findings.

In this chapter, I have only considered every individual deer separately. However, the interaction between individuals may also be included to understand their migration behaviours better as deer are social animals. For example, Figure 5.5 shows the relationship between the distance from a winter range centre to another and the difference between the dates on which the two individuals left their winter ranges. Figure 5.5 indicates that two deer staying close to each other in winter does not necessarily mean they leave their winter ranges about the same time. Therefore, they might begin to migrate by following other cues.

In summary, the inference approach performs well with the mule deer data as two or all of Models 1 to 3 can successfully fit the data of every individual except
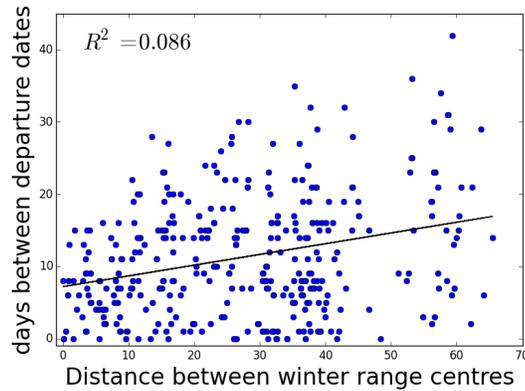
Figure 5.5: Distance between winter range centres and days between departure dates. Every dot indicates the distance between the winter range centres of two mule deer in km and the difference between the dates on which the two deer left their winter ranges.

for a deer, whose data was only fitted by Model 3. The results have shown that resource qualities have an impact on the decision of migration. However, to understand migration behaviours more thoroughly, more investigation is needed since the main purpose of the analysis here is to demonstrate the application of the models and inference method.

Chapter 6

Discussion and Conclusions

Through the study of animal movement, this thesis has aimed at gaining insight into the prediction of long-term space use patterns and selection of resources. To investigate the accuracy of estimating space use patterns from individual movement process by PDE methods, I have examined three such methods, namely the Hyperbolic Scaling method, the Moment Closure method and Patlak's approach. I applied these three methods to analyse various movement kernels and compared the resulting approximated distributions to the real long-term distributions given by the Master Equation. The results showed that when the process of relocation only involves smooth changes in mean velocity or environmental conditions, these methods would give closer approximations to the real distributions. In particular, the Moment Closure method usually outperforms others when considering a smooth kernel.

As for understanding the selection of resources by animals, the modelling framework introduced in Chapters 3 has been used to recognise the trade-off between resources and travelling cost, represented by distance, on movement decisions in a wide range of situations. This modelling framework is especially useful for inferring movement decisions made at large scales such as seasonal migration. In Chapter 4, I have presented another framework, which enables the inference of movement depending on local resources. In general, animals are assumed to decide their direction by assessing the resources either across the whole landscape or in the surrounding areas and then move following a process incorporating both re-

source selection and uncertainty. I have developed a Bayesian algorithm, allowing for decision changes at any time, to infer the parameters of these two modelling frameworks from data. This algorithm has successfully captured parameter values, including the one representing the relative importance of resources, with 95% central posterior distributions in most simulations.

§ 6.1   Comparison of three PDE approximation methods

PDE approximation methods are tools to scale up individual movement rules to population-level space use patterns. They are much more efficient in terms of computer time than applying the Master Equation. Nonetheless, the accuracy of estimation depends on the movement kernel and the approximation method used (Chapter 2). If the movement kernel has non-smooth changes in mean velocity or resource weighting functions, the PDE methods can lead to a large discrepancy between the approximations and the real distribution emerging from the kernel. A popular and older method, Patlak's approach, was proved to be unreliable for predictions even for very simple movement kernels (Figure 2.1). I also have applied the PDE methods to three 1-D central-place foraging models, defined by movement kernels bringing about discontinuous, continuous and differentiable mean velocity functions respectively. By analysing these models, I have shown that the PDE methods provided more accurate approximations when the mean velocity function was more smooth (Figures 2.3, 2.6, 2.9). Likewise, the PDE methods led to approximations closer to the real long-term distribution when a more smooth landscape was assumed (Figure 2.10). In particular, when the movement kernel is sufficiently smooth, the Moment Closure method usually outperformed the others. Possible reasons for the better performance of the Moment Closure method include that the drift might not be strong in these cases or the third and higher moments might be sufficiently small.

For non-smooth movement kernels, the results in Chapter 2 indicate that population density is often overestimated near points where abrupt changes in mean velocity happen. In other words, the range of frequent use might be underestimated. Therefore, one should be careful when applying PDE approximations

to predict space use patterns such as home ranges and be aware of a potential discrepancy between prediction and reality. This awareness would benefit decision-making on critical issues such as planning conservation areas as the size of home ranges often plays a central role in the determination of the scale of management (Allen and Singh, 2016).

Furthermore, by observing the diffusion terms for the three PDE methods in Equations (2.5) (the Hyperbolic Scaling and Moment Closure methods) and (2.15) (Patlak's approach), it is clear that Patlak's approach always leads to distributions with larger variance if the first moment of velocity is nonzero. This is because the second moment instead of variance is used for the diffusion term (Equation 2.15). The overestimation of variance found in Patlak's approach, however, often offsets the tendency of PDE methods to overestimate the density at points where the movement kernel has abrupt changes (e.g. Figure 2.3f). In consequence, approximations obtained by Patlak's approach seem to be reasonably close to real distributions. This makes Patlak's approach appear to perform relatively better than other approaches in cases of non-smooth movement kernels. Nonetheless, in general, when using any of these PDE methods to analyse non-smooth movement kernels, one may obtain a prediction which is a rough qualitative representation of the real distribution but quite different to the real distribution in terms of quantity. In addition, Patlak's approach relies on the assumptions that the whole system changes slowly and the environment alters gradually (Patlak, 1953; Turchin, 1991). My work reveals that these assumptions might also be required when using the Hyperbolic Scaling and Moment Closure methods.

PDE approximate methods have been widely used and the Hyperbolic Scaling and Moment Closure methods and Patlak's approach in this thesis, in fact, belong to wider classes of hyperbolic scaling methods, moment closure methods and correlated random walk models respectively. Despite the popularity of PDE methods, very few works investigated the accuracy of approximations derived from these methods. An advantage of a hyperbolic scaling method is that this method avoids movement in infinite speed, which could be brought about by another scaling limit technique, the parabolic scaling (Holmes, 1993; Hadeler, 2000;

Hillen and Stevens, 2000; Eftimie, 2012). Although Holmes (1993) and Hillen
and Stevens (2000) compared hyperbolic models with parabolic models, neither
of them examined the accuracy of approximations. Moment closure methods are
often used to estimate the formation of spatial patterns arising from the interac-
tions between individuals such as competition and predator-prey interactions (e.g.
Bolker and Pacala (1997); Law et al. (2003); Murrell and Law (2003); Murrell
et al. (2005)). Among these works, Bolker and Pacala (1997) did compare their
results of using a moment closure approach to realisations of stochastic models
to investigate the effect of neglecting higher order moments on approximations.
In addition, Lloyd (2004) compared two moment closure approaches for recur-
rent epidemics and examined the performance of approximations by comparing
the results to either analytic models or numerical simulations. Both Bolker and
Pacala (1997) and Lloyd (2004) observed that moment closure approaches could
fail when the population or the number of neighbours is small. However, they
also indicated that it remains unclear under which conditions moment closure
methods would be valid. Moreover, there is much difficulty in assessing the error
of approximations (North and Ovaskainen, 2007). Meanwhile, my work suggests
that examining the smoothness of the underlying movement kernel would be a
simple preliminary assessment of the accuracy of approximations.

In summary, even though only some simple cases have been examined in this
thesis, these examples provide a starting point of understanding the performance
of PDE approximate methods.

However, the movement models examined in Chapter 2 are in only 1-D and it
is much harder to analyse 2-D cases in general. What is unknown includes how
the directional component of a 2-D movement kernel would affect the accuracy
of approximations. Nevertheless, PDE methods might also give poor estimations
for 2-D models where the mean velocity changes radically, since these changes
will cause the estimate steady-state pattern to change abruptly.

§ 6.2   The modelling framework for analysing movement responses to resources

In Chapters 3-5, I turned my attention to understanding how to infer resource selection in continuous time. In Chapter 3, I developed a modelling framework based on an OU process, which describes a biased random walk towards an attraction centre in continuous time. A set of OU processes with different attraction centres were combined to form a switching OU model to allow for changes of destination at any time (Chapter 3). That is, the attraction centre was decided according to environmental conditions, which may vary over time, and the OU process used switches to another with a different attraction centre correspondingly. With this modelling framework, three types of simulations were constructed to consider situations where the resource qualities changed depending on seasons or the foraging paths of the animal (Sections 3.3-3.5). The inference from these simulations by an algorithm developed on the basis of Blackwell et al. (2016) has shown the ability of the algorithm to estimate the parameter values used in simulations. In general, the algorithm performed well by capturing real values of parameters with 95% central posterior intervals of the posterior distributions for a majority of cases.

Also, using a switching continuous-time movement model, Chapter 4 modelled a drift along the direction up local resource gradient. The movement direction, rather than an attraction centre, was determined at each point in the process of inference. The inference from simulated data was successful in approximating the drift and diffusion terms when the density of data point was sufficiently high. This enabled the uncovering of movement decisions directed by neighbouring environmental clues by considering movements in continuous time.

The model developed in Chapter 3 was fitted to movement data of mule deer (*Odocoileus hemionus*) to demonstrate the methodologies for understanding the mechanisms behind the migration scenarios. Results of using the migration model described in Section 3.3 showed that migration might stem from the trade-off between the cost of travel and the expected gain of better resources. This provides an explanation of migration and agrees well with two important hypotheses of

migration, namely the forage maturation hypothesis (Fryxell, 1991; Hebblewhite et al., 2008) and the green wave hypothesis (Drent et al., 1978; Bischof et al., 2012), which are briefly introduced in the following.

The forage maturation hypothesis states that the forage quality is defined by considering both quantity and quality of food because the nutrition value and digestibility decline as plants grow and mature. Therefore, animals should prefer plants at an intermediate stage of growth (Merkle et al., 2016). The green wave hypothesis assumes that the migration of herbivores is triggered in pursuit of high-quality food, which becomes available as the onset of spring progresses from places of lower latitude or altitude. Here, the onset of spring can be defined to be the time when plants reach the intermediate stage of growth according to the forage maturation hypothesis. These two hypotheses have been tested to provide explanations for the migration of birds and ungulate (van der Graaf et al., 2006; Bischof et al., 2012; van Wijk et al., 2012; Kölzsch et al., 2015; Merkle et al., 2016; Aikens et al., 2017).

In Chapter 5, the inference from the mule deer data implies that the expected resources in the rest of the season and IRG might better explain the migration of mule deer. In particular, the peak of IRG might be interpreted as the intermediate stage of maturation since it usually corresponds to the inflection points of the NDVI. These results agree with the green wave hypothesis and also the forage maturation hypothesis when using IRG.

6.2.1 Comparisons with previous work

Most movement models, including both discrete- and continuous-time models, have neglected the fact that animals may make a decision at points other than those being observed (e.g. Johnson et al. (2008)). By adopting the method of Blackwell et al. (2016), the inference algorithm in Chapters 3 and 4 enables the analysis of resource selection from movement data to consider points where changes of movement states might occur. Therefore, using this algorithm would avoid problems such as failing to identify key features such as a road (Thurfjell et al., 2014) and obtain a more appropriate interpretation of movement data.

RSA and SSA have been commonly used with applications ranging from studies of space use (Chetkiewicz and Boyce, 2009; Panzacchi et al., 2016; Viana et al., 2018), the use of dynamic resources (Lone et al., 2018), to the influence of human activities and constructions (Hebblewhite and Merrill, 2008; Gillies et al., 2011; Lendrum et al., 2012; Scrafford et al., 2018). However, choosing an appropriate spatio-temporal scale has always been a challenge (Boyce, 2006; Thurfjell et al., 2014). This is not only because selection of resources or steps can take place at multiple scales but also because analysis methods are confined to the assumption that the scale of decision making matches the fixed rate of observations (Zweifel-Schielly et al., 2009; Benhamou, 2014; Fleming et al., 2014a; Bastille-Rousseau et al., 2018). Attempting better understanding of movement motivated by resources both at small and large scales, the method built on a SSF in Bastille-Rousseau et al. (2018) considers the probability of moving towards habitats at a distance R away from the animal's current location and determines how large this distance R should be by a model selection procedure. The modelling framework introduced in Chapter 3, meanwhile, directly compares potential targets regardless of distance and thus is not restricted to a specific spatial scale.

Breed et al. (2017) used a similar switching OU process to the modelling framework in Chapter 3 to model movements between patches. They only use conditions in the source patch to determine the probability of leaving the patch. On the other hand, by considering conditions in the target patch as well as those in the source patch, my model would better represent movements depending on factors not only in the local area but also in the distance. This model is especially important in identifying places of interest and key factors driving movement at a large scale.

6.2.2 POSSIBLE FUTURE DIRECTIONS

Extensions to the model

Since the modelling frameworks in Chapters 3 and 4 for resource selection are based on a resource weighting function, they are ready to incorporate drivers of movement other than resource quality and travel distance, both of which are

external factors. As the movement ecology paradigm proposed by Nathan et al. (2008) describes, internal factors such as searching for food, safety or mate motivate movement and are essential to uncover the mechanism leading to the observed trajectories. Therefore, my model could also incorporate internal factors driving movement to achieve better understanding of the underlying mechanisms of movement. For example, since an individual seeks safety, predator pressure may result in a decrease in the probability of selecting a favourable resource (Forester et al., 2009). Memory also plays a fundamental role in movement decision making and the consequent space use patterns (Fagan et al., 2013; Merkle et al., 2014; Riotte-Lambert et al., 2015; Potts and Lewis, 2016a; Bracis and Mueller, 2017; Merkle et al., 2017). Other drivers include topography (Potts, Mokross, Stouffer and Lewis, 2014), interactions between animals (Vanak et al., 2013), barriers and corridors (Panzacchi et al., 2016). All of these factors can be included in the weighting function in my models. For example, one might use

$$w(\mathbf{z}(\mathbf{x})) = \exp(\beta_1 z_1(\mathbf{x}) + \beta_2 z_2(\mathbf{x}) + \cdots + \beta_k z_k(\mathbf{x})) \tag{6.1}$$

to include k factors z_1 to z_k.

Different behavioural states could be taken into account in my model as movement traits such as speed may change according to the habitat encountered or activities undertaken (Morales and Ellner, 2002). As well as modelling animal movement with the same drift and diffusion coefficients along the whole path, the switching OU model in Chapter 3 is ready for different values for these coefficients according to habitats and movement modes (Harris and Blackwell, 2013). For example, in addition to using an OU process to represent movement towards a specific centre, more OU process attracted to the same centre can be defined using different drift and diffusion coefficients to describe different movement modes such as encamped and exploratory states (Morales et al., 2004). By modelling different movement modes with different processes, my method would be able to detect behavioural changes and provide a more precise explanation for movement in a heterogeneous environment.

Improvements in the efficiency of inference

The datasets analysed in this thesis only contain hundreds or thousands of data points, which were sufficient in my study. However, the amount of data available nowadays can be large because of the advanced technology (Kays et al., 2015). As a consequence, to deal with datasets of large size efficiently, it is necessary to develop novel methods such as Kálmán filters to speed up the inference of OU models (Fleming et al., 2017), or alternatively, to rarefy data by strategies such as identifying areas of interest from location data (Potts et al., 2018). Using Kálmán filters may significantly reduce the execution time of the algorithm, as the examples given in Fleming et al. (2017). Nevertheless, further examination is required to understand if Kálmán filters can be applied to the inference of my model. There are other Bayes filters such as grid-based approaches and particle filters but it would be straightforward to use Kálmán filters here because the distributions considered are Gaussian (Fox et al., 2003; Fleming et al., 2017). Rarefying paths as an initial step of analysis would decrease the size of data and therefore speed up the main inference procedure. However, discarding some data may result in less accurate inference outcomes. Such techniques might enable my modelling framework to deal with data at a high temporal resolution such as seconds or smaller but further research is needed to investigate their applicability.

§ 6.3   Summary

This thesis contains a comparison between PDE approximation methods by examining different types of movement kernels. It reveals that one should be cautious when applying these PDE methods to analyse non-smooth movement kernels and the Moment Closure method would be a better choice than Patlak's approach in general. Starting with simple examples, this part of the thesis sheds light on how well these PDE methods perform.

I have also developed two flexible modelling frameworks for inferring resource selection by moving animals in continuous time for decision making depending on the resources across the whole landscape or in the local area. These have extended the usage of the Bayesian inference procedure in Blackwell et al. (2016)

and an application to real movement data has been demonstrated (Chapter 5). This chapter has discussed possible ways to extend the models, including the incorporation of drivers such as memory and the representation of different behaviours. The modelling frameworks bring together the strengths of continuous-time movement models and resource selection analysis to provide an advanced tool for understanding resource selection decisions.

Appendices

Appendix A

Measuring distance between distributions by Euclidean distance

Here, the difference between distributions is measured by Euclidean distance instead of Kullback-Leibler divergence (KL-divergence) as in Chapter 2. I calculated the Euclidean distance between steady-state distributions derived by the partial differential equation (PDE) methods and the numerical solutions obtained from the Master Equation (Equation 2.21).

When considering the movement kernel $k_\tau^1(z|x)$ in Equation (2.32), the contours in Figures A.1a,b show that the distance between analytic solutions derived from the PDE methods (Equations 2.37, 2.40) and numerical solutions grows as μ/σ becoming larger. For this movement kernel, the steady-state distribution derived from Patlak's approach is closer to the numerical solution than those using the Hyperbolic Scaling and Moment Closure methods (Figure A.1). For the movement kernel $k_\tau^2(z|x)$ in Equation (2.41), either the Moment Closure method or Patlak's approach provide approximations with the smallest Euclidean distance to the numerical solution. The former performs better in the region where μ/σ is larger, while the latter outperforms others for smaller μ/σ (Figure A.2d). The Moment Closure method gives the best approximations to the numerical solution

when analysing movement kernel $k_\tau^3(z|x)$ in Equation (2.52) (Figure A.4).

Figures A.1-4 show qualitatively similar patterns to those observed in figures in Chapter 2. This is revealed by comparing Figures A.1, A.2, A.3 and A.4 to Figures 2.3a-d, 2.5, 2.6a,b and 2.8d in the main text respectively. Therefore, the general implications given by the result are not an artefact of the metric adopted.

(a) Moment Closure

(b) Patlak's method

(c) $0.05 \leq \mu \leq 0.2$, $\sigma = 0.05$

(d) $\mu = 0.05$, $0.05 \leq \sigma \leq 0.2$
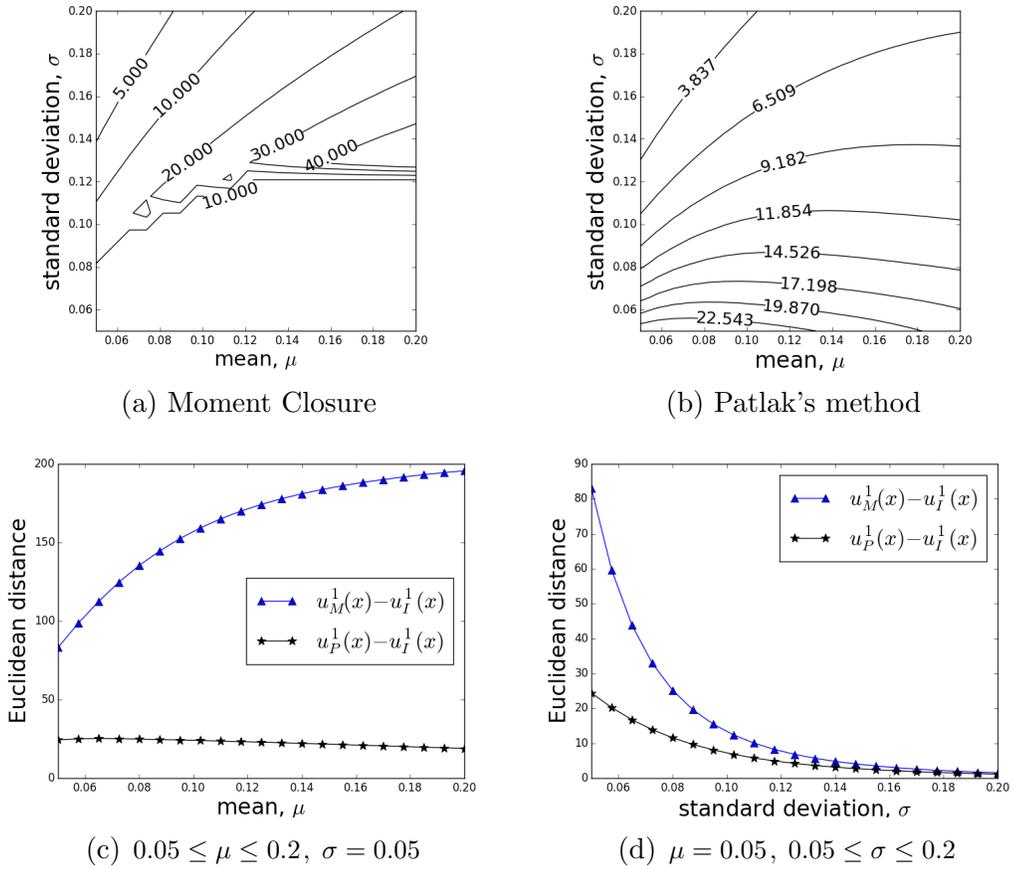
Figure A.1: Discontinuous mean velocity movement kernel $k_\tau^1(z|x)$ (Equation 2.32 in the main text) with μ the mean move length in one step and σ the standard deviation of move length: (a) The contours of the Euclidean distance between the numerical solution, $u_I^1(x)$, and the analytic solution, $u_M^1(x)$ (equals to $u_H^1(x)$ in Equation 2.37 in the main text), derived using a Moment Closure technique, $\mu, \sigma \in [0.05, 0.2]$. (b) The contours of the Euclidean distance between $u_I^1(x)$ and the analytic solution, $u_P^1(x)$ (Equation 2.40 in the main text), derived using Patlak's method, $\mu, \sigma \in [0.05, 0.2]$. (c) Euclidean distance between $u_M^1(x)$ and $u_I^1(x)$ (▲), and $u_P^1(x)$ and $u_I^1(x)$ (⋆) with $0.05 \leq \mu \leq 0.2$ and $\sigma = 0.05$. (d) Euclidean distance between $u_M^1(x)$ and $u_I^1(x)$ (▲) and, $u_P^1(x)$ and $u_I^1(x)$ (⋆) with $0.05 \leq \sigma \leq 0.2$ and $\mu = 0.05$.
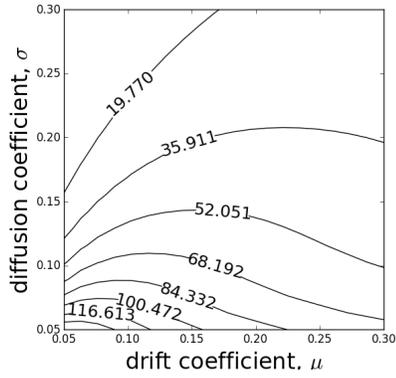
(a) Hyperbolic Scaling
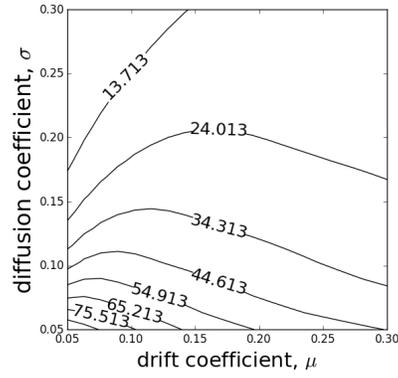
(b) Moment Closure
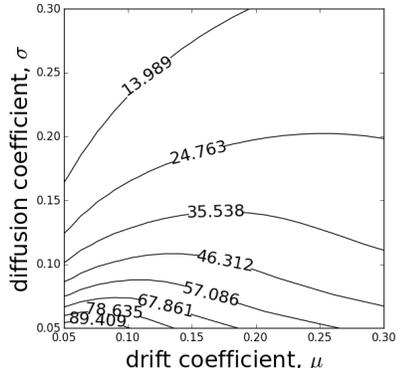
(c) Patlak's method
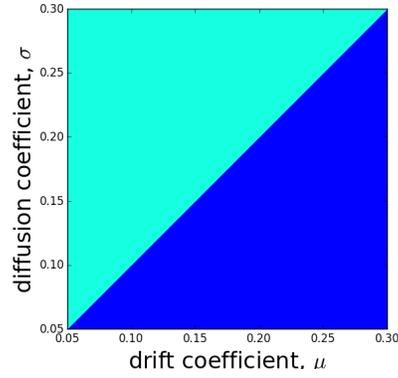
(d)

Figure A.2: Continuous mean velocity movement kernel $k_\tau^2(z|x)$ (Equation 2.41 in the main text) with μ (resp. $|x|$) the mean move length in one step for $|x| > \mu$ (resp. $|x| \leq \mu$) and σ the standard deviation of move length: The contours of the Euclidean distance between the numerical solution, $u_I^2(x)$, and (a) the analytic solution, $u_H^2(x)$ (Equation 2.44 in the main text), derived from a Hyperbolic Scaling method; (b) $u_M^2(x)$ (Equation 2.46 in the main text), derived from a Moment Closure technique; (c) $u_P^2(x)$ (Equation 2.50 in the main text), derived from Patlak's method. (d) Turquoise region: the Euclidean distance between $u_I^2(x)$ and $u_P^2(x)$ is smaller than from $u_M^2(x)$ or $u_H^2(x)$. Blue region: the distance between $u_I^2(x)$ and $u_M^2(x)$ is the smallest.

(a) $0.05 \leq \mu \leq 0.3$, $\sigma = 0.2$.

(b) $\mu = 0.2$, $0.05 \leq \sigma \leq 0.3$.
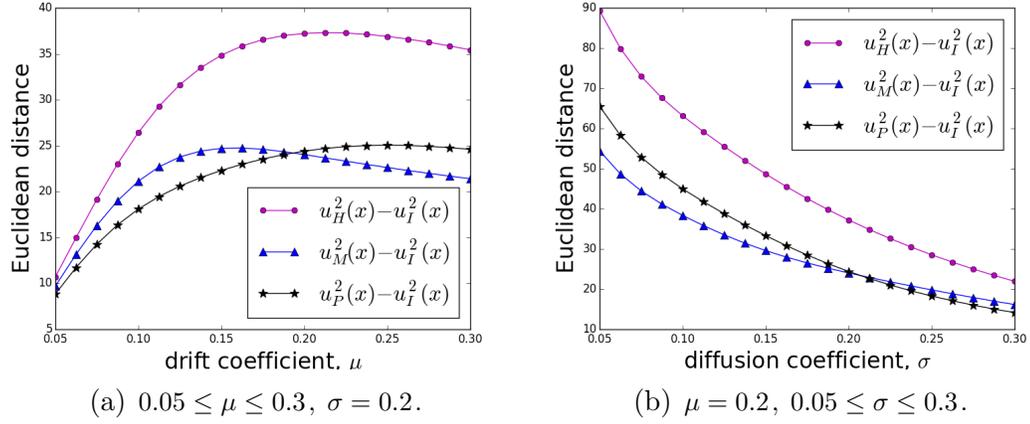
Figure A.3: Continuous mean velocity movement kernel $k_\tau^2(z|x)$ (Equation 2.41 in the main text) with μ (resp. $|x|$) the mean move length in one step for $|x| > \mu$ (resp. $|x| \leq \mu$) and σ the standard deviation of move length: (a) Euclidean distance between the numerical solution, $u_I^2(x)$, and the the analytic solution, $u_H^2(x)$ (Equation 2.44 in the main text), derived from a Hyperbolic Scaling method ($\bullet$); $u_I^2(x)$ and $u_M^2(x)$ (Equation 2.46 in the main text), derived from a Moment Closure technique ($\blacktriangle$); $u_I^2(x)$ ($\star$) and $u_P^2(x)$ (Equation 2.50 in the main text), derived from Patlak's method, with $0.05 \leq \mu \leq 0.3$ and $\sigma = 0.2$. (b) Euclidean distance between $u_H^2(x)$ and $u_I^2(x)$ ($\bullet$), $u_M^2(x)$ and $u_I^2(x)$ ($\blacktriangle$), and $u_P^2(x)$ and $u_I^2(x)$ ($\star$) for $\mu = 0.2$, $0.05 \leq \sigma \leq 0.3$.

(a) $0.05 \leq \mu \leq 0.5$, $\sigma = 0.1$
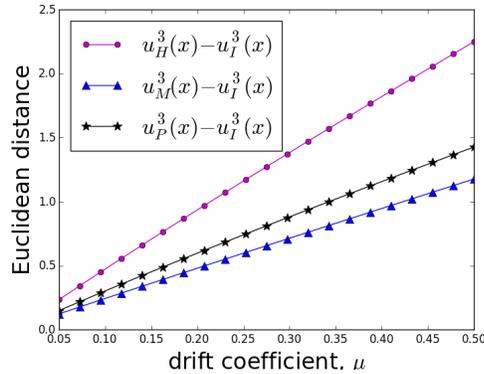
Figure A.4: Differentiable mean velocity movement kernel $k_\tau^3(z|x)$ (Equation 2.52 in the main text) with μx^2 the mean move length in one step and σ the standard deviation of move length: Euclidean distance between the numerical solution, $u_I^3(x)$, and the the analytic solution, $u_H^3(x)$ (Equation 2.55 in the main text), derived from a Hyperbolic Scaling method ($\bullet$); $u_I^3(x)$ and $u_M^3(x)$ (Equation 2.57 in the main text), derived from a Moment Closure technique ($\blacktriangle$); $u_I^3(x)$ ($\star$) and $u_P^3(x)$ (Equation 2.61 in the main text), derived from Patlak's method, with $0.05 \leq \mu \leq 0.5$ and $\sigma = 0.1$.

Appendix B

A comparison between continuous-time discrete-space models and gradient-following models

Both the continuous-time discrete-space (CTDS) model introduced in Hanks et al. (2015) and the gradient-following model from Chapter 4 consider continuous-time movement processes when given resource data in the form of a grid. Here, I draw a comparison between CTDS models (Hanks et al., 2015) and the gradient-following model introduced in Chapter 4 by applying the inference method described in Hanks et al. (2015) to analyse simulations of gradient-following movements. This comparison reveals that the inference method of Hanks et al. (2015) is able to infer the parameter related to resource selection in the gradient-following model (α in Equation 4.2 in Chapter 4) but not the other parameter describing the uncertainty of movement (σ in Equation 4.2). Although our method is much slower than Hanks et al. (2015), my algorithm simultaneously infer resource selection and the movement process. Furthermore, my inference method turns out to be much more accurate.

Hanks et al. (2015) uses a CTDS model to analyse animal movement and focuses on the transition from a cell in a space grid to one of its neighbouring cell. The transition rates between cells are determined by the selection of resources, depending on covariates affecting movement from a cell to another and the coefficients of these covariates. The likelihood of observing the animal staying in cell G_{i_t} for time τ_t and then moving to cell $G_{i_{t+1}}$ given $\boldsymbol{\beta}$, the coefficients of covariates, is (Hanks et al., 2015)

$$[G_{i_t} \to G_{i_{t+1}}, \tau_t | \boldsymbol{\beta}] = \lambda_{i_t i_{t+1}}(\boldsymbol{\beta}) \exp\{-\tau_t \lambda_{i_t}(\boldsymbol{\beta})\}, \tag{B.1}$$

where $\lambda_{i_t i_{t+1}}(\boldsymbol{\beta})$ is the transition rate from G_{i_t} to $G_{i_{t+1}}$ and $\lambda_{i_t}(\boldsymbol{\beta}) = \sum_{i_t \sim j_t} \lambda_{i_t j_t}(\boldsymbol{\beta})$ is the total transition rate from cell G_{i_t} with $i_t \sim j_t$ meaning G_{i_t} and G_{j_t} are directly connected. To apply the method of Hanks et al. (2015), I use R package `ctmcmove` (Hanks, 2018) to estimate $\boldsymbol{\beta}$ and calculate the likelihood of transition from a cell to another cell in Equation (B.1) given residence time in the source cell.

Meanwhile, the drift term α in our gradient-following model represents the strength of selection and can be related to $\boldsymbol{\beta}$ in the model of Hanks et al. (2015). However, Hanks et al. (2015) only estimates $\boldsymbol{\beta}$ but not the variance of movement processes, there is no direct link between the inference of the diffusion term σ in the gradient-following model and the method of Hanks et al. (2015). Therefore, when using the inference method of Hanks et al. (2015) to analyse simulations of the gradient-following model, I will only infer the drift term α and assume the diffusion term σ is known. Based on the gradient-following model and using the centre of a cell to represent the cell, the likelihood of moving from the centre of G_{i_t} to the centre of $G_{i_{t+1}}$ in time τ_t is

$$[\boldsymbol{\mu}_t \to \boldsymbol{\mu}_{t+1}, \tau_t | \alpha, \sigma] =$$
$$\frac{1}{2\pi\sigma^2\tau_t} \exp\left(-\frac{(x_{t+1} - x_t - \alpha\rho_1(t)\tau_t)^2 + (y_{t+1} - y_t - \alpha\rho_2(t)\tau_t)^2}{2\sigma^2\tau_t}\right), \tag{B.2}$$

where $\boldsymbol{\mu}_t = (x_t, y_t)$ and $\boldsymbol{\mu}_{t+1} = (x_{t+1}, y_{t+1})$ are centres of G_{i_t} and $G_{i_{t+1}}$, α and σ are the drift and diffusion coefficients of the movement process, and $\boldsymbol{\rho}(t) = (\rho_1(t), \rho_2(t))$ is the drift direction determined by a resource gradient.

Thus Equation (B.1) and Equation (B.2) are equivalent.

Defining

$$f(\boldsymbol{\beta}) := \lambda_{i_t i_{t+1}}(\boldsymbol{\beta}) \exp\{-\tau_t \lambda_{i_t}(\boldsymbol{\beta})\}, \tag{B.3}$$

then

$$f(\boldsymbol{\beta}) = \frac{1}{2\pi\sigma^2\tau_t} \exp\left(-\frac{(x_{t+1} - x_t - \alpha\rho_1(t)\tau_t)^2 + (y_{t+1} - y_t - \alpha\rho_2(t)\tau_t)^2}{2\sigma^2\tau_t}\right) \tag{B.4}$$

and $f(\boldsymbol{\beta})$ is obtained by inference from trajectory and resource data using R package `ctmcmove` (Hanks, 2018). Rearranging Equation (B.4) leads to

$$-2\sigma^2\tau_t \log\left[2\pi\sigma^2\tau_t \cdot f(\boldsymbol{\beta})\right] = (x_{t+1} - x_t - \alpha\rho_1(t)\tau_t)^2 + (y_{t+1} - y_t - \alpha\rho_2(t)\tau_t)^2 \tag{B.5}$$

Note that Equation (B.5) is a quadratic equation of α. To ease notation, I define

$$A = x_{t+1} - x_t \tag{B.6}$$

$$B = \rho_1(t)\tau_t \tag{B.7}$$

$$C = y_{t+1} - y_t \tag{B.8}$$

$$D = \rho_2(t)\tau_t \tag{B.9}$$

$$E = 2\sigma^2\tau_t \log\left[2\pi\sigma^2\tau_t \cdot f(\boldsymbol{\beta})\right]. \tag{B.10}$$

Then for $f(\boldsymbol{\beta}) \neq 0$, $\boldsymbol{\rho}(t) \neq (0,0)$ and $4(AB+CD)^2 - 4(B^2+D^2)(A^2+C^2+E) \geq 0$, solving Equation (B.5) gives

$$\alpha = \frac{2(AB + CD) \pm \sqrt{4(AB + CD)^2 - 4(B^2 + D^2)(A^2 + C^2 + E)}}{2(B^2 + D^2)} \tag{B.11}$$

Note that the transition likelihood $f(\boldsymbol{\beta})$ (Equation B.3) depends on locations rather than being fixed. Therefore, I calculate α by Equation (B.11) for every non-zero $f(\boldsymbol{\beta})$ between neighbouring cells with $\tau_t = 1$ if I also have $\boldsymbol{\rho}(t) \neq (0,0)$ and $4(AB + CD)^2 - 4(B^2 + D^2)(A^2 + C^2 + E) \geq 0$ to show how accurately α is estimated.

For example, I apply the method of Hanks et al. (2015) to analyse the simulated trajectory of movement following a local resource gradient (Figure B.1). Figure
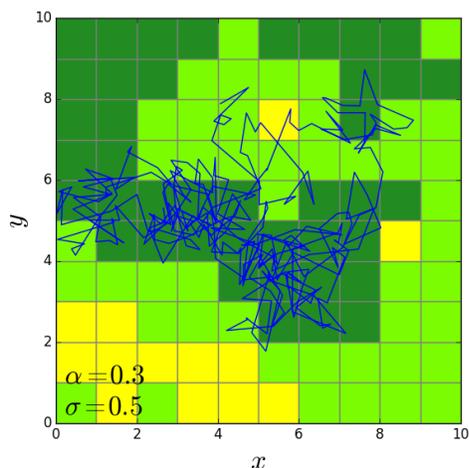
Figure B.1: Simulated trajectories of movement following a local resource gradient. The different colours in the landscape represent different resource qualities: yellow: low resource quality; green: medium resource quality; dark green: high resource quality.

B.2 shows the density of estimated α along with the real value indicated by a red line. The black dashed lines shows the 95% central interval. Although the real value lies within the interval, the variance is very large (Var = 1.26). The execution time was under 1 minute.

On the other hand, using the inference algorithm described in Chapter 4 to analyse the simulated trajectory in Figure B.1 captures real values of α and σ successfully with 95% central posterior intervals (Figure B.3). The variances of the posterior distributions of α and σ are 0.0013 and 0.0002 respectively, much smaller compared to the results using the method of Hanks et al. (2015). It took around 35 minutes to complete 200,000 iterations of MCMC sampling.
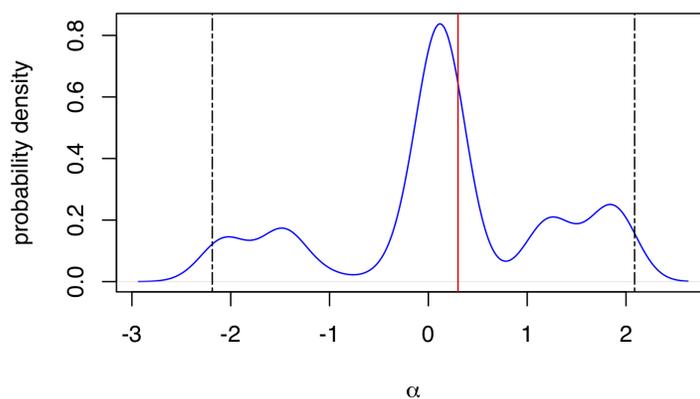
Figure B.2: The density of estimated α obtained using R package `ctmcmove` (Hanks, 2018) and Equation (B.11) to make inference from the simulated trajectory in Figure B.1. The red line indicates the real value of α used in simulation. Black dashed lines shows the 95% central interval.
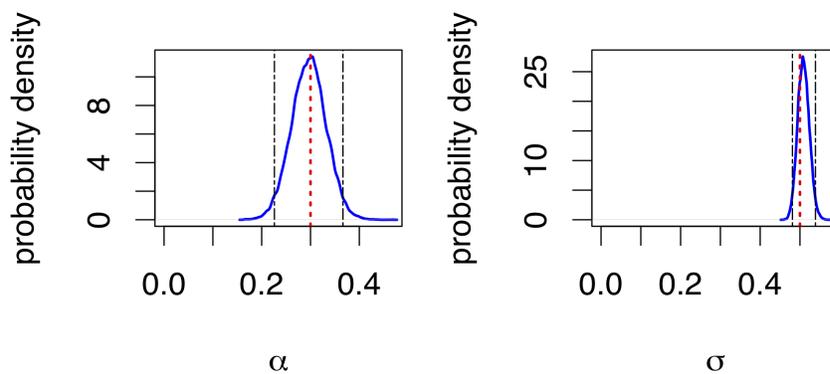
Figure B.3: The posterior distributions derived from analysing the simulated trajectory in Figure B.1 using the MCMC algorithm in Chapter 4. The red line indicates the real value of α used in simulation. Black dashed lines shows the 95% central interval.

References

Ager, A. A., Johnson, B. K., Kern, J. W. and Kie, J. G. (2003), 'Daily and Seasonal Movements and Habitat Use by Female Rocky Mountain Elk and Mule Deer', *Journal of Mammalogy* **84**(3), 1076–1088.
URL: *https://dx.doi.org/10.1644/BBa-020*

Aikens, E. O., Kauffman, M. J., Merkle, J. A., Dwinnell, S. P. H., Fralick, G. L. and Monteith, K. L. (2017), 'The greenscape shapes surfing of resource waves in a large migratory herbivore', *Ecology Letters* **20**(6), 741–750.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.12772*

Albert, J. (2007), *Bayesian Computation with R*, Bayesian Computation with R, Springer.
URL: *https://books.google.co.uk/books?id=kVk_WHFfIrMC*

Albertsen, C. M., Whoriskey, K., Yurkowski, D., Nielsen, A. and Flemming, J. M. (2015), 'Fast fitting of non-gaussian state-space models to animal movement data via template model builder', *Ecology* **96**(10), 2598–2604.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/14-2101.1*

Allen, A. M. and Singh, N. J. (2016), 'Linking movement ecology with wildlife management and conservation', *Frontiers in Ecology and Evolution* **3**, 155.
URL: *https://www.frontiersin.org/article/10.3389/fevo.2015.00155*

Alt, W. (1980), 'Biased random walk models for chemotaxis and related diffusion approximations', *Journal of Mathematical Biology* **9**(2), 147–177.
URL: *https://doi.org/10.1007/BF00275919*

Avgar, T., Lele, S. R., Keim, J. L. and Boyce, M. S. (2017), 'Relative selection strength: Quantifying effect size in habitat- and step-selection inference', *Ecology and Evolution* **7**(14), 5322–5330.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.3122*

Avgar, T., Potts, J. R., Lewis, M. A. and Boyce, M. S. (2016), 'Integrated step selection analysis: bridging the gap between resource selection and animal movement', *Methods in Ecology and Evolution* **7**(5), 619–630.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12528*

Barnett, A. H. and Moorcroft, P. R. (2008), 'Analytic steady-state space use patterns and rapid computations in mechanistic home range analysis', *Journal of Mathematical Biology* **57**(1), 139–159.
URL: *https://doi.org/10.1007/s00285-007-0149-8*

Bastille-Rousseau, G., Gibbs, J. P., Yackulic, C. B., Frair, J. L., Cabrera, F., Rousseau, L.-P., Wikelski, M., Kmmeth, F. and Blake, S. (2017), 'Animal movement in the absence of predation: environmental drivers of movement strategies in a partial migration system', *Oikos* **126**(7), 1004–1019.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/oik.03928*

Bastille-Rousseau, G., Murray, D. L., Schaefer, J. A., Lewis, M. A., Mahoney, S. P. and Potts, J. R. (2018), 'Spatial scales of habitat selection decisions: implications for telemetry-based movement modelling', *Ecography* **41**(3), 437–443.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.02655*

Bastille-Rousseau, G., Potts, J. R., Yackulic, C. B., Frair, J. L., Ellington, E. H. and Blake, S. (2016), 'Flexible characterization of animal movement pattern using net squared displacement and a latent state model', *Movement Ecology* **4**(1), 15.
URL: *https://doi.org/10.1186/s40462-016-0080-y*

Bateman, A. W., Lewis, M. A., Gall, G., Manser, M. B. and Clutton-Brock, T. H. (2015), 'Territoriality and home-range dynamics in meerkats, suricata

suricatta: a mechanistic modelling approach', *Journal of Animal Ecology* **84**(1), 260–271.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1365-2656.12267*

Benhamou, S. (2011), 'Dynamic approach to space and habitat use based on biased random bridges', *PLOS ONE* **6**(1), 1–8.
URL: *https://doi.org/10.1371/journal.pone.0014592*

Benhamou, S. (2014), 'Of scales and stationarity in animal movements', *Ecology Letters* **17**(3), 261–272.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.12225*

Bischof, R., Loe, L. E., Meisingset, E. L., Zimmermann, B., Van Moorter, B. and Mysterud, A. (2012), 'A migratory northern ungulate in the pursuit of spring: Jumping or surfing the green wave?', *The American Naturalist* **180**(4), 407–424. PMID: 22976006.
URL: *https://doi.org/10.1086/667590*

Blackwell, P. (1997), 'Random diffusion models for animal movement', *Ecological Modelling* **100**(1), 87 – 102.
URL: *http://www.sciencedirect.com/science/article/pii/S0304380097001531*

Blackwell, P. G. (2003), 'Bayesian inference for Markov processes with diffusion and discrete components', *Biometrika* **90**(3), 613–627.
URL: *http://dx.doi.org/10.1093/biomet/90.3.613*

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**(2), 184–195.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12460*

Bolker, B. and Pacala, S. W. (1997), 'Using moment equations to understand stochastically driven spatial pattern formation in ecological systems', *Theoretical Population Biology* **52**(3), 179 – 197.
URL: *http://www.sciencedirect.com/science/article/pii/S0040580997913319*

Börger, L., Dalziel, B. D. and Fryxell, J. M. (2008), 'Are there general mechanisms of animal home range behaviour? A review and prospects for future research', *Ecology Letters* **11**(6), 637–650.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1461-0248.2008.01182.x*

Boyce, M. S. (2006), 'Scale for resource selection functions', *Diversity and Distributions* **12**(3), 269–276.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1366-9516.2006.00243.x*

Boyce, M. S., Vernier, P. R., Nielsen, S. E. and Schmiegelow, F. K. (2002), 'Evaluating resource selection functions', *Ecological Modelling* **157**(2), 281 – 300.
URL: *http://www.sciencedirect.com/science/article/pii/S0304380002002004*

Bracis, C. and Mueller, T. (2017), 'Memory, not just perception, plays an important role in terrestrial mammalian migration', *Proceedings of the Royal Society of London B: Biological Sciences* **284**(1855).
URL: *http://rspb.royalsocietypublishing.org/content/284/1855/20170449*

Breed, G. A., Golson, E. A. and Tinker, M. T. (2017), 'Predicting animal home-range structure and transitions using a multistate ornstein-uhlenbeck biased random walk', *Ecology* **98**(1), 32–47.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/ecy.1615*

Brillinger, D. R. (2010), Modeling Spatial Trajectories, *in* A. E. Gelfand, P. J. Diggle, M. Fuentes and P. Guttorp, eds, 'Handbook of Spatial Statistics', 1 edn, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, CRC Press, the United States of America.

Brussard, P. F. (1992), 'The future of yellowstone', *Science* **255**(5048), 1148–1149.
URL: *http://science.sciencemag.org/content/255/5048/1148*

Bunnefeld, N., Brger, L., van Moorter, B., Rolandsen, C. M., Dettki, H., Solberg, E. J. and Ericsson, G. (2011), 'A model-driven approach to quantify

migration patterns: individual, regional and yearly differences', *Journal of Animal Ecology* **80**(2), 466–476.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2656.2010.01776.x*

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 of London B: Biological Sciences* **365**(1550), 2157–2162.
URL: *http://rstb.royalsocietypublishing.org/content/365/1550/2157*

Chetkiewicz, C.-L. B. and Boyce, M. S. (2009), 'Use of resource selection functions to identify conservation corridors', *Journal of Applied Ecology* **46**(5), 1036–1047.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2664.2009.01686.x*

Chib, S. and Greenberg, E. (1995), 'Understanding the metropolis-hastings algorithm', *The American Statistician* **49**(4), 327–335.
URL: *https://amstat.tandfonline.com/doi/abs/10.1080/00031305.1995.10476177*

Codling, E. A., Plank, M. J. and Benhamou, S. (2008), 'Random walk models in biology', *Journal of The Royal Society Interface* **5**(25), 813–834.
URL: *http://rsif.royalsocietypublishing.org/content/5/25/813*

Drent, R., Ebbinge, B. and Weijand, B. (1978), 'Balancing the energy budgets of arctic-breeding geese throughout the annual cycle: a progress report', *Verhandlungen Ornithologischen Gesellschaft in Bayern* **23**, 239–264.

Dunn, J. E. and Gipson, P. S. (1977), 'Analysis of radio telemetry data in studies of home range', *Biometrics* **33**(1), 85–101.
URL: *http://www.jstor.org/stable/2529305*

Eastern Greater Yellowstone Mule Deer Project (n.d.), `https://migrationinitiative.org/content/eastern-greater-yellowstone-mule-deer-project`. Accessed: 2019-03-04.

Eftimie, R. (2012), 'Hyperbolic and kinetic models for self-organized biological aggregations and movement: a brief review', *Journal of Mathematical Biology* **65**(1), 35–75.
URL: *https://doi.org/10.1007/s00285-011-0452-2*

Fagan, W. F., Lewis, M. A., Auger-Mth, M., Avgar, T., Benhamou, S., Breed, G., LaDage, L., Schlgel, U. E., Tang, W.-w., Papastamatiou, Y. P., Forester, J. and Mueller, T. (2013), 'Spatial memory and animal movement', *Ecology Letters* **16**(10), 1316–1329.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.12165*

Fleming, C. H., Calabrese, J. M., Mueller, T., Olson, K. A., Leimgruber, P. and Fagan, W. F. (2014*a*), 'From fine-scale foraging to home ranges: A semivariance approach to identifying movement modes across spatiotemporal scales.', *The American Naturalist* **183**(5), E154–E167. PMID: 24739204.
URL: *https://doi.org/10.1086/675504*

Fleming, C. H., Calabrese, J. M., Mueller, T., Olson, K. A., Leimgruber, P. and Fagan, W. F. (2014*b*), 'Non-markovian maximum likelihood estimation of autocorrelated movement processes', *Methods in Ecology and Evolution* **5**(5), 462–472.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12176*

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.
URL: *http://www.sciencedirect.com/science/article/pii/S1574954117301115*

Ford, R. G. (1983), 'Home range in a patchy environment: Optimal foraging predictions', *American Zoologist* **23**(2), 315–326.
URL: *http://dx.doi.org/10.1093/icb/23.2.315*

Forester, J. D., Im, H. K. and Rathouz, P. J. (2009), 'Accounting for animal movement in estimation of resource selection functions: sampling and data

analysis', *Ecology* **90**(12), 3554–3565.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/08-0874.1*

Fortin, D., Beyer, H. L., Boyce, M. S., Smith, D. W., Duchesne, T. and Mao, J. S. (2005), 'Wolves influence elk movements: Behavior shapes a trophic cascade in YELLOWSTONE NATIONAL PARK', *Ecology* **86**(5), 1320–1330.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/04-0953*

Fox, V., Hightower, J., , Schulz, D. and Borriello, G. (2003), 'Bayesian filtering for location estimation', *IEEE Pervasive Computing* **2**(3), 24–33.

Fryxell, J. M. (1991), 'Forage quality and aggregation by large herbivores', *The American Naturalist* **138**(2), 478–498.
URL: *https://doi.org/10.1086/285227*

Gillies, C. S., Beyer, H. L. and St. Clair, C. C. (2011), 'Fine-scale movement decisions of tropical forest birds in a fragmented landscape', *Ecological Applications* **21**(3), 944–954.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/09-2090.1*

Gurarie, E., Andrews, R. D. and Laidre, K. L. (2009), 'A novel method for identifying behavioural changes in animal movement data', *Ecology Letters* **12**(5), 395–408.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1461-0248.2009.01293.x*

Hadeler, K. (2000), 'Reaction transport equations in biological modeling', *Mathematical and Computer Modelling* **31**(4), 75 – 81. Proceedings of the Conference on Dynamical Systems in Biology and Medicine.
URL: *http://www.sciencedirect.com/science/article/pii/S0895717700000248*

Hanks, E. (2018), *ctmcmove: Modeling Animal Movement with Continuous-Time Discrete-Space Markov Chains*. R package version 1.2.9.
URL: *https://CRAN.R-project.org/package=ctmcmove*

Hanks, E. M., Hooten, M. B. and Alldredge, M. W. (2015), 'Continuous-time discrete-space models for animal movement', *The Annals of Applied Statistics*

9(1), 145–165.
URL: *https://doi.org/10.1214/14-AOAS803*

Harris, K. J. and Blackwell, P. G. (2013), 'Flexible continuous-time modelling for heterogeneous animal movement', *Ecological Modelling* **255**, 29 – 37.
URL: *http://www.sciencedirect.com/science/article/pii/S0304380013000471*

Hebblewhite, M. and Merrill, E. (2008), 'Modelling wildlifehuman relationships for social species with mixed-effects resource selection models', *Journal of Applied Ecology* **45**(3), 834–844.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2664.2008.01466.x*

Hebblewhite, M., Merrill, E. and McDermid, G. (2008), 'A multi-scale test of the forage maturation hypothesis in a partially migratory ungulate population', *Ecological Monographs* **78**(2), 141–166.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/06-1708.1*

Heffelfinger, J. (2018), *Deer of the Southwest: A Complete Guide to the Natural History, Biology, and Management of Southwestern Mule Deer and White*, Texas A&M University Press.

Hillen, T. and Painter, K. J. (2013), Transport and anisotropic diffusion models for movement in oriented habitats, *in* M. A. Lewis, P. K. Maini and S. V. Petrovskii, eds, 'Dispersal, Individual Movement and Spatial Ecology: A Mathematical Perspective', Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 177–222.
URL: *https://doi.org/10.1007/978-3-642-35497-7_7*

Hillen, T. and Stevens, A. (2000), 'Hyperbolic models for chemotaxis in 1-d', *Nonlinear Anal.: Real World Appl.* **1**(3), 409–433.
URL: *http://dx.doi.org/10.1016/S0362-546X(99)00284-9*

Holgate, P. (1971), Random walk models for animal behavior, *in* G. Patil, E. Pielou and W. Walters, eds, 'Statistical Ecology: Sampling and Modeling Biological Populations and Population Dynamics', number 2 *in* 'Pennsyvania State Statistics', Pennsylvania State University Press, pp. 1–12.

Holmes, E. E. (1993), 'Are diffusion models too simple? a comparison with telegraph models of invasion', *The American Naturalist* **142**(5), 779–795. PMID: 19425956.
URL: *https://doi.org/10.1086/285572*

Horne, J. S. and Garton, E. O. (2006), 'Selecting the best home range model: An information-theoretic approach', *Ecology* **87**(5), 1146–1152.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/0012-9658%282006%2987%5B1146%3ASTBHRM%5D2.0.CO%3B2*

Horne, J. S., Garton, E. O., Krone, S. M. and Lewis, J. S. (2007), 'Analyzing animal movements using brownian bridges', *Ecology* **88**(9), 2354–2363.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/06-0957.1*

Integral Calculator (2019), `https://www.integral-calculator.com`. Accessed: 2019-01-09.

Johnson, D. S., Thomas, D. L., Ver Hoef, J. M. and Christ, A. (2008), 'A general framework for the analysis of animal resource selection from telemetry data', *Biometrics* **64**(3), 968–976.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1541-0420.2007.00943.x*

Kays, R., Crofoot, M. C., Jetz, W. and Wikelski, M. (2015), 'Terrestrial animal tracking as an eye on life and planet', *Science* **348**(6240).
URL: *http://science.sciencemag.org/content/348/6240/aaa2478*

Kölzsch, A., Bauer, S., Boer, R., Griffin, L., Cabot, D., Exo, K.-M., Jeugd, H. P. and Nolet, B. A. (2015), 'Forecasting spring from afar? Timing of migration and predictability of phenology along different migration routes of an avian herbivore', *Journal of Animal Ecology* **84**(1), 272–283.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1365-2656.12281*

Kullback, S. and Leibler, R. A. (1951), 'On information and sufficiency', *The Annals of Mathematical Statistics* **22**(1), 79–86.
URL: *http://www.jstor.org/stable/2236703*

Law, R., Murrell, D. J. and Dieckmann, U. (2003), 'Population growth in space and time: Spatial logistic equations', *Ecology* **84**(1), 252–262.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/0012-9658*

Lendrum, P. E., Anderson, C. R., Long, R. A., Kie, J. G. and Bowyer, R. T. (2012), 'Habitat selection by mule deer during migration: effects of landscape structure and natural-gas development', *Ecosphere* **3**(9), art82.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/ES12-00165.1*

Lendrum, P. E., Anderson, Jr, C. R., Monteith, K. L., Jenks, J. A. and Bowyer, R. T. (2013), 'Migrating mule deer: Effects of anthropogenically altered landscapes', *PLOS ONE* **8**(5), 1–10.
URL: *https://doi.org/10.1371/journal.pone.0064548*

Lewis, M. A. and Murray, J. (1993), 'Modelling territoriality and wolf-deer interactions', *Nature* **366**(23), 738–740.
URL: *https://doi.org/10.1038/366738a0*

Lloyd, A. L. (2004), 'Estimating variability in models for recurrent epidemics: assessing the use of moment closure techniques', *Theoretical Population Biology* **65**(1), 49 – 65.
URL: *http://www.sciencedirect.com/science/article/pii/S0040580903001175*

Lone, K., Merkel, B., Lydersen, C., Kovacs, K. M. and Aars, J. (2018), 'Sea ice resource selection models for polar bears in the barents sea subpopulation', *Ecography* **41**(4), 567–578.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.03020*

Manly, B., McDonald, L., Thomas, D., McDonald, T. and Erickson, W. (2002), *Resource selection by animals: statistical analysis and design for field studies*, 2 edn, Springer Netherlands.

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.
URL: *https://doi.org/10.1186/s40462-014-0021-6*

Merkle, J. A., Cross, P. C., Scurlock, B. M., Cole, E. K., Courtemanch, A. B., Dewey, S. R. and Kauffman, M. J. (2018), 'Linking spring phenology with mechanistic models of host movement to predict disease transmission risk', *Journal of Applied Ecology* **55**(2), 810–819.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1365-2664.13022*

Merkle, J. A., Fortin, D. and Morales, J. M. (2014), 'A memory-based foraging tactic reveals an adaptive mechanism for restricted space use', *Ecology Letters* **17**(8), 924–931.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.12294*

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 of London B: Biological Sciences* **283**(1833).
URL: *http://rspb.royalsocietypublishing.org/content/283/1833/20160456*

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).
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/oik.03356*

Mitchell, M. S. and Powell, R. A. (2004), 'A mechanistic home range model for optimal use of spatially distributed resources', *Ecological Modelling* **177**(1), 209 – 232.
URL: *http://www.sciencedirect.com/science/article/pii/S0304380004001176*

Monteith, K. L., Bleich, V. C., Stephenson, T. R., Pierce, B. M., Conner, M. M., Klaver, R. W. and Bowyer, R. T. (2011), 'Timing of seasonal migration in mule deer: effects of climate, plant phenology, and life-history characteristics', *Ecosphere* **2**(4), art47.

URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/ES10-00096.1*

Moorcroft, P. R. and Barnett, A. (2008), 'Mechanistic home range models and resource selection analysis: A reconciliation and unification', *Ecology* **89**(4), 1112–1119.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/06-1985.1*

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.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/0012-9658%281999%29080%5B1656%3AHRAUAM%5D2.0.CO%3B2*

Morales, J. M. and Ellner, S. P. (2002), 'Scaling up animal movements in heterogeneous landscapes: The importance of behavior', *Ecology* **83**(8), 2240–2247.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/0012-9658%282002%29083%5B2240%3ASUAMIH%5D2.0.CO%3B2*

Morales, J. M., Haydon, 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 walks', *Ecology* **85**(9), 2436–2445.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/03-0269*

Murrell, D. J., Dieckmann, U. and Law, R. (2004), 'On moment closures for population dynamics in continuous space', *Journal of Theoretical Biology* **229**(3), 421 – 432.
URL: *http://www.sciencedirect.com/science/article/pii/S0022519304001523*

Murrell, D. J. and Law, R. (2003), 'Heteromyopia and the spatial coexistence of similar competitors', *Ecology Letters* **6**(1), 48–59.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1461-0248.2003.00397.x*

Murrell, D., Perrin, A. E. N. and Losos, E. J. B. (2005), 'Local spatial structure and predatorprey dynamics: Counterintuitive effects of prey enrichment', *The American Naturalist* **166**(3), 354–367.
URL: *http://www.jstor.org/stable/10.1086/432035*

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. and Smouse, P. E. (2008), 'A movement ecology paradigm for unifying organismal movement research', *Proceedings of the National Academy of Sciences* **105**(49), 19052–19059.
URL: *http://www.pnas.org/content/105/49/19052*

North, A. and Ovaskainen, O. (2007), 'Interactions between dispersal, competition, and landscape heterogeneity', *Oikos* **116**(7), 1106–1119.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0030-1299.2007.15366.x*

Okubo, A. (1980), *Diffusion and ecological problems : Mathematical models*, Berlin ; New York : Springer-Verlag. "An extended version of the Japanese edition, Ecology and diffusion, [Seitaigaku to kakusan] published in 1975".

Othmer, H. G., Dunbar, S. R. and Alt, W. (1988), 'Models of dispersal in biological systems', *Journal of Mathematical Biology* **26**(3), 263–298.
URL: *https://doi.org/10.1007/BF00277392*

Painter, K. J. (2014), 'Multiscale models for movement in oriented environments and their application to hilltopping in butterflies', *Theoretical Ecology* **7**(1), 53–75.
URL: *https://doi.org/10.1007/s12080-013-0198-0*

Panzacchi, M., Van Moorter, B., Strand, O., Saerens, M., Kivimki, I., St. Clair, C. C., Herfindal, I. and Boitani, L. (2016), 'Predicting the *continuum* between corridors and barriers to animal movements using Step Selection Functions and Randomized Shortest Paths', *Journal of Animal Ecology* **85**(1), 32–42.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1365-2656.12386*

Patlak, C. S. (1953), 'Random walk with persistence and external bias', *The bulletin of mathematical biophysics* **15**(3), 311–338.
URL: *https://doi.org/10.1007/BF02476407*

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O. and Matthiopoulos, J. (2008), 'Statespace models of individual animal movement', *Trends in Ecology*

& *Evolution* **23**(2), 87 – 94.
    URL: *http://www.sciencedirect.com/science/article/pii/S0169534707003588*

Pedersen, M. W., Patterson, T. A., Thygesen, U. H. and Madsen, H. (2011), 'Estimating animal behavior and residency from movement data', *Oikos* **120**(9), 1281–1290.
    URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2011.19044.x*

Platen, E. and Bruti-Liberati, N. (2010), *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*, number 64 *in* 'Stochastic Modelling and Applied Probability', Springer-Verlag Berlin Heidelberg.

Plummer, M., Best, N., Cowles, K. and Vines, K. (2006), 'Coda: Convergence diagnosis and output analysis for mcmc', *R News* **6**(1), 7–11.
    URL: *https://journal.r-project.org/archive/*

Potts, J. R., Bastille-Rousseau, G., Murray, D. L., Schaefer, J. A. and Lewis, M. A. (2014), 'Predicting local and non-local effects of resources on animal space use using a mechanistic step selection model', *Methods in Ecology and Evolution* **5**(3), 253–262.
    URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12150*

Potts, J. R., Brger, L., Scantlebury, D. M., Bennett, N. C., Alagaili, A. and Wilson, R. P. (2018), 'Finding turning-points in ultra-high-resolution animal movement data', *Methods in Ecology and Evolution* **9**(10), 2091–2101.
    URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13056*

Potts, J. R., Hillen, T. and Lewis, M. A. (2016), 'The "edge effect" phenomenon: deriving population abundance patterns from individual animal movement decisions', *Theoretical Ecology* **9**(2), 233–247.
    URL: *https://doi.org/10.1007/s12080-015-0283-7*

Potts, J. R. and Lewis, M. A. (2016*a*), 'How memory of direct animal interactions can lead to territorial pattern formation', *Journal of The Royal Society*

Interface **13**(118).
URL: *http://rsif.royalsocietypublishing.org/content/13/118/20160059*

Potts, J. R. and Lewis, M. A. (2016*b*), 'Territorial pattern formation in the absence of an attractive potential', *Journal of Mathematical Biology* **72**(1), 25–46.
URL: *https://doi.org/10.1007/s00285-015-0881-4*

Potts, J. R., Mokross, K. and Lewis, M. A. (2014), 'A unifying framework for quantifying the nature of animal interactions', *Journal of The Royal Society Interface* **11**(96).
URL: *http://rsif.royalsocietypublishing.org/content/11/96/20140333*

Potts, J. R., Mokross, K., Stouffer, P. C. and Lewis, M. A. (2014), 'Step selection techniques uncover the environmental predictors of space use patterns in flocks of amazonian birds', *Ecology and Evolution* **4**(24), 4578–4588.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.1306*

Potts, J. R. and Petrovskii, S. V. (2017), 'Fortune favours the brave: Movement responses shape demographic dynamics in strongly competing populations', *Journal of Theoretical Biology* **420**, 190 – 199.
URL: *http://www.sciencedirect.com/science/article/pii/S002251931730125X*

Preisler, H. K., Ager, A. A., Johnson, B. K. and Kie, J. G. (2004), 'Modeling animal movements using stochastic differential equations', *Environmetrics* **15**(7), 643–657.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/env.636*

Preisler, H. K., Ager, A. A. and Wisdom, M. J. (2013), 'Analyzing animal movement patterns using potential functions', *Ecosphere* **4**(3), art32.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/ES12-00286.1*

R Core Team (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

Rhodes, J. R., McAlpine, C. A., Lunney, D. and Possingham, H. P. (2005), 'A spatially explicit habitat selection model incorporating home range behavior', *Ecology* **86**(5), 1199–1205.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/04-0912*

Riotte-Lambert, L., Benhamou, S. and Chamaill-Jammes, S. (2015), 'How memory-based movement leads to nonterritorial spatial segregation', *The American Naturalist* **185**(4), E103–E116. PMID: 25811090.
URL: *https://doi.org/10.1086/680009*

Sawyer, H. and Kauffman, M. J. (2011), 'Stopover ecology of a migratory ungulate', *Journal of Animal Ecology* **80**(5), 1078–1087.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2656.2011.01845.x*

Sawyer, H., Kauffman, M. J., Nielson, R. M. and Horne, J. S. (2009), 'Identifying and prioritizing ungulate migration routes for landscape-level conservation', *Ecological Applications* **19**(8), 2016–2025.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/08-2034.1*

Sawyer, H., Lindzey, F. and McWhirter, D. (2005), 'Mule deer and pronghorn migration in western wyoming', *Wildlife Society Bulletin* **33**(4), 1266–1273.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.2193/0091-7648%282005%2933%5B1266%3AMDAPMI%5D2.0.CO%3B2*

Sawyer, H., Nielson, R. M., Lindzey, F. and McDonald, L. L. (2006), 'Winter habitat selection of mule deer before and during development of a natural gas field', *The Journal of Wildlife Management* **70**(2), 396–403.
URL: *http://www.jstor.org/stable/3803685*

Schick, R. S., Loarie, S. R., Colchero, F., Best, B. D., Boustany, A., Conde, D. A., Halpin, P. N., Joppa, L. N., McClellan, C. M. and Clark, J. S. (2008), 'Understanding movement data and movement processes: current and emerging directions', *Ecology Letters* **11**(12), 1338–1350.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1461-0248.2008.01249.x*

Schlägel, U. E. and Lewis, M. A. (2016), 'Robustness of movement models: can models bridge the gap between temporal scales of data sets and behavioural processes?', *Journal of Mathematical Biology* **73**(6), 1691–1726.
URL: *https://doi.org/10.1007/s00285-016-1005-5*

Scrafford, M. A., Avgar, T., Heeres, R. and Boyce, M. S. (2018), 'Roads elicit negative movement and habitat-selection responses by wolverines (gulo gulo luscus)', *Behavioral Ecology* **29**(3), 534–542.
URL: *http://dx.doi.org/10.1093/beheco/arx182*

Shariatinajafabadi, M., Wang, T., Skidmore, A. K., Toxopeus, A. G., Klzsch, A., Nolet, B. A., Exo, K.-M., Griffin, L., Stahl, J. and Cabot, D. (2014), 'Migratory herbivorous waterfowl track satellite-derived green wave index', *PLOS ONE* **9**(9), 1–11.
URL: *https://doi.org/10.1371/journal.pone.0108331*

Signer, J., Fieberg, J. and Avgar, T. (2017), 'Estimating utilization distributions from fitted step-selection functions', *Ecosphere* **8**(4), e01771.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/ecs2.1771*

Smith, B. J. (2007), 'boa: An r package for mcmc output convergence assessment and posterior inference', *Journal of Statistical Software* **21**(11), 1–37.

Smouse, P. E., Focardi, S., Moorcroft, P. R., Kie, J. G., Forester, J. D. and Morales, J. M. (2010), 'Stochastic modelling of animal movement', *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **365**(1550), 2201–2211.
URL: *http://rstb.royalsocietypublishing.org/content/365/1550/2201*

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), 583–639.
URL: *https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00353*

Statisticat and LLC. (2018), *LaplacesDemon Tutorial*. R package version 16.1.1.
URL: *https://web.archive.org/web/20150206004624/http://www.bayesian-inference.com/software*

Thurfjell, H., Ciuti, S. and Boyce, M. S. (2014), 'Applications of step-selection functions in ecology and conservation', *Movement Ecology* **2**(1), 4.
URL: *https://doi.org/10.1186/2051-3933-2-4*

Turchin, P. (1991), 'Translating foraging movements in heterogeneous environments into the spatial distribution of foragers', *Ecology* **72**(4), 1253–1266.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1941099*

Turchin, P. (1998), *Quantitative analysis of movement : measuring and modeling population redistribution in animals and plants*, Sunderland, Mass. : Sinauer Associates. Includes bibliographical references (p. [349]-383) and index.

van der Graaf, S. A., Stahl, J., Klimkowska, A., Bakker, J. P. and Drent, R. H. (2006), 'Surfing on a green wave - how plant growth drives spring migration in the Barnacle Goose *branta leucopsis*', *Ardea* **94**(3), 567–577.

Van Moorter, B., Visscher, D., Benhamou, S., Brger, L., Boyce, M. S. and Gaillard, J.-M. (2009), 'Memory keeps you at home: a mechanistic model for home range emergence', *Oikos* **118**(5), 641–652.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2008.17003.x*

van Wijk, R. E., Klzsch, A., Kruckenberg, H., Ebbinge, B. S., Mskens, G. J. D. M. and Nolet, B. A. (2012), 'Individually tracked geese follow peaks of temperature acceleration during spring migration', *Oikos* **121**(5), 655–664.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2011.20083.x*

Vanak, A. T., Fortin, D., Thaker, M., Ogden, M., Owen, C., Greatwood, S. and Slotow, R. (2013), 'Moving to stay in place: behavioral mechanisms for coexistence of african large carnivores', *Ecology* **94**(11), 2619–2631.
URL: *https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/13-0217.1*

Viana, D. S., Granados, J. E., Fandos, P., Pérez, J. M., Cano-Manuel, F. J., Burón, D., Fandos, G., Aguado, M. Á. P., Figuerola, J. and Soriguer, R. C. (2018), 'Linking seasonal home range size with habitat selection and movement

in a mountain ungulate', *Movement Ecology* **6**(1), 1.
URL: *https://doi.org/10.1186/s40462-017-0119-8*

Wang, Y.-S. and Potts, J. R. (2017), 'Partial differential equation techniques for analysing animal movement: A comparison of different methods', *Journal of Theoretical Biology* **416**, 52 – 67.
URL: *http://www.sciencedirect.com/science/article/pii/S0022519317300024*

Wilson, K., Hanks, E. and Johnson, D. (2018), 'Estimating animal utilization densities using continuous-time markov chain models', *Methods in Ecology and Evolution* **9**(5), 1232–1240.
URL: *https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12967*

Zweifel-Schielly, B., Kreuzer, M., Ewald, K. C. and Suter, W. (2009), 'Habitat selection by an alpine ungulate: the significance of forage characteristics varies with scale and season', *Ecography* **32**(1), 103–113.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0587.2008.05178.x*