UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

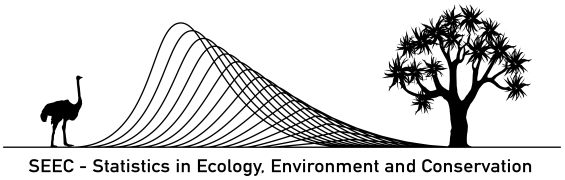
Using movement modelling to improve the design and analysis of vantage point surveys in bird and wind energy studies

Francisco Cervantes Peralta

Supervisors:
Dr. Birgit Erni, Dr. Theoni Photopoulou and Dr. Rob Simmons

Thesis presented for the degree of
DOCTOR OF PHILOSOPHY

in the Department of Statistical Sciences
UNIVERSITY OF CAPE TOWN



October 2020

Declaration

I, Francisco Cervantes Peralta, hereby declare that this thesis is my own unaided work, both in concept and execution, and that apart from the normal guidance from my supervisors, I have received no assistance except for:

- Dr Megan Murgatroyd kindly provided the Verreaux's Eagle GPS tracking data.
- Jon Smallie (WildSkies Ecological Services) kindly provided the wind farm monitoring data.

Every effort has been made to attribute any contributions of others, including references to the literature.

Neither the substance nor any part of the thesis has been in the past, or is being, or is to be submitted for any other degree at this University, or any other university.

Signed:

Francisco Cervantes Peralta
13th of October 2020

This thesis is dedicated to my grandparents.

Using movement modelling to improve the design and analysis of vantage point surveys in bird and wind energy studies

Francisco Cervantes Peralta

Abstract

Wind energy, although mostly a clean and increasingly efficient energy source, is known to affect communities of flying vertebrates. Mortality by collision with turbines is one of the main impacts on birds and bats associated with wind energy. Soaring birds are particularly vulnerable due to their collision prone behaviours, low manoeuvrability, and their slow population recovery rates. The focus of this thesis is on the identification of areas that are intensively used by soaring birds in order to inform wind turbine placement and minimize collision risk.

This thesis is particularly concerned with predictions of bird-use intensity that are based on flight trajectories mapped by observers from vantage points. This survey technique is standard practice during the environmental impact assessment of wind energy facilities, although its virtues and limitations are largely untested. Flight trajectories are counted, timed and mapped during these surveys. However, most assessments ignore the spatial information contained in the trajectories, and mappings are often reduced to metrics such as closest distance to a turbine or whether a particular habitat is visited. In this thesis, I use visual mappings of flight trajectories to estimate the long-term distribution of bird activity using: i) a kernel density estimator adapted to calculate the density of flight trajectories, and ii) modelling flights as being driven by a stochastic process under the influence of a potential field. Acknowledging the subjectivity introduced in the mapping of trajectories by field observers, I also study the discrepancy between mapped and true trajectories. Finally, I showcase the application of the various analytical techniques with a case study, in which I compare collision risk predictions with actual observed fatalities at a wind farm in South Africa.

Kernel density estimation proved to be a good exploratory technique, and the estimator designed to estimate trajectory density outperformed other methods that ignore the temporal structure in trajectory data. Nevertheless, kernel methods are limited by its inability to predict bird activity outside areas observed from vantage points. Potential-based models allowed predictions in unobserved areas based on landscape characteristics, and showed promising results identifying areas of high collision risk. I found that the

difference between true and mapped trajectories can be substantial, and it should be accounted for in any spatial analysis of vantage point observations.

Although based on a single study case, the results are promising and show that the spatial distribution of collision risk predicted with the suite of methods presented in this thesis correlates well with the distribution of observed fatalities. The framework proposed to predict collision risk improves existing procedures in that it uses movement and spatial information contained in the observed trajectories. In addition, it accounts for all known sources of uncertainty throughout the modelling process.

Acknowledgements

I feel sincerely privileged for having the opportunity to undertake this PhD journey. I have many people to thank and anything I write here will fall short. My supervisors Dr Birgit Erni, Dr Theoni Photopolou and Dr Rob Simmons combine a set of skills that uplifted my work in every step of the way. They patiently accompanied me as I decided to jump into all sorts of muddy waters. Far from leading nowhere, this enriched my understanding of science beyond what is captured by this thesis. Thank you so much for your guidance, support and patience.

My sincere gratitude goes to Prof. Res Altwegg and Prof. Francesca Little for their continuous support. Many thanks go to Dr Megan Murgatroyd for sharing with me her Verreaux's Eagle tracking data, her passion for the species and our friendly chats. I extend my gratitude to Sam Ralston-Paton (BirdLife South Africa) who has always gone out of her way to help me with my research. Jon Smallie (WildSkies) provided the wind farm monitoring data and has unconditionally been open to share his experience with me. Thank you. To Mark and Carol Duckitt (Rondeberg Private Nature Reserve), I need to thank their help with the drone experiments and also their kind hospitality.

Vicky Garía Pérez has been always been there for me, through innumerable outside and inside boas. Thank you for your support, love and sense of humour. I have a family that I don't deserve, and too many names should go here, thank you all. Special thanks go to my mother, always having my back, and without whom this would have just not been possible, to my grandmother María, who's been a second mother to me, and to my grandfather Pascual, who has been an unattainable role model. I cannot forget Wanda and Aston, my two loyal friends. To them I have to thank their company and routine beach walks; they are possibly responsible for any good ideas expressed in this thesis. Neither I should forget my friends and colleagues, thanks for being there — losing a few neurons to beer, has been as important as nurturing the survivors. A special mention goes to Marlei Martins who in early days fanned the flames and helped materialize this journey.

This thesis has been funded by the NRF through an Innovation Master's Scholarship and by a Science Faculty PhD Fellowship. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Contents

Chapter 1

Introduction

Understanding the distribution of organisms in space and time, as well as their movements, is a central issue in ecology and wildlife conservation (Levin *et al.*, 2009; Ovaskainen *et al.*, 2016). Being able to anticipate the location of organisms allows a better management of human activity, and reduces the negative effects of our technological achievements on ecological systems. Impacts associated with wind energy production on avian populations are an example of conflicting interactions between human development and wildlife. Collision with turbine blades and displacement (by either exclusion or obstruction to movement) from areas occupied by infrastructure are the main threats associated with wind energy production (Drewitt & Langston, 2006). In this context, understanding the distribution and abundance of birds, would contribute to the sustainable growth of a renewable and efficient energy source that has a low carbon footprint.

The distribution of wildlife populations can be studied at different levels. Individual animals make choices during their daily activities that produce emerging patterns at the population level. For example, Ovaskainen *et al.* (2016) show how different movement modes of individual butterflies result in butterfly populations to be concentrated in meadows, and not in forests. In other words, systematic behaviours of individuals produce emerging patterns at the population level. In general, both, extrapolating from individual behaviour to population patterns and *vice versa* is problematic, and the best approach to take depends on the question at hand. Research has shown that wildlife and wind energy interactions are driven, not only by bird or wind turbine abundance,

but also by specific behaviours that make certain species more vulnerable than others (Schuster *et al.*, 2015). In this sense, studying population level distribution or abundance of birds may provide only a partial picture of the mechanisms leading to collisions or displacement. In contrast, studying behaviour from the point of view of the individual bird, could target the right questions to understand how birds are exposed to wind energy impacts (Ferrer *et al.*, 2012; Sur *et al.*, 2018).

Although not the main focus of this thesis, telemetry is intimately linked to the study of animal movement and behavioural research, and therefore, the topic will be touched upon occasionally. There exist different bird-borne telemetry systems, such as: Global Positioning System (GPS), Argos or very high frequency radio (VHF). Throughout this text I will refer mostly to GPS technology when talking about telemetry, unless stated otherwise. Modern tracking technology allows the study of animal movement by acquiring location data with increasing accuracy and definition (higher acquisition rates). Telemetry is uncovering key aspects of bird movement that help managers design general strategies to minimize the exposure of birds to wind energy development (e.g. Katzner *et al.*, 2012; Reid *et al.*, 2015; Poessel *et al.*, 2018). This technological advent has also raised interest on a breadth of analytical techniques that can be applied to investigate animal movement and population redistribution (Ovaskainen *et al.*, 2016).

However, despite its rapid development, telemetry devices cannot always be deployed efficiently to investigate the potential negative effects of particular wind energy projects on local bird populations (Hebblewhite & Haydon, 2010; Katzner *et al.*, 2016a). The main reasons for telemetry being impractical at this scale are: i) the costs associated with tracking the multiple species that may be affected by a single wind energy project, and ii) the project development time frames not being long enough for this type of study. Notably, a different type of telemetry, radar, is a promising tool for accurately capturing bird flight paths (May *et al.*, 2017). However, its use has been limited to a few cases due to its high deployment costs (Katzner *et al.*, 2016a). Because it is believed that local environmental conditions are the main drivers of bird collisions and displacement, environmental impact practitioners resort to visual surveys of bird activity from vantage points to collect movement information at the wind farm level (Strickland *et al.*, 2011; Jenkins *et al.*, 2015; Katzner *et al.*, 2016a). Visually-observed flight trajectories contain movement and behavioural information, yet no attempts have been made to analyse these data within a movement ecology framework. This thesis aims at filling this gap.

Bird-borne telemetry devices typically collect data constantly (i.e. with a constant or nearly constant frequency) over longer periods of time than vantage point surveys do.

The accuracy of telemetry devices on the location of birds is also superior to that of visual observers. However, vantage point surveys are logistically simpler to implement, they are more flexible in terms of the species they target, multiple species may be studied at the same time, in the precise location where data is needed, and it is a survey method that is not invasive to the birds. In this manuscript, I explore the potential of visually mapped trajectories to predict bird behaviour and assess wind farm impacts on avian communities. Vantage point surveys are the industry standard method to collect bird movement information; and therefore, developing analytical methods for these surveys is critical for a more sustainable development of wind energy. Vantage point surveys are the main focus of this thesis, although the techniques presented here could easily be adapted for the analysis of telemetry data, should it be available.

1.1 Birds and wind energy

In 2019, renewable energy accounted for more than a quarter of the total electricity generation in the World (REN21, 2019). Technological advances have made wind and solar energy more competitive than ever before, now being cheaper than fossil fuel based alternatives. The share of renewable energy technologies is expected to grow even larger, as experts warn that the world is behind schedule to meet the target of limiting global temperature increase to less than 1.5 °C (UNFCCC, 2015). Wind energy is the second most important renewable energy source, in terms of capacity installed, only behind hydropower (REN21, 2019). Commercial-scale wind energy production developed in the 1980's in the United States (Kaldellis & Zafirakis, 2011) and in 2019 it accounts for around 5.5 % of the total energy produced in the World.

Being at the forefront of climate change mitigation, it is necessary to consider some ecological impacts associated with wind energy production. In the late 1980's the first bird fatalities produced by collisions with wind turbines were reported in the United States (see Erickson *et al.*, 2001, and references therein). Numerous fatalities related mostly to Red-tailed Hawk *Buteo jamaicensis*, American Kestrel *Falco sparverius* and Golden Eagle *Aquila chrysaetos* in Altamont Pass, California, lead to closer scrutiny of other wind farms worldwide. Shortly after, in the 1990's, high numbers of collisions of Griffon Vulture *Gyps fulvus* and Common Kestrel *Falco tinnunculus* with wind turbines were reported in Tarifa and Navarra, Spain (see Langston & Pullan, 2003, and references therein). Later, between 2000-2010 a significant reduction of the population of

White-tailed Eagles at the Smøla archipelago occurred after the construction of a wind energy facility (Bevanger *et al.*, 2010). From the 13 eagle pairs present in the vicinity of the Smøla wind power plant in 2005, only four remained in 2010. This reduction was produced by both collision mortality and displacement. Other important cases have been reported; for example collisions of Egyptian Vultures *Neophron percnopterus* in Tarifa (Carrete *et al.*, 2009) and Griffon Vulture in Castellón, both in Spain (Martínez-Abraín *et al.*, 2012). In South Africa, only the results of two years of monitoring of eight wind farms have been reported to date (Ralston-Paton *et al.*, 2017; Perold *et al.*, 2020), and they raise some concern about the number of fatalities of Verreaux's Eagle *Aquila verreaxii*, Black Harrier *Circus maurus* and Jackal Buzzard *Buteo rufofuscus*.

Despite these alarming numbers, most wind energy facilities around the world are considered relatively safe for birds (Drewitt & Langston, 2006). Put into perspective, the number of birds killed by wind turbines are small compared to other human activities (Loss *et al.*, 2015). These assertions, however, must be taken cautiously, because mortality data come mostly from some wind farms located in Europe and North America (Arnett *et al.*, 2007; Thaxter *et al.*, 2017; Loss *et al.*, 2019). There is also uncertainty about what population-level, cumulative and synergistic effects individual wind farms may have on bird populations (May *et al.*, 2019). However, the few cases analysed to date point towards populations of long-lived species being sensitive to even small number of fatalities produced by wind farms, most notably when these species are already endangered or under extraordinary environmental pressure (Carrete *et al.*, 2009; Martínez-Abraín *et al.*, 2012; Katzner *et al.*, 2016b). Raptors, vultures and other soaring birds (from now referred to as raptors), account for a small number of fatalities compared to other bird species; however, they are considered particularly vulnerable to impacts produced by wind energy (see Schuster *et al.*, 2015). Raptors are typically long-lived, mature slowly and produce few offspring, making their populations easy to destabilize by additional mortality.

Collision risk seems to be related to complex interactions between bird behaviour and particular characteristics of the environment (Schuster *et al.*, 2015). There is some agreement among experts that concentration of raptors around wind turbines favours collision risk, for example, along migratory routes (Desholm, 2009; Thaxter *et al.*, 2017) and roosting or breeding sites (Schuster *et al.*, 2015). However, bird abundance does not always correlate with number of collisions (de Lucas *et al.*, 2008; Ferrer *et al.*, 2012). Collisions seem to be associated with particularly distracting behaviours, such as hunting or interactions with conspecifics (Marques *et al.*, 2014; Schuster *et al.*, 2015). The

location of the turbine on the landscape may also play a role, because soaring birds rely on air currents, produced by thermal convection or slope updraughts to sustain their flights (de Lucas *et al.*, 2012; Marques *et al.*, 2014; Schuster *et al.*, 2015). It is hypothesised that weak updraughts, produced by low temperatures or slow wind speeds, may decrease manoeuvrability of soaring birds and, therefore, increase the risk of collision with turbine blades (Barrios & Rodríguez, 2004; Schuster *et al.*, 2015).

Several measures to reduce collision risk have been proposed, such as: stopping turbines on demand when target bird species approach (de Lucas *et al.*, 2012), painting turbine blades with certain patterns to increase their visibility (Hodos, 2003; Young *et al.*, 2003) or managing the location of food sources to direct birds' flights (Martínez-Abraín *et al.*, 2012). The effectiveness of these measures is site-specific and not extensively tested, and it is agreed that it is best to avoid placing turbines in those areas that are used by birds intensively to avoid collisions (Northrup & Wittemyer, 2013; Arnett & May, 2016). This is achieved at two levels: i) strategic regional plans for wind energy production should incorporate the distribution of particularly sensitive bird species into the development plans for the wind energy sector (i.e. at regional scale), and ii) at the wind farm level, the turbine layout should be optimized to avoid those locations in the landscape that are most likely to be used by sensitive bird species (i.e. at local scale). Although planning at the regional scale should prevent major impacts on bird populations, our understanding of bird distributions is limited, and in practice, bird conservation is not always the first priority of national or regional administrations. It is then likely that assessments at the local scale are needed to further prevent impacts on birds.

Data at the local scale are typically collected within the environmental impact assessment of the wind farms (Arnett *et al.*, 2007; Arnett & May, 2016). Visual surveys are conducted from vantage points to collect information on bird abundance and behaviour (e.g. Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). The basic idea is to place observers at points of good visibility to collect information about the movements of target bird species that cross the area surrounding the vantage point, typically up to a distance of 2 kilometres. The information recorded depends on the objective of the surveys, but typically includes number of birds, time spent flying around the vantage point and the mapping of flight paths on topographic maps or ortho-photographs. Because of their flexibility and relatively easy logistics, vantage point surveys are popular for studying the activity of large soaring birds. However, data captured in this way are subject to errors in the recorded position and imperfect detection of birds. In addition, flight paths are often reduced to some point metric of interest,

such as minimum distance to turbines or to some landscape element, leading to the loss of important spatial information.

To my knowledge, there are no published studies that analyse collision risk or displacement of birds based on the characteristics of visually recorded flight paths as a whole (but see Masden *et al.*, 2012, for a case study with radar). Understanding the drivers of bird movement, helps us understand the decisions made by the bird at each point in time and space. This mechanistic approach to study movement makes it possible to identify collision prone situations and avoidance behaviours. In addition, understanding the structure of errors associated with mapping of flights by human observers, allows us to account for the added uncertainty in inference drawn from vantage point data. Techniques used to model dynamical systems, such as stochastic differential equations or state-space models, widely used in animal movement modelling, prove to be key in unravelling these complex relationships between animal behaviour, landscape properties and human perception.

1.2 Vantage point surveys

Vantage point surveys are central to the environmental impact assessment of wind energy facilities around the world. These surveys are used to study the movements of large, soaring birds; species with broad home ranges that occur in rather low densities (Bibby *et al.*, 2000; Hardey *et al.*, 2009). Manuals and guidelines have been published describing the objectives and implementation of vantage point surveys to assess impacts associated with wind energy (e.g. Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). Broadly, from a point of good visibility, an observer records the presence and activity of all birds that belong to a list of target species. Surveys usually last for hours and the areas overlooked from the vantage points are large and often heterogeneous in terms of habitat and topography. Unlike regular point counts, which generally target small species, during vantage point surveys, large birds in flight can be tracked for several minutes, recording not only bird occurrence (counts or passage rates), but also movement paths and behaviours.

It is extremely difficult to keep track of all the vantage point sampling designs and analyses, because they largely depend on the objectives of each study. In addition, most environmental impact assessment studies are never published. Taking as a reference what

has been published in the scientific literature, we see that vantage points are sometimes used to record those birds that move close to or across turbine rows, without taking into account fine-scale spatial information about the actual flight trajectory (e.g. Barrios & Rodríguez, 2004; de Lucas *et al.*, 2008; Smallwood *et al.*, 2009). These studies focus on the occurrence of high risk flights when turbines are already built or at least their location is known. Alternatively, the focus of vantage point surveys may be in those birds that fly over some defined plot that might or might not contain turbines (e.g. around the observer). In this case, observers usually record birds up to a certain distance (usually less than 1-2 km) (e.g. Hull & Muir, 2013; Dahl *et al.*, 2013; Johnston *et al.*, 2014), but in some studies observation are conducted without a distance limit (e.g. Walker *et al.*, 2005; Garvin *et al.*, 2011). Within this approach it is common to record flight trajectories in addition to abundance, although in practice, trajectories are often reduced to points of interests for their analysis, such as the point closest to a turbine (Hull & Muir, 2013) or the point of first observation (Walker *et al.*, 2005).

Counting crossings through turbine rows or over a plot is simple, and the analysis of these data is typically straight forward. For example, we may use generalized linear models to associate abundance of flights with environmental conditions, behaviour type or location of the vantage point, etc (e.g. see Barrios & Rodríguez, 2004; Garvin *et al.*, 2011). However, this approach implicitly assumes a homogeneous distribution of the activity within observed plots and may be failing to capture information about processes occurring at finer scale driven by heterogeneous landscape (e.g. topographic slopes within the plots may be used more often than flat areas). Particularly, in those studies that aim at making predictions about where collisions may occur in the future, sometimes without a clear reference of where turbines will be erected, it is desirable to build activity intensity maps. For example, in pre-construction studies turbines are absent and sometimes their location is conditioned to the results of environmental impact assessments. Therefore, fine-scale analysis of space use and selection occurring within the typically large extent of vantage point survey plots are necessary. That I am aware of, the only published study that uses full flight trajectories to make a spatial analysis of bird activity was conducted by Walker *et al.* (2005). I find that this work illustrates well the methodologies that I have encountered in (unpublished) environmental impact assessment studies that use flight trajectories to inform turbine locations.

Walker *et al.* (2005) recorded flight trajectories of Golden Eagles during the pre-construction, construction and post-construction periods of a wind farm. A grid was super-imposed over the flight paths and the length of flights inside each cell was computed. As I un-

derstand it, they did not account for spatial or temporal autocorrelation in scores of contiguous grid cells produced by the same flight running across several cells. Given that their analysis was mostly exploratory, there is nothing wrong with the methodology; however, accounting for autocorrelation produced by the movement of the eagles could produce more complete inference on their use of space. In addition to the grid analysis, they also sub-sampled the trajectories to obtain locations separated at least 45 minutes. These locations were considered independent from each other and were analysed using minimum convex polygons and kernel density estimation. Sub-sampling the data to "make them independent" is a common approach that, although not wrong, does not use data to their full potential.

It is worth noticing that radar surveillance has also made its way into the research of birds and wind energy interactions as an alternative to, or to complement, vantage point surveys (Masden *et al.*, 2012; Liechti *et al.*, 2013; Cabrera-Cruz *et al.*, 2016; May *et al.*, 2017; Aschwanden *et al.*, 2018; Jenkins *et al.*, 2018). Radar systems can monitor three-dimensional volumes continuously for long periods of time (night and day) and are able to provide information on bird location with little error. However, the identification of species is not always possible without ground-truthing and the quantification of birds requires the prior calibration of the radar to match signal strength with bird size (Schmaljohann *et al.*, 2008; Gürbüz *et al.*, 2015). The method also suffers from certain bias, such as imperfect probability of detection changing with distance or interference with ground clutter and background signals, such as insects (Schmaljohann *et al.*, 2008; Dokter *et al.*, 2013; May *et al.*, 2017). Depending on specific objectives, field deployment of the equipment might be logistically and economically challenging (Gürbüz *et al.*, 2015). However, technology is advancing quickly and we may soon see more radar studies. With them there would be a greater volume of movement data, which could benefit from similar analytical principles to those expressed in this thesis.

1.3 Animal movement modelling

Biological processes are fundamentally dynamic and mobile organisms move to respond and adapt to changes in the environment (Nathan *et al.*, 2008). Thus, understanding movement and the drivers behind movement processes is fundamental to construct a complete picture of the biological needs and preferences of wildlife. The large volumes of telemetry data produced in the last decades have made it possible to approach ecological

questions from the perspective of the moving individual. The result is the establishment of a branch of ecological research called movement ecology (Nathan *et al.*, 2008), which is a valuable asset in conservation of biodiversity and environmental impact assessment (Barton *et al.*, 2015).

Animal movement bouts can be represented by paths, also called trajectories (Calenge *et al.*, 2009; Turchin, 2015). Movement models investigate the properties of these trajectories and due to our limited understanding of animal behaviour, invariably incorporate a stochastic component (Turchin, 2015; Ovaskainen *et al.*, 2016; Hooten *et al.*, 2017). Although animal movements describe continuous trajectories, often we can only observe a sequence of discrete locations at certain times (e.g. GPS tracking devices provide locations with certain frequency). Thus, for modelling purposes and for convenience, movement trajectories are often represented by a series of steps (Turchin, 2015; McClintock *et al.*, 2014; Hooten *et al.*, 2017). Thus, movement models can be broadly classified as describing movement in continuous time or in discrete time, depending on whether they describe the movement process as evolving continuously in time or at a finite number of steps.

Modelling these data as a process that occurs in discrete time has some advantages. Discrete-time models are considered more intuitive to specify and analyse, partly because practitioners may use well developed time series analytical techniques that are readily implemented in many statistical packages (McClintock *et al.*, 2014; Hooten *et al.*, 2017). Another technical advantage of discrete-time models is that the specification of switching movement behaviours is relatively easy to associate with discrete locations, compared to considering continuous behavioural changes (Langrock *et al.* 2012; McClintock *et al.* 2014; see Blackwell 1997; Johnson *et al.* 2008; Parton & Blackwell 2017; Michelot & Blackwell 2019 for examples of continuous-state switching in continuous-time). On the flip side, movement steps are usually described by step lengths and turning angles and these are affected by the time between locations. Therefore, inference drawn from discrete-time models is affected by the sampling frequency of the observation process and extra care is necessary to ensure that the time scale of the ecological process of interest matches that of the sampling rate (McClintock *et al.*, 2014; Hooten *et al.*, 2017). In addition, the observation of locations along trajectories often occurs at irregular time periods, which requires regularizing the data before fitting discrete time models.

Continuous-time models are often regarded as being technically more challenging to understand than discrete-time models, because they are based on stochastic differential equations, which may be elusive to ecologists and conservation practitioners (McClintock

et al., 2014). However, continuous-time models offer important advantages. Perhaps, the most important one is that inference from these models is robust to the process by which movement trajectories are observed (McClintock *et al.*, 2014; Hooten *et al.*, 2017). Continuous-time models accommodate irregularly-sampled data, without the need of pre-processing and facilitate the comparison of inference drawn from different studies. Continuous-time models are also considered to run faster than their discrete-time counterparts (McClintock *et al.*, 2014); however, in practice, I find this to be highly dependent on the purposes of the modelling, the software available and the coding skills of the modeller. Another characteristic attributed to continuous-time models is that they usually impose heavier correlation between speed and directional persistence than discrete-time models (McClintock *et al.*, 2014). This is because discrete-time models consider the distribution of step lengths to be independent from the distribution of turning angles, while in continuous time models speed and directional persistence are usually correlated (McClintock *et al.* 2014, but see Parton & Blackwell 2017).

Although identification of different behaviours would certainly be an interesting avenue to explore, in this thesis, I will be particularly concerned with inference about the intensity with which different areas are used by birds. Perhaps the most popular discrete-time models that link animal movement to the distribution of resources in the landscape are step selection functions (SSF, e.g. Fortin *et al.* 2005; Avgar *et al.* 2016). In essence, these models compare the locations selected by the animal (i.e. observed locations) with a sample of locations taken in the vicinity of those observed (i.e. available locations). Comparing the resources present at observed vs. available locations we may learn about resource selection. In continuous-time, selective use of space (e.g. resource selection) may be investigated using potential functions (Brillinger *et al.*, 2002; Preisler *et al.*, 2013; Russell *et al.*, 2018). These models consider that the animal moves on a surface, which is referred to as potential surface, and the motion is governed by Newtonian physics (i.e. gravity pulls the bird downwards). Thus, the gradient of the potential surface influences the movement of the animal, who is directed towards points of low potential. The objective is to understand whether the position of certain elements of the landscape corresponds to areas of low potential, indicating selection, or to areas of high potential, indicating avoidance.

Throughout this thesis I will be particularly concerned with potential-based models. I choose to model movement in continuous-time because integrating data coming from different sources, sampled at different, and potentially irregular intervals, will be important. Trajectories observed from vantage points can be captured with an arbitrarily

small frequency and very often raw data are in the form of continuous trajectories (i.e. trajectories physically drawn on paper maps). Some form of discretization is necessary when digitizing data, but otherwise we may consider these trajectories to be continuous. Therefore, a continuous-time description of the movement process seems more natural.

Potential-based movement models

Potential-based movement models are a type of diffusion models, which are characterized by incorporating a stochastic process called Brownian motion (Nelson, 2001; Särkkä & Solin, 2019). A Brownian motion is a stochastic process with state $\mathbf{B}_t$, at time t, and evolves continuously in time, with independent increments that have a Gaussian distribution, such that the increment from time s to time t has distribution $\mathbf{B}_t - \mathbf{B}_s \sim \mathrm{N}(\mathbf{0}, (t-s)\mathbf{I})$, where the vector $\mathbf{0}$ and the identity matrix $\mathbf{I}$ have the appropriate and matching dimensions. Throughout the thesis, I assume that $\mathbf{B}_t$ is two-dimensional and so are the zero vector and identity matrix. In the context of animal movement, if an animal moves according to a Brownian motion it implies that its location can be represented by a random vector $\mathbf{X}_t$. Once we observe a location $\mathbf{X}_s$ at time s, a new location $\mathbf{X}_t$ at time time has a Gaussian distribution with mean $\mathbf{X}_s$ and variance $\sigma^2(t-s)\mathbf{I}$ (note the scaling parameter for the variance that I denote by σ). This formulation assumes no correlation between the two dimensions of motion, although this can be relaxed by specifying a covariance matrix instead of a scaling parameter. I assume independence between dimensions throughout the thesis (at least independence conditional on some model structure, as we will see), so I will not elaborate on this any further. Because $\mathbf{X}_t \sim \mathrm{N}(\mathbf{X}_s, \sigma^2(t-s)\mathbf{I})$, the longer it takes for us to observe the new location (i.e. at time t), the larger it is our uncertainty on where it is going to happen. Note that we could equivalently define a distribution for the increment in $\mathbf{X}_t$ as $\mathbf{X}_t - \mathbf{X}_s \sim \mathrm{N}(\mathbf{0}, \sigma^2(t-s)\mathbf{I})$. In the literature, Brownian motion is also referred to as Wiener process, with notation $\mathbf{W}_t$. I will use the term *Brownian motion* throughout this text, to emphasize its relationship with animal movement.

Brownian motion is a useful building block for more complicated models, which are usually specified as stochastic differential equations (Nelson, 2001; Brillinger *et al.*, 2002; Särkkä & Solin, 2019). For example, by adding a deterministic component to a Brownian motion, we can describe a systematic drift in the movement of an animal that is otherwise random:

$$\mathrm{d}\mathbf{X}_t = \mu \ \mathrm{d}t + \sigma \ \mathrm{d}\mathbf{B}_t. \tag{1.1}$$

In expression 1.1, a small change in the state of the process $\mathbf{X}_t$ (e.g. animal location) is given by a random Gaussian increment with variance $\sigma^2 \ \mathrm{d}t$ and a deterministic increment $\mu \ \mathrm{d}t$. Both the deterministic component and the variance need not be fixed and may change with time and the state of the process, such that,

$$\mathrm{d}\mathbf{X}_t = \mu(\mathbf{X}_t, t) \ \mathrm{d}t + \sigma(\mathbf{X}_t, t) \ \mathrm{d}\mathbf{B}_t. \tag{1.2}$$

These constructs are extremely flexible and useful for modelling the mechanics of animal movement. Broadly, the idea is that the drift component $\mu(\mathbf{X}_t, t)$ induces bias to the motion that can explain preferential directions of movement, while the diffusion term $\sigma(\mathbf{X}_t, t)$ controls how stochastic inputs behave over time and state space. One may then formulate the these terms to test particular hypothesis about animal movement. A popular approach and one that is used repeatedly in this thesis is the mean-reverting drift. This functional form considers the system to have a "preferred" state $\boldsymbol{\theta}$ and therefore it tends to stay in, or otherwise revert to, this state. This model is known as the Ornstein-Uhlenbeck process (Uhlenbeck & Ornstein, 1930) and is formulated as

$$\mathrm{d}\mathbf{X}_t = \beta(\boldsymbol{\theta} - \mathbf{X}_t) \ \mathrm{d}t + \sigma(\mathbf{X}_t, t) \ \mathrm{d}\mathbf{B}_t, \tag{1.3}$$

where β is a parameter that controls the time it takes for the system to revert to its preferred state, and can be considered to control the attraction towards it. This model has been used, for example, to represent animal home-range behaviour (Dunn & Gipson, 1977). Under this model the animal moves randomly, but systematically drifts towards a central point within its territory, defining a Gaussian distribution of activity, centred at a preferred location (e.g. den, nest, centre of the herd, etc.). More complicated forms can be defined for the drift term in equation 1.2. For example, the landscape within which the animal moves may contain different points of attraction or repulsion, defining a potential surface, with areas of low potential being more attractive than areas of high potential (Brillinger *et al.*, 2002; Preisler *et al.*, 2013; Russell *et al.*, 2018). These concepts will also be central throughout this thesis and details will be provided in the corresponding chapters.

Diffusion models are appealing, firstly because they describe the mechanics involved in

animal movement and therefore they allow us to understand (and even quantify) the drivers of movement. In addition, they provide the tools to investigate the properties of the long-term distribution of animal activity, either by analytically (if possible) or numerically finding the solution to the stochastic differential equation 1.2, or by simulation.

When animals move, they change their location continuously in time and space. This is not to say that they move all the time, but that when they do, they move through time and space continuously. In practice, trajectories can be viewed as sequences of consecutive locations observed at a certain rate. Conceivably, this rate could be infinitely small (i.e. continuous movement), although this is often impractical. Instead, continuous trajectories are discretized into a finite number of consecutive locations and finite difference approximations are used, rather than the exact solutions.

Kernel density estimation

Another popular way of looking into long-term activity distributions of moving animals is by using kernel density estimation. This set of techniques considers that areas more frequently visited in the past, are more likely to be visited in the future (Silverman, 1986b). Therefore, looking at the density of observed locations, we may predict where future locations are likely to occur. The issue then becomes how to calculate the relative density of observed locations at each point within our study area. To avoid an uninformative uniform density throughout the whole study area, kernel density estimation considers that each observation contributes to the density only in its vicinity. Furthermore, the density contribution is maximum at the exact location of the observation and decreases with distance to it. This decrease is given by a kernel function, and the final density is calculated by averaging over the contributions of the kernels of all observations. The method is appealing due to its simplicity and because, in principle, no previous knowledge about the ecological system is needed. However, these virtues also limit inference to areas where data were observed and extrapolation to other ranges becomes complicated.

Traditionally, kernel density estimation assumes that data are independent and identically distributed (Silverman, 1986b; Worton, 1989; Laver & Kelly, 2008). This assumption is likely violated by animal movement data, which are temporally autocorrelated, especially when location acquisition rates are small. (Keating & Cherry, 2009) describe a

method to incorporate temporal variation in kernel density estimates, but do not explicitly model temporal autocorrelation. Horne *et al.* (2007) introduced the use of Brownian bridges to analyse animal telemetry data. The technique considers the long-term distribution of animal activity to be given by integrating a "running" kernel that moves along and between observed locations and has larger variance the further away it is from an observed location. Although this approach solves the problem of serial autocorrelation in the observations, it aims at estimating the distribution of activity in the past, during the tracking period, rather than projecting into the future. In other words, it estimates where the animal could have been during the study period, when it was not observed (i.e. between observations), but neither the original Brownian Bridge model, nor its extensions, (e.g. Benhamou, 2011; Kranstauber *et al.*, 2012) project these estimates to times before or after the study period.

The only kernel density estimator, that I am aware of, that incorporates temporal autocorrelation and that can be used to project activity distribution estimates into the future is the Autocorrelated Kernel Density Estimator (AKDE), which was proposed by Fleming *et al.* (2015). The AKDE works under the assumption that the animal moves according to a stationary stochastic process. This means that the average distance between two locations increases with time, up to a point where it stabilizes. This "variance function" is given by the semi-variogram of the process (or equivalently by the autocorrelation function). Although the method is rigorous and highly applicable in the context of home-range studies (Noonan *et al.*, 2019), it relies on being able to estimate a stationary semi-variance function that reaches a plateau after certain time, indicating that the long-term distribution of animal activity is spatially constrained (see Fleming *et al.* 2015 Appendix C). Vantage point data present several characteristics that might be at odds with the stationarity requirements: i) we only observe a portion of the range within which birds move, and therefore, even if their behaviour over the entire range is spatially constrained, it might not appear so over the reduced observed range, ii) we typically observe several trajectories that may correspond to different birds with different ranges, and iii) the individual trajectories are too short to estimate an asymptotic covariance.

In this thesis, I propose a kernel density estimator especially tailored for visually observed flight paths. It considers each trajectory to be a single observation in the space of all possible trajectories to circumvent the problem of temporal autocorrelation.

1.4 Bayesian inference

In statistics, there are two main streams of thought with respect to how uncertainty is defined: the frequentist and the Bayesian. In this thesis, I mainly use the Bayesian paradigm. An exhaustive and rigorous analysis of the differences between the two paradigms is outside of the scope of this document, and rather, I expose the reasons behind the choice of paradigm in this particular study (for further details the reader is referred to Gelman *et al.*, 2014; McElreath, 2019).

The choice of paradigm is based mostly on the use of prior and posterior distributions within the Bayesian paradigm. Before analysing any data, prior distributions allow expressing our beliefs about parameter values in the form of a probability distribution. This probability distribution is "combined" with the likelihood of the data - probability of our data given parameter values - to produce parameter "estimates" in the form of posterior distributions - probability of parameter values given our data. While for some, this is a reason for critique, due to the subjectivity introduced in the estimation process, it has very important regularizing properties (Hooten & Hobbs, 2015). This means that parameter estimates are constrained by the prior to lie within reasonable ranges, given what we previously knew about the process we are studying. In addition, when the data are informative (there is little uncertainty in the likelihood), the likelihood tends to override the prior. Therefore, when strong evidence can be drawn from the data, our prior beliefs lose weight in the final inference. In the context of animal movement, we may express what is known about the ecology of the species under study in the form of priors and include it in the models. This is particularly appealing when dealing with vantage point observations, from which we typically collect a limited number of short flights that might offer limited information about the movement process.

The concepts of prior and posterior distributions are used frequently throughout the thesis, so I provide a quick explanation of their relationship. Assume that model parameters are captured by the random variable (random vector in the case of multiple parameters) Θ and that the probability of observing a certain value θ is given by the probability distribution $f(\Theta = \theta)$, or to simplify notation $f(\theta)$. If it is further assumed that the possible set of observations of a given experiment is captured by the random vector Y and the probability of observing one particular set is $f(Y = y) \equiv f(y)$, then by Bayes' theorem we derive that

$$f(\theta|y) = \frac{f(y|\theta)f(\theta)}{f(y)}, \tag{1.4}$$

which states that the probabilistic distribution of parameter values, given our observed data ($f(\theta|y)$; the posterior distribution) is proportional to the likelihood of our data given the parameters ($f(y|\theta)$), times by the prior distribution of the parameters ($f(\theta)$). The distribution of the data $f(y)$ acts as a normalizing constant that ensures that the posterior distribution integrates to one.

Although under uniform (uninformative) priors one may argue that the likelihood and the posterior should express very similar things, there are important conceptual differences. Notably the likelihood need not integrate to one and therefore is not a pdf. It does not express the probability of observing certain parameter values and so intervals constructed from it, do not represent the probability of the interval containing the "true" parameter value. The probability of an interval containing the true parameter value refers to the technique used for building the interval - e.g. 95% of the intervals contain the true value. Bayes' posterior distributions are pdf and can be handled as such. Therefore, one may integrate the posterior to find the probability of observing a parameter value within an interval (note that we do not consider that there is a "true" parameter value within the Bayesian framework). These concepts will be used frequently throughout this thesis.

1.5 Structure of the thesis

The overarching objective of this thesis is to estimate the long-term activity distribution of soaring birds to inform exposure to collision with wind turbines. I base these estimates on movement trajectories visually recorded from vantage points. However, telemetry data sets are used throughout the thesis to illustrate concepts and also to support statistical analysis.

In chapter 2, I proceed to describe the methods proposed for estimating the expected long-term activity distribution of soaring birds using kernel density estimation. I formulate an estimation procedure adapted to work with samples of short flights, observed during survey sessions, over a defined region of space, while acknowledging that only observations, and not necessarily bird activity, are confined to this region. Then, I study the performance of the estimator on simulated data. These data are simulated from a

known theoretical long-term distribution of activity. As such, it is possible to evaluate the performance of the proposed procedure, and also compare it to that of other existing kernel density estimators.

In chapter 3 I propose a movement model that describes the dynamics of soaring bird movement using stochastic differential equations driven by potential functions. The key issue in this chapter is to investigate how accurately model parameters can be estimated using visually observed flights, as opposed to telemetry data sets. Short flights observed from vantage points offer a limited view of a movement process that potentially spans a much larger area. In addition, the movement process may be driven by factors that are unobservable from a vantage point (e.g. the location of the bird's nest). Working under the assumption that data acquired from a bird-borne GPS transmitter is not subject to these limitations, I use parameters obtained from a data set of two Verreaux's eagles, as a reference to compare the performance of models fitted to simulated vantage point observations.

In chapters 2 and 3, for simplicity, and to tackle one issue at a time, I ignore the fact that visually observed flights are subject to positional errors introduced by the observers. In chapter 4, I present the results of an experiment conducted using a drone as a surrogate for a bird, to investigate the magnitude and dynamics of this error. Generalized linear models and stochastic differential equations are used to model the magnitude of the error at the time of detection of a bird and the evolution of the error with time, respectively.

In chapter 5 I present a case study, where I showcase the application of the methods described in chapters 2 to 4. I analyse a real vantage point data set, recorded during the monitoring programme of a wind farm, in the Eastern Cape Province, in South Africa. Vantage point data were collected during the pre-construction studies of the wind farm. These flight data are coupled with observed mortality recorded during the first year of operation of the wind farm, which opens the opportunity to validate the predictive power of the proposed methods.

Chapters are written, for the most part, as stand-alone documents. This means that concepts might be repeated several times throughout the thesis, notably in the various introductions. This thesis is meant to serve as a compendium of ideas and techniques that can be readily applied to real situations by analytically-oriented practitioners in conservation and environmental impact assessment. Efforts have been made to keep models and ideas as simple as possible, although it seems unavoidable that the reader should be familiar with basic concepts of linear models, Bayesian analysis, calculus and perhaps, at

least a superficial understanding of differential equations. Although formulations might seem intimidating for readers who are not well versed in mathematics, there are only a handful of pivotal concepts that are used repeatedly in the different chapters. The contents of this thesis represent an advance towards a more analytical approach to environmental impact assessment of wind energy projects, and will hopefully promote the application of similar techniques to other problems related to biodiversity management, as well as the use of the often limited data to their full potential.

Chapter 2

Kernel-based estimation of bird activity distribution using vantage point observations

2.1 Introduction

Wind energy is one of the main players in the renewable energy sector at present and it is expected to become even more popular in the future (REN21, 2018). However, there are concerns about the effects this type of energy production might have on the receiving environment, and particularly on flying wildlife. Birds and bats are known to interact with wind turbines, resulting in either collision with turbine blades or displacement due to the reduction of habitat quality (Schuster *et al.*, 2015). Large birds that use air draughts for soaring or that are airborne hunters, as well as those that are threatened, range restricted or with narrow ecological niches are particularly affected by these impacts (Thaxter *et al.*, 2017). Thus, understanding where these species live and fly greatly contributes towards the sustainable development of the wind energy sector.

The best approach to reduce the negative effects of wind energy production on wildlife is placing wind turbines away from areas preferred by sensitive species (Arnett & May, 2016). Planning of wind energy development starts at a strategic level, considering

which geographical regions are preferred based on wind resources, and also amongst other things, on the distribution of sensitive species. However, wind resources are often attractive for both, soaring birds and the installation of wind farms, which results in the need for impact assessment at a finer-scale (e.g. Carrete *et al.*, 2012; Liechti *et al.*, 2013; Ralston-Paton, 2017). In addition, research shows that wildlife mortality is often concentrated around certain turbines, which further motivates considering characteristics of the local landscape when studying collision risk (Marques *et al.*, 2014; Schuster *et al.*, 2015).

At the local scale, to optimize the turbine layout as to avoid particular landscape features that are intensively used by birds, it is standard practice to undertake bird activity surveys in areas proposed for wind energy development. In this respect, vantage point surveys are recommended to investigate the use of space by soaring birds (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). During vantage point surveys, an observer is placed in a point of good visibility overlooking the surrounding area (viewshed). Guidelines suggest observing circular or semi-circular plots of between 800 metres and 2 kilometres in radius, centred on the observer - although the topography of the landscape could obscure certain areas, resulting in irregularly-shaped viewsheds. Observers capture data on bird activity for hours at a time (e.g. Jenkins *et al.*, 2015, suggest not spending more than three hours at a single vantage point session). Levels of activity can be evaluated by counting the number of flights observed at each vantage point and by timing these flights. Counting and timing activity gives an idea of the intensity with which the area surrounding the vantage point is used. This is equivalent to assuming that bird activity is distributed uniformly within the vantage point viewshed. Considering that vantage point viewsheds span several square kilometres, this measure of intensity might be at a too coarse scale. In heterogeneous landscapes, it would be desirable to investigate whether the assumption of uniformly distributed activity holds or the distribution of bird activity is concentrated at locations that could be more problematic in terms of collision risk (Ferrer *et al.*, 2012; Schuster *et al.*, 2015).

During vantage point surveys, observers also sketch on topographic maps the flights they observe to gain some understanding of the flight behaviour of the birds (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). Although the spatial distribution of activity within a viewshed may be studied using these mapped flights, there is no agreement on how to conduct this type of fine-scale analysis. Lacking a theoretical framework, practitioners resort to *ad-hoc* methods that often entail subsampling or ignoring spatial information and the temporal autocorrelation in the data

(e.g. Walker *et al.*, 2005; Dahl *et al.*, 2013; Jenkins *et al.*, 2018). For example, Walker *et al.* (2005) conducted a kernel density analysis of vantage point observations but they used the location of first sighting of each trajectory only, to make data independent. Dahl *et al.* (2013) characterized flight trajectories observed from vantage points by dividing them into sections of different behaviour or height, but then aggregated these data by vantage point to analyse the influence of environmental variables. Jenkins *et al.* (2018) combined radar observations with vantage points to obtain a larger volume of, and more accurate, data than with regular vantage point surveys. They then analysed the total number of flights passing through a wind farm rather than building a spatially explicit model for activity intensity. While these analytical approaches are valid and help dealing with spatially or temporally autocorrelated data, sub-sampling or aggregating data results in the loss of information (Fieberg *et al.*, 2010). If we wish to retain and exploit all data in our analysis in the presence of autocorrelation, we need techniques that accommodate this type of structure.

Recognizing that flight trajectories plotted on topographic maps share some similarities with tracking data, I frame the problem within the context of animal movement modelling, which in the last decade has seen an advent of analytical procedures (Patterson *et al.*, 2017). Ideally, one would aim at understanding the mechanisms behind observed movement patterns (I discuss this approach in chapter 3). In practice, it is not always possible to formulate a model structure that describes bird movement, either because of our ignorance about the movement behaviour of the species studied or because there are not enough data to estimate the model parameters reliably. In these cases, and also for exploratory purposes, I propose using kernel density estimation to investigate the spatial distribution of bird activity. The principle underlying kernel density estimation is relatively simple and the method is familiar to ecologists (Worton, 1989; Horne *et al.*, 2007; Fleming *et al.*, 2015). Kernel densities are used extensively in the estimation of utilization distributions, which inform about the relative frequency with which different locations are used by an animal (Van Winkle, 1975). These relative frequencies are based on a smoothed version of the frequency distribution of past observations. Although kernel density estimators offer a limited understanding of the mechanisms driving the observed spatial patterns, they are useful exploratory tools and provide a first approximation to general patterns suggested by the data.

Here, I use the same logic underpinning the calculation of utilization distributions to describe the distribution of activity within a vantage point viewshed. To frame the problem, I briefly describe the characteristics of vantage point data, understood as flight

trajectories mapped from vantage points. I then move on to outline the principles behind kernel density estimation, and the importance of choosing the right amount of smoothing to target the desired patterns in the data. Then, I propose a kernel density estimation scheme that accommodates temporally autocorrelated data, in the form of flight trajectories. Finally, I test the theoretical advantages of the proposed estimator on simulated data, and show the improvement in performance against other popular kernel density estimators that ignore temporal autocorrelation.

2.2 Methodology

2.2.1 Description of vantage point data

During vantage point surveys an observer maps the movements of one or more large bird species (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). The objective is to describe the intensity with which different areas are used by these birds, in order to predict possible impacts associated with human infrastructure, such as collision with wind turbines and power lines or avoidance of altered habitats. The area observed from a vantage point is limited to a roughly circular (360 degrees) or semi-circular (180 degrees) region of radius between 0.8 and 2 kilometres around the observer. Each observed flight is represented by a line on a map that tries to capture, as accurately as possible, the trajectory followed by the bird. Vantage points are visited several times a year and each visit may last between one and eight hours (see Strickland *et al.* (2011); Scottish Natural Heritage (2014); Jenkins *et al.* (2015) for methodological details and suggestions). Perhaps one of the most important characteristics of vantage point surveys is that the observed areas are relatively small compared with the areas birds traverse on their daily activities. Therefore, we only have access to a partial picture of the activity of these birds, and it is common to have no birds flying over any given vantage point during a survey. When there are flights, these are relatively short; they are rarely longer than five minutes. Therefore, vantage point surveys often produce a limited amount of data.

In terms of data processing, some discretization of the trajectories in series of locations is necessary for digitization purposes. However, the observation occurs continuously in time and flight paths may be represented by lines on the maps. For this reason, I

formulate models in continuous time. To justify this decision further, it might be useful to distinguish between discretizing a trajectory in a series of contiguous locations from the thinning of this sequence of locations. In the first case, one transforms a continuous path into a sequence of locations, and one would expect the sequence to retain most of the information about movement behaviour. The additional thinning of the location sequence entails removing some of the locations so that the remaining ones are further apart. Although this procedure reduces temporal autocorrelation, it might lead to the loss of certain features in the data. For example, the persistence in movement direction or turning angles is known to depend on the frequency at which locations are acquired (McClintock *et al.*, 2014). Excessive thinning might also obscure visitation rates to certain habitats or landscape elements. The optimal degree of thinning that simplifies the analysis without loosing information is difficult to judge. Although digital data might be necessarily discrete, bird movement takes place in continuous time and, therefore, this type of framework allows a more faithful description of the underlying data-generating process. In addition, models that are formulated in continuous-time are robust to the frequency at which samples were acquired (or thinned). On the other hand, continuous-time formulations often impose correlations between speed and directional persistence (McClintock *et al.*, 2014). Therefore, it is necessary to evaluate whether these properties are adequate for describing our movement data. In flying animals, and particularly in soaring birds, efficiency in flight and stability are intimately linked to flight speed and therefore flights are typically smooth (see chapter 6 in Biewener & Patek, 2018). I believe that in smooth flights, where lift and stability depend on speed, correlation in, and between, speed and direction are sensible assumptions.

In summary, vantage point observations are typically represented by a relatively small set of short flight trajectories, embodied by sequences of locations that are close enough in time to closely resemble a line (see figure 2.4). In the following sections, I describe some of the principles that underpin kernel density estimation, and based on these principles, I propose an estimator that accommodates data with the characteristics just described.

2.2.2 Kernel-based estimation of activity distribution

Kernel density estimation is arguably the most popular choice of non-parametric density estimator, and has attracted a lot of attention in ecological applications (e.g. Worton, 1989; Horne *et al.*, 2007; Fleming *et al.*, 2015). The movement of an animal within some portion of the plane defines a spatial distribution of use intensity (i.e. activity time). If

we were to predict where the animal is at any time, based on past locations, our best choice would be that area used most intensively in the past. This concept matches that of a probability distribution. In fact, by normalizing the intensity distribution so that it integrates to one, we obtain a probability density distribution. Therefore, the problem of estimating the activity distribution of an animal over a region of the plane is analogous to estimating a density function of the probability of finding the animal at any point within that region.

The shape of the activity distribution of an animal is determined by complex interactions with the surrounding environment and is likely to take complicated forms. The main appeal of non-parametric estimators is the flexibility they provide in terms of the shape of the distributions they can estimate. While parametric estimators are restricted to some class of distributions and have limited freedom to accommodate the distribution suggested by the data, non-parametric estimators are unrestricted to specific shapes. In addition, parameterizing a model entails some degree of understanding of the ecology of the animal, and thus, making some assumptions about the shape of the activity distribution. Non-parametric estimators are free from these assumptions, yet they rely largely on the amount and "quality" of the data and also on the optimization of some tuning parameter, as we will see shortly.

Lets assume that the activity of an animal is not observed continuously and only random locations $\mathbf{X}_t = (x_t, y_t)$ are available, where x_t is Easting y_t is Northing of the location observed at time t, for $t = 1, 2, ..., n$. We are interested in estimating the probability density distribution f that underlies the observations. The basic idea behind kernel density estimation is that the probability density in a region of the plane is proportional to the number of observations in that region. To avoid cutting the plane into arbitrary regions to calculate density, we define the area of influence of each observation using a kernel function K. Kernel functions are typically symmetrical, positive and integrate to one. We centre a kernel at each observation, to make their contribution to the probability density be at its maximum at the observed locations and decrease with distance to them. Different kernel functions may be used and the choice is mostly irrelevant (Silverman, 1986a). Using the Gaussian kernel is common practice, and I will use it throughout this thesis. In contrast, choosing an appropriate method to optimize the parameters shaping the kernel functions is critical (e.g. standard deviation of the Gaussian). This tuning parameter controls how density changes with distance to the observations and therefore determines the smoothness of the density function. This parameter is commonly referred to as bandwidth or smoothing parameter.

The kernel density estimator is given by the average density contributed by all the kernels centred at the observations (figure 2.1). There are different and equivalent formulations for the kernel density estimator. I will use the product kernel in this study (Scott, 1992). To generalize to more than two dimensions let's define a point in a D-dimensional space $\mathbf{x} = (x_1, x_2, ..., x_D)$, where $x_1, x_2, ..., x_D$ represent coordinates in each dimension. Let's now consider n observations of a random vector $\mathbf{X}_t$ with $t = 1, 2, ..., n$ of the same dimension as $\mathbf{x}$, such that $\mathbf{X}_t = (X_{t,1}, X_{t,2}, ..., X_{t,D})$. Assuming that the smoothing parameter h is equal in all dimensions, and a Gaussian kernel, the product kernel is defined as:

$$\begin{aligned} \hat{f}(\mathbf{x}) &= \frac{1}{nh^D} \sum_{t=1}^{n} \left\{ \prod_{d=1}^{D} K\left(\frac{x_d - X_{t,d}}{h} \right) \right\} \\ K(z) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}. \end{aligned} \tag{2.1}$$

The above expression places a Gaussian kernel K with standard deviation h at each observation, and calculates the density at the point $\mathbf{x}$ in each dimension $d = 1, 2, ..., D$. The densities computed for each dimension are multiplied and the average of the products of all dimensions of all observations, gives the kernel density at the point $\mathbf{x}$. Note that the kernel functions are circular, but this does not imply that the resulting estimated density will be circular, only that the influence of observed data points decays at the same rate in every direction of space.

2.2.3 Choosing a smoothing parameter

The issue now is to estimate an appropriate smoothing parameter (h in expression 2.1). The subject has been studied extensively and a plethora of methods are available (see for example, reviews by Gitzen *et al.*, 2006; Heidenreich *et al.*, 2013; Noonan *et al.*, 2019). A common approach is to select a smoothing parameter that minimizes the Mean Integrated Squared Error (MISE) of the estimate

$$\text{MISE}(h) = E\left[\int \left\{ \hat{f}(x) - f(x) \right\}^2 \, \mathrm{d}x \right], \tag{2.2}$$

where $\hat{f}$ is the estimated density and f is the target density to estimate. Typically f is unknown, which motivates the search for methods that minimize approximations to

MISE optimally. The details are quite technical and outside the scope of this study, but the idea is that the minimizer of MISE depends on f'', the second derivative of f (the smoothness of f), and finding a good substitute for this second derivative leads to making certain assumptions about f. Different assumptions define different families of kernel density estimators. reference-function bandwidths assume that the target density has some parametric form, and use sample-based estimates of the parameters shaping the target density to optimize the smoothing parameter analytically (Silverman, 1986a; Fleming *et al.*, 2015). Other, so-called plug-in methods, use more sophisticated substitutes for f'' (see Heidenreich *et al.*, 2013, and references therein).

Popular methods that circumvent the need for *a priori* assumptions about the shape of the target distribution fall into the category of cross-validation. These methods aim at minimizing the Integrated Squared Error (ISE) rather than MISE.

$$\text{ISE} = \int \left\{ \hat{f}(x) - f(x) \right\}^2 \, \mathrm{d}x. \tag{2.3}$$

Dropping the reference to the variable x to simplify notation and expanding, we obtain

$$\text{ISE} = \int (\hat{f} - f)^2 = \int \hat{f}^2 - 2 \int \hat{f} f + \int f^2. \tag{2.4}$$

Since the last term does not depend on $\hat{f}$, ISE is minimized by minimizing

$$\int \hat{f}^2 - 2 \int \hat{f} f \tag{2.5}$$

It can be shown (e.g. Silverman, 1986a) that

$$E\left[\int \hat{f}(x) f(x) \right] = E\left[\frac{1}{n} \sum_i \hat{f}_{-i}(X_i) \right], \tag{2.6}$$

where $\hat{f}_{-i}(X_i)$ is the estimated density computed at the observation X_i using all observations X_j such that $j \neq i$. This leads to the least squares cross-validation (LSCV) target function (Silverman, 1986a; Worton, 1989):

$$\min_h \text{LSCV}(h) = \int \hat{f}(x)^2 - \frac{2}{n} \sum_i \hat{f}_{-i}(X_i). \tag{2.7}$$

Likelihood cross-validation (CV), is another cross-validation technique. In this case, the objective is to minimize the Kullback-Leibler (KL) distance rather than the MISE or ISE (Silverman, 1986a; Horne & Garton, 2006). Essentially, the Kullback-Leibler distance measures how optimal a density estimate is at predicting a new observation, and it is defined as

$$\mathrm{KL}(f, \hat{f}) = \int f(x) \log \left\{ \frac{f(x)}{\hat{f}(x)} \right\} \, \mathrm{d}x \tag{2.8}$$

Based on the definition of the KL distance, it can be proven (e.g. Silverman, 1986a) that the quantity to be maximized by CV is

$$\max_h \mathrm{CV}(h) = \frac{1}{n} \sum_i \log \hat{f}_{-i}(X_i) \tag{2.9}$$

This method has shown good results in home range estimation studies (Horne & Garton, 2006), although theoretically it is very sensitive to observations that lie in the tails of the target distribution (Schuster & Gregory, 1981; Bowman, 1984). When such observations are present, likelihood cross-validation will over-smooth the density estimate. In fact, when using this criterion the choice of kernel function becomes relevant. For example, kernels with bounded support may produce $\hat{f} = 0$ resulting in bandwidth estimates to blow up. These lurking issues deter practitioners from using this technique, which is otherwise simple to implement and has the appeal of making intuitive sense.

2.2.4 Estimating densities from autocorrelated data

All the methods discussed in section 2.2.3, assume that observations are independent and identically distributed (iid). However, locations sampled from an animal movement trajectory may be autocorrelated if the sampling frequency is high. The distribution of data with temporal dependence is conditional on past observations and therefore violates iid assumptions. When using estimators that require iid data, we assume that clumps of locations arise by chance, whereas location data with temporal autocorrelation are clumped because they are close in time. If this temporal structure is ignored and inappropriate kernel density estimators are used, smoothing parameters tend to shrink and densities are under-smoothed (e.g. Fleming *et al.*, 2015; Noonan *et al.*, 2019).

Kernel density estimators that deal with temporal autocorrelation have been developed for the estimation of activity distribution from telemetry data. Tracking devices typically capture the location of the animal with certain frequency (from a few seconds to days). To my knowledge the most popular kernel density estimators that have been designed to account for temporal autocorrelation are Brownian bridge movement models (Horne *et al.*, 2007; Benhamou, 2011; Kranstauber *et al.*, 2012) and the autocorrelated kernel density estimator AKDE (Fleming *et al.*, 2015; Noonan *et al.*, 2019).

Although not strictly a kernel density estimator, a Brownian bridge model estimates the probability of the animal being at a point between two observed locations based on the time difference between the locations and the mobility of the animal (Horne *et al.*, 2007). These models assume diffusive movement between consecutive observed locations. This is equivalent to considering a Gaussian kernel with a time dependent mean and variance. The method accounts for temporal autocorrelation in the data; however, it does not attempt to estimate a long-term activity distribution of the animal, but rather the distribution of past, unobserved locations — what Fleming *et al.* (2015) and Noonan *et al.* (2019) named *occurrence distribution*.

Fleming *et al.* (2015) proposed an approach to target the long-term activity distribution. Their objective is to estimate the home range based on a utilization distribution — i.e. an activity distribution around an activity centre. They also focus on autocorrelated telemetry data and consider the observed trajectory being a realization of a stochastic process described by a mean function and a covariance function. The mean function is a constant and is located at the activity centre of the animal, while the form of the covariance function is estimated from the data using the autocorrelation function. With these two moments of the stochastic process, they re-derive the reference optimal bandwidth (Silverman, 1986a). Instead of assuming a Gaussian distribution as the underlying target density they assume a Gaussian process and use the covariance function instead of the sample variance in the derivation. The method is rigorous and mathematically tractable, and the bandwidths selected are larger than those estimated by the traditional methods, yielding larger home range estimates. The method is reported to work well at estimating home ranges of different species (Noonan *et al.*, 2019).

As pointed out in section 2.2.1, vantage point surveys typically provide a limited amount of data, and localized in the surroundings of the vantage points. Flights observed during these surveys are conceivably represented by continuous lines on a map, or as sequences of discrete locations close in time. Therefore, all locations along a trajectory are considered to be known with similar precision, down to observation error (see chapter 4

for a discussion on observation error) and there is no need for interpolating unobserved locations. For this reason, Brownian bridges, would not be useful for analysing vantage point survey data. If field protocols involved capturing sequences of location spaced by some time interval, Brownian bridges could be useful to estimate a distribution of possible trajectories between locations. AKDE might not be the best choice for this type of data either for several reasons: i) from a vantage point, observers typically do not cover the home range of the birds they observe, therefore it might not be possible to find the mean the stochastic process reverts to or to observe enough mean reversions to estimate a semi-variogram model, and ii) because several birds can be observed from a vantage point, data could arise from different stochastic processes, which would require a joint, population-level AKDE estimator; however, a population-level semi-variance function estimated from an often limited amount of short flights would be difficult to estimate.

For these reasons, I propose a different approach to computing kernel density estimation of flight trajectories observed from vantage points. The objective is to target the correct, long-term activity distribution, and at the same time, to relax the conditions imposed on the underlying process that generates the data. I will now briefly describe the characteristics of vantage points observations, and then explain the proposed estimator and why it is expected to provide good results for these data.

2.2.5 Proposed density estimator for vantage point activity

To define the kernel estimator that accommodates continuous, short trajectories, typical of vantage point surveys, I start from the estimator for discrete observations (i.e. location based). From this, I derive a new type of kernel that may be used on flight paths, possibly recorded in continuous-time.

Suppose we observe n locations of a bird $\mathbf{X}_t = (X_t, Y_t)$ or more generally, $\mathbf{X}_t = (X_{t,1}, X_{t,2})$, for $t = 1, 2, ..., n$. Assuming equal bandwidth in all dimensions, and iid data, we may estimate the density of locations using the the product kernel defined in equation 2.1. If instead of observing n iid locations, we were able to observe a path of length L and duration T, the product kernel density estimator at a point $\mathbf{x} = (x_1, x_2)$ could be re-defined as,

$$\hat{f}(\mathbf{x}) = \frac{1}{Lh^D} \int_T \prod_{d=1}^{D} K\left(\frac{x_d - X_{t,d}}{h}\right) \mathrm{d}t$$
$$K(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \tag{2.10}$$

where d indexes the different dimensions — in the two-dimensional case $D = 2$. I refer to expression 2.11 as the *path* kernel (figure 2.1), which is still the average density contribution of each location along the trajectory at the point $\mathbf{x}$, based on a kernel function.

Should m trajectories be observed from a vantage point, the density would be estimated by taking the average over the m path kernels:

$$\hat{f}(\mathbf{x}) = \frac{1}{m} \sum_{j=1}^{m} \left\{ \frac{1}{L_j h^D} \int_{T_j} \prod_{d=1}^{D} K\left(\frac{x_d - X_{j,t,d}}{h}\right) \mathrm{d}t \right\}. \tag{2.11}$$

Therefore, we are effectively assuming that each path contributes as one independent observation to the total density. However, each trajectory has a kernel function that fits its particular shape, given by equation 2.10 (figure 2.1).

So far, I have worked out an expression for estimating the probability density of observing a bird at any point in space, with equal bandwidth (smoothing) in all directions. The next step is to design a procedure to determine the appropriate amount of smoothing. As mentioned earlier, the focus is on estimating the long-term activity distribution imposed by the underlying process. I also mentioned in section 2.2.4 that in the presence of autocorrelation, density estimates tend to be under-smoothed. Therefore we need a bandwidth selection method that is robust to autocorrelation.

It is not immediately apparent how to perform traditional leave-one-out cross validation, when dealing with trajectories in continuous-time. Therefore, I propose making use of the fact that different trajectories could be considered independent from each other and conduct cross-validation leaving out one full flight at a time, in agreement with procedures suggested for data with dependencies imposed by hierarchical or blocking structures (Roberts *et al.*, 2017). I will refer to the procedure based on leaving out flights at a time as leave-group-out cross-validation (LGOCV), to differentiate it from leave-k-out, in which all subsets of k elements are left out at a time, or from k-fold cross-validation in which the data are split into groups of k elements randomly, and then

one group is left out at a time (see Hastie *et al.*, 2009; Arlot & Celisse, 2010).

In this study, I propose using a target function that maximizes the likelihood of unseen data. The most immediate procedure suggested by this premise is likelihood cross-validation. Under the leave-group-out scheme, a full flight is left out of the sample, a density estimate is calculated without the left out flight and the likelihood of this flight is calculated using the estimated density. The bandwidth that maximizes the average likelihood of all flights is then selected. In a continuous time framework and under a leave-group-out scheme the likelihood cross-validation target function becomes:

$$\max_{h} \text{LGOCV}(h) = \frac{1}{m} \sum_{j=1}^{m} \log \left(\frac{1}{L_j} \int_{T_j} \hat{f}_{-j}(\mathbf{X}_t^{(j)}) \, \mathrm{d}t \right), \tag{2.12}$$

where $\frac{1}{L_j} \int_{T_j} \hat{f}_{-j}(\mathbf{X}_t^{(j)}) \, \mathrm{d}t$ substitutes $\hat{f}_{-i}(X_i)$ in equation 2.7 and represents the average density calculated for flight j from a density distribution estimated leaving out flight j. This procedure searches for a bandwidth that maximizes the likelihood of flights that were not used for estimation. We can say it targets the long-term distribution of flights, because the focus is on predicting trajectories unknown to the model. I expect the procedure to reduce the sensitivity of likelihood cross-validation to observations located in the tails of the distribution by averaging the likelihood of left-out data over entire trajectories. To investigate the performance of this bandwidth selection method and also the path kernel, I perform a series of simulations that I detail next.

As a note on implementation, trajectories need to be discretized into a series of locations, and therefore, the integrals in expressions 2.10, 2.11 and 2.12 must be approximated by sums. If we estimate densities using the discrete locations then, the calculation of the path kernel is similar to the product kernel in equation 2.1, although locations are grouped into trajectories for optimizing the bandwidth. However, we may resort to estimating an exact convolution integral for each segment connecting contiguous locations and then adding the resulting densities of all segments within a trajectory. This algorithm is implemented in the R package `spatstat` (Baddeley *et al.*, 2015) by the function `density.psp` and it is this function that was adapted for the calculations in this chapter.

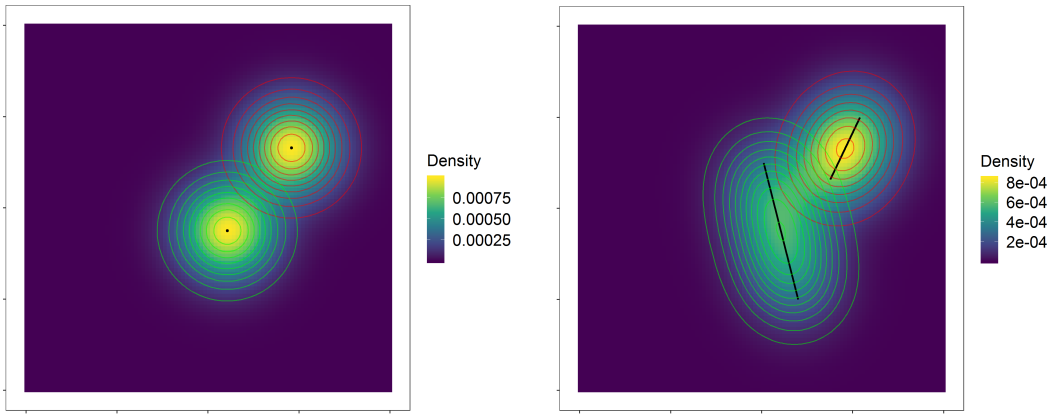


Figure 2.1: Illustration of kernel density estimation for points (left) and for paths (right). The red and green contours represent kernel functions centred at each observation (either point or path). The colour gradient represents the resulting density estimate obtained by averaging the kernel functions at each of the image pixels.

2.2.6 Simulations

To test the performance of the proposed estimator, I simulate data with similar characteristics to flight trajectories observed from vantage points. I represent flight trajectories by sequences of locations separated by a short time interval, grouped forming different trajectories. The idea is to generate data in such a way that in the long-term, if many flights are simulated, the distribution of locations is well described by some target density that we can use as a reference. We may then compare densities estimated from fewer data, resembling what is really observed during vantage point surveys, to study the performance of the estimator.

To make sure that in the long-term, the spatial distribution of simulated trajectories conforms to the desired target density, I simulate the flight trajectories as if the animal was moving randomly within a potential field (Brillinger *et al.*, 2002; Preisler *et al.*, 2013; Russell *et al.*, 2018). Under the potential field model, an animal moves with a long-term velocity driven by the gradient of a potential function $H(\cdot)$. As such, the animal tends to move from areas of high potential to areas of low potential. Consequently, in the long term, animal locations concentrate around areas of low potential. The potential function may be designed so that, in the long term, the distribution of flights conforms to the desired target density. The systematic drift in direction imposed by the potential field is disturbed by stochastic forces, adding a short-term random motion that is characteristic

of animal movement. Although the underlying movement process evolves in continuous time, the observed realizations come in the form of sequences of locations $\mathbf{X}_t = (x_t, y_t)$, being $t \in [0, T]$. To impose temporal persistence not only on consecutive locations but also on consecutive velocities, I simulate from a potential-based velocity model, such that:

$$\begin{aligned} \mathrm{d}\mathbf{X}_t &= \mathbf{V}_t \ \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= -\beta(\nabla_H + \mathbf{V}_t) \ \mathrm{d}t + \sigma \ \mathrm{d}\mathbf{B}_t. \end{aligned} \tag{2.13}$$

The velocity $\mathbf{V}_t$ tends to revert to an long-term velocity given by $-\nabla_H$, the negative of the gradient of the potential function, at a rate controlled by the parameter β. The stochastic term is given by the expression $\sigma \ \mathrm{d}\mathbf{B}_t$, which is a Brownian motion with distribution $N(\mathbf{0}, \sigma^2 \mathrm{I} \ \mathrm{d}t)$, where I is the 2x2 identity matrix. In practice, simulations are generated from a discretized version of model 2.13, where discrete times are separated by a fixed time interval τ (see appendix D for an example of simulation code),

$$\begin{aligned} \Delta\mathbf{X}_t &= \mathbf{V}_t\tau \\ \Delta\mathbf{V}_t &= -\beta(\nabla_H + \mathbf{V}_t)\tau + \sigma\epsilon_t, \end{aligned} \tag{2.14}$$

where τ is a fixed time increment between locations and $\epsilon_t \sim \mathrm{N}(\mathbf{0}, \tau\mathrm{I})$.

Three different potential functions are used in this study (see figure 2.2):

i) a potential with a single centre of activity is given by the distance to the centre of activity μ:

   $H(\mathbf{x}) = \|\mathbf{x} - \mu\|$;

ii) a potential with two centres of activity is given by the average distance to each activity centre μ_1 and μ_2:

   $H(\mathbf{x}) = 0.5 \, \|\mathbf{x} - \mu_1\| + 0.5 \, \|\mathbf{x} - \mu_2\|$;

iii) low potential is concentrated at some distance r from the origin, resulting on a ring of low potential:

   $H(r) = r - \mu_r$,

   where $r = \|\mathbf{x}\|$ is the radial distance from $\mathbf{x}$ to the origin and μ_r is the radius around which activity concentrates.

To approximate the long-term target distribution of locations imposed by the potential-based models, I simulate (using expression 2.14) 50 trajectories, one day long each, with one sampled location every 0.1 minutes ($\tau = 0.1$ min). Because the system is driven by a potential field, one may use the fact that in the long term, the density of locations is expected to concentrate around each point of minimum potential and decrease with distance to them, following a Gaussian distribution (e.g. Brillinger *et al.*, 2002). The variance of these Gaussians may be estimated by fitting a parametric model to the simulated locations.

The aforementioned potentials impose three different densities: i) a bivariate Gaussian distribution, ii) a mixture of two bivariate Gaussian distributions that result in a bimodal shape and iii) a ring-shaped distribution constructed by centring a bivariate Gaussian on a ring (radial Gaussian). In ecological terms, the first distribution provides a simple approximation to an activity pattern centred at a single point in the landscape, the second adds some complexity considering two activity centres, and the third distribution represents a complicated structure that combines activity concentrated around linear features with multi-modality (figure 2.2).

Once the target density is calculated, I proceed to simulate different vantage point scenarios that could be encountered during the assessment of impacts of wind energy facilities on birds. I generate different number of flight paths of different duration to gain understanding of how these two ways of increasing the amount of data affect the estimation of the target densities. Five iterations of scenarios with 2, 4, 6, 8 and 10 trajectories, of duration: 2, 5, 10, 50 and 100 minutes, were simulated. In total there were five simulations at each of the 25 different combinations of number of trajectories and trajectory duration (see figure 2.4 for an example).

Simulation of flight paths was conducted in R (R Core Team, 2019), with the added functionality of the `tidyverse` (Wickham, 2017), the package `sf` (Pebesma, 2018) for handling spatial data and `adehabitatLT` (Calenge, 2006) for managing trajectory objects. The variance of the Gaussians conforming the long-term distributions imposed by the potential fields were modelled in a Bayesian framework using the package `rstan` (Stan Development Team, 2018a).

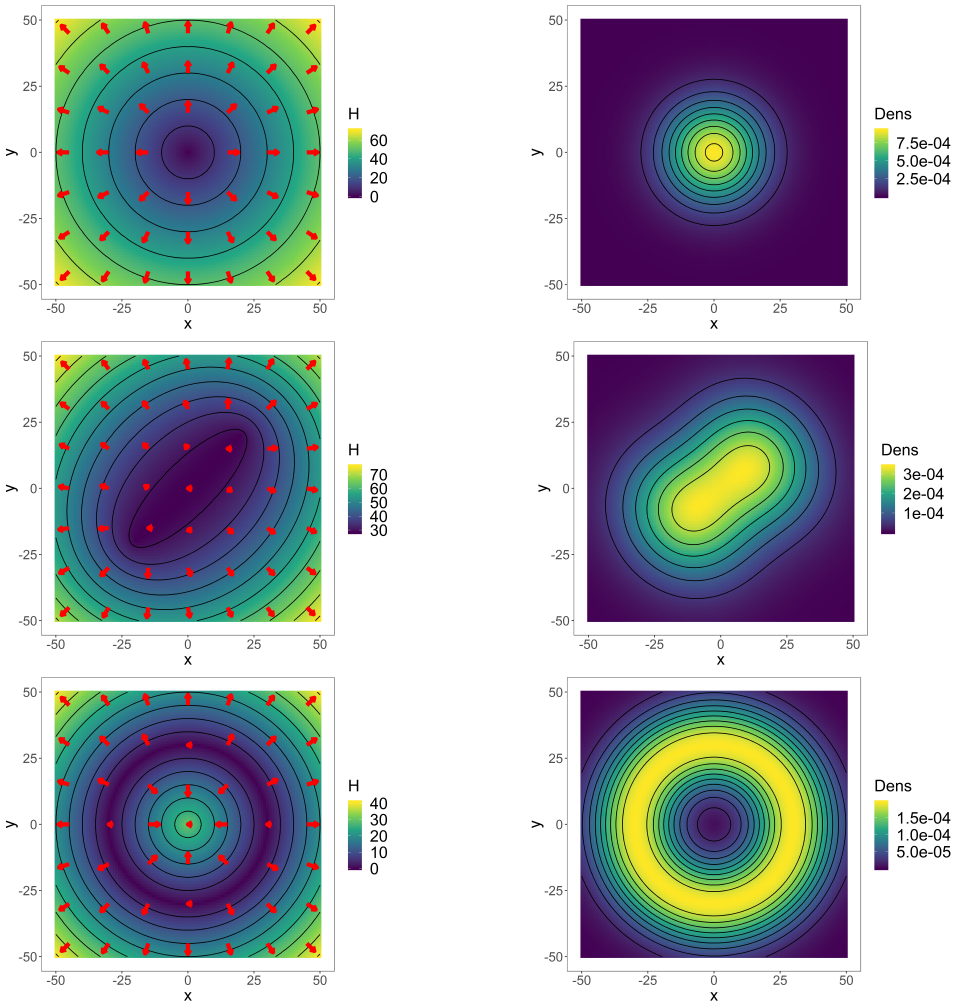


Figure 2.2: Left panels: potential fields defined by the potential functions H used to simulate flight trajectories. The red arrows represent the direction of the gradient. Right panels: densities induced by the potential fields used as target distributions to test the performance of different kernel density estimators. From top to bottom: unimodal, bimodal and ring potentials corresponding to a single bivariate Gaussian, bivariate Gaussian mixture and radial Gaussian densities, respectively. In potential-based models animals tend to move from areas of high potential to areas of low potential. We can see, by the colour gradients, that areas of low potential (blue in the left panels), correspond with areas of high density of locations (yellow in the right panels).

2.2.7 Accuracy of the estimator

I evaluated the accuracy of the proposed estimator by comparing its performance with some traditional methods developed for independent and identically distributed (iid) data. Among the large variety of estimators for iid data, I chose the reference-function bandwidth, one plug-in and one cross-validation method (see section 2.2.2 for a description of the different types of estimators): reference-function bandwidth (Ref), 2-stage plug-in (PI) and smoothed cross-validation (SCV) (Duong, 2007). The choice of methods was motivated by their popularity due to their good performance (Duong, 2007, and references therein) and them being available in the R package `ks` (Duong, 2019). In addition, a visual preliminary assessment revealed good performance of these estimators compared with other methods designed for iid data.

As explained in sections 2.2.2 and 2.2.5, current methods designed to account for autocorrelation in kernel density estimation would be unsuitable for the analysis of vantage point data for two main reasons: i) the observation process occurs in continuous-time, so no interpolation is needed, and ii) because only fragments of trajectories that are potentially independent from each other are observed rather than a long trajectory of the same process — the observed processes might not even be stationary at the scale flight trajectories are observed. Therefore, I do not compare the proposed estimator with the methods designed for autocorrelated data.

I used the overlapping coefficient (OVL) (Weitzman, 1970; Clemons & Bradley, 2000) between the target density and the estimated density to measure the accuracy of the different estimators. The overlapping coefficient calculates the area of overlap between two distributions f and $\hat{f}$. Being these probability density functions (they integrate to one) a proportion of overlap is calculated by finding the minimum between f and $\hat{f}$ throughout their domain and integrating (figure 2.3). The higher the overlap between target and estimated densities the higher the similarity between the two and the more accurate the estimator. The OVL is calculated as

$$\text{OVL} = \int \min\{f(\mathbf{x}), \hat{f}(\mathbf{x})\}\, \mathrm{d}\mathbf{x}. \tag{2.15}$$

In order to summarize the results, I fit a Generalized Additive Model (GAM) to the OVL scores as response variable, and fitting smooth terms for the effects of the number of trajectories and the duration of the trajectories, and linear terms for the effect of the density estimation method. All analyses were conducted in R (R Core Team, 2019).

Kernel estimates using PI and SCV methods were computed using the package `ks` and GAMs were fitted using the package `mgcv` (Wood, 2017).

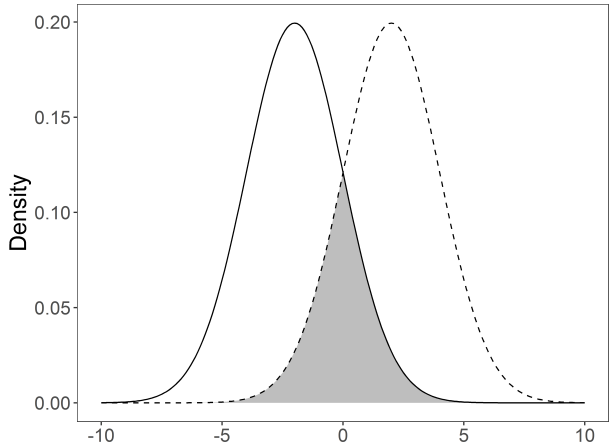


Figure 2.3: Representation of the overlapping coefficient (grey area) of two Gaussian density functions (solid and dashed lines). The overlapping coefficient is calculated by integrating the minimum value of the two functions at each point of their domain.

2.3 Results

To investigate the performance of the proposed kernel density estimation scheme, the path kernel with LGOCV bandwidth selection, I simulated movement trajectories within purposely designed potential fields (see figure 2.4 for an example of simulated flight paths). The objective is to understand how the shape of the target density, the number of observed flights and the duration of these flights affect estimation accuracy. Agreement between true and estimated density distributions (accuracy) was measured by the overlapping coefficient (OVL). OVL was calculated for estimates of three different target densities: unimodal, bimodal and ring distributions (figure 2.2). A comparison with other popular kernel density estimators was also conducted: reference-function bandwidth (Ref), two-stage plug-in (PI) and the smoothed cross-validation (SCV).

OVL values were modelled using a GAM with number of flights available for density estimation, duration of these flights and estimation method, as covariates. The different kernel density estimation methods were included in the GAM as parametric (linear)

regression terms and their coefficients are presented in table 2.1. Overlapping coefficients tend to be smaller for Ref, SCV and PI bandwidth selection methods than they were for LGOCV, as shown by their negative regression coefficients. Number of flights and flight duration were included in the GAM as non-linear (smooth) terms (figure 2.5). Flight duration had a larger positive effect on OVL than number of flights as seen in the steeper slope along the Time (flight time) axis compared to the Num. (number of flights) axis. This is particularly true for plug-in and SCV bandwith selection methods. The proposed LGOCV bandwidth selection performed well compared to other methods even in situations with fewer data, and therefore the effect of number of flights and flight duration was smaller.

In data-poor situations, methods designed for iid data tend to under-smooth estimated densities more than the proposed LGOCV method (figure 2.6). Of the iid-data methods, the reference-function bandwidth performed best in scenarios with small sample sizes, although plug-in and smoothed-cross-validation capitalized on increasing sample size and performed better than reference-function bandwidth in data-rich situations (figure 2.5).

In terms of target distributions, the best results were obtained when estimating a unimodal underlying density (figures 2.5 and 2.6). The bi-modal density was estimated slightly worse than the unimodal distribution and the ring distribution was the most difficult to estimate accurately.

Table 2.1: GAM parametric (linear) coefficients (with standard error in brackets). Overlapping coefficient (OVL) was the response variable and bandwidth selectors the explanatory variables (*LGOCV*: leave-group-out cross-validation, *Ref*: reference-function bandwidth, *Plug-in*, *SCV*: smoothed cross-validation). In the columns, the parameters estimated for the different target distributions (*Unimod*: unimodal Gaussian, *Bimod*: bimodal Gaussian mixture, and *Ring*: Gaussian distribution centred on a ring around the origin). Larger OVL means better agreement between target and estimated densities. *LGOCV* is the reference level against which other bandwidth selection methods are compared, and the method that obtained the largest OVL. The coefficients for the other methods are differences in OVL with respect to the reference LGOCV.

Variable	Unimod	Bimod	Ring
LGOCV	0.687 (0.008)	0.588 (0.007)	0.605 (0.006)
Ref	-0.282 (0.011)	-0.170 (0.010)	-0.191 (0.009)
Plug-in	-0.372 (0.011)	-0.355 (0.010)	-0.333 (0.009)
SCV	-0.352 (0.011)	-0.310 (0.010)	-0.317 (0.009)

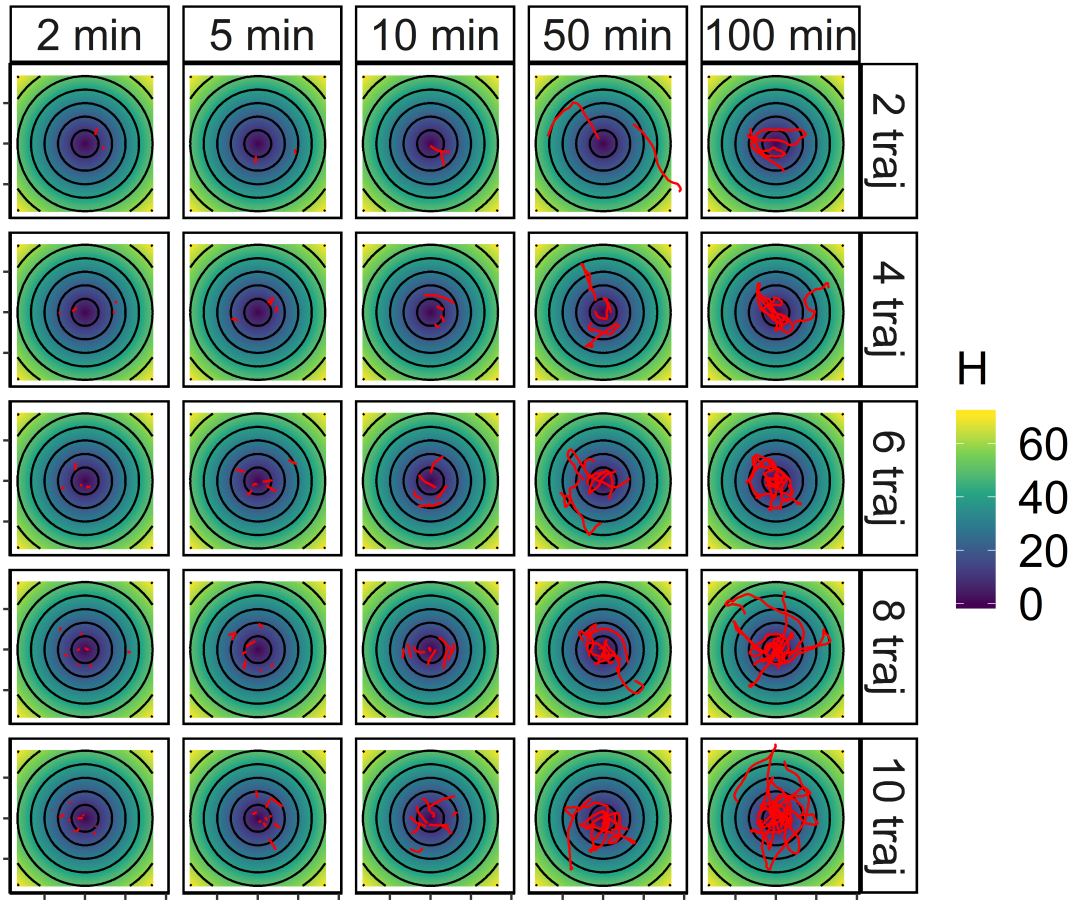


Figure 2.4: One iteration of different number of simulated flight trajectories (rows) of different duration (columns). The trajectories are represented in red, whereas the background colour represents the value of the potential function (H) that defines the field within which the bird is moving.

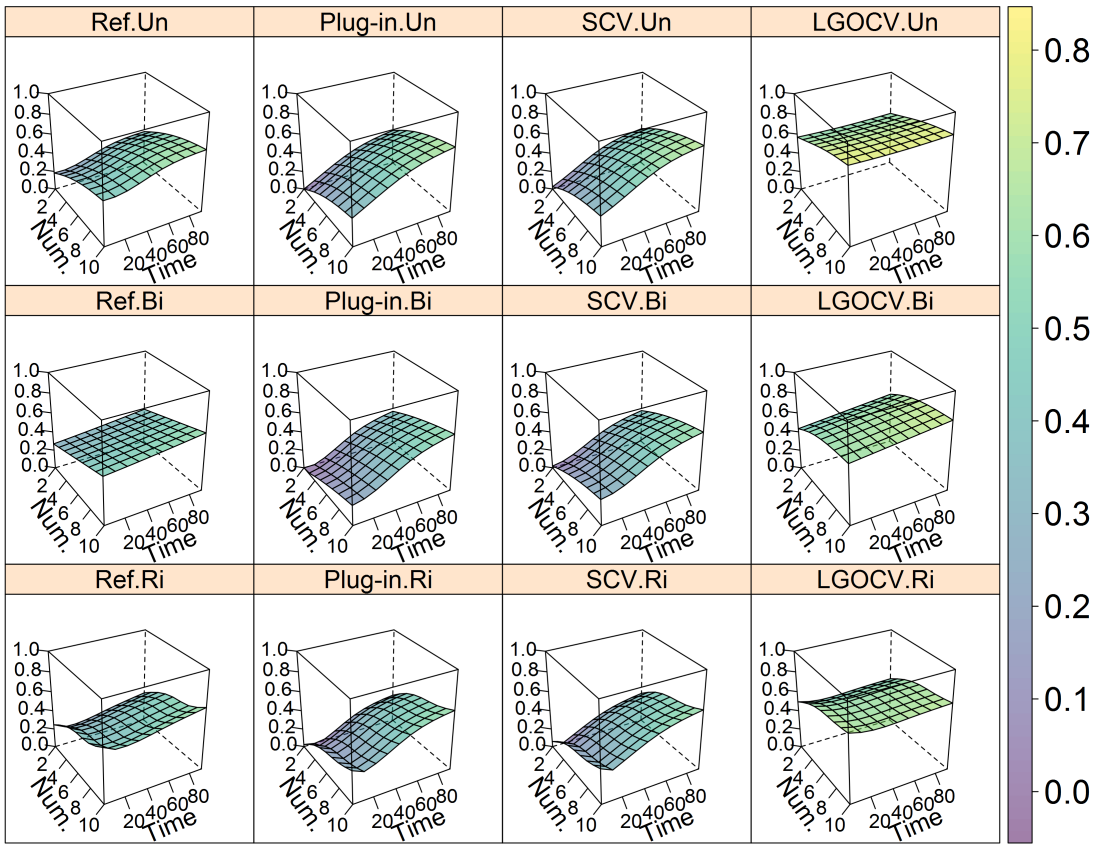


Figure 2.5: GAM predictions of the overlapping coefficient (OVL - vertical axis and colour gradient) between the true probability density and the density estimated from movement trajectories using different kernel density estimation schemes. OVL is given as a function of the number of trajectories (*Num.* axis) and the duration of these (*Time* axis, in minutes). The main difference in the estimators lies in the computation of the kernel bandwidth. From left to right panel columns: *Ref* stands for reference-function bandwidth (Silverman, 1986a), *Plug-in* is the two-step plug-in method and *SCV* stands for smoothed cross-validation (both suggested by Duong, 2007), finally *LGOCV*, for leave-group-out cross-validation, is the estimator proposed in this study. From top to bottom panel row, the estimators are tested in different target densities: *Un*, stands for uni-modal Gaussian, *Bi* for bi-modal Gaussian and *Ri* for ring (or radial) Gaussian density. Increasing the duration of the observed flights has a stronger effect on OVL than increasing number of flights, for most estimation methods. LGOCV performs better (higher OVL) than other bandwidth estimation methods; although in data-rich situations — where many flights of long duration are available — differences are smaller.

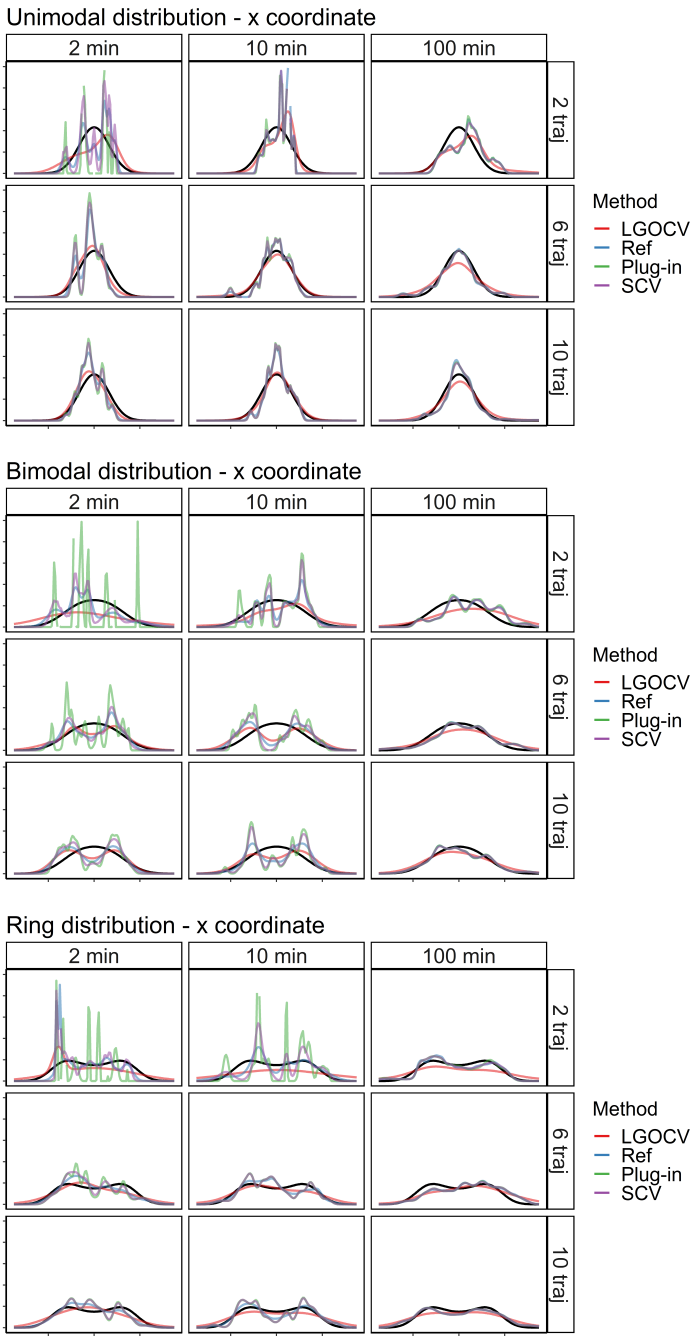


Figure 2.6: Marginal target distribution along of the x-coordinate (in black) and the marginal kernel estimates along the same axis calculated using different bandwidth selection methods. Only a selection of data-poor and data-rich configurations, out of all the simulated scenarios, is shown. In data-poor situations reference-function bandwidth (Ref), plug-in and smoothed cross-validation (SCV) tend to under-smooth more than leave-group-out cross-validation LGOCV.

2.4 Discussion

Predicting where birds are likely to fly is a recurring question when it comes to finding the best location for wind turbines, power lines and other such infrastructure (e.g. Katzner *et al.*, 2012; Jenkins *et al.*, 2018; Murgatroyd *et al.*, 2018). Mapping bird movements using telemetry tags (e.g. GPS, VHF, Argos) or automated detection devices (e.g. radar, video cameras) is not always feasible due to logistic and budget constraints (Walls *et al.*, 2009; Hebblewhite & Haydon, 2010; Katzner *et al.*, 2016a). Vantage point surveys provide an alternative to these methods that is versatile, relatively simple to implement and very popular to, for example, assess collision risk of soaring birds with wind turbines and power lines (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). However, analytical procedures that account for the particular characteristics of this survey method in a tractable and coherent way, are lacking. Here, I propose estimating the activity distribution within a vantage point viewshed using a kernel density estimator.

Technically, what we want to obtain is an approximation to the expected long-term activity distribution arising as a result of movement patterns of individual birds. Kernel density estimators use smoothing functions to estimate the distribution underlying sampled observations. The amount of smoothing is controlled by a parameter called the bandwidth or just smoothing parameter. Kernel density estimators work under the premise that future observations are more likely to occur in areas close to past observations. These foundations make kernel-based estimators very sensitive to temporal autocorrelation in the data. Because traditional bandwidth selection methods ignore temporal structure in the data, they are expected to perform poorly with flight trajectories observed from vantage points (Horne *et al.*, 2007; Fleming *et al.*, 2015; Noonan *et al.*, 2019).

The estimation scheme proposed in this study — the path kernel in combination with LGOCV — takes both spatial and temporal autocorrelation into account and outperforms some of the most popular methods for iid data (figure 2.5). With large sample size, both iid and autocorrelated data should converge to the same limiting long-term distribution (e.g. Kulkarni, 2011). The results of this study also show that kernel estimates benefit, not only from having a larger sample of flights, but also from observed flights being of longer duration. Indeed, all methods analysed seem to estimate with similar accuracy when working with large samples; however, increases in performance were much quicker when using the proposed leave-group-out cross-validation for auto-

correlated data. As figure 2.5 shows, methods designed for iid data perform poorly with small samples, and the leave-group-out cross-validation scheme proposed for autocorrelated data is much less sensitive to sample size. Therefore, estimates using the cross-validation procedure specifically designed for autocorrelated, grouped data (LGOCV), provide more accurate estimates when data are limited. Noonan *et al.* (2019) also observed a much faster convergence to the target density when using AKDE (designed for autocorrelated movement data) over estimators designed for iid data. AKDE might not always be suitable for the analysis of vantage point data, because it is sensitive to estimating a converging long-term variance. This is not to say that estimates from small samples are reliable representations of the target distributions. As for any statistical procedure, it is recommended to work with representative samples to obtain reliable estimates.

Distributions that are less smooth were more difficult to estimate by all estimators studied. Bi-modality was not a major issue but all estimators struggled to accurately represent the ring distribution. These results suggest that if the distribution of bird activity has hard boundaries (i.e. areas that are inaccessible to the bird) or very concentrated activity spots, these features could be misrepresented in the kernel estimates. In soaring birds, I would anticipate this not being an important problem, because flying species typically move without hard barriers. However, certain questions require the detection of potentially sharp changes in activity levels. For example, flight activity could drop suddenly within a certain threshold distance around wind turbines. How accurately we could detect this type of pattern remains unknown. A similar simulation approach as the one used in this study could be used to test the performance of kernel estimates at detecting specific changes in the underlying distribution in a Before-After Impact situation.

Cross-validation requires partitioning the data set into groups that are assumed to be independent. In the vantage point analysis setting, it is natural to start by considering different flights as being independent from each other. However, practitioners are advised to judge whether this assumption is reasonable and whether two different flights that were observed close in time should be treated as being independent. For example, two consecutive observations of the same individual bird separated only a few minutes, are likely correlated even though there was a brief gap in time between the two observations. Occasionally, considering activity on different days to be independent and use all flights observed on a particular day as cross-validation groups might be a better option. The subjectivity of this choice is the most evident weakness of the proposed method and also

the step that is not fully automatic and data-driven.

It is also important to remember that kernel density estimation is primarily an exploratory technique. In other words, these estimators are good for summarizing features of the data and try to estimate what distribution might underlie a set of observations. However, these techniques do not provide a mechanistic understanding of the process that generated the observations. For example, whether it was the weather, certain habitat preferences or behavioural states that motivated the observed activity pattern. In general, a consequence of this is that we cannot predict the behaviour of the system in a different environment to the one the observations where collected at — although some attempts have been done to investigate habitat preferences based on utilization distribution estimates (Millspaugh *et al.*, 2006). We must also assess whether our data are a good representation of the system we want to study — we may produce a kernel density estimate with as few as two short flights, which will very likely be a very poor representation of other flights we may observe in the future.

Because birds typically move without hard boundaries, I have not addressed this issue in details. However, hard boundaries are a recurrent problem in kernel density estimation and previous research has approached it in different ways (e.g. Scott, 1992; Benhamou & Cornélis, 2010; Tarjan & Tinker, 2016; Péron, 2019). Early work focused on correcting estimated densities close to the hard boundaries by modifying the support of the kernel functions near the boundary to have zero density beyond the boundary (Scott, 1992). Within this framework there are different choices of boundary kernels, and which one is appropriate depends on the target density. Because the target density is generally unknown, practitioners often choose methods that balance generality and consistency. For example, Benhamou & Cornélis (2010) use the reflection technique, which works by generating virtual locations at both sides of the boundaries and manipulating the resulting densities to remove density "spilled" outside of the boundary allocating it to the inside. More recent methods use information on habitat preferences of the studied species to fit semi-parametric models and avoid allocating use intensity to areas that are inaccessible or unattractive to the animal (Tarjan & Tinker, 2016; Péron, 2019). Although beyond the scope of this chapter, applying similar principles to the proposed path kernel method could expand its applicability to problems that require hard boundaries.

I have mentioned earlier in the chapter that Brownian bridges would not be adequate for analysing trajectories observed from vantage points. However, if instead of continuous trajectories observers recorded sequences of locations with certain frequency, then, Brownian bridges could be used to construct the path kernel and account for the

extra uncertainty between locations. The leave-group-out cross-validation bandwidth procedure would then proceed normally. This method would actually generalize better to other "automated" data capture methods, such as radar. Another possible research avenue would be to implement some kind of optimally weighted kernel density estimation scheme whereby locations towards the start and end points of the trajectory are considered to be less affected by temporal autocorrelation (Fleming *et al.*, 2018). This method improves home range estimates under sub-optimal sampling schemes (i.e. irregularly sampled data and few observed home range crossings). Although it is unknown how it would benefit density estimates from visually observed trajectories, I would anticipate limited gains given the relatively short duration of the flights. Yet, this remains an open question.

Another important topic that I have not addressed in this chapter is that of estimating uncertainty. Confidence intervals and confidence bands for kernel density estimates could be constructed using bootstrap samples (Efron, 1979; Neumann & Polzehl, 1998). Confidence intervals refer to intervals that cover the true density at a particular point, whereas confidence bands aim at covering the true density across the whole spatial domain of the kernel density estimate. The basic idea is to re-sample trajectories with replacement and estimate kernel densities from these "bootstrap" samples. Then, compute the empirical cumulative distribution function of the deviations between bootstrap estimates and their mean to define the limits of the confidence interval. If deviations in point-wise estimates at a particular location are used, we obtain a confidence interval for that particular location, whereas using the maximum deviation at each point, we obtain confidence bands.

Based on the results of this study, kernel density estimators that ignore the temporal structure of the data should be avoided when estimating the distribution of bird activity from flight trajectories. These conclusions can be extrapolated to other non-parametric methods such as histograms (the two-dimensional grid version). These are constructed by counting the number of flights traversing or measuring flight time within a grid cell. Histograms are perfectly valid as an exploratory technique; however, without smoothing the convergence with the distribution underlying the data is very slow. Histograms can also be severely affected by spatial and temporal structures in the data (see Silverman, 1986a; Scott, 1992, for discussions on the similarities of histograms and kernel density estimation).

This study sets the bases for the analysis of flight trajectories observed from vantage points using kernel density estimation. I have discussed the advantages and disadvant-

ages of these methods and developed a procedure that uses autocorrelated data effectively, providing better estimates than traditional kernel density estimators that ignore autocorrelation. I have also shown that activity distributions that are rugged and change abruptly might be over-smoothed by the proposed estimation procedure, yet, in general, over-smoothed estimates reflect more honestly the uncertainties about the underlying activity distribution. The practical value of this method will only become apparent after it has been tested in the field, yet it provides an intuitive, versatile and relatively easy to use tool to obtain estimates of animal activity distributions.

Chapter 3

Analysis of vantage point observations using potential-based velocity models to estimate exposure to collision with wind turbines

3.1 Introduction

Wind energy is now considered one of the main renewable alternatives to fossil fuels (REN21, 2018). However, wind farms are often built away from urban centres, on relatively undisturbed habitat, leading to impacts on wildlife. Mortality by collision with wind turbines and displacement of flying vertebrates stand out as the most concerning impacts associated with wind energy production (Drewitt & Langston, 2006; Dai *et al.*, 2015). Raptors and other soaring birds are particularly vulnerable; due to their low population densities and reproduction rates, a sustained loss of even a few individuals to wind turbines may result in severe population-level impacts (Smallwood *et al.*, 2009; Marques *et al.*, 2014; Watson *et al.*, 2018; May *et al.*, 2019). Furthermore, soaring birds use orographic and thermal updraughts, and display hunting and courtship flight beha-

viours at the height of the turbine blades, making them prone to collision (Drewitt & Langston, 2006; Marques *et al.*, 2014).

Despite extensive research on bird collisions with wind turbines, predicting where and when these collisions may occur is still challenging, and it is apparent that factors leading to impacts are species- and site-specific (Marques *et al.*, 2014). As a result, to understand and minimize collision risk, monitoring programmes of wind energy facilities are regularly implemented both, pre- and post-construction. Pre-construction studies focus on identifying risk factors leading to the aforementioned impacts and tailor mitigation measures; whereas post-construction monitoring aims to quantify the impacts of operational wind farms, validate pre-construction predictions, and adapt project management to mitigate its impacts on the environment (e.g. Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015).

Surveys of raptor activity from vantage points is a recommended field technique, routinely employed to assess exposure to collision during wind farm monitoring. Designs are manyfold (e.g. Barrios & Rodríguez, 2004; Garvin *et al.*, 2011; Johnston *et al.*, 2014), but generally, during these surveys, an observer maps the flights of birds entering an area of interest, and records the number of individuals and their flight time (Katzner *et al.*, 2016a). These quantities are used as a measure of exposure to impacts, and as inputs in collision risk models (Band, 2012; New *et al.*, 2015). Vantage point surveys may be thought of as a particular type of point count (Sur *et al.*, 2018). Point counts are typically used to count small birds, and consist of short surveys, covering small areas (< 1 km^2) that aim to sample bird abundance (Bibby *et al.*, 2000). However, vantage point surveys are adapted to study large, far-ranging species. Due to the low population density of these birds, and in order to collect data effectively, vantage points surveys may go on for hours at a time, and the area covered from them (viewshed) may span several square kilometres. These adaptations present new associated assumptions and limitations that need to be addressed; yet, studies on their implications for data collection and analysis are lacking.

Data collected during vantage point surveys are often used to assess collision risk; however, fatalities predicted using these data do not always correlate with actual observed mortality (de Lucas *et al.*, 2008). Research shows that bird abundance estimates calculated using vantage point surveys are influenced by the number of observation hours, and the spatial coverage of the surveys (Douglas *et al.*, 2012; Sur *et al.*, 2018). It has been repeatedly suggested that the mismatch between bird abundance and fatalities may stem from the scale at which abundance is estimated (de Lucas *et al.*, 2012; Ferrer *et al.*,

2012; Sur *et al.*, 2018). Areas overlooked from vantage points may potentially include a number of turbines. Therefore, considering that mortality tends to concentrate at a few turbines (e.g. de Lucas *et al.*, 2012), bird abundance at the wind farm, or even vantage point level, might be a metric that is too coarse for studying collision risk.

Here, I investigate the accuracy of vantage point surveys to study characteristics in bird movement as a way to study the spatial distribution of bird activity, instead of their abundance in absolute numbers. In other words, the main focus of this chapter is not on the number of birds or the time they spend flying over the viewshed, but on where they fly - an aspect that has not yet been validated. By explicitly analysing the trajectories captured by vantage point observers, we may investigate associations between movement characteristics and environmental covariates. These, in turn, can be used to estimate the expected long-term spatial distribution of bird activity and use it as a proxy for collision exposure. Note that by studying bird activity levels, we learn about how likely it is to find a bird near a turbine location (collision exposure); not about the probability of collision with it. Existing collision risk models, such as Band (2012) or New *et al.* (2015), can be used to calculate collision risk, given certain exposure (see Masden & Cook, 2016, for a review of available models).

In this chapter, I investigate the accuracy with which parameters used to characterize bird movement behaviour can be estimated from vantage point observations. I use a movement model that describes the dynamics of bird movements based on the reactions of the birds to certain elements of the landscape (e.g. steep slopes or nests). These reactions translate into systematic drift towards, or away from, these landscape elements. Systematic drifts in movement are considered to be driven by potentials, such that areas of low potential are more attractive than areas of high potential (Brillinger *et al.*, 2002; Preisler *et al.*, 2013; Russell *et al.*, 2018). Which particular landscape features should be included in the analyses depends on the ecology of the species of interest and the information available.

I focus on the Verreaux's Eagle, an African raptor affected by wind energy development in South Africa (Simmons, 2005; Ralston-Paton, 2017; Perold *et al.*, 2020). To understand how accurately we can estimate movement parameters using trajectories observed from vantage points, I use simulations. I simulate flight trajectories with a similar structure to those observed from vantage points and resembling the behaviour of the Verreaux´s Eagle. Then, I fit a model to these simulated flights to study how accurately movement parameters can be recovered. To add realism to the simulations, I use parameters extracted from an existing Verreaux's Eagle GPS data set (Murgatroyd *et al.*, 2016b,

2018). I also investigate how different vantage point survey designs affect parameter estimates. Finally, I show how conducting the analysis in a Bayesian framework helps us incorporate existing information about the ecology of the species of interest when estimating model parameters.

3.2 Methodology

In this section, I will briefly outline the historical development of continuous-time movement models and justify why modelling bird movement in a continuous-time framework. Next, I explain how potential-based movement models can be used to describe habitat preferences of an animal, and how these models can be formulated to accommodate autocorrelated velocity. Then, I illustrate how data collected on movement trajectories are affected by the observation process (characteristics of the surveys; essentially, discontinuous observation) during vantage point surveys. This will set the foundation for describing the simulation of flight trajectories that resemble vantage point observations. Finally, I explain how I evaluate the accuracy of movement parameters estimated from flight trajectories observed during vantage point surveys.

3.2.1 Continuous-time movement models

Before estimating exposure to collision, I need to define a model that provides some mechanistic understanding of bird movements and some guidance on where they are likely to fly. I assume that any decision the animal makes to fulfil its physiological needs, will somehow reflect on a movement response (Nathan *et al.*, 2008). More specifically, I assume that birds will be attracted to those locations that are beneficial (e.g. in terms of foraging, resting or saving energy), while they will avoid unfavourable areas (e.g. competitors' territories, predators or dangerous infrastructure). I will refer to these locations as *target* locations from now on. To specify a model that captures the action of these behavioural responses, I will use a *correlated velocity model* (Johnson *et al.*, 2008; Gurarie *et al.*, 2017). Such models describe changes in velocity in continuous time and take the form of stochastic differential equations. In particular, I use a formulation that describes systematic drifts in motion as a result of the action of potential fields (Brillinger *et al.*, 2002; Preisler *et al.*, 2013; Russell *et al.*, 2018).

Animal movement is a process that evolves continuously in time and space; therefore, it is naturally modelled within a continuous-time, continuous-space framework. However, movement trajectories cannot always be observed continuously. Often the observation mechanism used to study animal movement only captures some locations along movement trajectories, separated by some time interval, which motivates their analysis in a discrete-time framework. However, model parameters estimated by discrete-time models are linked to the temporal resolutions of the observed locations (Patterson *et al.*, 2017). For example, step-length and turning angles are dependent on the time separation between locations. In a continuous-time framework, it is acknowledged that the process underlying the discrete observations evolves continuously in time; and therefore parameters try to describe properties that are independent of the rate at which observations are acquired. This is the main appeal of modelling movement in a continuous-time framework. Furthermore, in the analysis of vantage point surveys it will be beneficial to combine information obtained from different data sources, such as telemetry devices and visual surveys. Modelling movement as a continuous-time process will facilitate the integration of data that were acquired with different temporal resolution.

Dunn & Gipson (1977), introduced the Ornstein-Uhlenbeck (OU) process (Uhlenbeck & Ornstein, 1930) to model animal movement with attraction towards an activity centre. Under this model the movement of the animal is driven: i) by a diffusion term, or Brownian motion, that describes a random motion, and ii) a systematic drift towards an activity centre. The resulting model resembles animal home-range behaviour as demonstrated by Dunn & Gipson (1977), in which an individual moves unpredictably, while undertaking its daily activities around an activity centre, such as its den, nest, flock, etc. This model is particularly attractive, since it is regarded as the continuous-time analogue of the first order auto-regressive model in discrete time and it is stationary - i.e. the long-term mean and variance of the process are constant over time (e.g. Schach, 1971). This property allows the calculation of a stationary state that describes the probabilistic spatial distribution of the activity at a far time horizon. The differential form of an OU process reads as follows:

$$\mathrm{d}\mathbf{X}_t = \beta(\boldsymbol{\theta} - \mathbf{X}_t)\ \mathrm{d}t + \sigma\ \mathrm{d}\mathbf{B}_t. \tag{3.1}$$

For sake of readability I use subscript notation instead of parenthesis for time indexing, such that, $\mathrm{d}\mathbf{X}_t \equiv \mathrm{d}\mathbf{X}(t)$. The deterministic term, defining the systematic drift towards the activity centre, is described by the distance between the current position of the animal

$\mathbf{X}_t$ and the activity centre, $\boldsymbol{\theta}$. I will refer to the parameter β as the *drift parameter*. It controls the rate at which the movement trajectory moves towards the activity centre (i.e. modulates the attractive force). The random motion is captured by $\mathbf{B}_t$, which is a Brownian motion, such that $\mathbf{B}_t - \mathbf{B}_s \sim \mathrm{N}(0, (t-s)\mathbf{I})$ for any $t > s$ (see Nelson, 2001). In other words, the variance of the process depends on the time difference between two observations, but not on the time these were observed. Lastly, the coefficient σ, or *diffusion parameter*, scales the random displacements.

Further extensions build upon the concept of potential functions (Brillinger *et al.*, 2002; Preisler *et al.*, 2004, 2013), which shape the attractive forces, captured by the drift term, in the stochastic differential equation. A later generalization of the OU process by Harris & Blackwell (2013) incorporates a more flexible framework for linking spatial heterogeneity to movement patterns by allowing the specification of different model parameters for different habitats.

Johnson *et al.* (2008) pioneered the use of the OU velocity process (Uhlenbeck & Ornstein, 1930) to describe the velocity at which an animal travels, rather than its location.

$$\begin{aligned} \mathbf{X}_t &= \mathbf{X}(0) + \int_0^t \mathbf{V}(\tau)\ \mathrm{d}\tau \\ \mathrm{d}\mathbf{X}_t &= \mathbf{V}_t\ \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= \beta(\gamma - \mathbf{V}_t)\ \mathrm{d}t + \sigma\ \mathrm{d}\mathbf{B}_t. \end{aligned} \tag{3.2}$$

In this case γ represents a preferred, or long-term, velocity (and it may be a function of time and space, as we will see shortly), rather than a location, and $\mathbf{V}_t$ the animal velocity at time t. The main advantage of this approach is that, once the velocity is integrated, the model describes a continuous, smooth path rather than a trajectory with sharp corners produced by sudden changes in direction typical of equation 3.1. This is a desirable property for models that describe bird movement in continuous time, since flight trajectories sampled at short time intervals should appear smooth for the most part, due to the connection between speed and stability in flight (Biewener & Patek, 2018). Gurarie *et al.* (2017) coined the term *correlated velocity models* to describe continuous-time models for velocity, based on an OU process and described some of the most common forms of such models. I will mention a broader class of diffusion processes introduced by Blackwell (1997): with the term "mixed Ornstein-Uhlenbeck process" he defines a generalization of the basic model in which the animal may take on different states, which in turn, influence the parameters of the OU process. More recently, Michelot & Blackwell

(2019) extended correlated velocity models to accommodate different movement modes or states. These, two classes of models consider the movement process to switch between discrete movement modes. In this chapter, I will consider a model that features continuous states, in line with that proposed by Johnson *et al.* (2008). Continuous state changes are captured by the influence of continuous covariates on movement parameters, as I explain in section 3.2.3.

3.2.2 Potential-based velocity model

I use the term *potential-based velocity model* to refer to a particular case of a correlated velocity model (expression 3.2) in which the velocity γ is given by the gradient of a potential function. Brillinger *et al.* (2002) introduced the notion of using potential fields to describe the attraction or avoidance reactions of animals towards certain locations. This concept was further developed by Preisler *et al.* (2004, 2013). The idea is to consider areas attractive to the animal, having lower potential than areas that are less attractive (or repulsive). Then, the animal is expected to move from areas of high potential to areas of lower potential (figure 3.1).

In the original formulation of the potential-based movement model, Brillinger *et al.* (2002) describes the equations of motion in a deterministic system as:

$$\begin{aligned} \mathrm{d}\mathbf{X}_t &= \mathbf{V}_t \, \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= -\beta(\nabla H + \mathbf{V}_t) \, \mathrm{d}t. \end{aligned} \tag{3.3}$$

Form this expression, it is implicit that $\mathbf{X}_t$ is the integral of the velocity $\mathbf{V}_t$ with respect to time (see expression 3.2). In expression 3.3 the change in velocity is opposite to ∇H, the gradient of the potential function $H(\mathbf{X}_t)$, which is a vector formed by the partial derivatives of $H(\cdot)$ evaluated at the point $\mathbf{X}_t$. The change in velocity is also opposite to the current velocity $\mathbf{V}_t$. Therefore, the drift parameter β modulates the rate at which velocity relaxes to $-\nabla H$. Thus, for $\nabla H = 0$, β modulates the rate at which the velocity declines to zero. That, $\nabla H \neq 0$ implies that velocity relaxes to $-\nabla H$ instead of zero, yet other properties of the system remain unchanged. Regardless of the value of ∇H, for large values of β, the contribution of the velocity term vanishes quickly and it can be shown that the system of equations 3.3 is well approximated by (see Smoluchowski approximation in Nelson (2001)):

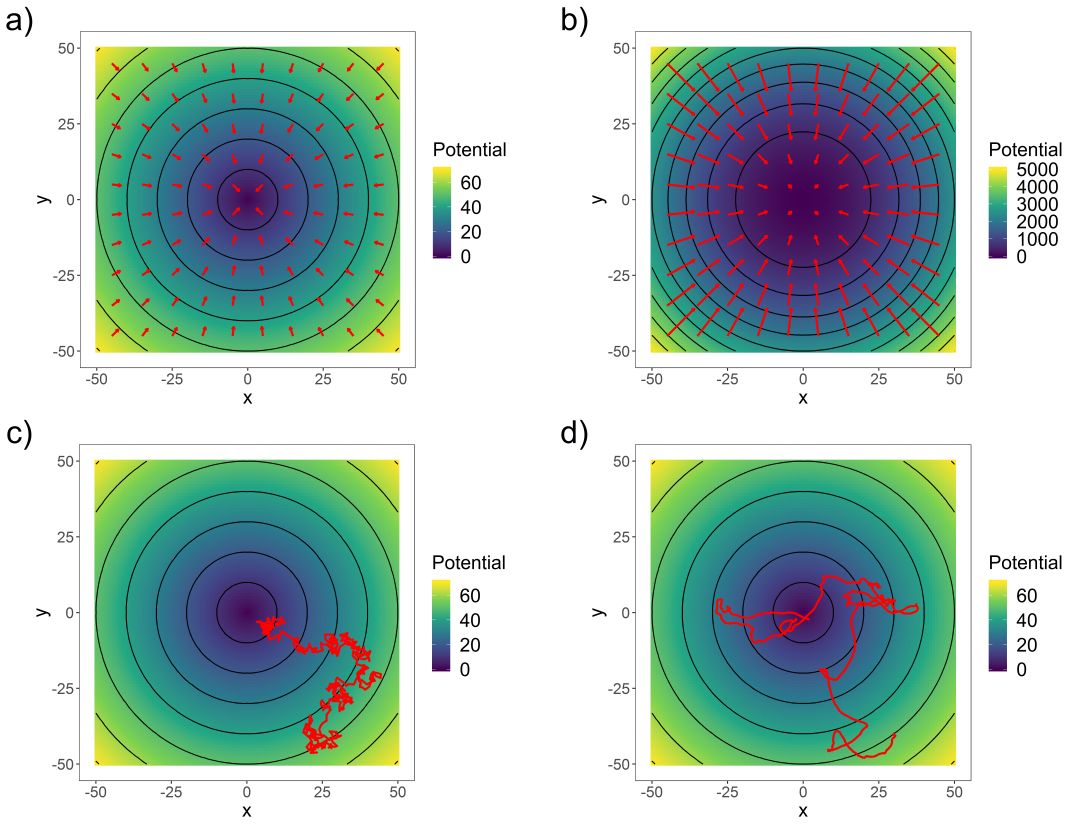


Figure 3.1: Motion within potential fields. Top panels: potential surfaces describing attraction towards the point $\boldsymbol{\theta} = (0, 0)$. a) The potential functions correponding to the potential fields $H_1 = \sqrt{(x-\theta_x)^2 + (y-\theta_y)^2}$ (potential equal to the distance from $\boldsymbol{\theta}$). b) $H_2 = (x-\theta_x)^2 + (y-\theta_y)^2$ on the right (potential equal to the square of the distance from $\boldsymbol{\theta}$). The red arrows show the magnitude (up to a scaling constant) and the inverse direction of the gradient. Bottom panels: trajectories simulated from c) Ornstein-Uhlenbeck process with independent increments in velocity and d) potential-based velocity process with autocorrelated velocity.

$$\mathrm{d}\mathbf{X}_t = -\beta \nabla H \ \mathrm{d}t \tag{3.4}$$

To complete the model, a stochastic component is added to this equation in the form of a Brownian motion modulated by a diffusion parameter σ, in a similar way to equations 3.1 and 3.2:

$$\mathrm{d}\mathbf{X}_t = -\beta \nabla H \ \mathrm{d}t + \sigma \ \mathrm{d}\mathbf{B}_t \tag{3.5}$$

This formulation is appealing for its simplicity and can be a good approximation to systems in which friction is high (drift parameter β is high). High friction means low persistence in velocity, apart from that induced by the potential (figure 3.1). In other words, conditioning on the potential, there is little autocorrelation in consecutive velocity measurements. Whether this is a reasonable assumption to make when analysing animal movement depends on each particular case. In general, if velocity measurements are obtained far apart in time (e.g. hourly measurements), they can possibly be assumed to be independent from each other. However, with new high-resolution tracking technologies, measurements can be acquired at increasing rates (e.g. few minutes or even seconds), at which measurements are likely correlated. From vantage points, flight trajectories are observed in continuous-time, exacerbating the autocorrelation issue (despite having to discretize the observed trajectories into sequences of locations for analytic purposes).

To account for persistence in the velocity of movement, I will use a potential-based velocity model (figure 3.1). This model combines the potential field framework with the autocorrelation properties of the Ornstein-Uhlenbeck velocity process shown in equations 3.2 and 3.3 (Nelson, 2001; Johnson *et al.*, 2008; Russell *et al.*, 2018).

$$\begin{aligned} \mathrm{d}\mathbf{X}_t &= \mathbf{V}_t \, \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= \beta(\gamma(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t) \, \mathrm{d}t + \sigma \, \mathrm{d}\mathbf{B}_t. \end{aligned} \tag{3.6}$$

In this model, the velocity tends to relax to a preferred velocity $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ at a rate given by the drift parameter β. Note that the preferred velocity γ in equation 3.2 is a special case of 3.6, where $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ equals a constant. Similarly, the term ∇H in 3.3 is a special case of 3.6 where $\gamma(\cdot)$ depends on the location of the animal $\mathbf{X}_t$ but not on its velocity $\mathbf{V}_t$. Therefore the proposed model 3.6 is more general than the previous ones. This kind of model can be conceived to be a form of advection-diffusion model for velocity, in which the drift is the response of the velocity to the pull by advective forces, while the Brownian motion adds diffusive behaviour the velocity (Ovaskainen *et al.*, 2016).

The particular form of γ that I will use for this study is

$$\gamma(\mathbf{X}_t, \mathbf{V}_t) = -\nabla H |\mathbf{V}_t|. \tag{3.7}$$

The preferred velocity $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ is given by the negative of the gradient of a potential function ∇H times the norm of the velocity (i.e. speed) at time t. Therefore, the animal is still considered to move within a potential field $H(\mathbf{X}_t)$, where areas of low potential are

more attractive than areas of high potential. For this application, I define the potential to be equal to the distance to attractive or repulsive locations $\boldsymbol{\theta} = (\theta_x, \theta_y)$ (see figure 3.1), such that

$$H(\mathbf{X}_t) = \sqrt{(x_t - \theta_x)^2 + (y_t - \theta_y)^2}, \tag{3.8}$$

where $\mathbf{X}_t = (x_t, y_t)$ and therefore,

$$\nabla H = \begin{bmatrix} \frac{\partial H}{\partial x} \\ \frac{\partial H}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{x_t - \theta_x}{r_t} \\ \frac{y_t - \theta_y}{r_t} \end{bmatrix}, \tag{3.9}$$

where r_t stands for the Euclidean (or radial) distance between $\mathbf{X}_t$ and $\boldsymbol{\theta}$, which makes the norm of ∇H unitary for all $\mathbf{X}_t$. A unitary norm implies that the preferred velocity has magnitude proportional to the speed at which the animal is travelling at the current time, and direction towards locations of low potential, as specified by H. The reason for the choice of this particular form is that only the direction of movement, and not the movement speed, is affected by the potential. This means that the bird steers towards points of attraction, but does not necessarily change its travelling speed. We could also use other forms of potential and normalize the gradient by dividing its components by its norm to make the norm of ∇H unitary. This characteristic is particularly attractive to use with vantage point surveys because instantaneous speed is typically impossible to measure. Using a potential that influences only direction but not speed makes models fitted to telemetry data (which is time-referenced) and vantage point data more comparable.

If there are K target locations (attractive or repulsive) at time t, then,

$$\mathrm{d}\mathbf{V}_t = \sum_{k=1}^{K} \beta_k(\gamma_k(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t)\ \mathrm{d}t + \sigma\ \mathrm{d}\mathbf{B}_t, \tag{3.10}$$

where the resulting drift of the animal (expected change in velocity) is given by the sum of the actions of all K potentials, $\sum_{k=1}^{K} \beta_k \gamma_k(\mathbf{X}_t, \mathbf{V}_t)\ \mathrm{d}t$, and by the effect of friction $\sum_{k=1}^{K} -\beta_k \mathbf{V}_t\ \mathrm{d}t$.

To find the velocity at which the animal is travelling at time t we may solve equation 3.10. To do so, it is convenient to consider the fact that $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ is likely to change slowly compared to the speed of movement. More precisely, this is to say that for small

time horizons δt, the points of attraction are likely to not to change. For example, if the species is attracted to steep slopes, the particular slope the bird is attracted to is likely to be the same one for some time. Because of the stochastic term in the equation 3.10, the solution has the form of a stochastic process. And because the stochastic term is Gaussian (a Brownian motion), the solution can be approximated by Gaussian process (it is an approximation because of the non-linear dynamics). The stochastic process is represented by a transitional density, that describes the transition from one state $\mathbf{V}_0$ to a new state $\mathbf{V}_t$, t time units away. Details on the derivation of this expression can be found in appendix A. Because we assume that $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ does not change in δt time units, the approximate transition density is only valid at short time horizons. Therefore, to investigate the long-term behaviour of the system we may numerically integrate small steps in time using expression 3.10 iteratively and investigate the distribution of locations t time units away from $\mathbf{V}_0$ (where $t > 0$ is a multiple of δt; see appendix D for a simulation code example).

3.2.3 A movement model for the Verreaux's Eagle

The Verreaux's Eagle is a widespread African raptor (Simmons, 2005). The species is of conservation concern in South Africa (Taylor *et al.*, 2015), prone to collision with wind turbines, and its range is likely to overlap substantially with wind energy development in this country (Ralston-Paton *et al.*, 2017; Perold *et al.*, 2020).

Verreaux's Eagles are considered specialist hunters of Rock Hyrax (Simmons, 2005), yet Murgatroyd *et al.* (2016a) described a flexible prey consumption in transformed habitats (e.g. croplands). Like other soaring raptors, Verreaux's Eagles use thermal and slope uplift depending on topography and atmospheric conditions (Murgatroyd *et al.*, 2018). Eagles living in low, flat terrain, are more prone to thermal soaring, than those inhabiting areas where steep slopes are more abundant. Verreaux's Eagles are central place foragers that maintain their territories year round, spending most nights near the nest and making excursions of longest duration towards midday (Murgatroyd *et al.*, 2016b).

In this study, movement parameters to generate the simulations are estimated from two data sets that contain GPS locations of two Verreaux's Eagles tagged in the Western Cape of South Africa during the pre-breeding period - a total tracking time of three months for each eagle (see Murgatroyd *et al.*, 2016b, 2018). One eagle, a female, was tagged in the Sandveld, which is a rather flat region, dominated by croplands, with some

remnants of natural vegetation. I refer to this individual as *eagle1* and use the parameter values estimated for her to simulate flights observed from vantage points. The other eagle *eagle2*, a male, was tagged in the Cederberg, a mountainous area of more rugged terrain, covered in natural vegetation. I use the parameters fitted to this individual only to define prior distributions. As I explain in section 3.2.5, models are fitted to vantage point simulations in a Bayesian framework, and having a different individual to generate priors will be useful for studying the effect of priors on model inference.

University of Amsterdam Bird Tracking System (UvA-BiTS) GPS-loggers (Bouten *et al.*, 2013) were attached to the eagles in 2012 (Cederberg) and 2013 (Sandveld), mounted on a backpack harness (see Murgatroyd *et al.*, 2016b, for tagging details). Locations were recorded at a variable frequency of up to one location every three seconds. High resolution tracking data ($>$ 1 location per minute) were more abundant for the Cederberg data. To homogenize the number of locations analysed for both eagles, I ran a moving window, removing locations less than 1 minute apart. Since, I am most interested in modelling flight behaviour, I excluded all night-time locations, when the activity of the eagles is minimal. I excluded the first day after the eagles were trapped, since the behaviour may not be representative of their regular activity. Finally, I eliminated days in which the average time between fixes was greater than an hour, since long periods without information may impact the model's ability to associate changes in behaviour with environmental conditions. This is because the direction of travel between start and end locations in the long term may be very different from the direction of travel in the short term, and the association between long- and short-term directions with the slope of the potential might differ. The choice of one hour threshold was subjective and based on exploratory observations revealing that most data ($>$ 99%) were way below this threshold.

In addition to the general model structure described in expression 3.10, we can build more realistic models tailored to a particular bird species. In this study, I develop a movement model for the Verreaux's Eagle, based on what is known for the species, and fit this model to the GPS data of the Sandveld and Cederberg eagles. The posterior distribution of the parameters of these models will be used for either simulating flights observed from vantage points (*eagle1*), or as priors when fitting models to these simulated flights (*eagle2*), as we will see in sections 3.2.4 and 3.2.5.

For the Verreaux's Eagle, I hypothesise a systematic drift towards steep slopes, towards its own nest and possibly a drift either towards to, or away from, a neighbours nest. I incorporate the fact that drift and diffusion contributions to the movement of the eagles

may change in different days, as well as within the same day. To model changes in drift towards any type of target location k (steep slopes, own nest or neighbour's nest) with time of day, I model drift parameters β in expression 3.10 as:

$$\beta_{kt} = \zeta_k + \alpha_k t'^2, \tag{3.11}$$

where t' is a variable representing time to noon in hours. The value of the drift associated with type of location k at midday ($t' = 0$) is given by ζ_k. At other times, the change in the value of the drift is proportional to the square time difference with midday in hours, by a factor of α. This term produces a drift towards location k that changes with time of day, and is either maximum or minimum towards midday depending on the sign of α.

Similarly, I define a model that imposes an effect of time of the day on the diffusion parameter σ. The diffusion parameter at time t has the form:

$$\sigma_t = \nu e^{t'^2 \psi^{-2}}, \tag{3.12}$$

where t' represents time to noon in hours. The coefficient ψ controls the variability in diffusion as day progresses (similar to the standard deviation of the Gaussian distribution). The parameter ν is a scaling factor that modifies the maximum value of the diffusion for the day.

The effects just explained describe the change in drift and diffusion within the same day. The proposed model also contemplate changes in drift and diffusion between different days. Thus, I model the parameters that define drift and diffusion terms as being randomly sampled each day from a population of parameters, such that for day j,

$$\begin{aligned} \beta_{kjt} &= \zeta_{kj} + \alpha_{kj} t'^2 \\ \zeta_{kj} &\sim \mathrm{N}(\mu_{\zeta_k},\ \sigma^2_{\zeta_k}) \\ \alpha_{kj} &\sim \mathrm{N}(\mu_{\alpha_k},\ \sigma^2_{\alpha_k}), \end{aligned} \tag{3.13}$$

and also,

$$\begin{aligned} \sigma_{jt} &= \nu e^{t'^2 \psi_j^{-2}} \\ \psi_j &\sim \mathrm{N}(\mu_\psi,\ \sigma^2_\psi). \end{aligned} \tag{3.14}$$

Therefore, the terms that form the drift and diffusion parameters for each day j, ζ_{kj}, α_{kj} and ψ_j are sampled each day from a distribution of possible parameters with mean μ_{ζ_k}, μ_{α_k} and μ_ψ and standard deviation σ_{ζ_k}, σ_{α_k} and σ_ψ, respectively. Note that this formulation implies that there is no correlation between days (daily parameters are independent from each other). A more complicated version of the model could consider daily parameters being indexed by time and incorporate temporal autocorrelation in consecutive days. I will assume independence for simplicity, and because the focus on this chapter is on the simulation of trajectories that resemble eagle behaviour, and not on providing a detailed description of the movement ecology of the Verreaux's Eagle.

Once some principles to model bird movement are defined, we can use data to find parameter values that best describe observed movement patterns. By the reverse process, we may simulate movements that are consistent with movement patterns of our choice. Details on the simulation of Verreaux's Eagle movements observed from vantage points are provided in section 3.2.4.

3.2.4 Vantage point simulations

The main objective in this chapter is to investigate how accurately movement parameters can be estimated from vantage point observations. To do so, I use the movement parameters fitted to the GPS tracking locations of two Verreaux's Eagles (section 3.2.3), to simulate flight trajectories observed from vantage points. Then, I try to recover these parameters by fitting a similar model to the simulations. To produce simulations that resemble flight trajectories recorded during vantage point surveys, we need to understand: i) how the birds targeted by these surveys move and then, ii) how the observation process influences the data we capture. In section 3.2.3, we have just defined a model structure for the movement of the eagles. Now, I explain how vantage point survey data are simulated.

Different types of surveys allow us to collect different types of data. For example, attaching a GPS device to a bird allows us to observe the location of a single bird (per GPS installed) with a certain frequency (figure 3.2). There are no boundaries to the observable area and, as long as the GPS tracker is transmitting, we can receive the location of the bird, no matter where it is. On the other hand, by placing an observer on a vantage point, we can potentially observe in continuous-time the flights of all birds crossing an observable area; which is limited to one or two kilometres around the vantage point, and

only while the observer is present (figure 3.2). The more often we visit a vantage point, the more flights we observe at that particular location - provided that birds cross the observable area. On the other hand, the more different vantage points we visit, the more varied the conditions we observe, and the better we understand the general behaviour of the birds. Observers will surely miss some birds and introduce some error on the location, shape, timing of the flights, etc. when capturing data. But for now, to tackle one issue at a time, I ignore these errors and focus only on temporal and spatial constrains associated with the location and timing of the vantage point surveys (see chapter 4 for more details on observer error).

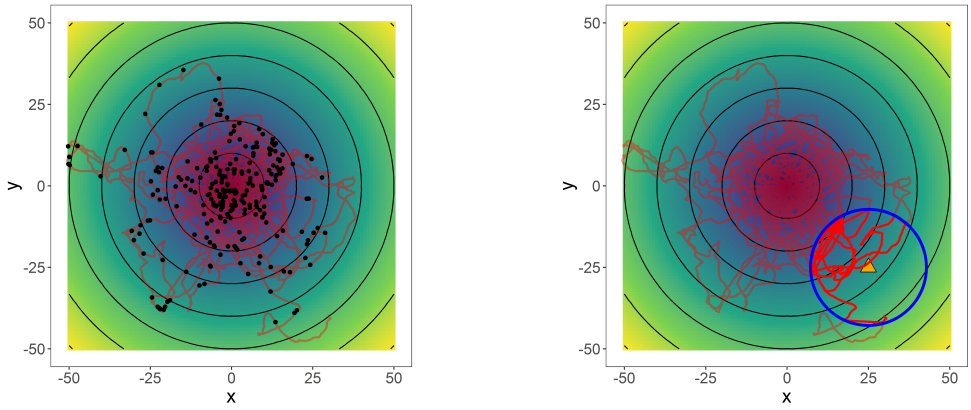


Figure 3.2: Illustration of how the characteristics of different survey methods affect the observation of flight trajectories of a bird moving on a potential field (yellow areas have higher potential than blue areas). In dark red, the full trajectory traced by the bird - hidden to the observer. On the left panel, the black points represent the locations made available by a GPS tracking device. On the right panel, the orange triangle represents a vantage point; the area enclosed by the blue circle is the portion of space visible to an observer and the bright red path is the portion of the flight trajectory visible to the observer.

These principles lead to at least two ways of approaching the simulation of flights observed from vantage points to evaluate the effectiveness of this survey method at capturing bird movement patterns. A set of common preparation steps are needed for both approaches: i) define a landscape within which bird movements and vantage point surveys will be simulated, ii) define a structure for the movement model that will be used for the simulations, and iii) choose sensible parameters to input into the models to generate movements that resemble real bird flights.

Once the basic elements for the simulation of bird flights are in place, we may proceed in two ways to recreate vantage point surveys. We may simulate a long, continuous flight

trajectory and then take sections of it by randomly overlaying circular areas representing vantage point viewsheds, in a similar way as figure 3.2 was generated. This approach is appealing because it resembles closely what happens in reality. However it is quite inefficient from the point of view of computing power, because some surveys simulated this way might not collect any activity at all. Also, this approach provides some insight on how many vantage points we need to deploy, and how long observers need to survey for to obtain a certain amount of data, but it is difficult to allocate a pre-defined number of flights (or flight time) to each survey. Therefore, it is difficult to disentangle the effect of number of vantage points and flight time per vantage point on activity estimates, or how these effects are related to particular landscape configurations (spatial distribution of covariates related to movement behaviour).

A more effective approach, and the one I use in this study, is to first simulate a number of areas observed from vantage points (circular areas of certain radius), within the defined broader landscape. Then, I simulate flights of certain duration within the vantage point observable areas only. In this way we control the spatial coverage of the landscape through the number of vantage points generated, and we also control the amount of data collected from each vantage point, by specifying the duration of the flights simulated within each vantage point viewshed. However, we do not address the issue of how many vantage points we need to deploy and for how long we need to survey them for in order to obtain certain amount of flight time. Yet, it is reasonable to assume that these quantities will depend on the intensity of bird activity in the study region, the characteristics of the landscape and the spatial distribution of the vantage points. Therefore, these questions should be addressed separately for each particular study.

As mentioned earlier, model parameters used in the simulation of flights observed from vantage points are obtained from a movement model fitted to the GPS tracking locations of two Verreaux's Eagles (section 3.2.3). However, the model used for the simulations, and the model fitted to the simulated flights to recover the parameters, are slightly different. It was purposefully done like this to reproduce situations often faced in practice, when bird movement is driven by variables that we cannot observe from vantage points. More specifically, the model fitted to the GPS data set incorporates systematic drift towards steep slopes, the eagle's own nest and the nest of a neighbouring eagle. Simulations of flights observed from vantage points were generated considering all these variables. However, the model fitted to the simulations only attempts to recover the parameter related to drift towards steep slopes, because in real vantage point situations we often ignore the location of eagles' nests or we are unable to identify individual eagles.

Simulation procedure

To simulate bird movement, I use the potential-based velocity model described in sections 3.2.2 and 3.2.3. Sensible parameter values to input into the model and to generate flights that resemble real bird activity, are obtained from a GPS data set of two Verreaux's eagles (section 3.2.3).

The simulations proceed as follows:

1. *Define landscape.* I define the broader region to simulate bird movements by the minimum convex polygon (see Worton, 1987) enclosing the GPS locations of one of the tracked eagles, *eagle1*. This ensures that the landscape roughly contains the basic elements that the eagle needs in its daily activities.

2. *Generate areas observed from vantage points.* To investigate how survey design influences parameter estimates, I simulated different scenarios, randomly choosing 2, 4, 6, 8 and 10 vantage point locations within the defined landscape. Around these vantage points, I generate a circular viewshed of 2 km radius. The only constraint when sampling vantage point locations is that their viewsheds do not overlap.

3. *Simulate bird movements.* To investigate how the amount of data collected from each vantage point affects estimates of movement parameters, I change the flight duration per vantage point. I simulated accumulated flight durations of 3, 5, 10, 15 and 30 minutes per vantage point. The simulation of each flight runs as follows (see appendix D for an example of simulation code):

   (a) Sample all necessary parameters to build time-varying drift and diffusion parameters (see section 5.2.4 and 3.2.3) from the posterior distributions obtained by fitting a model to the eagle's GPS data set.

   (b) Randomly select a time of day, between 7 am and 7 pm (when eagles are most active), at which the flight starts.

   (c) Randomly select an initial position and velocity within the corresponding vantage point viewshed. The initial velocities were sampled from a distribution centred at zero with the variance estimated for the stationary distribution of the velocity process (see appendix A).

   (d) Extract the value of the gradient of the potential at the current location.

(e) Input parameter and gradient values into expression 3.10 to simulate a velocity value, considering a time step of 0.5 minutes.

(f) Update position.

(g) Iterate steps (d) to (f) until the desired duration is reached or the trajectory leaves the vantage point viewshed. In the latter case, this flight is finished, the duration of the flight is subtracted from the total target duration, and a new flight commences from step (a).

Points two and three above were repeated five times, resulting in five repetitions of each combination of flight time and number of vantage points.

3.2.5 Model fitting

Models were fitted in a Bayesian framework, which has the desirable property of naturally accommodating prior knowledge into the modelling process (Ellison, 2004; Gelman *et al.*, 2014; McElreath, 2019). Prior distributions (or just *priors*), inform the model of what are reasonable values for the model parameters to take on. Once data is analysed, our prior knowledge is updated with information contained in the data and we obtain what is called a posterior distribution for the parameters (or just *posterior*). This idea opens the possibility of combining several data sources to improve inference.

For example, we may use movement parameters estimated from telemetry data as priors for fitting movement models to vantage point data. Vantage point surveys typically provide small data sets, compared to telemetry methods, but they can be designed to collect information from a specific area of interest. Also, the collection of vantage point data is relatively simple and non-invasive to the birds. On the other hand, telemetry provides large location data sets, yet there is no guarantee of the tagged birds using the area of interest - or at least behave in the same way as the birds in this area. In addition, the deployment of tracking devices is complex and often invasive to the birds (Hebblewhite & Haydon, 2010; Kays *et al.*, 2015). Within a Bayesian framework, we may use existing telemetry data for understanding the general behaviour of the species, and then, update this prior knowledge with local information from vantage point surveys.

Posterior parameter distributions combine the information contained in the prior distribution and in the data (see Gelman *et al.*, 2014; McElreath, 2019, for at-length discussions on the use of priors). When data is limited, as it often happens with vantage

point surveys, and priors are informative (have small variance, small uncertainty), the posterior distribution is dominated by the prior. In contrast, when data are informative and priors are weak (large variance, large uncertainty), the posterior is dominated by the data. The model for the eagle's GPS data was given relatively weak priors (for the standardized covariates) $\mathrm{N}(0, 0.5)$ for hyper-parameters μ_ζ and μ_α, $\mathrm{half-Cauchy}(0, 0.2)$ for hyper-parameters σ_ζ and σ_α, $\mathrm{N}(5, 1)$ for hyper-parameters μ_ψ, $\mathrm{half-Cauchy}(0, 1)$ for hyper-parameters σ_ψ, and $\mathrm{half-Cauchy}(0, 2.5)$ for ν.

To understand the effect of different types of priors on parameter estimated from vantage point data, I fit the same potential-based velocity models to the vantage point simulations using three different priors and compare the resulting posterior distributions:

i) Vague priors (same as for GPS data), that let the likelihood of the data dominate the posterior. Using vague priors, we consider that there is no previous information about the bird species we are studying. Although using vague priors is a common choice, there is almost always information available that can help us choose better priors and we should make use of it (see Gelman *et al.*, 2014; McElreath, 2019).

ii) Informative priors, based on the posterior distribution of the parameters used to simulate the data (*eagle1*). Choosing as priors the distributions that actually generated the parameter values for the simulation, represents a scenario where we know, prior to fitting a model to the vantage point data, what parameters values govern the movement of the birds. This scenario is very unlikely in reality, but it helps us understand what the priors are doing in the other scenarios, thus, working similar to a control.

iii) Informative but biased priors, based on the posterior distribution of parameters fitted to a GPS data set, but in this case from a different eagle (*eagle2*). Since these priors are biased in relation to the data-generating parameters, they will help us understand to what degree, parameter values obtained from a different bird may help inference, and also to what degree prior knowledge can be updated with vantage point observations. I add some extra variance to the biased priors to capture the uncertainty arising from using parameter values of a different eagle. This third scenario is probably the most common we may find during wind farm monitoring. Under this circumstance, we have information available about what the movement parameters for a certain species look like, but we are unsure to what degree this information can be extrapolated to the individuals we observe in our study area. For example, we may have GPS tracking data from a different eagle

living in a different environment, or we may have to adapt plausible values from the literature.

Models were fitted in R (R Core Team, 2018), using the package `rstan` (Stan Development Team, 2018b). This package implements Hamiltonian Monte Carlo sampling which is an efficient Markov Chain Monte Carlo (MCMC) algorithm that samples the posterior distribution of the parameters using Hamiltonian equations (Neal, 2011). The general idea behind MCMC is that it is an iterative sampling procedure used to approximate probability distributions without closed form (Gelman *et al.*, 2014). In Bayesian model fitting, the objective is to create a sequence of parameter samples that in the long-term converges to the posterior distribution of the parameters. Parameter values are sampled sequentially from a transitional density typically centred a the previous sample, forming a chain with the Markov property (i.e. samples depend only on the previous draw). The posterior is evaluated at each step (i.e. for each sampled parameter value) and parameter values are accepted or rejected following some pre-defined rule. Hamiltonian Monte Carlo incorporates information about the gradient of the target posterior to accept or reject new samples. Thus, it improves the efficiency of the algorithm by sampling more intensively from regions of high density (Neal, 2011). Because the starting point of the chains are somehow arbitrary the values at the beginning of the chains are not necessarily representative of the long-term stationary distribution, and therefore, it is good practice to discard them. In addition, to explore the parameter space effectively, it is recommended to produce several chains with different starting values and make sure that they all converge to similar distributions.

I ran four Markov chains with 4000 iterations each, discarding the first 2000, which are used for adapting the chains. Convergence of the model was assessed visually by inspecting that all the chains explore a similar distribution of parameter values. I also rely on the $\hat{R}$ statistic, which provides a measure of the expected reduction in posterior variance, should posterior sampling carry on to infinity (Gelman *et al.*, 2014). In practice, although Hamiltonian Monte Carlo is an effective algorithm to sample from the posterior, samples along the Markov chains are autocorrelated to some extent. High autocorrelation reduces precision in posterior estimates, and therefore I used the number of effective samples as a guide to assess posterior precision (Gelman *et al.*, 2014).

3.2.6 Analysis of simulation results

The objective in this study is to investigate how effective vantage point surveys are for learning about movement behaviour and habitat preferences of soaring birds. To do so, I generate synthetic data (simulations) that resemble Verreaux's eagle flights observed from vantage points and use them to estimate movement parameters. These flights are generated in such a way that their trajectories evolve according to some known habitat-driven behaviour that I based on GPS tracking studies on the species (see section 3.2.4 for details on flight simulations).

I simulate 25 scenarios with different combinations of number of vantage points and flight duration per vantage point. To have replicates of each scenario, each of the 25 combinations are repeated five times. Then, movement models are fitted to the flights simulated under the different survey scenarios, using different priors. Finally, the posterior distributions of the model parameters are compared with the distributions that were used to generate the data. The difference between the data-generating parameters and the parameters estimated from the simulated data provides a measure of how effective vantage point observations are at capturing behavioural information that relates to habitat preferences and use of space. I also investigate how number of vantage points, flight duration per vantage point and type of prior, affect posterior distributions.

I use two metrics to compare the parameter distributions from which the simulations were generated and the posterior distributions produced by the model fitted to the simulated vantage point data: the relative bias (RB) and the overlapping coefficient (OVL) (Weitzman, 1970; Clemons & Bradley, 2000).

The relative bias provides a measure of how different a parameter posterior distribution is with respect to the data-generating parameter distribution on average, and relative to the parameter value:

$$RB(\hat{\theta}_s) = \frac{E[\theta|y_s] - E[\theta]}{E[\theta]}. \tag{3.15}$$

The term $\hat{\theta}_s = E[\theta|y_s]$ is the expected value of the posterior distribution of a model parameter θ fitted to the simulated data y_s and $E[\theta]$ is the expected value of the distribution from which θ was sampled to generate the simulations (this is the data-generating distribution, not to be confused with the prior of θ). The subscript s indexes the different repetitions of a particular vantage point survey scenario (characterized by a number

of vantage points and a flight time per vantage point). To make bias in the posterior distribution of different parameters comparable, we divide by $E[\theta]$. In Bayesian statistics inference is usually not based on point estimates, such as $\hat{\theta}$, but rather on the posterior distribution $f(\theta|y_s)$. However, the concept of bias is still useful to calculate average differences between two distributions.

The second metric I use to compare posterior and data-generating distributions is the overlapping coefficient. OVL represents the overlapping area between the distribution from which parameter values were sampled to simulate data and the posterior distribution of the parameters given the simulated data. It provides a measure of agreement between the data-generating distribution and the posterior distribution based on how much they overlap each other. An overlap of zero means that posterior and data-generating distributions are completely disjoint, and on the contrary; an overlap of one means that posterior and data-generating distributions agree perfectly. The overlapping coefficient is calculated as:

$$OVL(y_s) = \int min\{f(\theta|y_s), f(\theta)\}\ \mathrm{d}\theta, \tag{3.16}$$

where $f(\theta|y_s)$ is the posterior distribution of a parameter θ estimated from the simulated data y_s and $f(\theta)$ is the probability density function from which θ was sampled to generate the simulated data (again not to be confused with the prior distribution of θ). As in the formulation of RB, the subscript s indexes the different repetitions of a particular vantage point survey scenario. When there is no agreement between posterior and data-generating distributions, we still need a measure of how different posterior values are from data-generating values and that is what RB is used for.

To summarize the results I fitted a Generalized Additive Model (GAM) to the RB and OVL values obtained for different combinations of number of vantage points, flight duration per vantage point, and types of prior. Number of vantage points and flight duration per vantage point were explained by smooth (non-linear) terms, while the type of prior was included as a linear term. I used the software R (R Core Team, 2018) to run all the analyses and simulations, with the added functionality of the `tidyverse` packages (Wickham, 2017). GAMs were fitted using the package `mgcv` (Wood, 2017).

3.3 Results

3.3.1 Summary of data-generating model

In this section, I will briefly describe the results obtained after fitting the potential-based velocity model (see sections 3.2.2 and 3.2.3) to the GPS tracking locations of the two Verreaux's Eagles. The model parameters described in this section will be later used for simulating Verreaux's Eagle flights observed from vantage points (see section 3.3.2). In particular, the parameters estimated for the Sandveld eagle *eagle1* are used for simulating flights. The parameters estimated for the Cederberg eagle *eagle2* were only used as priors in models fitted to simulated vantage point data, and not to simulate flights. These priors allowed me to investigate how fitting a model to vantage point observations of one eagle, using priors coming from a different eagle, affect parameter posterior distributions.

A 90% posterior predictive interval constructed from the drift component of the models fitted to the eagles' data covered 91% of both the Sandveld and Cederberg eagle data (figure 3.3), showing that model predictions are constrained within sensible limits. However, the model fails to capture periods with no change in velocity, which probably corresponds to periods of eagle inactivity. To assess model fit for the diffusion term, I computed the "residuals" by subtracting the median posterior drift from the observed change in velocity (which was the response variable), then, I compared these residuals with the posterior distribution of the diffusion (figure 3.5). The model seems to capture well the diffusion observed for the Sandveld eagle, whereas the diffusion for the Cederberg eagle seems to be shifted towards the end of the day, in relation to model estimates (figure 3.5). Overall, the posterior predictive checks suggest some misalignment between model predictions and observed eagle behaviour, but the model captures the general behaviour well enough to produce the rough representation needed for the simulations (see discussion in section 3.4). Data simulated from the model again shows some similarities with the observed data, such as the reversions to a central point typical of central place foragers (figure 3.6). However, the model produces trajectories that are smoother than those featured by the data, suggesting that eagle behaviour would probably better captured by model that incorporate multiple movement modes, such as resting and flying, for example.

Under the potential-based model specified for the Verreaux's Eagle, the Sandveld indi-

vidual showed some systematic drift towards steep slopes (see $\mu_{\zeta_{slp}}$ and $\mu_{\alpha_{slp}}$ terms in table 3.1 and figure 3.4). This attraction is more variable at the beginning and end of the day. Drift towards nesting sites, both own and neighbours' was positive for the most part (see $\mu_{\zeta_{nest}}$ and $\mu_{\alpha_{nest}}$ for drift towards own nest, and $\mu_{\zeta_{ngh}}$ and $\mu_{\alpha_{ngh}}$ for neighbours nest in table 3.1). The drift towards the eagle's own nest was consistently positive throughout the day, although weaker towards midday (figure 3.4). The drift towards the neighbour's nest also tends to decrease around midday; however, the effect of time of day was greater for the drift towards this nest, than it is for the eagle's own nest (see $\mu_{\alpha_{ngh}}$ compared to $\mu_{\alpha_{nest}}$ in table 3.1). Such was the effect of time of day on the drift towards the neighbour's nest, that it became negative (repulsion) around midday in some days (figure 3.4).

The model for the Sandveld eagle also showed relatively large variability in attraction towards steep slopes in different days (see $\sigma_{\zeta_{slp}}$ in table 3.1 and figure 3.4). Notably, the change in drift towards steep slopes as the day progresses was quite variable in different days (see $\sigma_{\alpha_{slp}}$ in table 3.1). Drift towards nests was also variable (see $\sigma_{\zeta_{nest}}$ and $\sigma\zeta_{ngh}$ terms in table 3.1). However, the effect of time of day on drift towards nests varied less in different days, than the drift towards steep slopes (see $\sigma_{\alpha_{nest}}$ and $\sigma_{\alpha_{ngh}}$ terms in table 3.1).

The Cederberg eagle, showed similar patterns to those observed for the Sandveld eagle in terms of movement drift (figure 3.4 and table 3.1). However, all coefficients tended to be more variable, particularly at the beginning and end of the day. This was especially true for attraction towards steep slopes, which also tended to be larger on average for the Cederberg eagle than for the Sandveld eagle (see different $\mu_{\zeta_{slp}}$, and particularly $\mu_{\alpha_{slp}}$, between Sandveld and Cederberg eagle in table 3.1).

The diffusion parameter for the Sandveld eagle was larger than that for the Cederberg eagle (see ν in table 3.1 and figure 3.5). Diffusion in the Sandveld eagle was also more concentrated around midday, than it was for the Cederberg eagle (see differences in μ_{ψ} and σ_{ψ} between Sandveld and Cederberg eagle in table 3.1, and figure 3.5).

Simulations of Verreaux's Eagle flights observed from vantage points incorporated all the characteristics exposed above. The objective then, is to investigate whether a model fitted to the simulated flights can recover the parameters used for the simulations.

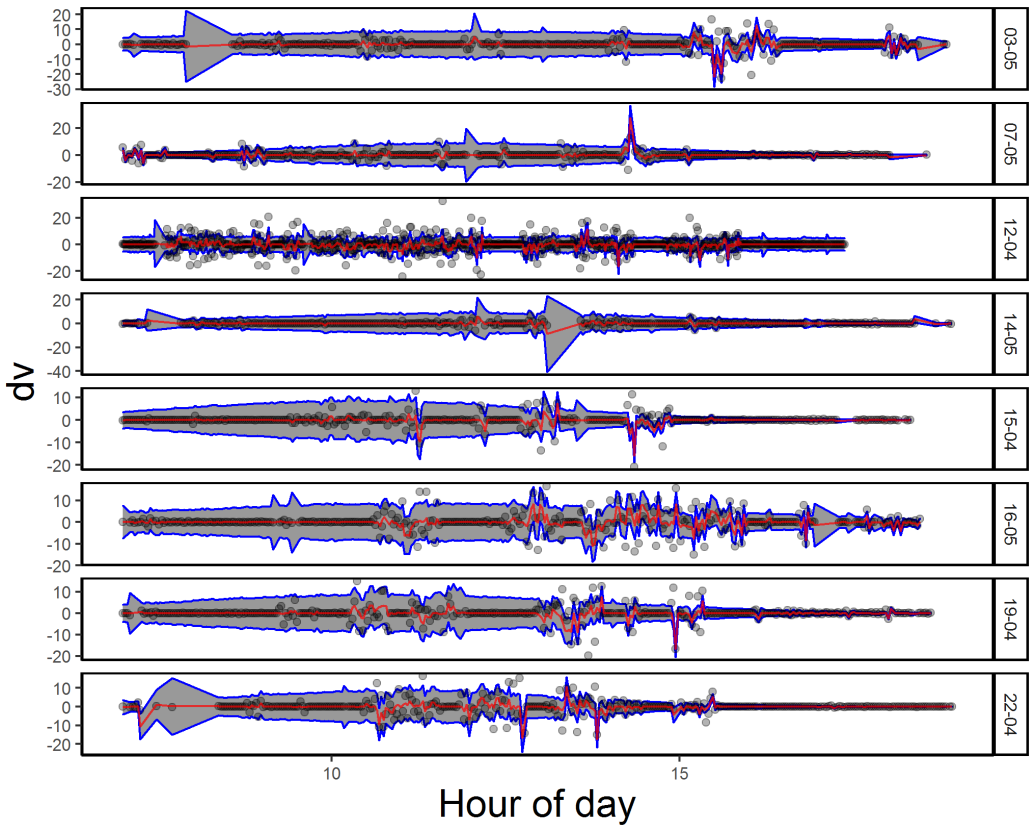


Figure 3.3: 90% posterior predictive band (grey band enclosed by blue line) computed for changes in velocity (dv) in eight random tracking days from the model fitted to the Sandveld eagle. The actual measurements are represented by dots and the posterior median is represented by the red line. The model captures the general increase of activity towards midday, although it fails to capture the periods of constant velocity (bird possibly inactive) interspersed through the day. Note that prediction's uncertainty grows before longer time intervals with no observations. This is because we are predicting change in velocity and not velocity itself and I decided to associate the change with the beginning of the interval. Predictions of velocity would be uncertain after the long time interval, as expected.

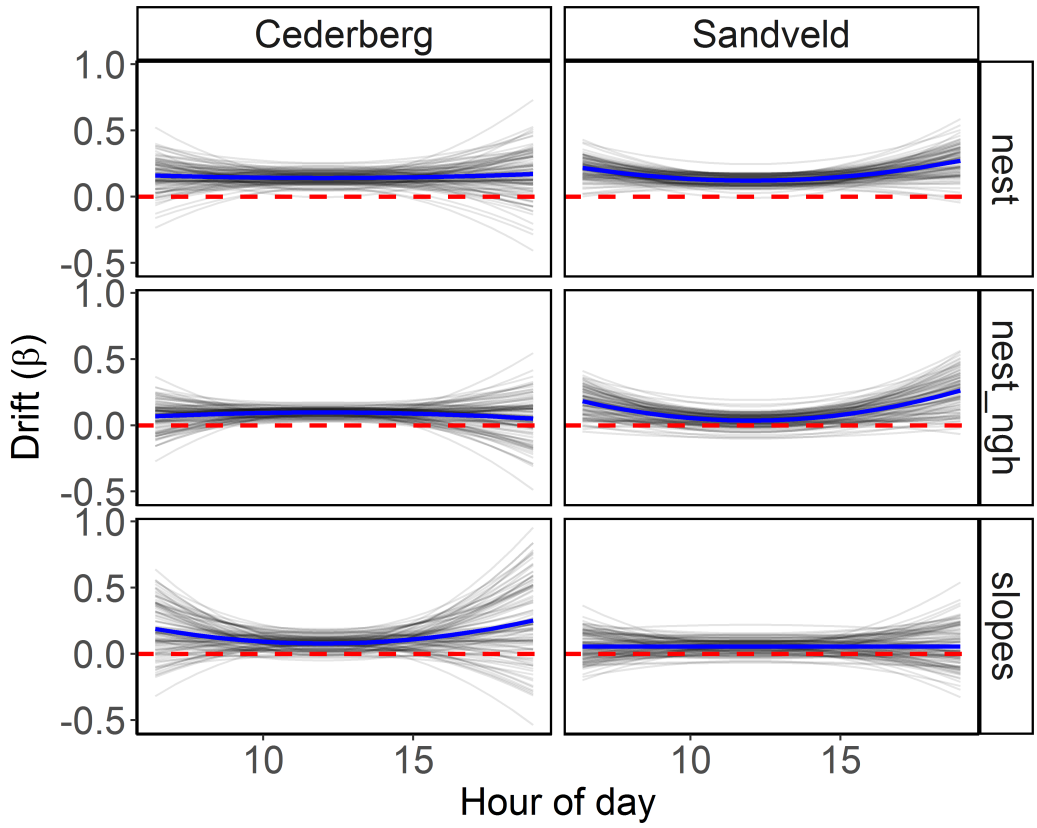


Figure 3.4: Posterior simulations of drift parameters towards different target locations. Under the potential-based velocity model fitted to the GPS locations of the Sandveld and Cederber eagles, the drift towards steep slopes (slopes), the eagles' own nest (nest) and the neighbours' nest (nest_ngh) change quadratically with the time difference with noon. Movement drift also changes in different days. In the plot, each line corresponds to one realization of the drift for a random day based on the posterior distribution of the drift parameters (100 realizations are shown), and it shows how drift changes with time of day. Positive drift parameters translate into systematic drift towards the corresponding target locations (slopes and nests); whereas negative drift translates into drifting away from target locations. Zero drift is represented by the red, dashed line and the average drift is shown by the blue line. We observe more variable drift in the morning and evenings, and more consistent towards nests than towards slopes. All drift parameters are more variable for the Cederberg eagle than for the Sandveld eagle. The drift towards slopes of the Cederberg eagle is remarkably more variable, in early and late hours, than that of the Sandveld eagle.

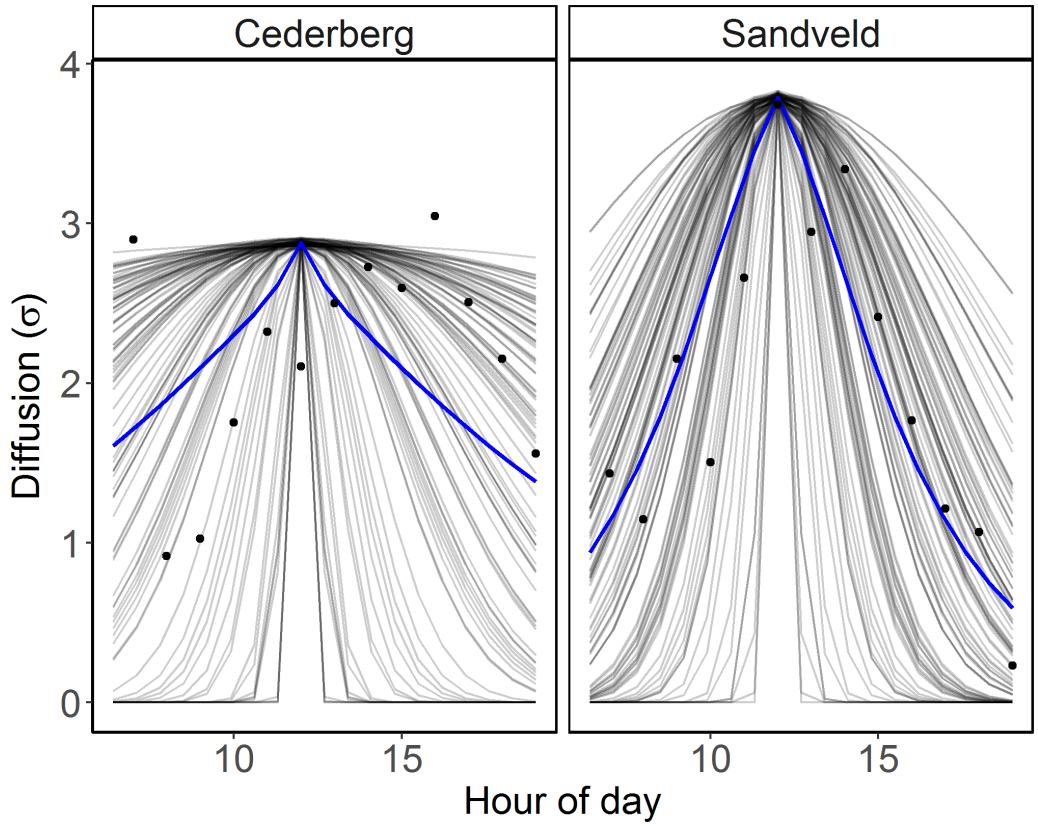


Figure 3.5: Posterior simulations of diffusion parameters. Under the potential-based velocity model fitted to the GPS locations of the Sandveld and Cederber eagles, diffusion is maximum at midday and decreases following a Gaussian curve towards early and late hours of the day. In the plot, each line corresponds to one realization of the diffusion for a random day based on the posterior distribution of the diffusion parameters (100 realizations are shown), and it shows how diffusion changes with time of day. The average diffusion is shown by the blue line. Under this model diffusion is the main driver of movement, and therefore more diffusion translates into more activity. The Sandveld eagle shows an activity pattern that is more concentrated towards midday than the Cederberg eagle. The black points represent the standard deviation of the residuals computed by subtracting the median posterior model prediction at each hour of the day from the actual observations. While the Sandveld residuals follow the estimated diffusion pattern, the residuals computed for the Cederberg eagle suggest that the actual diffusion is shifted towards later hours than the model estimates.

Table 3.1: Summary of parameter posterior distributions of the potential-based velocity model fitted to the GPS locations of the Sandveld and Cederberg eagles. These posteriors were also used for simulating vantage point observations (Sandveld eagle) and as prior distributions (Cederberg eagle). Under this model, eagle movements may drift towards (or away from) steep slopes, their own nest and a neighbour's nest. Drift changes with the time difference with noon. Thus, ζ parameters represent drift at midday; whereas α, modulates the change in drift with time of day. The model also incorporates a diffusion term that also changes with time of day. The parameter ν controls diffusion at midday; whereas ψ changes the diffusion parameter with time of day. All parameters may change in different days, and are given Gaussian distributions with mean μ and standard deviation σ (e.g. $\mu_{\zeta_{slp}}$ stands for mean across days for the drift towards slopes at noon). The column *area* shows which eagle the parameters come from: S - Sandveld and C - Cederberg. The columns with percentages represent credible areas. *n_eff* stands for number of effective Markov Chain Monte Carlo samples used to define the posteriors and *Rhat*, provides a measure of convergence (Rhat = 1 is good convergence).

	mean	area	sd	2.5%	50%	97.5%	n_eff	Rhat
$\mu_{\zeta_{slp}}$	0.05	S	0.01	0.02	0.05	0.07	2204.00	1.00
	0.07	C	0.01	0.05	0.07	0.09	3552.00	1.00
$\mu_{\zeta_{ngh}}$	0.04	S	0.01	0.02	0.04	0.06	1855.00	1.00
	0.10	C	0.01	0.08	0.10	0.11	3974.00	1.00
$\mu_{\zeta_{nest}}$	0.12	S	0.01	0.11	0.12	0.14	3434.00	1.00
	0.14	C	0.01	0.12	0.14	0.16	2644.00	1.00
$\mu_{\alpha_{slp}}$	0.01	S	0.04	-0.07	0.01	0.09	2611.00	1.00
	0.19	C	0.06	0.08	0.19	0.31	1270.00	1.00
$\mu_{\alpha_{ngh}}$	0.22	S	0.03	0.16	0.22	0.27	2413.00	1.00
	-0.06	C	0.03	-0.12	-0.06	0.01	1871.00	1.00
$\mu_{\alpha_{nest}}$	0.15	S	0.02	0.11	0.15	0.20	2809.00	1.00
	0.04	C	0.03	-0.02	0.04	0.10	2187.00	1.00
$\sigma_{\zeta_{slp}}$	0.05	S	0.01	0.03	0.05	0.07	1671.00	1.00
	0.05	C	0.01	0.03	0.05	0.07	2047.00	1.00
$\sigma_{\zeta_{ngh}}$	0.05	S	0.01	0.03	0.05	0.07	1578.00	1.00
	0.03	C	0.01	0.02	0.03	0.05	1586.00	1.00
$\sigma_{\zeta_{nest}}$	0.03	S	0.01	0.02	0.03	0.05	1716.00	1.00
	0.05	C	0.01	0.03	0.05	0.06	2285.00	1.00
$\sigma_{\alpha_{slp}}$	0.16	S	0.03	0.10	0.16	0.22	1855.00	1.00
	0.32	C	0.05	0.24	0.32	0.43	2112.00	1.00
$\sigma_{\alpha_{ngh}}$	0.12	S	0.02	0.09	0.12	0.17	2078.00	1.00
	0.17	C	0.02	0.12	0.16	0.22	2135.00	1.00
$\sigma_{\alpha_{nest}}$	0.11	S	0.02	0.07	0.11	0.15	1677.00	1.00
	0.16	C	0.02	0.11	0.16	0.21	2541.00	1.00
μ_{ψ}	4.55	S	0.41	3.80	4.55	5.36	8030.00	1.00
	6.71	C	0.97	4.85	6.72	8.53	6319.00	1.00
σ_{ψ}	2.74	S	0.39	2.09	2.70	3.58	4820.00	1.00
	11.88	C	2.37	8.04	11.61	17.12	3057.00	1.00
ν	3.79	S	0.02	3.74	3.79	3.84	3880.00	1.00
	2.88	C	0.02	2.85	2.88	2.91	4308.00	1.00

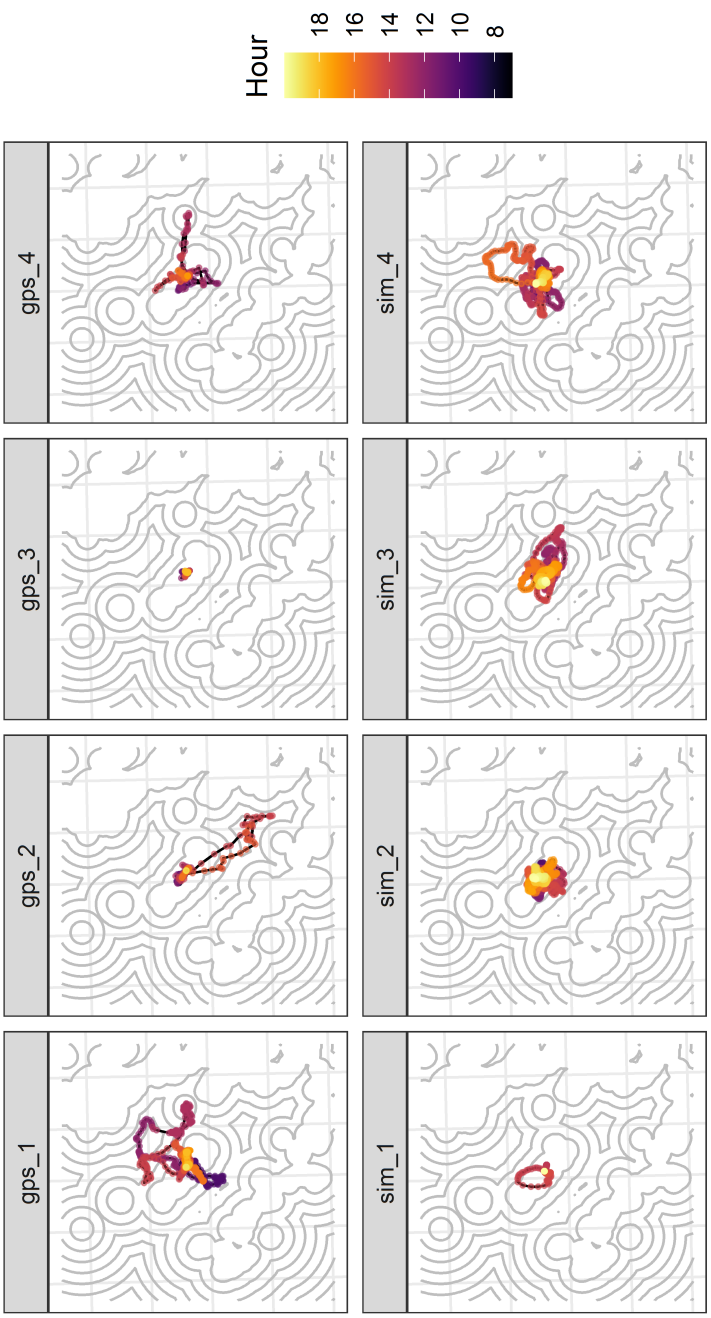


Figure 3.6: Four randomly selected days of Verreaux's Eagle activity (recorded by GPS), versus four flights simulated using the model fitted to the Verreaux's Eagle GPS data. Numerical indices do not represent any particular relationship between observed and simulated flights. We see that the model captures some features of the data such as central place foraging. However, observed flights seem to be more discontinuous than those represented by simulations, suggesting that eagles display a resting behaviour that was not incorporated into the model.

3.3.2 Simulation results

Flights observed from vantage points were simulated using the model described in section 3.2.3, the procedure outlined in section 3.2.4 and the parameter values for the Sandveld eagle presented in section 3.3.1. I simulated accumulated flight durations of 3, 5, 10, 15 and 30 minutes per vantage point, in scenarios with 2, 4, 6, 8 and 10 different vantage point locations, which amounts to a total of 25 settings. Each of these simulations was repeated five times.

To recover movement parameters from the simulated data, I fitted a model similar to the one fitted to the eagles' GPS locations - also used to generate the simulations. However, I excluded from the model variables related to distance from nests, because often observers are not able to identify individual eagles during vantage point surveys. Thus, drift towards slopes became the main parameter of interest. To investigate the effect of different types of priors on the posterior distributions fitted to the vantage point data, I used three different types priors: i) vague; ii) the data-generating distributions (posterior distributions fitted to *eagle1*); and iii) posterior distribution of a different eagle than the one used to simulate the flights (posterior distributions fitted to *eagle2*). I compared the distribution of the parameters used for the simulations with the posterior distribution of the parameters fitted to the simulated data using the relative bias (RB) and the overlapping coefficient (OVL) (see section 5.2.4 and 3.7). Relative Bias and overlapping coefficients were calculated for the posterior distributions of all parameters fitted using all combinations of number of vantage points, flight time per vantage point and type of prior. The results were summarized using GAMs with RB and OVL as response variables and number of vantage points, flight time and type of prior as explanatory variables.

In the model used for simulating Verreaux's eagles flights observed from vantage points, drift parameters changed with time of day and also between days (see section 3.2.3 and 3.3.1). Each day has a maximum or minimum drift towards midday (midday drift, given by the parameter $\mu_{\zeta_{slp}}$ in table 3.1). Drift was, in general, over-estimated in the model fitted to the simulated vantage point observations, with bias of up to twice the true (data-generating) value on average, when vague priors were used (figure 3.8). Based on the GAM fitted to the RB values, posterior distribution bias in drift parameters tended to decrease as number of vantage points and observed flight time per vantage point increased. This effect was most evident when informative priors were used to fit the models (see top central panel of figure 3.8). When vague priors were used for model fitting, bias in the posterior of drift parameters was mostly unaffected by increasing number of

vantage points or fight time per vantage point, over the range of scenarios tested (see top left panel of figure 3.8). When using the data that generated the flights as priors, posterior distributions became more biased as sample size increases, particularly when increasing flight time per vantage point (see top right panel of figure 3.8). Note that, as I mentioned at the beginning of this section, models were purposefully misspecified, by ignoring some important attractors (i.e. nests).

Relative bias in most posterior drift parameters presented a similar behaviour (see appendix E), with one exception: the parameter that modulates the change in drift with time of day (see $\mu_{\alpha_{slp}}$ with *eagle2* priors in figure 3.8). According to the GAM, the average bias obtained for this parameter was up to 20 times the mean of the data-generating distribution. Increasing the number of vantage points or flight time per vantage point had little effect on this bias, over the range of tested scenarios. Using vague priors, relative bias was smaller but still around five times the mean of the data-generating distribution (figure 3.8).

The model that simulates eagle's movement also incorporates a diffusion term, which is also the main driver of movement (it tell us *how much movement* there is). Diffusion also changes with time of day, being maximum at midday (midday diffusion, given by the parameter ν in table 3.1). According to the GAM fitted to RB values, bias in the posterior distribution of diffusion parameters was less than 100% on average (figure 3.8). Midday diffusion tended to be under-estimated, while the change in diffusion with time of day tended to be over-estimated. Similar to what I observed for the drift parameters, change in diffusion with time of day was more biased than midday diffusion. Also similar was the effect of increasing number of vantage points and observed flight time. The reduction in bias in diffusion when increasing sampling efforts is substantial only when using informative priors for model fitting (figure 3.8).

Posterior distributions of all parameters fitted to the simulated vantage point data showed very small overlapping coefficients with the data-generating distributions (see figure 3.9 and appendix F). Overlap only improved substantially using the data-generating distributions as priors (see figure 3.9). Increasing the number of vantage points, and especially, observed flight time per vantage point, produced little increase in overlap, which stayed low (< 0.5) even for the scenarios with the most abundant data (see appendix F). Overlap between data-generating distributions and posterior estimates were close to one when using data-generating distributions as priors (figure 3.9).

A graphical representation of the relative bias and overlapping coefficients between all

estimated parameters and data-generating distributions, as well as the effect of increasing number of vantage points and flight time per vantage point, is presented in appendices E and F.

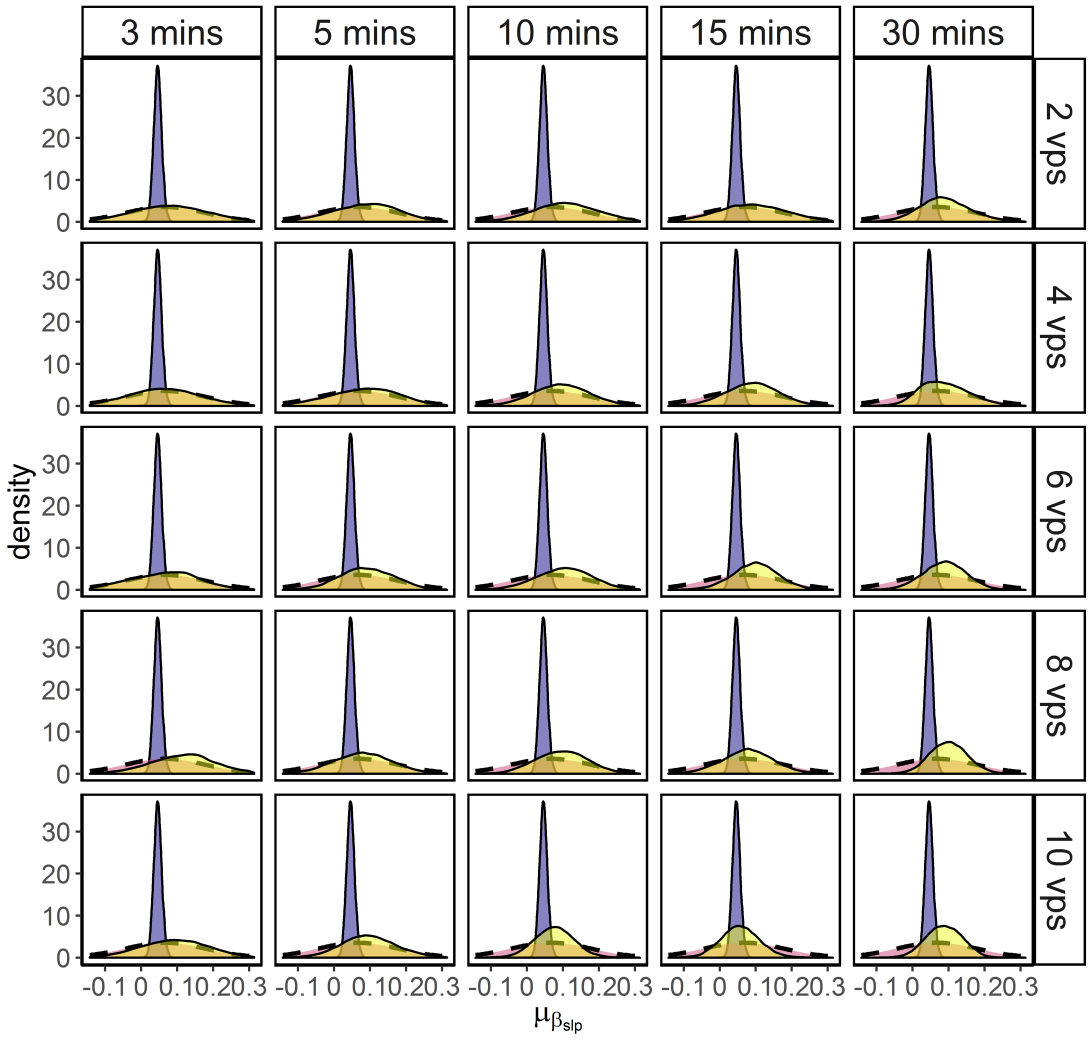


Figure 3.7: Graphical representation of the comparison between data-generating distribution (in blue), the prior (in red with dashed line) and the posterior (yellow with solid line). Different simulated vantage point scenarios are presented, with different number of vantage points (rows) and different observed flight time per vantage point (columns). This figure shows the posterior of the mean drift at midday towards steep slopes across days ($\mu_{\zeta_{slp}}$). The posterior agrees with the prior in scenarios with little data (top-right panels) and becomes more similar to the data-generating distribution in scenarios with more data (bottom-left panels).

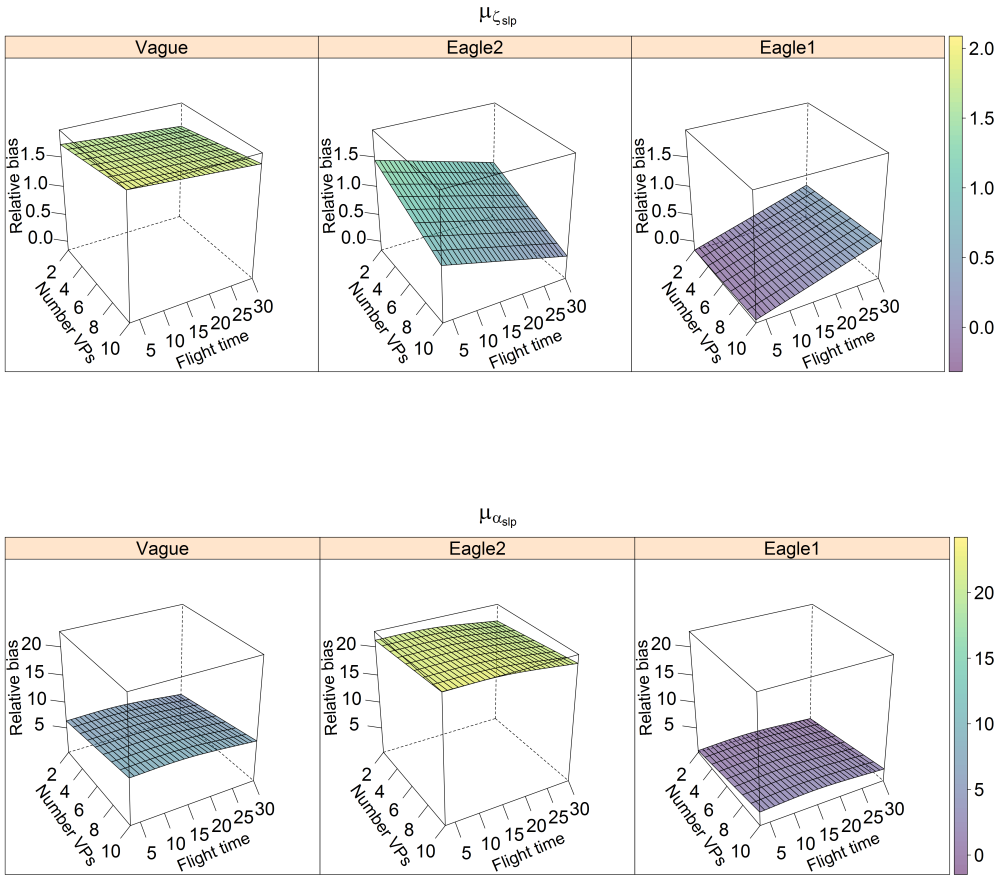


Figure 3.8: Mean relative bias (in the vertical axis and colour gradient) of the posterior distribution of the parameters fitted to the simulated vantage point observations and effect of number of vantage points and flight duration per vantage point, as predicted by the GAM. A relative bias = 0 means that the posterior is unbiased. A relative bias = 1 means that bias is as large as the parameter value that generated the data. A relative bias = 2, the bias is twice as large as the data-generating parameter, etc. Posterior relative bias is compared for different priors: vague (left), informative (centre - these priors come from parameters of a different eagle than the one use to generate the data), and data-generating distributions (right column - these are the distributions that were used to generate the data). The surface shows the effect of increasing the number of vantage points and the observed flight time per vantage point, on the relative bias. $\mu_{\zeta_{slp}}$ is the mean drift towards slopes at noon across different days. $\mu_{\alpha_{slp}}$ is the mean change in drift with time of day, across different days.

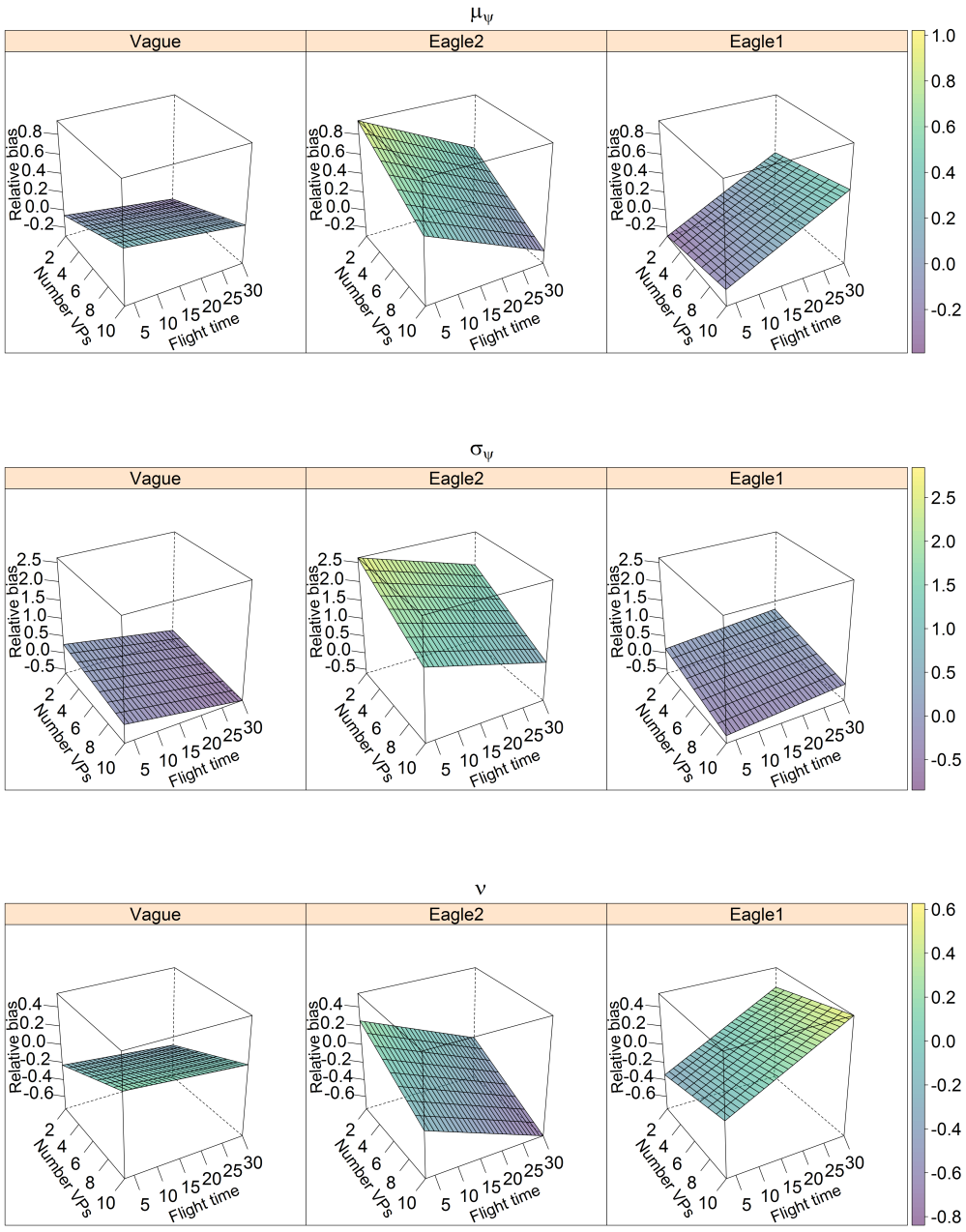


Figure 3.8: (Continued) μ_ψ corresponds to the mean change in diffusion with time of day across different days, whereas σ_ψ is the standard deviation of the change in diffusion with time of day across different days. ν corresponds to the maximum diffusion any day.

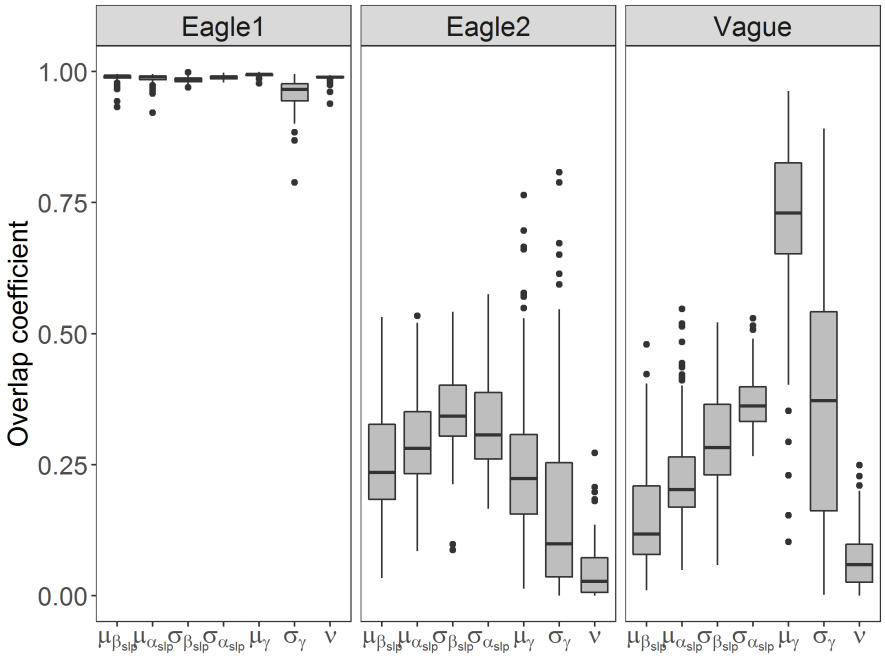


Figure 3.9: Overlapping coefficients between the posterior of the movement parameter estimated from the simulated vantage point observations and in relation to the data-generating parameters. The boxes represent the lower quantile, median and upper quantile, while the whiskers cover one and a half times the interquantile range. Eagle1 priors correspond to the data-generating distributions, Eagle2 priors are constructed from the movement parameters of a different individual of the same species with some added variance and vague priors are priors centred at zero with large variance.

3.4 Discussion

Vantage point surveys are standard practice in environmental impact assessments of wind energy facilities on raptors and other soaring birds (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). However, their reliability at characterizing bird activity is largely untested. Here, I investigate how the analysis of flights observed from vantage points using movement modelling techniques, may help us understand bird behaviour. Assuming that bird activity could be a good indicator of exposure to collision with wind turbines, understanding the drivers of movement should provide useful insights into where collisions are likely to occur. The question then becomes how accurately can we estimate movement parameters using vantage point observations? To answer this question, I simulated Verreaux's Eagle's flights as if they were observed from vantage points. To add realism to the simulations, they were based on the parameters obtained by fitting a similar model to a GPS data set of the same species. Then, I analyse how accurately data-generating parameters can be recovered from the simulated data. In addition, I study how different survey designs and priors used to fit these models to the synthetic data, influence the accuracy of the estimated parameters.

Potential-based velocity models are appealing descriptors of movement patterns, because they connect flight characteristics of individual birds with long-term spatial distribution of bird activity (Brillinger *et al.*, 2002; Preisler *et al.*, 2013; Russell *et al.*, 2018). Movement strategies of Verreaux's Eagles (or any other soaring bird) likely involve long-term decisions, such as staying within their territory or moving towards hunting or resting grounds (Simmons, 2005; Murgatroyd *et al.*, 2016b). These tendencies form the structural part of potential-based velocity models (drift components). On the other hand, more sporadic and erratic movements in response to unpredictable stimuli, such as local wind currents, interactions with other individuals or disturbances, are pooled in the random component of these models (diffusion component). Although potential-based models are build upon local behaviour (e.g. instantaneous changes in velocity) the potential surface is constructed at the landscape level, allowing flexible estimation of far-ranging long-term behaviours effects and not just local decisions (e.g. Russell *et al.*, 2018). However, constructing potential surfaces that are too flexible (e.g. using thin-plate splines) makes it difficult to understand what elements of the landscape are driving the behaviour of the animals. In this chapter, I sacrifice flexibility by explicitly defining potential points of attraction based on landscape elements, such as steep slopes or nests. This approach makes it easier to understand the drivers of movement but it is more restrictive

and subjective in the sense that it is the analyst that defines the points of attraction *a priori*. Thus, points of attraction produced by factors other than those controlled might be completely missed. For example, complex territorial interactions or preferred hunting grounds could be easily missed by the researcher, while they might play a central role in the animal's activity.

As I mentioned, I fitted potential-based velocity models taking an approach that interrogates the data about whether selected elements of the landscape could play a role in the Verreaux's Eagle activity (i.e. steep slopes, own nest and neighbouring nest). Although model predictions were constrained within sensible limits, the model failed to capture periods with no change in velocity, which probably represent periods when the eagles were not active. This suggests that a model with at least an additional "resting" state would be more appropriate to represent eagle behaviour. However, the objective of the models fitted to the eagle data was to provide an approximation to flights an observer might record during vantage point surveys, and not a detailed representation of eagle behaviour. In addition, a vantage point observer would not record perched eagles, and therefore, I favour the simplicity of a single-state model. It must be noted that this decision could introduce bias in the estimation of both drift and diffusion and these problems should be addressed if the primary objective was to infer eagle behaviour.

The movement of two Verreaux's Eagles captured by GPS trackers, was described by a positive drift of flight trajectories towards their nest, neighbouring nests and steep slopes. The drift towards nests was greater and more consistent than the drift towards steep slopes. This suggests a territorial behaviour and a variable selection of steep slopes by this species, possibly for hunting and slope soaring (Simmons, 2005; Murgatroyd *et al.*, 2016b, 2018). A diffusion term based on a Brownian (random) motion was added to the model to account for sporadic, unpredictable stimuli. Parameters associated with both systematic drift and diffusion varied between different days and also changed with time of day. Diffusion, which is the main driver of movement in the proposed model, was maximum towards midday and decreased towards the morning and evening. These results are consistent with what is known about the species. The Verrreaux's Eagle is a diurnal raptor and its activity is known to peak at midday (Simmons, 2005; Murgatroyd *et al.*, 2016b). Drift towards steep slopes and nests was greater, although more variable during mornings and evenings than it was during midday. Since the species is a central place forager, that stays around the nest year round (Simmons, 2005), movement in early and late hours of the day are likely to occur near the nest. These morning and evening flights are likely to be shorter and more erratic, than the longer trips towards midday.

The location and intensity of air currents produced by wind speed and temperature, together with energy management strategies by the birds, are also likely to contribute to these different movement modes (see Shepard *et al.*, 2013; Murgatroyd *et al.*, 2018). For example, thermal uplift forms when land is heated by the sun, and therefore it is more abundant towards midday. These conditions may favour longer trips (e.g. Duerr *et al.*, 2015), and translate into greater diffusion parameters, at midday. Therefore, in the presence of thermal uplift, birds might be discouraged from approaching slopes (e.g. during midday), translating into lower drift parameters towards these features.

The two eagles showed similar behaviours, despite them living in different environments; however, there were some differences. Attraction towards steep slopes was greater for the Cederberg eagle, probably due to a much larger availability of these features in this mountainous region, and prey being more concentrated around the abundant rocky outcrops. The activity of this eagle was more evenly distributed throughout the day than that of the Sandveld individual. The smoother relief in the Sandveld region, probably lacks orographic uplift, making eagles favour thermal soaring towards midday. It is during midday that the drift parameters of the two eagles are most similar. This suggest that the Cederberg eagle may start its activity earlier in the day, using slope currents, while the Sandveld individual needs to wait for thermal convection to perform excursions. Towards midday, both eagles probably tend to favour thermal over slope uplift.

These movement patterns were inferred from the models fitted to the GPS data of the two Verreaux's Eagles, but must be taken with caution due to the aforementioned model misspecification. However, these movement patters provide a useful structure for simulating eagle movement behaviour and test the inferential properties of vantage point data. To this end, I used simulations generated using the posterior distributions of the parameters fitted to the eagles' GPS data, and that therefore, recreate eagle behaviour. From a vantage point, an observer can only collect samples of flights restricted to the vantage point viewshed. It is unclear whether these samples contain information about long-term movement strategies by the birds, or the scale at which the samples are collected is too small in relation to the scale at which movement patterns emerge. In addition, it is not always possible to identify all factors that affect movement behaviour (e.g. individual identification or location of nests and roosting sites). Although I knew the location of the nests of the eagles tracked by the GPS transmitters, and the simulated flights incorporated a drift towards nests, the model fitted to these simulations to recover the generating parameters considered only steep slopes and not nesting sites. Having variables included in the generating model that are not estimated from the simulations

adds realism to the exercise by introducing some structure in the data that is unaccounted for, and that may affect model estimates.

Based on the present study, the agreement between movement parameters estimated from simulated vantage point data and the data-generating parameters depend largely on the type of prior used for the modelling. Vague priors are typically centred at zero and have large variance. These priors produced biased posteriors with little overlap with the data-generating parameters. Parameters related to unstructured movement (diffusion) were estimated more accurately than those capturing systematic behaviours (drift). Increasing sample size, by either increasing the number of vantage points or the flight time per vantage point, did not improve the agreement between posteriors and data-generating distributions, when using vague priors. These results suggest that vantage point data contain little information pertaining systematic movement behaviour.

Using informative priors obtained from movement parameters of a different eagle, was key for reducing bias in parameters estimated from simulated vantage point data. These informative priors were not enough to obtain unbiased estimates but they facilitated bias reduction in scenarios with good spatial coverage and sample size (accumulated observed flight time). Furthermore, using informative priors also helped with inference by constraining parameter estimates to stay within reasonable limits in scenarios with poor spatial coverage and small sample size. Increasing number of vantage points and increasing observed flight time per vantage point had a similar effect on estimated parameter bias. This similarity might be explained by the fact that the Sandveld region, used as a landscape for the simulations is homogeneous in terms of habitat and topography. In this situation, the spatial coverage of the area (number of vantage points) might have relatively little importance due to all vantage points covering a similar type of habitat. In more heterogeneous habitats, I would expect spatial coverage to be more important than shown in this study. Parameters that capture complex behaviours, such as change in drift and diffusion with time of day, were most biased, even with good spatial coverage and large sample size. This suggests that, when working with vantage point data, models should be kept simple, at least within the context of potential-based velocity models.

Two eagles of the same species are likely to share similar behavioural characteristics, and that is why informative priors based on the parameters of an eagle different than the one used for the simulations, improved parameter estimates. However, we must acknowledge potential differences amongst individuals, and incorporate this uncertainty into the modelling process, by for example, increasing the variance of the priors. In general it is recommended to choose priors based on existing literature or personal believes

(e.g. Ellison, 2004; Gelman *et al.*, 2014; McElreath, 2019). When data are informative, the choice of prior is mostly irrelevant. However, in the case of modelling movement behaviour using flights observed from vantage points, given the low information content of these data, I recommend choosing priors very carefully, and if possible, obtained from models similar to the one being fitted.

I must also acknowledge that, although simulations were based on a single bird, in real vantage point surveys, observers collect data from multiple individuals. There might be several eagles around a wind farm, and because individual identification is usually impossible, all individuals are pooled together as one species. In this situation, we estimate movement parameters for a population of eagles, rather than for a single individual. As shown in this study, I would expect all individuals of the same species to share some similar general movement behaviours. Therefore, I anticipate the result of this study to be applicable to a situation in which multiple individuals are being studied from the vantage points. However, I would expect parameter distributions to have larger variance. To what extent it is possible to pool data from multiple species with similar characteristics to make more powerful inference remains an open question. Similar experiences with this and other species with different habitat preferences would be valuable to understand to what degree the results presented here can be generalized to other situations.

Potential-based velocity models could be naturally extended to incorporate height as a third dimension if the final objective of the analysis is to predict exposure to collision with wind turbines. Nevertheless, capturing three-dimensional data in the field becomes a challenge. Observers can only capture information continuously in two dimensions (i.e. by plotting trajectories on a map). Capturing information on a third dimension would entail other means of recording (perhaps with a voice recorder). Perhaps practitioners would benefit from discretizing these measurements in time. In other words, capture annotated sequences of bird locations with some frequency (e.g. every 30 seconds) - similar to the way GPS trackers send information, instead of drawing continuous trajectories. If locations are captured with high enough frequency, this type of data could be used to fit potential-based velocity models.

It is also worth mentioning that vantage point observers will likely introduce positional error in the recorded flight trajectories. I acknowledge that these errors might introduce extra uncertainty in the estimated parameters, making it more difficult to infer clear movement patterns. I have ignored these errors in this chapter, although I deal with them extensively in chapter 4 and I incorporate them into the modelling process in chapter 5.

Once a movement model is specified, we may study the relative frequency with which different parts of the landscape are used by the birds in the long term by simulation. This long-term distribution should reveal which locations are best to place wind turbines and which should be avoided. In chapter 5, I show how this is done in practice and how useful activity distribution estimated using potential-based velocity models is for predicting exposure to collision.

3.5 Conclusion

In this study, I use potential-based velocity models to learn about the movement behaviour of the Verreaux's Eagle, as a way to predict exposure to collision with wind turbines. For the first time, I investigate how accurately model parameters that describe the movement of the eagles can be estimated from vantage point observations. The results suggest that potential-based velocity models are useful to estimate systematic selection of landscape features. This approach represents an improvement over methods that base collision exposure on estimated flight abundance in that collision exposure based on movement behaviour is spatially explicit. Therefore, we may evaluate collision risk at each individual turbine independently. However, fitting these models to vantage point observations produced reliable inference only when abundant data was available. In this study, about 6 vantage points and 15 minutes of flight time observed per vantage point, provided estimates with bias that were, on average, smaller than the parameter value used to simulate the data; and only when informative priors were used in the model. What constitutes a large enough sample or how much spatial coverage is needed, should be assessed in each particular case, depending on the complexity of the landscape and the behaviour of the species. The use of carefully selected, informative priors proved to be fundamental to obtain accurate parameter estimates and capitalize on increasing sampling efforts. Although only potential-based models were tested, I would expect similar results when using discrete-time analogues like step-selection functions, as suggested by Sur *et al.* (2018). In conclusion, if movement models are to be used to study the spatial distribution of bird activity using vantage point observations, it is desirable to have at least one or two hours of observed flights, covering of a variety of landscape configurations, to obtain reliable estimates of movement parameters. Furthermore, I found that the use of Bayesian priors is critical to facilitate the estimation process, and make a coherent use of existing information and collected data.

Chapter 4

Error in the mapping of flights observed from vantage points

4.1 Introduction

In an effort to mitigate global warming, energy production is changing rapidly worldwide. Renewable energies are considered key elements to control carbon emissions to the atmosphere, and they represent a larger proportion of the energy mix than ever before (REN21, 2018). Wind energy stands out as one of the more efficient and prevalent of these technologies. However, wind energy production presents its own environmental challenges. Generally installed away from populated areas, wind energy facilities produce environmental impacts on relatively untransformed habitats and on species sensitive to human disturbance (Gasparatos *et al.*, 2017). Birds and bats collide with wind turbines and ancillary infrastructure, and tend to be displaced from wind farms due to habitat degradation (Schuster *et al.*, 2015). Collision-prone raptors and soaring birds are particularly sensitive to these impacts due to their relatively low population densities and low reproduction rates (Marques *et al.*, 2014; Watson *et al.*, 2018). It is apparent that collision risk is not only project- and species-specific, but it is also concentrated at particular turbines (Marques *et al.*, 2014). Therefore, to improve our understanding of bird collisions and also to minimize this risk to biodiversity, it is necessary to conduct collision

risk studies at the local scale (i.e. within a wind farm).

Bird collision risk studies use samples of activity captured from vantage points to investigate exposure to collision (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). Collision risk models may then be used to translate exposure into predicted number of fatalities (Masden & Cook, 2016). Vantage point surveys occupy a prominent role in predicting potential impacts of wind energy on bird populations and defining effective mitigation strategies. Therefore, it is crucial that we understand the limitations of this field technique and their potential consequences during data analysis.

From a vantage point an observer records activity of soaring birds over some portion of the landscape. To be effective at detecting birds that occur at low densities, but that are visible from certain distance, the areas sampled from vantage points (viewshed) are relatively large, often enclosing heterogeneous habitats (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014). To model and extrapolate bird activity to areas not covered during surveys, it is convenient to use environmental variables as predictors (e.g. vegetation, topography, weather, etc.). Thus, during vantage point surveys it is recommended to plot the observed flight trajectories on a map to keep track of birds movements, as well as the habitats and landscape features they visit. The correct association of an observed flight with a particular landscape feature or habitat, will depend on how accurately the flight was recorded. Although it is acknowledged that flights could be inaccurately mapped (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014), studies on the accuracy with which observers are able to perceive and capture the location and shape of birds flights are lacking.

Field surveys are affected by visual perception, which is a classic research field in the psychological sciences (Cutting & Vishton, 1995). It is clear that visual perception of space deviates from Euclidean geometry (Wagner, 1985; Cutting & Vishton, 1995; Todd *et al.*, 2001; Erkelens, 2015). The proportionality of egocentric distance or depth (i.e. distance in a radial axis from the observer) is perceived differently from that in lateral distance or frontal depth (i.e. perpendicular to the line of sight), yet a precise geometry of the visual space is still elusive (Wagner, 1985; Todd *et al.*, 2001; Proffitt, 2006; Erkelens, 2015). Yang & Purves (2003) described visual perception of distance from a probabilistic, Bayesian perspective, in which prior assumptions about a"typical" or idealized structure of the environment, would bias the perceived environment, more so in absence of reliable spatial cues. This statistical approach, found support for several of the previously observed traits in visual biases, not only in distance perception, but also in other aspects of visual research (see review by Geisler, 2008).

In this chapter, I investigate some of the basic properties of the error made by observers when plotting trajectories of a flying bird on a map. Henceforth, I will refer to this error as *mapping* error. Typically, it is impossible to know how accurately flight trajectories were captured, since the mapped trajectories are the only information available about the *true* trajectories. To overcome this problem, I compare the flight trajectories of a remotely controlled drone as provided by its on-board GPS, with the same trajectories recorded and plotted on a map by three observers tracking the drone visually. I do not attempt to find a particular formula for converting mapped flight trajectories to the true trajectories, but rather I strive to understand (and eventually predict), in probabilistic terms, what the true trajectory might have been, given a mapped trajectory. With this objective in mind, I: i) estimate the magnitude of the mapping error made by the observers; ii) investigate whether the error is affected by certain variables, such as distance to the flying object or the topography of the landscape; and iii) explore the dynamics under which the mapping error evolves, as the flight progresses.

Understanding the magnitude and dynamics of the mapping error will allow us to make probabilistic statements about where birds flew in reality. Incorporating mapping error into our analysis of flight trajectories will help us make more powerful inference by separating uncertainty due to the observation process from that due to the complexity in bird behaviour.

4.2 Methodology

4.2.1 Field methods

Field trials were conducted between the 20th and the 25th of August 2016, to determine the accuracy of three observers at visually tracking and plotting flight trajectories onto topographic maps. The observers were all experienced in bird monitoring, including the recording of bird flights from vantage points.

Observers were asked to sketch a representation of the flights traced by a drone on a topographic map, an aerial photo or a combination of both. They could choose the map they felt more comfortable with. All three observers recorded every flight. They were kept at a certain distance from each other to avoid them copying each other's mappings. The drone, a DJI Phantom 3 Professional model, had an on-board positioning system

(GPS) that tracked its location continuously, providing a location every second. This allowed me to compare the observers' mappings with the GPS trajectories.

Two areas of different topography were selected for the experiment to investigate the effect visual clues could have on the accuracy of the observers: i) a flat area with few topographical references, and ii) a rugged area where topographical elements could be more easily identified on the maps (figure 4.2). Three different points were defined for the take-off of the drone in each of these two areas (i.e. three take-off points in flat terrain and three in rugged terrain, see figure 4.1). Each take-off point had an associated vantage point, from which all flights were observed. In other words, all flights taking off from the same point were recorded from the same associated vantage point.

The drone performed flights of an approximate duration of two minutes. Three types of flights of different sinuosity were defined: *straight* – a straight flight with one turn of maximum 90 degrees; *sinuous* – flight with at least three turns in different directions and *circular* – a flight tracing a 360 degrees arc of approximately constant radius. Two repetitions of these flights were performed from each of the different take-off points. In one of the repetitions the flight was kept at a low altitude, of approximately 20 meters above the take-off point. In the other repetition, the flight was higher, at an altitude of approximately 70 meters above the take-off point (figure 4.1). Note that the height above the ground of the drone is different from the height above take-off point, especially when flying over rugged terrain.

Flights were conducted either from the take-off point and away from the observers or from the distance and towards the observers. First, the drone would take off from a point close to the vantage point and perform a flight moving away from the observers. Once the flight was over, the drone would remain in place, while the observers were allowed to plot the flight they just observed. Once they finished, they would locate the drone again and were allowed to reposition it on the map, if so they considered. Then, the drone would performed another flight that they also plotted onto a map, back to the take-off point and towards the observers. This flight back to the take-off point was considered to be a different flight, independent from the first one.

In total, the drone performed 36 flights. The order in which flights of different heights and shapes were conducted, was randomized. Consequently, whether the drone moved towards and away from the observers in any particular flight was also random. For logistical reasons, six consecutive flights were conducted from an area with a specific topography (flat or rugged) before moving to the other one. In total, each area was

visited three times (2 areas x 3 visits each x 6 flights in each visit = 36 flights in total, see figure 4.1).

The observers were not informed of the type of flight that was being performed, or even of the fact that there existed different flight types.

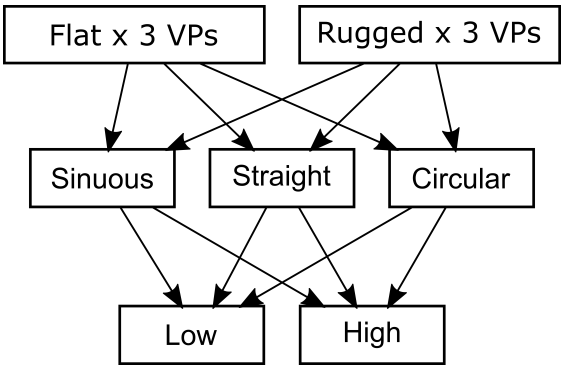


Figure 4.1: Schematic representation of the experimental design used to investigate the mapping error made by three observers visually tracking a flying drone. Three vantage points were set in a flat region and three vantage points in a rugged region. From each of the vantage points, the three observers mapped three types of flights: sinuous, straight and circular. Two flights of each type were mapped: one at a low height and one at a high height. Whether the drone moved towards or away from the observer was randomized. In total the drone performed 36 flights.

4.2.2 Data processing

Flight trajectories captured by field observers and by the drone's GPS were transferred, firstly, to QGIS (QGIS Development Team, 2017), then to R (R Core Team, 2019) using the `sf` package (Pebesma, 2018) to handle spatial objects.

In the field, observers drew continuous lines representing the flight paths they observed. The digitizing process imposes the discretization of these lines into sequences of (arbitrarily many) locations. In a first digitization process, I transferred the scanned flight drawings into vectors of locations. The amount of, and spacing between, these locations depended largely on the length and sinuosity of the flight, as well as on the desired level of detail of the digital trajectory. Once the first vector of locations was in the computer, and to homogenize all trajectories, I transformed the vectors of arbitrarily-spaced loca-

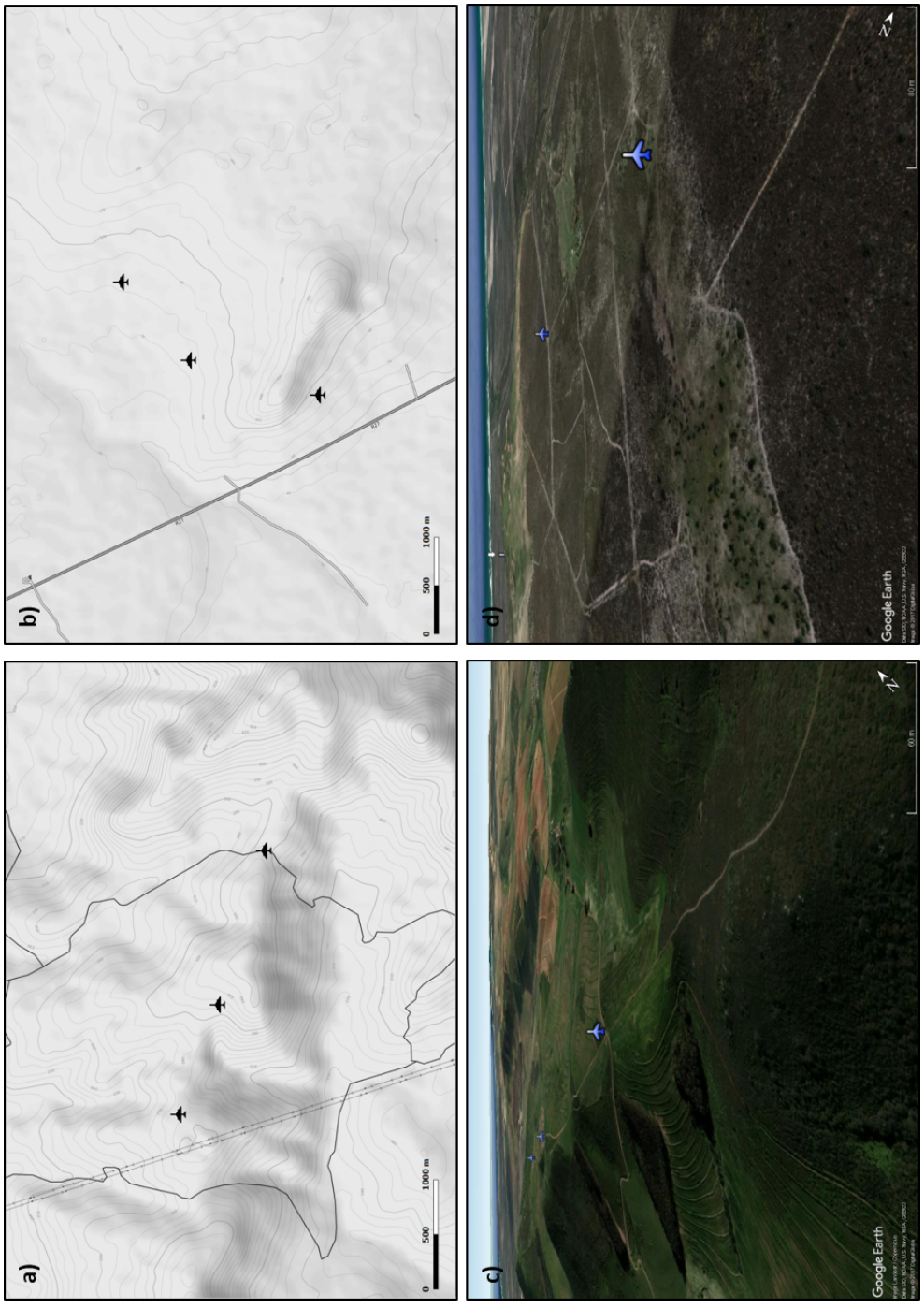


Figure 4.2: The two types of terrain used in the experiment. Figures a) and c) show a topographic and a 3D view of the rugged terrain respectively; Figures b) and d) show similar views for flat terrain. Take-off points are represented by airplane icons.

tions into vectors of equally-spaced locations. I did this by interpolating one equidistant location for every second of flight. Then, I removed the original (arbitrarily-spaced) locations. I used the function *redisltraj* of the R package `adehabitatLT` (Calenge, 2006) to do the rediscretization (interpolation).

The choice of equally spaced intervals was motivated by the fact that observers could only capture time references at the beginning and end of the flights, with the consequent loss of information about changes in speed during the flights. Dividing the trajectory in segments of equal length and assuming constant time intervals, effectively results in assuming that flights were performed at a constant speed. Although efforts were made in the field to perform flights at a constant speed, this condition was not always met, which could introduce some error in the analysis. This setback is common in visual tracking, since observers are unable to record time continuously during their observations.

The GPS in the drone provided location data in the form of Easting, Northing and time stamps (x, y and t) at a rate of one location per second. To be able to compare GPS and observed flights location by location, I interpolated equally spaced locations along the trajectories recorded by the GPS, in the same way I did with the trajectories drawn by the observers (figure 4.3).

4.2.3 Model specification

The data consist of drone flight trajectories, each of them embodied by a series of equally spaced locations. These locations were captured in the field using two methods: i) a telemetry device installed in the drone, which is supposed to have negligible mapping error, and ii) three different observers, who drew the flights on cartographic maps (see figure 4.3 and for more details see section 4.2.1). The locations captured by the drone will be referred to as *telemetry* or *true* locations from now on, and will be denoted by $\mathbf{r}_t$ in the equations. The locations captured by the observers will be referred to as *observed* locations and will be denoted by $\mathbf{Z}_t$. Both $\mathbf{r}_t$ and $\mathbf{Z}_t$ are two-dimensional vectors with components (x_t, y_t), with x_t measuring Easting, y_t Northing and the subscript t ordering locations in time. The observed location $\mathbf{Z}_t$ is considered to be a stochastic process and hence it is represented by a capital letter. The mapping error $\mathbf{D}_t$ is also a stochastic process, such that:

$$\mathbf{Z}_t = \mathbf{r}_t + \mathbf{D}_t \tag{4.1}$$

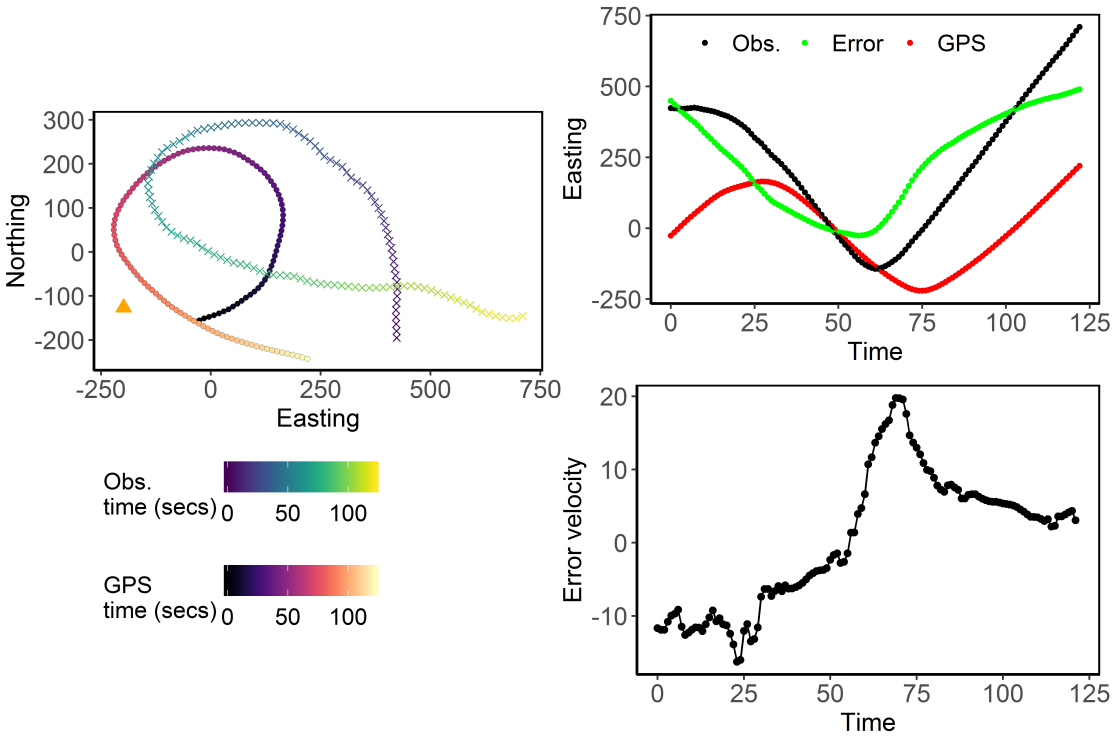


Figure 4.3: On the left panel: one of the drone flights, represented by a sequence of locations (dots) that go from black at time 0, to red, to yellow at time 125 seconds; and the same flight mapped by one of the observers, represented by a sequence of locations (crosses) that go from blue at 0, to green, to yellow at time 125 seconds; the orange triangle represents the vantage point from where the flight was observed. On the right top, the values of the Easting coordinate of the same flight over time as recorded by the drone GPS (red) and by the observer (black); the error (difference between GPS and observed) is plotted in green. On the right bottom panel, the velocity of the mapping error in the Easting coordinate of the same flight over time.

Note that the x and y dimensions of the stochastic processes $\mathbf{Z}_t$ and $\mathbf{D}_t$ are considered independent from each other. This means that I assume mapping error in the Northing and Easting coordinate to be interchangeable, in the sense that their probability distribution is the same. However, mapping errors are autocorrelated in time, as we will see shortly.

To understand the properties of the mapping error $\mathbf{D}_t$, I propose a model based on a stochastic differential equation that models its magnitude and dynamics. I start by modelling the magnitude of the mapping error when the drone is first detected ($t = 0$), before it starts moving. I shall refer to the mapping error at $t = 0$ as $\mathbf{D}_0$ and to this model as the *initial error model*. This model is key to understand the magnitude of mapping error at any given time, since the initial error marks the starting point from which the error evolves as flight progresses. Then, I define an *error dynamics model*, which describes the evolution of the mapping error in continuous time. The error dynamics model describes the dynamics of the error, capturing the correlation structure of consecutive error measurements along a flight trajectory. Accounting for this autocorrelation allows to keep track of the uncertainty in mapping error estimates more accurately. With the two models combined we may estimate an expected true location given an observed location at any given time. Therefore, the model captures the relationship between the shape of the true trajectory and the shape perceived by the observers, and it constitutes a data-generating model - i.e. it provides an understanding of the error dynamics with which we may simulate full error trajectories.

Initial error model

To start building the model for the initial error (at the time the drone is detected), I assume that in a homogeneous landscape, and by symmetry, mapping error is equally likely to take any direction. I also assume that the Easting and Northing components are independent from each other. Therefore, mapping error is modelled as two independent Gaussian random variables (one for Northing and one for Easting) centred at zero and with a common variance. This variance captures the variation in mapping error, but since the error is centred at zero, it will ultimately determine its magnitude.

The initial error $\mathbf{D}_0$ provides the initial condition for the mapping error at any other time $\mathbf{D}_t$, and it is modelled as a random variable. More precisely, since $\mathbf{D}_0$ has Northing and Easting components, it is a random vector (multivariate generalization of a random

variable). The variation in the initial mapping error is expected to be affected by topographical cues in the environment surrounding the drone, the distance between the drone and the observer, and by the height above the ground at which the drone is flying (Witt *et al.*, 2007; Eckstein, 2011).

I measured the amount of topographical cues available in a given flight using the Vector Ruggedness Measure (VRM), which accounts for variation in terrain slope and aspect (Sappington *et al.*, 2007). I used the average VRM in an area enclosing 500 meters around each flight (i.e. all locations along the same flight share the same VRM value). This measure was calculated using the SAGA package in QGIS. The distance between the drone and the observer, as well as the height of the drone above ground, were calculated from the GPS locations of the drone. The drone only provides information about its height with respect to the take-off point. To find the height of the drone above ground, I used the altitude of the take-off point, which I obtained from the Shuttle Radar Topographic Mission digital elevation model (NASA JPL, 2013).

To understand the variability of the mapping error, how large it might potentially become, and how it is affected by the aforementioned variables, I fit the following model to $\mathbf{D}_0$:

$$\begin{gathered} \mathbf{D}_{0(jv)} \sim \mathrm{N}(\mathbf{0}, \sigma^2_{0(jv)}\mathbf{I}) \\ \log(\sigma_{0(jv)}) = \theta_0 + \theta_1 \mathrm{VRM} + \theta_2 \mathrm{height} + \theta_3 \mathrm{dist} + \xi(\mathrm{obs}_j) + \eta(\mathrm{VP}_v) \\ \eta(\mathrm{VP}_v) \sim \mathrm{N}(0, \sigma_\eta) \end{gathered} \tag{4.2}$$

I model the log of the standard deviation of the random vector $\mathbf{D}_{0(jv)}$, which is centred at $\mathbf{0} = (0, 0)$ and has diagonal covariance matrix $\sigma_0^2\mathbf{I}$. $\mathbf{D}_{0(jv)}$ represents the initial mapping error of observer j at vantage point v. This model considers a baseline standard deviation θ_0 in the mapping error made by an average observer when estimating the spatial location of an object situated on a flat terrain, at ground level and at the same location as the observer. This parameter could be regarded as a measure of how precisely the average observer would situate her position on a map. The parameters θ_1, θ_2, and θ_3 control how characteristics of the environment affect the accuracy in locating the drone: VRM, drone height above the ground and distance of the drone from the observer, respectively. The parameters $\xi(\mathrm{obs}_j)$ account for the particular abilities of each observer j, adding or subtracting variability in error from the baseline. The effects of all observers sum up to zero ($\sum_{j=1}^{J} \xi(\mathrm{obs}_j) = 0$, for $J = 3$ observers). Finally, a random effect $\eta(\mathrm{VP}_v)$ accounts for uncontrolled factors associated with each vantage point v that may affect the spatial

perception of their surrounding areas.

Error dynamics model

The error dynamics model describes the evolution of the mapping error over time, as flight unfolds. The objective of this model is to determine how the observed trajectory is related to the true trajectory traced by the drone. As mentioned at the beginning of section 4.2.3, the observed flight is represented by the stochastic process $\mathbf{Z}_t = \mathbf{r}_t + \mathbf{D}_t$, where $\mathbf{r}_t$ is the true flight and $\mathbf{D}_t$ is the mapping error; another stochastic process. Each realization of the mapping error results in a trajectory that distorts $\mathbf{r}_t$ to produce an observed trajectory $\mathbf{Z}_t$. Realizations of the mapping error are themselves trajectories (see figure 4.3), and therefore, we may study properties, such as: position, velocity and acceleration of the mapping error. In this section, I compare several models that associate properties of mapping error $\mathbf{D}_t$ with properties of the true trajectories traced by the drone $\mathbf{r}_t$.

I compare several models to determine which one best captures the dynamics of the mapping error (see section 4.2.4 for details on model selection). Models take the from of second order stochastic differential equations, represented as systems of first order equations. They incorporate stochastic terms in the form of Brownian motion $\mathbf{B}_t$. Brownian motion is a stochastic process that is centred at zero, with independent increments, such that $\mathbf{B}_t - \mathbf{B}_s \sim \mathrm{N}(0, (t-s)\mathbf{I})$. Therefore, between time s and time t a trajectory defined by the process $\mathbf{B}_t$ changes by a random quantity that is proportional to the time difference $t - s$. It follows that no matter how small $t - s$ is, a trajectory of $\mathbf{B}_t$ is nowhere differentiable (see Nelson 2001 and figure 4.4).

However, flight trajectories $\mathbf{Z}_t$ captured by observers, as well as the true trajectories $\mathbf{r}_t$ traced by the drone are relatively smooth. Therefore, mapping error trajectories $\mathbf{D}_t$ must also be similarly smooth. Some smoothness in mapping error trajectories is achieved by making the stochastic process entering the model at the acceleration level. A so defined process produces trajectories that appear relatively smooth, since the stochastic "shocks" to changes in velocity are integrated at the position level (see figures 4.4 and 4.5). The models that are compared for describing the mapping error are defined as follows:

- Model 0 represents the acceleration of the mapping error evolving as a Brownian motion (figure 4.4). This implies that the behaviour of the mapping error trajectories is unrelated to that of the true trajectories. The rate of change of the mapping

error with respect to time $\mathrm{d}\mathbf{D}_t/\mathrm{d}t$ is equal to $\mathbf{V}_t$, the velocity of the mapping error. $\mathbf{V}_t$ therefore indicates the direction and magnitude of the change in mapping error per unit time. The change in $\mathbf{V}_t$ with respect to time, $\mathrm{d}\mathbf{V}_t/\mathrm{d}t$ can be conceived as the acceleration of the mapping error. In this case, the acceleration is modelled as a stochastic process, where the Easting and Northing increments in error velocity are independent from each other, and evolve according to Brownian motion $\mathbf{B}_t$ with a variance scale of σ:

$$\begin{aligned} \mathrm{d}\mathbf{D}_t &= \mathbf{V}_t \ \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= \sigma \ \mathrm{d}\mathbf{B}_t \\ \mathrm{d}\mathbf{B}_t &\sim \mathrm{N}(\mathbf{0}, \mathrm{d}t\mathbf{I}) \end{aligned} \tag{4.3}$$

- Model 1 considers the mapping error to behave like a harmonic oscillator, in which the error $\mathbf{D}_t$ tends to oscillate around zero (figure 4.5). In such a system, the acceleration of the mapping error towards zero at any given time, is proportional to the error at that time, by a factor of β. In other words, the larger the error, the faster the acceleration brings it back to zero. I refer to β as the "drift" parameter, because it controls the drift of the error towards zero. A "damping" term is included in the model by the parameter ϕ, to allow the amplitude and frequency of oscillations to change over time. In mechanics, the damping term represents a force that acts in opposite direction to the movement, reducing the amplitude of the oscillation. In this model, this term could do the opposite (if negative) and amplify the amplitude of the oscillations with time. The ϕ coefficient may take on any real value, and its sign will determine whether there is damping or amplification of the oscillations with time.

$$\begin{aligned} \mathrm{d}\mathbf{D}_t &= \mathbf{V}_t \ \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= -\beta\mathbf{D}_t \ \mathrm{d}t - \phi\mathbf{V}_t \ \mathrm{d}t + \sigma \ \mathrm{d}\mathbf{B}_t \\ \mathrm{d}\mathbf{B}_t &\sim \mathrm{N}(\mathbf{0}, \mathrm{d}t\mathbf{I}) \end{aligned} \tag{4.4}$$

- Model 2 is structured in a similar way to model 1, but now the equilibrium point is the initial error, instead of zero (figure 4.5). Observers believe that the position where they placed the drone at the beginning of the flight is accurate, and therefore any error in this location is likely to be propagated throughout the flight. Thus, mapping error will oscillate around $\mathbf{D}_0$ instead of zero. A damping/amplifying term is again included in the model through the parameter ϕ, to allow the amplitude

and frequency of oscillations to change over time.

$$\begin{aligned} \mathrm{d}\mathbf{D}_t &= \mathbf{V}_t \ \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= -\beta(\mathbf{D}_t - \mathbf{D}_0) \ \mathrm{d}t - \phi\mathbf{V}_t \ \mathrm{d}t + \sigma \ \mathrm{d}\mathbf{B}_t \\ \mathrm{d}\mathbf{B}_t &\sim \mathrm{N}(\mathbf{0}, \mathrm{d}t\mathbf{I}) \end{aligned} \tag{4.5}$$

- Under model 3, the observed trajectory would tend to match the shape of the true trajectory, but not necessarily its position (figure 4.5). Considering that the shape of a flight is given by the sequence of velocity vectors associated with it, the velocity of the observed trajectory would tend to match the velocity of the true trajectory. Considering that $\mathbf{Z}_t = \mathbf{r}_t + \mathbf{D}_t$, then, $\mathrm{d}\mathbf{Z}_t/\mathrm{d}t = \mathrm{d}\mathbf{r}_t/\mathrm{d}t + \mathrm{d}\mathbf{D}_t/\mathrm{d}t$ and if $\mathbf{V}_t = \mathrm{d}\mathbf{D}_t/\mathrm{d}t$, it follows that $\mathbf{V}_t = \mathrm{d}\mathbf{Z}_t/\mathrm{d}t - \mathrm{d}\mathbf{r}_t/\mathrm{d}t$, the velocity of the error equals the difference between the velocity of the true trajectory and the velocity of the observed trajectory. Therefore, under this model it is the mapping error velocity that drifts towards zero and not the error itself. Model 3 is formulated as:

$$\begin{aligned} \mathrm{d}\mathbf{D}_t &= \mathbf{V}_t \ \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= -\beta\mathbf{V}_t \ \mathrm{d}t + \sigma \ \mathrm{d}\mathbf{B}_t \\ \mathrm{d}\mathbf{B}_t &\sim \mathrm{N}(\mathbf{0}, \mathrm{d}t\mathbf{I}). \end{aligned} \tag{4.6}$$

- Under model 4, the mapping error (not its velocity) tends to drift towards zero by changing the direction of its trajectory, but not necessarily its travelling speed (figure 4.5). The dynamics of this model have an unstable equilibrium when the velocity of the error is zero. When there is no change in mapping error, the observed trajectory traces the same shape as the true trajectory, although possibly at a distance. A trajectory may stay in this state for some time regardless of how large the mapping error is. Therefore, some observed trajectories tend to match the position and shape of the true trajectory, while others might just match the shape but not necessarily the position.

$$\begin{aligned} \mathrm{d}\mathbf{D}_t &= \mathbf{V}_t \ \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= -\beta(\mathbf{V}_t + \omega_t) \ \mathrm{d}t + \sigma \ \mathrm{d}\mathbf{B}_t \\ \omega_t &= \hat{\mathbf{D}}_t|\mathbf{V}_t| \\ \mathrm{d}\mathbf{B}_t &\sim \mathrm{N}(\mathbf{0}, \mathrm{d}t\mathbf{I}) \end{aligned} \tag{4.7}$$

Where the hat notation means *unit vector* and the vertical lines, the norm or magnitude of a vector (i.e. $\hat{\mathbf{D}}_t = \mathbf{D}_t/|\mathbf{D}_t|$). Therefore, ω_t is a vector with direction from the current error towards zero and with magnitude equal to the current error speed. This preferred velocity brings the error to zero by shifting the direction in which the error is changing.

By solving these systems of stochastic differential equations we find the transitional density of the stochastic process. The transitional density defines the probability of transitioning from one state at time s to another state at time t for $t \geq s$ (see appendix B). Taking the limit as $t \to \infty$ of the transitional density, gives the stationary distribution of the process (if it exists). This is the state of equilibrium the process reaches if it evolves for long enough. The transitional density has mean and variances that change over time. The standard deviation function gives us an idea of how similar errors at different points along the trajectory are expected to be. At short intervals, consecutive errors are expected to be similar, and the variance of the process at short intervals small. On the contrary, consecutive errors separated by longer times are expected to deviate further from each other. This speaks about the autocorrelation in the mapping error along an observed trajectory.

If we are to estimate errors from the stochastic process at any given time, we need to find an initial value for the mapping error, an initial velocity at which the error changes, and then use the transitional density to find a distribution of possible error values at the desired time. An initial value for the mapping error may be sampled from the data itself or from the initial error model fitted to the data. A starting value for the mapping error velocity may be sampled either directly from the data, or from the stationary distribution of the velocity process $\mathbf{V}_t$ (see appendix B). As a way of assessing model fit, I use simulations to compare data generated by the mapping error models with actual data captured in the field. Eventually, simulations will be key to understand how the true flights might have looked like, given a sample of observed trajectories and a model for the mapping error.

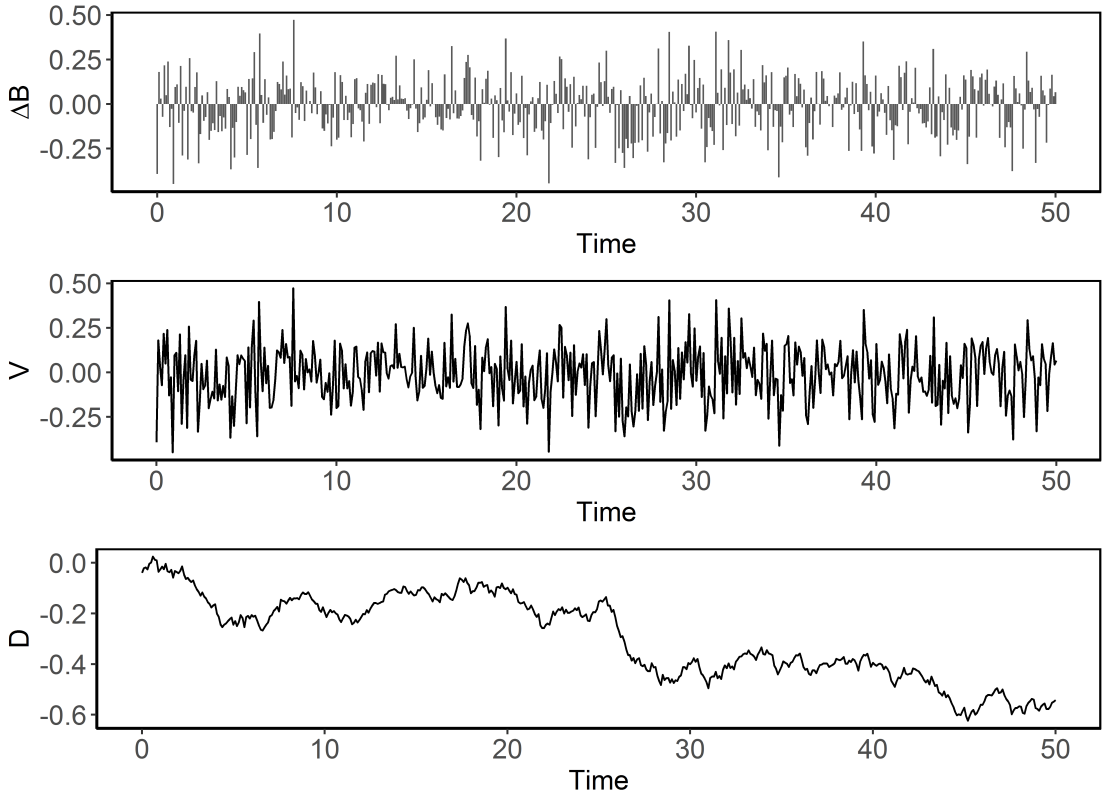


Figure 4.4: The figure represents the one-dimensional evolution of one realization of the mapping error $\mathbf{D}_t$ over time, when its velocity changes according to the increments of a Brownian motion $\Delta\mathbf{B}_t$ with variance $0.5^2\delta t$ ($\delta t = 0.1$). In the top panel, independent increments of the Brownian motion $\Delta\mathbf{B}_t$; in the central panel, the velocity of the mapping error process $\mathbf{V}_t$; and in the lower panel, the state of the mapping error process $\mathbf{D}_t$.

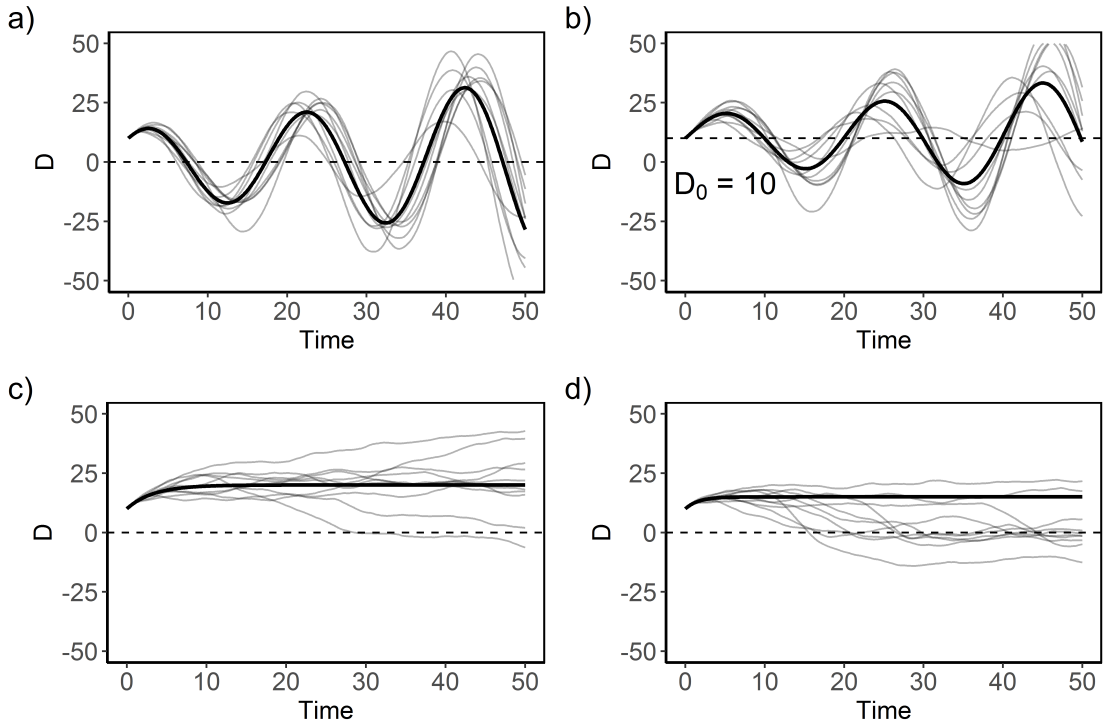


Figure 4.5: One-dimensional representation of the four models tested for describing the dynamics of the mapping error made by observers plotting the flights of a drone on a map. The black line represents the deterministic dynamics imposed by the models, and the grey lines are ten realizations incorporating the stochastic terms. All trajectories start with the same initial conditions of mapping error $\mathbf{D}_0 = 10$ and velocity of the mapping error $\mathbf{V}_0 = 3$. The stochastic changes in velocity are driven by a Brownian motion with variance $0.5^2\delta t$ ($\delta t = 0.1$). In panel a) model 1, represents oscillations around a mapping error $\mathbf{D}_t = 0$. In panel b) model 2, has similar dynamics to model 1, but the oscillations are around the initial error $\mathbf{D}_0$, rather than around $\mathbf{D}_t = 0$. In panel c) model 3, describes a tendency to stabilize the mapping error velocity. Without change in mapping error, the observed and true trajectories should have a similar shape, although they might be some distance apart. In panel d) the mapping error tends to drift towards zero, although if the mapping error velocity stabilizes, the error may stay in the current state for some time.

Hierarchical structure for the error dynamics model

The hierarchical structure of the error dynamics model arises from the drift β_i and damping ϕ_i coefficients being characteristic of each flight i. Thus, the mapping error of each flight has its own dynamics given by these parameters. However, there are some general dynamics that drive the "population" of mapping error trajectories. These general dynamics can be thought of as the general behaviour of the drift and damping of the mapping error, should we observe many flights. These general dynamics are described by a model of the drift and damping coefficients such that:

$$\begin{aligned} \beta_i &= \mu_\beta + \eta_i \\ \eta_i &\sim \mathrm{N}(0, \sigma_\beta^2), \end{aligned} \tag{4.8}$$

or more compactly,

$$\beta_i \sim \mathrm{N}(\mu_\beta, \sigma_\beta^2). \tag{4.9}$$

Similarly, I give the damping terms ϕ_i a distribution

$$\phi_i \sim \mathrm{N}(\mu_\phi, \sigma_\phi^2), \tag{4.10}$$

which describes the general behaviour of the damping in a population of error trajectories.

Since I am interested in the general dynamics of the mapping error, my general concern is with the estimation of μ_β, σ_β, μ_ϕ, and σ_ϕ, rather than with the estimation of drift and damping in the mapping error of particular flights.

4.2.4 Model fitting and selection

Models were fitted in R (R Core Team, 2019), in a Bayesian framework, using Hamiltonian Monte Carlo sampling provided by the statistical software Stan through the R package `rstan` (Stan Development Team, 2018c). Hamiltonian Monte Carlo is a Markov Chain Monte Carlo algorithm that uses information contained in the gradient of the posterior distribution to characterize it effectively (Neal, 2011). Markov Chain Monte Carlo

algorithms form sequences of parameter values by sampling parameter values and computing posterior densities (Gelman *et al.*, 2014). There are different MCMC algorithms and their review is out of the scope of this work. I will just point out that sequences of parameter values form chains with the Markov property (i.e. each sampled value only depends on the previous one) and that in the long run, the distribution of sampled values arrives at a stationary distribution that converges to the parameter posterior distribution.

Because chains might be sensitive to the starting point of the sequence, I ran four Markov chains of 4000 iterations each with a warm-up phase of 2000 iterations to sample the parameter posterior distributions. The initial values contained in the warm-up phase were discarded because it may take a number of iterations for the chains to arrive to the stationary distribution and the first "transient" phase may not be representative of the posterior distribution. Covariates were standardized by dividing them by their standard deviation before fitting the models. Parameters were given relatively weak priors for the initial error model $f(\theta) \sim \mathrm{N}(0, 3)$ for $\theta \in \{\theta_0, \theta_1, \theta_2, \theta_3, \xi\}$. The random effects of vantage points were given a prior $f(\sigma_\eta) \sim \mathrm{Half} - \mathrm{Cauchy}(0, 3)$.

Convergence of the error dynamics model required more informative priors. The mean population drift and damping, μ_β and μ_ϕ were given priors $f(\theta) \sim \mathrm{N}(0, 0.5)$ for $\theta \in \{\mu_\beta, \mu_\phi\}$ and the dispersion in population drift and damping σ_β and σ_ϕ were given priors $f(\theta) \sim \mathrm{Half} - \mathrm{Cauchy}(0, 0.2)$. The scale of the Brownian motion σ had a prior $f(\sigma) \sim \mathrm{Half} - \mathrm{Cauchy}(0, 2.5)$.

Model convergence was assessed by visually inspecting that all Markov chains defined a similar distribution (the chains "mixed" well), and by the convergence diagnostic $\hat{R}$, which tells us about the reduction in the dispersion of the posterior that we could expect should we produce many more posterior samples (Gelman *et al.*, 2014). Thus, values of $\hat{R}$ that are close to one indicate good convergence, and values under 1.1 may be considered acceptable. A visual inspection of simulations from the model posteriors was also used to assess how the models capture the underlying structure of the data.

Candidate models are compared using cross-validation. The objective of cross-validation - intimately related to information criteria selection methods (see Stone, 1977) - is to favour models that capture structure in the data that generalizes well to other observations arising from the same underlying process, while penalizing models that fit to spurious patterns in the data. The procedure consists of dividing the data into several sets. Some are used as training sets and some as validation sets. The training sets

are used to estimate model parameters that are used to make "predictions" about the observations in the validation set.

Some authors recommend 5- or 10-fold procedures (Kohavi, 1995; Hastie *et al.*, 2009). Other authors (e.g. Zhang & Yang, 2015) suggest that allocating more observations to the validation set, increases the power to distinguish between closely competing models and at the same time, having fewer training cases, accentuates the effect of important variables. Under this premise, delete-$n/2$ cross-validation in which half of the data is used for training and half for prediction, would be preferred. In general, it is accepted that whichever procedure is chosen it must be repeated several times with different random allocation of the data to training and validation sets (Kohavi, 1995; Hastie *et al.*, 2009; Zhang & Yang, 2015).

Another important aspect in the allocation of data to training and validation sets is whether selection must be stratified. When data is structured, be it in blocks, in space or in time, allocating individual data or entire blocks to the cross-validation sets have different effects (Roberts *et al.*, 2017). Generally speaking, if we were interested in evaluating the predictive power of the model on unobserved cases within observed groups, we could allocate data randomly for training and validation. On the other hand, if we were concerned with predictive power on unobserved groups, then the groups allocated for training should be absent from the validation sets.

In this study, I use 5-fold cross-validation, dividing the data into five sets, and using four of them for training and the other one for validation. Then, I iterate the sets until all have been used for training and validation. Since, the models are fitted using Hamiltonian Monte Carlo sampling, which is computationally intensive, a 5-fold procedure balances computing time with variability in training and validation sets. Data is allocated to training or validation sets in blocks. This is, entire flights are either allocated for training or validation, because we are not concerned with interpolating predictions within partially observed flights, but rather with predicting whole flights. The training sets are used to estimate model parameters, which in turn are used to calculate the log-likelihood of the observations in the validation set. I select the model that provides the largest average log-likelihood of data unseen by the model.

4.3 Results

The drone performed a total of 37 flights. One of the flights had to be repeated due to lack of communication with observers 1 and 2, making the total number of flights 37, instead of the original 36. A total of 101 flights were successfully recorded by the observers ($n_{obs1} = 30$, $n_{obs2} = 34$, $n_{obs3} = 37$). The average duration of the flights was 140 seconds with a standard deviation of 32 seconds. The average Euclidean distance between locations recorded by the GPS in the drone and locations recorded by the observers was 422 meters, with a standard deviation of 336 meters. The distance between the observers and the drone ranged from 70 meters to 1581 meters during the flights.

Initial error model

At take off, the Euclidean mapping error made by the observers was on average 259 meters (standard deviation = 283; interquartile range = 295). Easting and Northing errors were modelled as independent random variables. Therefore, all Easting and Northing values were pooled together as a one-dimensional error for model fitting. The one-dimensional error had a mean of 16 meters and a standard deviation of 271 meters (figure 4.6). The drone was at a distance of between 121 meters and 1291 meters away from the observers when it was detected at the beginning of the flights. The height of the drone above the ground at the beginning of the flights ranged from 20 to 215 meters.

The *initial error model* estimated a mean baseline standard deviation in mapping error of approximately $e^{4.202} \approx 67$ meters (table 4.1). Assuming that the distribution of mapping error is centred at 0 and that most of its distribution lies within three standard deviations from the mean, the baseline mapping error is expected to lie roughly within the interval (-200, 200) meters - both Easting and Northing errors. Observer 1 (Obs1) had a larger dispersion in mapping error (table 4.1 and figure 4.6), leading to potentially larger errors (by a factor of $e^{0.53} \approx 1.7$, table 4.1), while observers 2 (Obs2) and 3 (Obs3) were more accurate (table 4.1 and figure 4.6), making smaller errors than average (correction factors of $e^{-0.19} \approx 0.8$ and $e^{-0.34} \approx 0.7$, respectively; table 4.1).

Terrain ruggedness did not have a clear effect on the initial mapping error. The model suggests that as ruggedness of the terrain (VRM) increases, the standard deviation of the mapping error also increases (by a factor of $e^{0.155} \approx 1.18$ for a unit increase in VRM; table 4.1). VRM values ranged from 0.06 to 1.25 and considering that the estimated

effect is for one unit increase, this covariate is not expected to influence mapping error importantly. In addition, the variance estimated for this effect is very large compared to its mean - the 95 % credible interval for this effect is ($e^{-0.824} \approx 0.4, e^{1.147} \approx 3.1$) -, suggesting that the effect of terrain ruggedness on mapping error is not only small on average, but also very variable.

The vantage point from which the flights were observed influenced the mapping error. This reflects in the random effect of vantage point on the standard deviation of the mapping error (see $\eta(\mathrm{VP}_1)$ to $\eta(\mathrm{VP}_6)$ and σ_η in table 4.1), suggesting that uncontrolled variables characteristic of the areas observed from the vantage points influence mapping error.

The height of the drone above ground was negatively correlated with the standard deviation of the mapping error (table 4.1), meaning that greater flight height leads to smaller mapping errors (more concentrated around zero). The estimated effect is small, but consistently negative (95 % credible interval for this effect is ($e^{-0.009}, e^{0}$), table 4.1). Despite this effect being small, it is a multiplicative effect for every meter above ground, which may lead to potentially important effects on mapping error (height ranged from 20 to 215 meters). Adding an *ad-hoc* quadratic effect on height had virtually no effect on model fit.

According to the model, the standard deviation of the mapping error is directly proportional to the distance between the observer and the drone (table 4.1). For each meter between them, the baseline standard deviation of the mapping error (given by θ_0 in table 4.1), would increase on average by a factor of 1.002 (0.2 % increase). Considering that the largest distance measured between observer and drone was 1291 meters, the increase in error standard deviation could be up to 13 times that at zero distance.

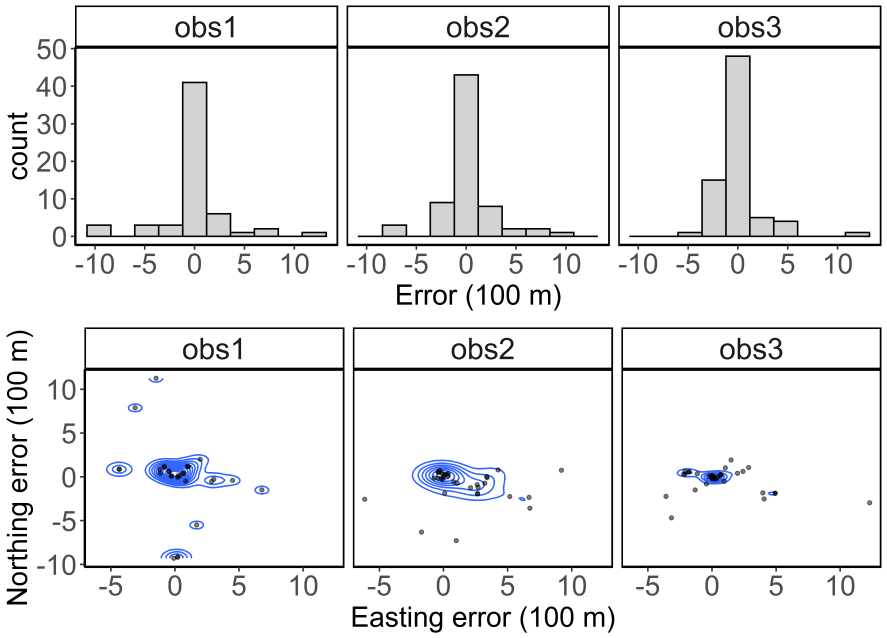


Figure 4.6: One-dimensional histogram of the mapping error at time $t = 0$ (top panels) and two-dimensional mapping errors (bottom panels). In the top panels, the Northing and Easting errors are pooled together. The bottom panels show the two-dimensional error, with Northing in the y-axis and Easting in the x-axis. The blue contours represent levels of kernel density estimated by the R package `MASS`. In both top and bottom panels, error are separated for each observer. Observer 1 has the greatest errors; observer 2 has larger errors than observer 3; and observer 3 has the smallest errors.

Table 4.1: Summary of the parameter posterior distribution of the initial error model. The table shows: mean, standard deviation (sd), 2.5^{th}, 50^{th} and 97.5^{th} quantiles, the number of effective Hamiltonian Monte Carlo samples (n_eff, roughly measures the precision of parameter estimates) and the convergence diagnostic $\hat{R}$. The corresponding covariate and the notation defined in the text for the parameters are shown in the first column. VRM stands for Vector Ruggedness Measure and Obs1-Obs3 are the estimated effects of each observer on mapping error. VP1-VP6 are estimates of the particular effects of the vantage points on mapping error, while σ_η is the standard deviation of the random effect for vantage point. The model estimates the log standard deviation of the mapping error. This means that for one unit increase in a covariate, the standard deviation of the error is multiplied by a factor of e^x, where x is the coefficient estimate for the covariate.

	mean	sd	2.5%	50%	97.5%	n_eff	Rhat
Intercept (θ_0)	4.202	0.608	3.065	4.185	5.484	1891.710	1.001
VRM (θ_1)	0.155	0.497	-0.824	0.155	1.147	4158.868	1.000
Height (θ_2)	-0.005	0.002	-0.009	-0.005	-0.000	6459.562	1.000
Distance (θ_3)	0.002	0.000	0.002	0.002	0.003	5853.147	1.000
Obs1 ($\xi(\text{obs}_1)$)	0.531	0.083	0.368	0.531	0.697	6473.855	1.000
Obs2 ($\xi(\text{obs}_2)$)	-0.195	0.075	-0.341	-0.196	-0.047	6259.787	1.000
Obs3 ($\xi(\text{obs}_3)$)	-0.337	0.078	-0.488	-0.337	-0.181	10106.598	1.000
VP1 ($\eta(\text{VP}_1)$)	-1.647	0.557	-2.837	-1.622	-0.614	1850.706	1.001
VP2 ($\eta(\text{VP}_2)$)	-0.195	0.515	-1.239	-0.188	0.801	1716.437	1.001
VP3 ($\eta(\text{VP}_3)$)	0.624	0.549	-0.524	0.638	1.654	1785.428	1.001
VP4 ($\eta(\text{VP}_4)$)	0.338	0.544	-0.729	0.338	1.428	1930.778	1.000
VP5 ($\eta(\text{VP}_5)$)	0.337	0.540	-0.707	0.335	1.414	1874.397	1.000
VP6 ($\eta(\text{VP}_6)$)	0.313	0.510	-0.695	0.316	1.352	1705.191	1.001
σ_η	1.104	0.482	0.539	0.990	2.350	2630.229	1.001

Error dynamics model

The 101 observed trajectories had between 21 and 51 locations each, with a mean of 29.3 and a standard deviation of 10.4. The mean distance between observers and drone was 695 meters and ranged between 70 meters and 1580 meters. The drone flew on average 100 meters high and within a range of 5 meters and 188 meters above ground level.

The model that found best support by the cross-validation procedure, based on the log-likelihood of the validation sets, was model 2 (figure 4.7). Under model 2, the mapping error tends to drift towards its initial state with an acceleration proportional to the difference between the initial and current error. This produces a mapping error that oscillates around the initial error (figures 4.5 and 4.8).

As explained in section 4.2.3, each observed flight i had a drift parameter β_i that described the dynamics of that particular flight. However, the interest was on the general dynamics of the mapping error. These dynamics are imposed by the distribution of β_is, which was given by $\beta_i \sim \mathrm{N}(\mu_\beta, \sigma_\beta^2)$. The posterior mean of the drift in velocity μ_β at which mapping error accelerated towards its initial value in an average flight was estimated to be 0.061 times the mapping error at any given time (table 4.2). This rate is measured in m/s per minute. This is effectively an acceleration that causes the mapping error to drift back towards initial value, faster as the error grows larger. The variation (standard deviation) in drift of the mapping error for different flights was captured by the parameter σ_β and had a posterior mean of 0.049 (table 4.2). This value is relatively large compared to the mean (μ_β), and it means that part of the distribution of the drift parameters had negative values. Negative drift values produce a drift away from the initial mapping error. According to the model this could happen roughly in 8 % of flights ($P(\beta < 0) = 0.08$).

The model also incorporated a damping term that modified the rate at which the mapping error changed. Similar to drift parameters, damping parameters ϕ_i may change for each flight i. Damping parameters also have a distribution that describes the general behaviour of a randomly chosen trajectory, such that $\phi_i \sim (\mu_\phi, \sigma_\phi)$. Model 2 estimated a posterior mean of 0.012 for the average damping in a trajectory (μ_ϕ; table 4.2). The standard deviation for the average damping was 0.004 (σ_ϕ; table 4.2). This positive damping effect acts counteracting the acceleration towards the initial error and reducing the amplitude of the oscillations. The parameters β and ϕ are found to be mostly uncorrelated ($r = 0.058$). Therefore, large acceleration towards the initial error in a flight is not necessarily associated with large damping or viceversa.

The model described a mapping error that was shocked by random perturbations that break the regular oscillatory dynamics around the initial error. Under this model, random perturbations in velocity are represented by a Brownian motion with standard deviation scaled by 13.262 (σ, table 4.2). According to the error dynamics model, the variance function of this process grows sharply during the first 2 minutes and then continues to increase more slowly until approximately 5 minutes (figure 4.9). After this period it stabilizes at around 350 meters.

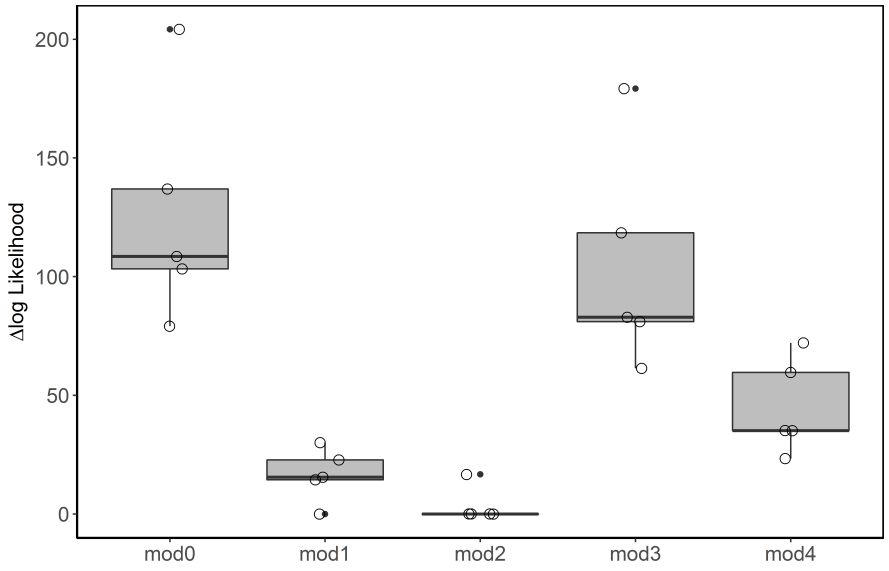


Figure 4.7: Difference between the maximum log-likelihood obtained in each cross-validation (CV) iteration and log-likelihood of each candidate error dynamics model ($\Delta log - Likelihood$). The model with the largest log-likelihood is considered to fit the data best. Therefore, a good model has a small $\Delta log - Likelihood$. The boxes capture the median and interquartile range (IQR), while the whiskers extend no further than 1.5 IQR. The solid points represent values further than 1.5 IQR and the empty circles are the actual $\Delta log - Likelihood$ values (note the horizontal jitter). Model 2 ($mod2$) fitted the data best in all but one CV iteration.

Table 4.2: Summary of the parameter posterior distributions of the error dynamics model (model 2). The table shows: mean, standard deviation (sd), 2.5^{th}, 50^{th} and 97.5^{th} quantiles, the number of effective Hamiltonian Monte Carlo samples (n_eff, roughly measures the precision of parameter estimates) and the convergence diagnostic $\hat{R}$. Model 2 considers the general behaviour the mapping error to be described by a drift towards the initial error. The drift for any flight i has a distribution $\beta_i \sim \text{N}(\mu_\beta, \sigma_\beta)$. The motion of the mapping error is damped at a rate proportional to the velocity at which error is changing. The damping parameter ϕ_i for any flight i has a distribution $\phi_i \sim \text{N}(\mu_\phi, \sigma_\phi)$. Finally, random inputs in the form of a Brownian motion with standard deviation scaled by σ, disturb the change in mapping error.

	mean	sd	2.5%	50%	97.5%	n_eff	Rhat
μ_β	0.061	0.006	0.050	0.061	0.073	1219.981	1.003
μ_ϕ	0.012	0.001	0.010	0.012	0.015	7660.905	1.000
σ_β	0.049	0.005	0.041	0.049	0.059	1680.412	1.002
σ_ϕ	0.004	0.002	0.000	0.004	0.007	1688.892	1.003
σ	13.224	0.055	13.117	13.224	13.332	11366.277	1.000

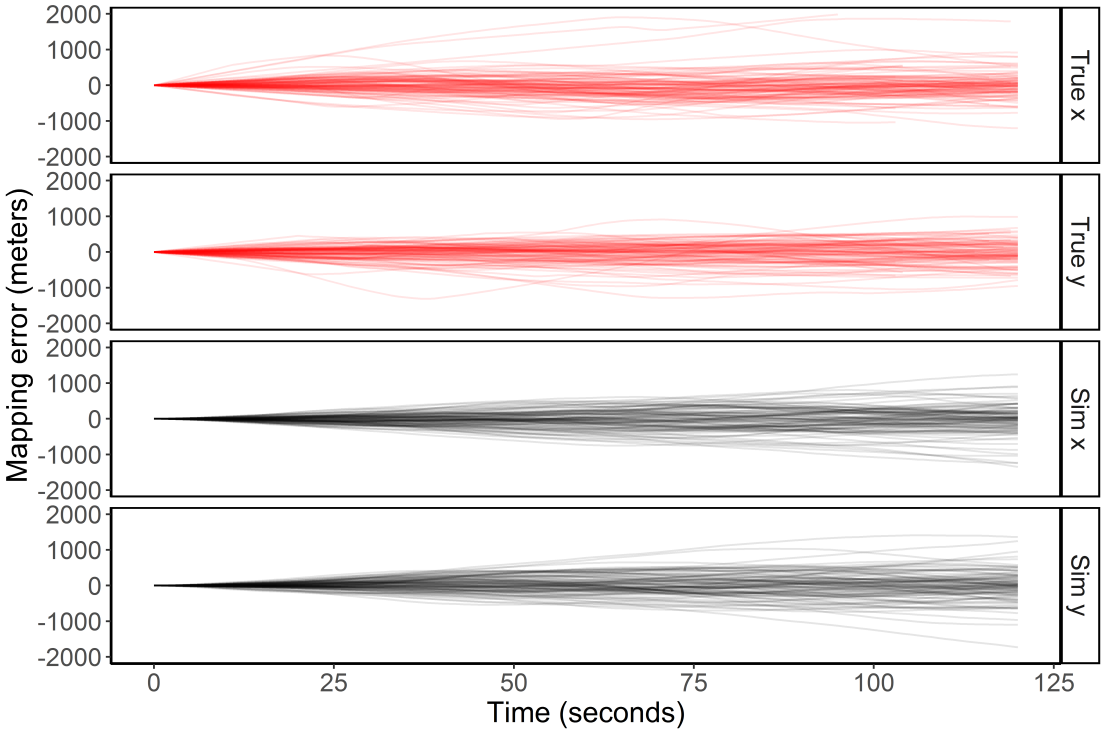


Figure 4.8: Comparison of the observed mapping error trajectories (top panels, in red) with simulated error trajectories using the error dynamics model, model 2 (bottom panels, in black). The initial mapping error (at $t = 0$) is subtracted from the error at every other time, to make all trajectories start (and centred) at zero.

4.4 Discussion

Vantage point surveys are often used to assess collision risk and avoidance of wind turbines, and other man-made structures by raptors and other soaring birds (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). During these surveys observers map flight trajectories of birds they observe to investigate their use of space. However, it is unclear how accurately these trajectories are mapped. In this study, I modelled the error made by three observers mapping the flights of a drone in two dimensions to approximate the error we would expect on the mapping of bird flights. The model considers the mapping error at the initial time (time of detection of the drone) being a random variable centred at zero. The variability in this error, given by its standard deviation, may be interpreted as a measure of how large the error is; the larger the standard deviation, the farther away from zero the error can be. However, it is important to remember that the error when the bird is first detected is considered to be zero on average (errors may take different directions, averaging each other out). The change in mapping error as the drone moved was modelled as a stochastic process that oscillates around the initial error value.

At detection time, I found the standard deviation of the error in location to be on average approximately 66 meters in both Easting and Northing directions. This results in an Euclidean distance between recorded and real position of approximately 130 meters. If we consider that 99 % of the error should lie within three standard deviations from the mean, in occasions, observed locations may be as far as 400 meters from the true locations. According to the initial error model, the variability in mapping error depends on the observer and also on the surroundings of the flying drone. For this study there were only three observers available and therefore it is difficult to define a rule to generalize to other observers. However, differences among observers were important. For example, the average error made by the least accurate observer more than doubled the error of the most accurate observer.

The effect of terrain ruggedness on the mapping error estimated by the initial error model is small and very variable. Flights performed around vantage points VP1 and VP2, were more accurately recorded than those performed at other vantage points. VP1 and VP2 were located in the flat areas, with less topographical clues. This finding contrasts with previous research on visual perception, which found that environmental cues favoured more accurate perception of the spatial layout (e.g. Sinai & He, 1998; Witt *et al.*, 2007). In fact, VP3 was also in a flat area and flights there had an associated error larger than

those flights performed at the more rugged vantage points. Therefore, I found no clear pattern associated with terrain ruggedness. Instead, variations in mapping error are better described as being produced by some uncontrolled characteristic of the vantage points (i.e. a random effect). This could be the result of a combination of factors, such as: ruggedness, flight height, distance, or the relative position of the sun with respect to the drone.

According to the initial error model, the height at which the drone flies has a negative effect on mapping error - the standard deviation of the error in location tend to decrease with increasing flight height. This effect should be extrapolated with caution out of the range of observed flights, since it seems unlikely that the mapping error will continue to decrease at indefinitely higher heights. Yet, within the range of the data, the decrease in mapping error was linear, as shown by the negligible effect of the quadratic effect of height on mapping error. The better accuracy at locating the drone in higher flights might be explained by the background against which the drone was observed (e.g. Eckstein, 2011). At lower heights observers need to detect the drone against the ground; whereas, at higher altitudes the drone's silhouette will be more clearly defined against the sky. This was indeed one of the remarks made by the observers themselves, who pointed out how they struggled to picture the drone when it was flying low, against a cluttered background. This may also explain why in some of the flat areas observers were more accurate at locating the drone than in more rugged ones.

At an initial time, when the drone was detected, greater distances between drone and observers produced an increase in the dispersion of the mapping error. The effect is multiplicative (mapping error is modelled in the log scale), which may result in potentially large error at far distances. For example, when the drone is first detected at a distance of one kilometre, on average conditions - height = 80 m, VRM = 0.5, no observer or vantage point correction - the standard deviation of mapping error is estimated to be around 500 meters (Euclidean distance with true location).

The error dynamics model suggests that the mapping error made at the time of detection of the drone, will be carried forward for the rest of the flight. Mapping error changes as flight progresses but it tends to oscillate around the initial error. Modelling mapping error as a stochastic process, governed by a mean function and a covariance function both dependent on time, I was able to capture the time dependence of consecutive mapping errors along flight trajectories. Thus, the model is appropriate to estimate the mapping error at an arbitrary point along an observed flight trajectory, and also to simulate error trajectories (figures 4.8 and 4.10). Furthermore, with the model proposed here, we may

filter (in a probabilistic sense) the mapping error out of the observed trajectories to estimate true trajectories.

The error dynamics model also provides information on how accurately the velocity of the flight was recorded. This is because the velocity of the error is equal to the difference between observed and true velocities. For example, when the velocity of the error is zero the error is not changing, and therefore the observed trajectory captures the velocity of the true trajectory accurately. This is helpful to understand the accuracy of assessments concerned with the direction of the flights. For example, whether flight paths tend to head towards certain environmental features, and therefore whether there is active selection of habitats, topographical features, etc. In chapter 5 I show how a data-generating model for the mapping error can be used to improve estimates of exposure to collision based on bird flights observed from vantage points.

Although the model captured the general dynamics of the mapping error, it seems to underestimate the variance of the error at short times, while overestimating it at longer times (see figure 4.8). This effect is probably produced by the choice of initial error velocity. With larger initial error velocities the errors would grow faster at the beginning of the flight. In this study I used the stationary distribution of the error velocity to sample initial velocity values. Perhaps this stationary distribution does not capture the variability in initial conditions adequately. I hypothesise that a different model specification with error velocities being represented by distributions with thicker tails than the Gaussian may be more appropriate (such as the Student-t distribution).

The dispersion of the mapping error is expected to grow up to 350 meters after the initial detection of the drone. Therefore, mapping errors may grow larger in longer flights, although according to the error dynamics model, it tends to stabilize after about five minutes. This also means that according to the fitted model, consecutive points along the same trajectory must be as far apart as five minutes to be regarded independent from each other. However, as mentioned earlier, extrapolations over periods longer than two minutes (the duration of most flights in the data) must be done cautiously.

The data used in this study was restricted to a reduced geographic area, flights of around two minutes in duration and to three observers. More research is needed to determine the extent parameters values change in different conditions and to test how robust the general structure of the model is to different observation conditions. Also, since the objective of this study is to understand the accuracy of observers at plotting bird flights onto maps, it is worth noticing the fact that the observers were concerned with how

difficult to detect the drone was compared to birds. They also commented on how birds give more cues on the direction of their flights, based on their morphology and colouration (e.g. different colouration in back and underparts). For these reasons, it is likely that mapping error when plotting drone flights is an overestimation of the error when plotting bird flights. Therefore, using the results from this study as a reference for the mapping error associated to bird flights records, should fall on the conservative side, with errors likely to be smaller than estimated by the model presented here. Another fact that supports the idea of the errors presented here being larger than errors in regular vantage point surveys, is that in this study observers could not practice or familiarize themselves at leisure with the study area. During wind farm monitoring programmes survey areas are visited repeatedly, for extended periods of time, whereas in this experimental setting observers only visited twice each vantage point for a short period time. I believe that the identification of landmarks that are familiar to the observer should reduce the errors estimated in this study.

I have presented empirical measurements of the mapping error three observers made when visually tracking and mapping a flying drone. I have also defined a model for the magnitude and dynamics this error. The proposed model provides a framework from which it is possible to evaluate the accuracy of visually mapped flights. It provides the means to obtain objective estimates of the mapping error at any point along a flight trajectory. Therefore, this model may be used to make probabilistic inference about the true trajectory traced by the drone, given an observed trajectory. Figure 4.10 provides an example of how simulations may be used to estimate possible true trajectories from observed trajectories using the mapping error model (see appendix D for a sample of simulation code - although the model is different, the simulation principles are the same). The results of this study suggest that the error in mapped flights can be substantial; and therefore, I advise against the use raw mappings for the assessment of infrastructure placement (e.g. turbine placement at wind farms). Rather, I suggest considering, in probabilistic terms, where the true flights might have occurred given the mapped flights.

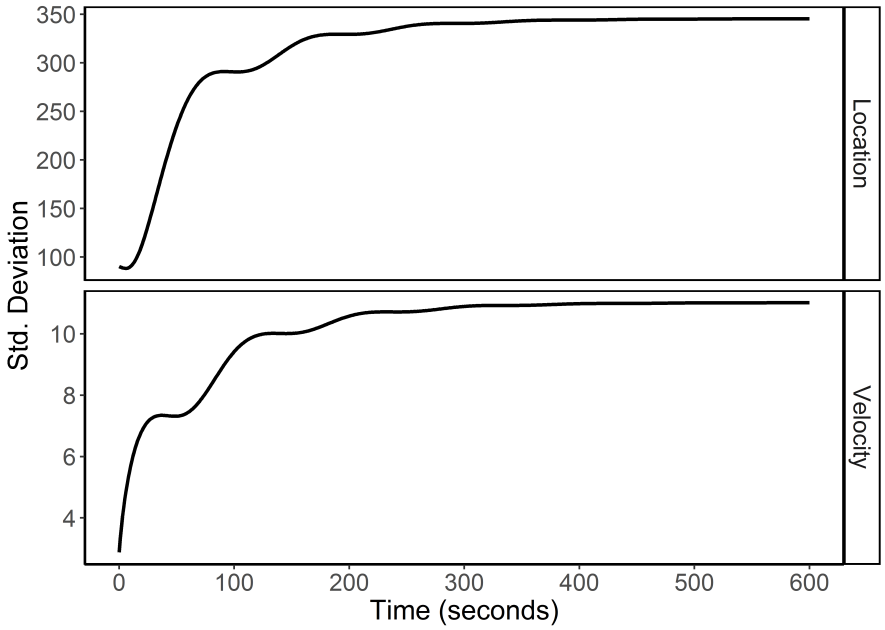


Figure 4.9: Change in standard deviation of the mapping error (top panel) and velocity of the mapping error (lower panel) processes with time.

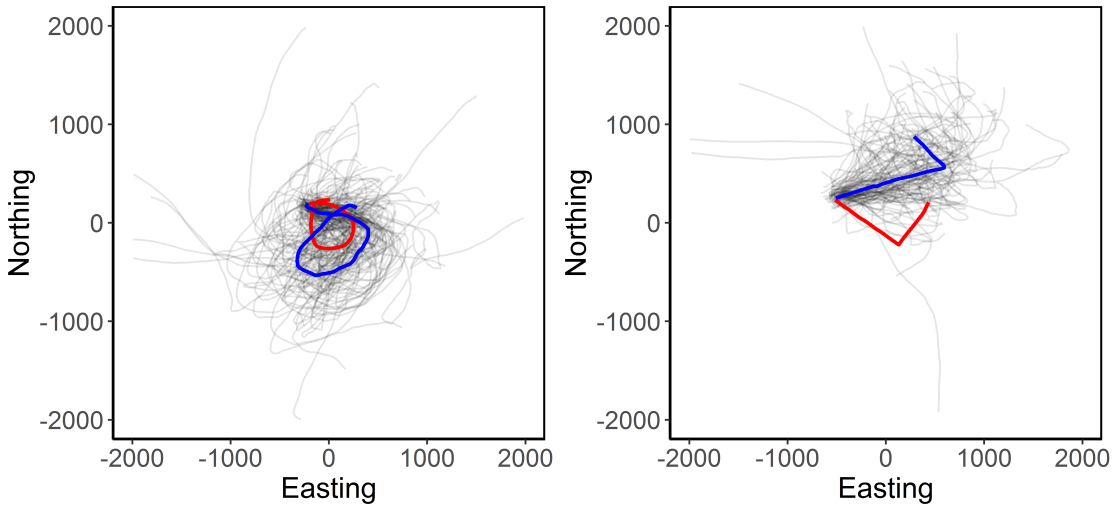


Figure 4.10: Trajectories simulated from an observed trajectory according to the mapping error model. In red, the true flight trajectories performed by the drone. In blue, the trajectories recorded by observer 3 in the left panel and observer 2 in the right panel. In grey, 100 simulated trajectories based on the observed trajectories and the mapping error model. On the left panel the true flight lies within the area occupied by the simulated flights. On the right panel the true flight, although still in the general area is not in the main area defined by the simulations. This suggests that simulations may be sensitive to the initial conditions - initial error velocity in this case.

Chapter 5

Predicting collision risk of the Verreaux's Eagle using vantage point observations at a wind farm in South Africa

5.1 Introduction

Wind energy is a major player in the renewable energy industry, and therefore it is at the forefront of climate change mitigation (REN21, 2018). However, the urgent need to reduce carbon emissions to the atmosphere must be balanced against other environmental concerns. From an ecological perspective, collision of flying vertebrates with turbine blades is one of the main impacts associated with wind energy production. Raptors and other large soaring birds are considered particularly sensitive to this impact due to their low abundance, low reproduction rates and collision-prone flight behaviours (Drewitt & Langston, 2006; Marques *et al.*, 2014).

It is for this reason that environmental impact assessment studies of wind farms incorporate especially designed bird monitoring programmes. Vantage point surveys very often provide critical information to assess the best turbine layout configuration, or even

the viability of a wind farm (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). During these surveys, observers sitting in a location with good visibility, collect data on abundance and flight behaviour of soaring birds within and around the wind farm. The way in which these data are collected is quite flexible, but in general, observers plot the flights they observe into topographic maps and collect ancillary information such as: number of individuals, duration of the flight or displayed behaviour (e.g. hunting, courtship, etc.). Vantage point surveys deviate from other well established survey techniques - such as point counts -, and their assumptions and limitation are still not well understood. For example, due to the small sampling radius of point counts (rarely more than 100 metres), they usually cover a single habitat, which allows making clear associations between bird activity and habitat characteristics (Bibby *et al.*, 2000; Matsuoka *et al.*, 2014). On the contrary, the large areas observed from vantage points (which span several squared kilometres) are usually heterogeneous, which complicates linking activity levels to the different habitats. To complicate things further, birds typically move after they are detected, and visit several locations consecutively, with potentially different habitats.

Very often observers map the flights they observe to keep track of the locations visited by the birds, the direction of the flights, or bird behaviour (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). However, these trajectories are not analysed explicitly; the analysis is either qualitative, or based on metrics such as abundance of flights, flight time over the vantage point, or even over the whole wind farm (e.g. de Lucas *et al.*, 2008; Garvin *et al.*, 2011; Jenkins *et al.*, 2018). In addition, results of monitoring programmes are rarely published in the scientific literature, and validation of methods designed to predict collision risk are rare (e.g. Marques *et al.*, 2014; Green *et al.*, 2016). Finally, although observers are likely to introduce errors when mapping the flight trajectories they observe (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014), to my knowledge, there are no studies that look at the characteristics of this error, or that account for it during the analysis.

In this case study, I use flights trajectories observed from vantage points to model the long-term activity distribution of the Verraux's Eagle *Aquila verreauxii* at a wind farm in the Eastern Cape of South Africa. Assuming that activity intensity correlates with collision risk, at least within the same species, I compare areas of high use intensity predicted by the model, with those areas where fatalities were actually observed. I predict activity intensity using vantage point observations collected in the pre-construction phase of the wind farm; whereas fatality data were collected during the first year of the project

operation. Movement data on the Verreaux´s Eagle were also available from telemetry-based studies on the species. Although these tracking data had no relation with the wind farm, and the tracked individuals lived far from the project, they provide valuable information pertaining the movement behaviour of the eagles.

I use two different approaches to estimate the long-term activity distribution: i) using a non-parametric model based on kernel density estimation (see chapter 2) and ii) using a mechanistic model based on attraction towards certain landscape features (see chapter 3). Kernel density estimators use the number of observations in each region of space to estimate the underlying probabilistic distribution that gives rise to the patterns we observe in the data (Silverman, 1986b). The underlying probabilistic distribution can be thought of as the long-term distribution of observations, should we observe the data-generating process repeatedly for a long time. This is a non-parametric method that requires little understanding of the data-generating process. At the same time, inference is limited to the distribution of activity without further insights on the mechanisms leading to this distribution. In addition to calculating a kernel-based density, I estimated the long-term activity distribution of Verreaux's Eagle at the wind farm by modelling the mechanisms driving the instantaneous velocity (heading and speed) of the flights observed from vantage points. I use a potential-based velocity model, which is a continuous-time movement model that considers drift in movement velocity to be produced by the action of attraction forces towards certain elements of the landscape (Johnson *et al.*, 2008; Gurarie *et al.*, 2017; Russell *et al.*, 2018). The advantage of this method is that it offers some degree of understanding of the behaviour of the eagles, which allows us to make predictions out of the range of the observations. However, as implemented in this study, we need to make assumptions in terms of what elements of the landscape my trigger eagle reactions.

These methods allow me to investigate the distribution of activity from two different perspectives, while accounting for the temporal dependence in the data. Finally, I acknowledge that observers may incorporate mapping error to their recorded trajectories, and that the recorded flights are the product of just one of many possible realizations of this error. Therefore, the same flights could have been mapped differently (see chapter 4). I model this error and incorporate it into the analysis to improve inference about the movement behaviour of the eagles.

5.2 Methodology

5.2.1 Study area

The wind farm is situated in the Eastern Cape of South Africa. It comprises 40 turbines with a generating capacity of 2.5 MW each, a hub height of 80 metres and a rotor diameter of 100 metres.

The vegetation on site is represented by a Dry Highveld Grassland (Mucina & Rutherford, 2010). The area is located at an altitude around 1700 metres above sea level. Hills and small mountains (> 200 m in elevation) covered by grasslands form a predominantly smooth relief, with the notable exception of the Great Escarpment. These large cliffs separate the Southern African plateau to the North from the lower grounds of South Africa towards the South. The wind farm is located 1 kilometre North-West from the border of the Great Escarpment, on the plateau side. The escarpment reaches elevations of between 200 m and 600 m in the area, forming a long linear topographical feature running in a South-West to North-East direction. Some other steep slopes are also found towards the South-West boundary of the wind farm.

Precipitation is low compared to other grassland regions (> 600 mm per year). Game, sheep and cattle farms present natural and semi-natural dams and water bodies, which are particularly concentrated towards the North-West boundary of the wind farm.

5.2.2 Field surveys

Field surveys followed the general guidelines suggested by Jenkins *et al.* (2012) for bird monitoring at wind farms in South Africa (later updated to Jenkins *et al.* 2015). Flight trajectories of soaring birds were recorded from five vantage points in 2011 and 2012. During the pre-construction surveys, the final layout of the turbines was not defined. For this reason, surveys spanned an area larger than the actual footprint of the wind farm, although the distribution of vantage points left some uncovered gaps within the study area. As a result, some of the actual turbine locations were not covered by any of the vantage points (figure 5.1). The observable area from the vantage points was defined as a circular region of 2 kilometre radius centred at each vantage point. A total of 48 hours of observation were conducted from each vantage point. To take into account seasonal

variation in bird activity, vantage points were visited in four different seasons, covering a full annual cycle. To capture variation in activity due to time of day, four observation sessions of three hours, were conducted at each vantage point, in each season. Therefore, a total of 16 observation sessions of three hours each, were conducted per vantage point, covering four different seasons and four different times of the day.

Systematic bird carcass searches were conducted in a sample of 32 turbines out of the 40 that form the wind farm, in 2014 and 2015 — first year after construction. Search areas were defined as squared plots of side 210 metres, centred at each turbine. Searches were structured in 35 parallel transects 6 metres apart. The search team consisted of two persons walking in parallel, each one on alternate transects. Turbines were searched periodically every one to two weeks. The eight turbines that were not selected for systematic searches were visited once per month and searched in a similar fashion by a single searcher with the objective of finding large bird carcasses, as these are easier to detect and are known to remain detectable for long periods of time (Barrientos *et al.*, 2018).

5.2.3 Study species: Verreaux's Eagle

I chose to study the collision risk of Verreaux's Eagle because it is a large raptor whose range largely overlaps with wind resources in South Africa. Only six wind farms shared their data publicly in the first report on the status of wind farm impacts on birds in South Africa (Ralston-Paton *et al.*, 2017; Perold *et al.*, 2020). A total of five Verreaux's Eagle fatalities have been reported in two of these wind farms. In the wind farm analysed in this study, four eagle collisions occurred in the first year after construction.

The Verreaux's Eagle is distributed throughout South and East Africa, and also in Chad, Sudan and the Central African Republic (BirdLife International, 2016). It inhabits rocky areas, favouring steep slopes (Murgatroyd *et al.*, 2016a), and preferring inaccessible cliff faces for nesting and roosting (Simmons, 2005). Rocky substrates are also attractive to the Verreaux's Eagle because they host colonies of rock hyraxes (genuses *Procavia* and *Heterohyrax*), which are its main prey. In absence of hyraxes the eagles adapt well to eating other small vertebrates (e.g. mole-rats, hares and tortoises — Simmons 2005; Murgatroyd *et al.* 2016b). They are also known to scavenge (Simmons, 2005).

Verraux's Eagles retain territories and concentrate their activity around their nests, year-round. Their excursions away from the nest are longest towards midday (Murgatroyd

et al., 2016a), although they often hunt during morning and evening (Simmons, 2005). Rock hyraxes are hunted mainly in low flights over rocky outcrops (Simmons, 2005).

The Verreaux's Eagle is, in general, considered to be of Least Concern throughout its range (BirdLife International, 2016). However, in South Africa it is considered locally vulnerable (Taylor *et al.*, 2015) due to its population declining in some areas. In 2017, BirdLife South Africa released a monograph on guidelines for impact assessment, monitoring, and mitigation focused exclusively on the Verreaux's Eagle, to encourage compatible wind farm management within Verreaux's Eagle ranges (Ralston-Paton, 2017).

5.2.4 Data analysis

I used flight trajectories of Verreaux's Eagles observed during the pre-construction phase of the wind farm to estimate the spatial distribution of collision risk for this species within the footprint of the project. Assuming that collision risk correlates with activity levels, I approached this problem in two different ways: i) analysing the spatial distribution of flight intensity using a kernel density estimator (see chapter 2), and ii) finding landscape features that might be attractive for the species, using a potential-based velocity model (see chapter 3). To make more accurate inference, I incorporate a model for the observer mapping error that aims at disentangling uncertainty due to this error from that regarding eagle behaviour (see chapter 4). All models used here are discussed at length in previous chapters, and therefore the explanations in the next sections are kept relatively concise.

Flight trajectories observed from vantage points were digitized using QGIS (QGIS Development Team, 2017). The resulting spatial objects were transferred into R for further analysis (R Core Team, 2019). Each trajectory was transformed into a sequence of locations, with a rate of one location every five seconds of flight. The duration of each flight was recorded in the field and flight speed was calculated dividing the length of the digitized trajectory by the flight duration. This provides an average flight speed for the whole flight, but it results in the loss of instantaneous speed information.

In the following sub-sections, I briefly explain the methodological details of the kernel analysis, potential-based model structure and also how I incorporate measurement error into the modelling process.

Kernel density estimation

Traditional kernel density estimators require data to be independent and identically distributed (Silverman, 1986b; Worton, 1989). However, sequences of locations along a flight trajectory are correlated in time. This means that the observation of one location restricts the observation of the next location, more so the closer they are in time, violating the independence assumption. There are two main kernel density estimators that have been developed for autocorrelated data, namely: Brownian bridges (Horne *et al.*, 2007, and later extensions) and the autocorrelated kernel density estimator (AKDE, (Fleming *et al.*, 2015)). Brownian bridges estimate a probabilistic distribution of activity between two observed locations (see Horne *et al.* (2007) and chapter 2). It is used to estimate unobserved activity over a period of time for which only a sample of observed locations are available. It is particularly useful when the time between observations is long, and therefore the uncertainty about where the process was between observations is large (e.g. the location of the animal between two GPS positions). When locations are acquired with high resolution (locations along trajectories are close in time), uncertainty about the location of the process between observations is small and the Brownian bridges technique loses its applicability.

The AKDE is a useful technique to estimate the long-term activity distribution of an animal around an activity centre based on observed movement trajectories (see Fleming *et al.* 2015 and chapter 2). Under this estimator, we assume the activity of the animal to be represented by a continuous-time stochastic process that is stationary - i.e. the variance of the process stabilizes after some time. AKDE requires that we observe the process for long enough to be able to estimate this long-term variance. It is unclear how AKDE behaves when estimating a semi-variogram from a set of independent, short trajectories that may or may not belong to the same bird. Estimating an AKDE from several

A third option, and the one I adopt here, is the path kernel proposed in chapter 2, selecting the bandwidth of the kernels using a leave-group-out cross-validation procedure. The path kernel integrates a kernel function along a continuously, or nearly continuously, observed trajectory, similar to those captured by vantage point observers, such that

$$\hat{f}(\mathbf{x}) = \frac{1}{m}\sum_{j=1}^{m}\left\{\frac{1}{L_j h^D}\int_{T_j}\prod_{d=1}^{D} K\left(\frac{x_d - X_{j,t,d}}{h}\right)\mathrm{d}t\right\}. \tag{5.1}$$

where d indexes the different dimensions — in the two-dimensional case we are concerned with $D = 2$. The index j represents each flight trajectory, L its length and T its duration. The term h is the bandwidth of smoothing parameter. The expression inside the summation term in 5.1 is referred to as the *path kernel* and has the same role as the traditional kernel functions for independent observations. Should m trajectories be observed from a vantage point, the density would be estimated by taking the average over the m path kernels. Effectively, each trajectory contributes as one observation towards the computation of the kernel density (see chapter 2 for details on this method).

To optimize the bandwidth h, I use a leave-group-out procedure (Roberts *et al.*, 2017). This procedure finds the optimal bandwidth by calculating a kernel density estimate leaving out one full trajectory at a time, computing a kernel estimate iterating different bandwidths and calculating the average likelihood of the locations along the unobserved trajectory for each bandwidth value. The bandwidth that maximizes the likelihood of unobserved flights is selected.

$$\max_h \text{LGOCV}(h) = \frac{1}{m}\sum_{j=1}^{m} \log\left(\frac{1}{L_j}\int_{T_j} \hat{f}_{-j}(\mathbf{X}_t^{(j)})\ \mathrm{d}t\right), \tag{5.2}$$

where $\mathbf{X}_t^{(j)}$ represents the locations along flight j and $\frac{1}{L_j}\int_{T_j} \hat{f}_{-j}(\mathbf{X}_t^{(j)})\ \mathrm{d}t$ the average density calculated for flight j from a density distribution estimated leaving out flight j.

Potential-based velocity models

I use a potential-based velocity model to explore the preferential heading of the eagles towards steep slopes and water bodies, which are the most prominent landscape features found in the study area. Previous research showed that steep, rocky slopes were used by Verreaux's Eagles for nesting and hunting (Simmons, 2005; Murgatroyd *et al.*, 2016a,b). I found no information regarding the effect of water bodies on the eagles, but I assumed that they could be attractive for their prey (and indirectly for the eagles), and possibly also as a source of water for the eagles.

A detailed description of the model is presented in chapter 3, and therefore, here I only explain the basic features. The model is formulated as follows:

$$\begin{aligned} \mathbf{X}_t &= \mathbf{X}_0 + \int_0^t \mathbf{V}_s \, \mathrm{d}s \\ \mathrm{d}\mathbf{V}_t &= \beta(\gamma(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t) \, \mathrm{d}t + \sigma \, \mathrm{d}\mathbf{B}_t. \end{aligned} \tag{5.3}$$

In expression 5.3, $\mathbf{X}_t$ is a stochastic process that represents the location of the bird at time t, whereas its movement velocity is represented by the stochastic process $\mathbf{V}_t$. The velocity $\mathbf{V}_t$ tends to relax to a preferred velocity given by the function $\gamma(\cdot)$, at a rate given by the parameter β. The larger β, the faster the velocity reverts to the preferred velocity. This preferred velocity may change in space, time or both. In this case, it depends on the relative position of the animal with respect to landscape features that attract or repel the bird, and also on the current travelling speed, such that:

$$\gamma(\mathbf{X}_t, \mathbf{V}_t) = -\nabla H |\mathbf{V}_t|, \tag{5.4}$$

where ∇H represents the gradient of the Euclidean distance function to points of attraction from location $\mathbf{X}_t$. The gradient of the Euclidean distance always has unit magnitude and direction away from the location distance is measured to. $|\mathbf{V}_t|$ is the travelling speed at time t. Therefore, the preferred velocity $\gamma(\cdot)$ has magnitude equal to the current travelling speed and direction towards the point of attraction (note the negative sign of the gradient). In other words, having a unitary gradient multiplied by the current speed results in the drift in velocity to affect its direction but not its magnitude (speed).

The term $\sigma \, \mathrm{d}\mathbf{B}_t$ represents a Brownian motion, which can be viewed as the continuous-time analogue of a random walk (e.g. Turchin, 1998). $\mathrm{d}\mathbf{B}_t$ is a stochastic process with distribution $\mathrm{N}(0, \mathrm{d}t)$, and the parameter σ scales the variance of the process.

Since there are two types of target locations (slopes and water bodies) the resulting drift in velocity is given by the sum of the actions of the two potentials such that,

$$\mathrm{d}\mathbf{V}_t = \sum_{k=1}^{2} \beta_k(\gamma_k(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t) \, \mathrm{d}t + \sigma \, \mathrm{d}\mathbf{B}_t, \tag{5.5}$$

where $k = 1$ stands for slopes and $k = 2$ stands for water bodies.

The drift parameters, β, operate at the trajectory level and may be different for different trajectories. To capture the variability in drift parameters between different trajectories, I add an extra layer to the model that captures the distribution of drift parameters at

the population level (i.e. population of trajectories). Thus,

$$\beta_j \sim \mathrm{N}(\mu_\beta, \sigma_\beta^2). \tag{5.6}$$

The hyper-parameter μ_β will ultimately define whether, on average (as opposed to any particular trajectory), the species tends to drift towards certain elements of the landscape. The hyper-parameter σ_β informs of how different drift is expected to be in different trajectories.

Models are firstly fitted using vague priors (table 5.1). However, due to the complexity of the model and the few vantage point observations available, I also used informative priors to investigate how the choice of priors affects the final inference. The informative prior distributions were extracted from the posterior parameter distributions of a potential-based velocity model fitted to a GPS telemetry data set of two Verreaux's Eagles (see chapter 3). Although these individuals live far away from the wind farm, it is assumed that birds of the same species should share similar habitat preferences. To account for the uncertainty with regards to the similarity between eagles living in different environments, the variance of the prior distributions was increased with respect to the original posteriors. There is no published information regarding preferences of Verreaux's Eagle with respect to water bodies, that I am aware of. As stated at the beginning of this section, I assume that there might be an attraction towards towards water bodies, and therefore, the drift towards these features was considered similar to that towards slopes, although with larger uncertainty. Details of the model fitted to the GPS data, and about the GPS data themselves, can be found in chapter 3 and in Murgatroyd *et al.* (2016a). The values of the prior parameter distributions used for fitting the models can be found in table 5.1.

Potential-based velocity models were fitted in a Bayesian framework using Hamiltonian Monte Carlo sampling in Stan through the R package `rstan` (Stan Development Team, 2018c). I ran four Markov chains of 4000 iterations each, with a warm-up phase of 2000 iterations. Convergence was assessed by visually inspecting that the distribution of the samples of all chains was similar, and also by the $\hat{R}$ diagnostic (Gelman *et al.*, 2014). This metric represents the expected reduction in the posterior variance, as the number of samples from the Markov chains goes to infinity. Values close to 1.0 indicate good convergence, while as a general rule of thumb, values under 1.1 are consider acceptable (Gelman *et al.*, 2014).

Table 5.1: Prior distributions used to fit the potential-based velocity model to the vantage point data. The two first columns show the mean and standard deviation (sd) of the informative priors obtained from previous telemetry studies of Verreaux's Eagle. The two last columns show the vague priors given when ignoring previous telemetry work on the species. In both cases, mean drift parameters were given a normal prior distribution. Variation parameters were given a half-Cauchy distribution. The diffusion parameter was also given a half-Cauchy distribution.

	Informative		Vague	
	mean	sd	mean	sd
Mean slope drift ($\mu_{\beta_{slp}}$)	0.067	0.212	0	1
Mean water body drift ($\mu_{\beta_{wb}}$)	0.067	0.312	0	1
Variation slope drift ($\sigma_{\beta_{slp}}$)	0.050	0.211	0	1
Variation water body drift ($\sigma_{\beta_{wb}}$)	0.050	0.311	0	1
Diffusion (σ)	2.879	1.116	0	10

Observer mapping error

Vantage point observers are likely to introduce mapping errors when they capture flight trajectories in the field. Assuming that this mapping error is also a stochastic process, different realizations of the mapping error would produce different mappings of the same bird trajectories. Consequently, drift and diffusion parameters of a potential-based velocity model fitted to the trajectories with different mapping errors would also be different. The objective then should be to find parameters that describe the general behaviour of the trajectories with different mapping errors, not only with the particular realisation of the mapping error that was observed. To this end, I simulated 100 realisations of the mapping error using the model described, and the parameters estimated in chapter 4. These realised errors were subtracted from the observed flight trajectories and the resulting trajectories consitute "alternative" flight trajectories arising from different realizations of the mapping error process. This is conceptually similar to the imputation distribution of missing data discussed in Hooten & Hobbs (2015). I fitted a potential-based velocity model, like the one just described in the previous section, to each of the data sets with different mapping error realizations separately, so that the parameters are always estimated for the original number of observations. However, the different realizations of the mapping error provide, given the error model, an approximation to the posterior distribution marginalized over error realizations and not conditional on just one realization of the mapping error. This will be useful for exploring how predictions differ between a model that accounts for the mapping error and one that ignores it, as we will see in the next section.

An approximation to the posterior parameter distribution conditional on the mapping error simulations is obtained by (see Hooten & Hobbs, 2015)

$$E(\theta|Y) \approx E_{S|Y}\left[E(\theta|S)\right]$$
$$\text{Var}(\theta|Y) \approx E_{S|Y}\left[\text{Var}(\theta|S)\right] + \text{Var}_{S|Y}\left[E(\theta|S)\right]$$

where θ represents the model parameters, Y is the observed data set and S is an imputed data set with a mapping error realization. Note that this approximation is only used to investigate how variable inference is across mapping error realizations and not to simulate trajectories. The simulation of trajectories is done by sampling from one of the posteriors fitted to a data set with a randomly-chosen error realization at a time (see next section on flight simulations). The main reason for doing this is that the variance parameters could become negative if estimated using equation 5.2.4.

In chapter 4, the mapping error is modelled as a stochastic process that evolves with time. The results indicate that this error oscillates throughout the flight around the initial error, made when positioning the bird on the map at the time of detection. The amplitude of these oscillations of the mapping error increases with time, and tends to reach a maximum where it stabilizes. The model for the mapping error was formulated as a continuous-time movement model such that:

$$\begin{aligned} \mathrm{d}\mathbf{V}_t &= \mathrm{d}\boldsymbol{\upsilon} + \ \mathrm{d}\mathbf{U}_t \\ \mathrm{d}\mathbf{D}_t &= \mathbf{U}_t \ \mathrm{d}t \\ \mathrm{d}\mathbf{U}_t &= -\beta(\mathbf{D}_t - \mathbf{D}_0) \ \mathrm{d}t - \phi\mathbf{U}_t \ \mathrm{d}t + \sigma \ \mathrm{d}\mathbf{B}_t \\ \mathrm{d}\mathbf{B}_t &\sim \mathrm{N}(\mathbf{0}, \mathrm{d}t\mathbf{I}), \end{aligned} \tag{5.7}$$

where $\boldsymbol{\upsilon}$ is the true velocity of the bird (which is unobserved), $\mathbf{V}_t$ is the observed velocity and $\mathbf{U}_t$ is the velocity at which the error changes at time t. $\mathbf{D}_t$ is the error in the location captured by an observer in the field at time t and $\mathbf{D}_0$ is the error at the initial time (when the bird is first detected, time $t = 0$). Therefore $\mathbf{D}_t - \mathbf{D}_0$ is the difference between the error at time t and the initial error. The initial error needs its own model, which I describe in chapter 4 and also briefly below.

The above stochastic differential equations describe the evolution of the mapping error over time. However, to estimate the magnitude of the error, we need to define a starting value for it. In chapter 4, I show that when a bird is first detected, the error observers

make when mapping its position can be modelled by:

$$\begin{aligned} \mathbf{D}_{0(jv)} &\sim \mathrm{N}(\mathbf{0}, \sigma^2_{0(jv)}\mathbf{I}) \\ log(\sigma_{0(jv)}) &= \theta_0 + \theta_1 \mathrm{VRM} + \theta_2 \mathrm{height} + \theta_3 \mathrm{dist} + \eta_{vp} \\ \eta_{vp} &\sim \mathrm{N}(0, \sigma_\eta) \end{aligned} \tag{5.8}$$

Here I model the log of the standard deviation of the two-dimensional error $\mathbf{D}_0$, which is centred at the $\mathbf{0} = (0, 0)$ vector and has diagonal covariance matrix $\sigma_0^2\mathbf{I}$. Therefore, the scale of the variance is the same in both Easting and Northing components. Note that being the error centred at zero, the standard deviation of the error may be indirectly interpreted as a measure of how large it may be.

The intercept e^{θ_0} gives the average error expected on a flat terrain at zero distance from the observer. The parameter θ_1 represents the effect of the Vector Ruggedness Measure (VRM), which provides an idea of how variable the topography of the landscape is. In chapter 4, I hypothesised that this variable could be influential for the mapping error. However, it did not have an important impact on the mapping error, and therefore was left out from the simulations in the present study. The parameters θ_2 and θ_3 control the change in initial error with increasing flight height and distance from the observer, respectively. In chapter 4, I was not able to estimate a random variation due to different observers since there were only three observers to analyse. Therefore, the effect of particular observers was ignored and I used the error expected for the average observer. The parameter η_{vp} adds some extra variation to account for uncontrolled variables associated with the different vantage points.

With a model for the initial error, I simulate 100 initial errors that serve as initial points for the simulation of 100 realisations of the mapping error using the dynamics described by equation 5.7. The simulated error trajectories are then subtracted from the observed flights and used as alternative (imputed) data sets.

Relationship between estimated activity and mortality distributions

Once the models are estimated, I compute the expected long-term activity distribution for the Verreaux's Eagle based on the potential-based velocity model fitted to the flight trajectories observed from vantage points in the pre-construction phase of the wind farm (augmented with different realisations of the mapping error). To do so, I simulate

Table 5.2: Parameters used for the simulation of initial mapping errors. Parameter values for the simulations were sampled from a Gaussian distribution with mean showed in the second column and standard deviation showed in the third column. The corresponding covariate and the notation defined in the text for the parameters are shown in the first column. σ_η is the standard deviation of the random effect for vantage point. The model estimates the log standard deviation of the location error. The parameters shown in the table are in the linear scale and therefore, a one unit increase in a covariate x changes the response by a factor of $e^{\theta x}$.

	mean	sd
Intercept (θ_0)	4.198	0.609
Height (θ_2)	-0.005	0.002
Distance (θ_3)	0.002	0.000
VP random effect (σ_η)	1.120	0.49

Table 5.3: Parameters used for the simulation of mapping error trajectories. Parameter values for the simulations were sampled from a Gaussian distribution with mean showed in the second column and standard deviation showed in the third column. The corresponding covariate and the notation defined in the text for the parameters are shown in the first column. The model considers error to drift towards the initial error at a rate given by β. The mean drift rate across trajectories is given by μ_β and the standard deviation of drift in different trajectories by σ_β. The rate of change of the error is damped at a rate proportional to ϕ. The mean damping across trajectories is given by μ_ϕ and the standard deviation of damping in different trajectories by σ_ϕ. Finally, random inputs disturb the error trajectories. The magnitude of these random inputs is given by a Brownian motion with standard deviation scale σ.

	mean	sd
Mean drift towards initial error (μ_β)	0.061	0.006
Mean damping (μ_ϕ)	0.012	0.001
Variation in drift towards initial error (σ_β)	0.049	0.005
Variation in damping (σ_ϕ)	0.004	0.002
Error diffusion (σ)	13.224	0.054

flight trajectories to investigate the long-term behaviour of the location process $\mathbf{X}_t$ by computing the relative frequency with which different regions within the study area are visited by the eagles. I simulated 5000 hours of flights, consisting of sequences of locations sampled at a rate of 0.1 minutes.

To simulate flights, I followed the following procedure:

1. I first sampled μ_{β_k}, σ_{β_k} and σ parameters from their respective posterior distributions. The subscript k represents steep slopes and water bodies,
2. I then simulate β_k from $\mathrm{N}(\mu_{\beta_k}, \sigma_{\beta_k})$,
3. Choose a random initial point for the flight within the study area, with uniform probabilities for easting and northing,
4. with β_k and σ, I simulate a sequence of locations spaced 0.1 minutes, until the trajectory leaves the study area,
5. then, repeat points 1-4 until the total flight time set is reached (i.e. 5000 hours).

To account for mapping error in the predictions, I added an extra step at the beginning of the simulation: the posterior distributions for μ_{β_k}, σ_{β_k} and σ are randomly sampled from one of the 100 imputed data sets.

The objective then is to investigate how the long-term activity distribution correlates with the observed distribution of Verreaux's Eagle fatalities recorded during the first year of wind farm operation. To summarize the relative frequency with which we would expect the different locations in our study area to be visited in the long-term, I overlaid a grid of square cells of 100 x 100 metres and counted the number of simulated points falling within each grid cell. To smooth the resulting grid of counts to some extent, each grid cell count was averaged over its eight immediate neighbours (resulting in an average of nine numbers including the target cell). I defined four categories of activity intensity in which grid cells of 100 x 100 metres could fall into. Activity intensity was ranked from lower to higher and divided in four equal quartiles. Grid cells in the lower quartile represent those areas that are used the least by the eagles, and therefore placing turbines here should be preferred. Grid cells within the second and third quartiles are considered of medium risk. These are the most common activity levels in the area. Whether turbines should be placed in these areas should be carefully analysed based on the total amount of eagle activity observed. Finally, grid cells in the highest quartile represent locations where eagle activity is particularly concentrated and turbine placement in these locations

should be avoided. Although categorizing activity intensity is useful for communication purposes, we must keep in mind that use intensity is a continuous variable and that the estimation of its precise value at the turbine locations is more informative than categorizing larger regions. If fatalities observed during the first year of operation are located in areas of high (or at least medium-high) activity intensity, it would be indicative that a spatially-explicit model of activity intensity based on pre-construction flight data could be useful for predicting where collisions are likely to occur.

To further investigate the predictive power of potential-based velocity models using different priors, I used the Area Under the Curve (AUC) of a Receiver Operating Characteristic (ROC) curve (Hanley & McNeil, 1982). The curve of a ROC describes how the rate of true positives (turbines where fatalities were recorded and were classified as likely to produce collisions) relates to the rate of false positives (turbines where fatalities were not recorded but were classified as likely to produce collisions) as we vary the classification threshold (the activity intensity value after which a turbine is classified as likely to produce collisions). The area under the ROC curve (AUC) is a number between 0 and 1. With an AUC of 0 a model only predicts false positives, with a value of 1 a model only produces true positives and with a value of 0.5 the rate of false positives equals the rate of true positives, and therefore, the method has no better predictive power than classifying by chance. Consequently, AUC values between 0.5 and 1 are desirable for effective predictions. The ROC analysis was performed with help from the R package `pROC` (Robin *et al.*, 2011).

5.3 Results

Of the total 40 turbines that comprise the final wind farm layout, 25 were within two kilometres from a vantage point and therefore, covered by the pre-construction surveys (figure 5.1). A total of 14 Verreaux's Eagle flight trajectories were observed from the vantage points during these surveys. Flights were on average 854 metres long, their average duration was 125 seconds and the average flight speed was 18.5 metres per second. In the first year of operation, there were four Verreaux's Eagle fatalities, and they were all found in the sample of turbines that were searched every one to two weeks. Three of the four fatalities occurred in the West corner of the wind farm and two happened one shortly after the other at the same turbine (figures 5.3 and 5.2). Only one of the turbines that registered some mortality during the first year of operation was

covered by any of the vantage points, and only one short Verreaux's Eagle flight was observed from it. The remaining two turbines that registered mortality were beyond the view of any vantage point.

Kernel density estimation of long-term Verreaux's Eagle activity distribution was only conducted around those vantage points that recorded more than three flight trajectories. Only two vantage points, those located in the South and South-West of the wind farm, met this threshold. These two vantage points allowed me to analyse the potential for collision of seven turbines. Most observed flight trajectories concentrate around the West facing slopes in the South-most group of turbines (figures 5.1 and 5.2). No fatalities were observed in any of the turbines for which kernel density was calculated.

Potential-based velocity models allowed me to extrapolate to areas beyond the view of vantage points, making use of habitat preferences for prediction (figure 5.3). A positive posterior mean drift towards steep slopes suggests preference of Verreaux's Eagle for this type of habitat (table 5.4). The mean drift towards water bodies was also positive suggesting attraction towards these features by the species (table 5.4). Posterior mean drifts represent expected drift values; however, posterior distributions of both, drift towards slopes and drift towards water bodies, span positive and negative values. Therefore, the drift of any particular flight can be positive (attraction) or negative (repulsion). Posterior simulations from the potential-based velocity model allowed me to estimate a long-term activity distribution for the Verreaux's Eagle based on the full posterior distribution of the behavioural parameters (figure 5.3).

When overlaid over the long-term distribution of eagle activity distribution estimated from pre-construction data, all fatalities in the first year of operation of the wind farm occurred at turbines considered to be in areas of either high or medium-high activity intensity (figures 5.3 and 5.4). The predictive power of the models was very similar regardless of whether vague or informative priors were used for model fitting (figure 5.3, 5.4 and 5.5). All models had AUC values around 0.73 (figure 5.5). When in addition to using informative priors, mapping error was explicitly modelled, the correlation between activity prediction and observed fatalities was only slightly better (figure 5.5). As an *ad-hoc* test, to understand whether vantage point data provided any local context to the predictions or using prior predictions only would suffice to achieve the observed predictive power, I simulated activity using the prior distributions in the same way I did with the posteriors. Although prior predictions also achieved AUC values close to 0.7, they performed the worst of all methods analysed at predicting the locations of fatalities

(figure 5.3, 5.4 and 5.5).

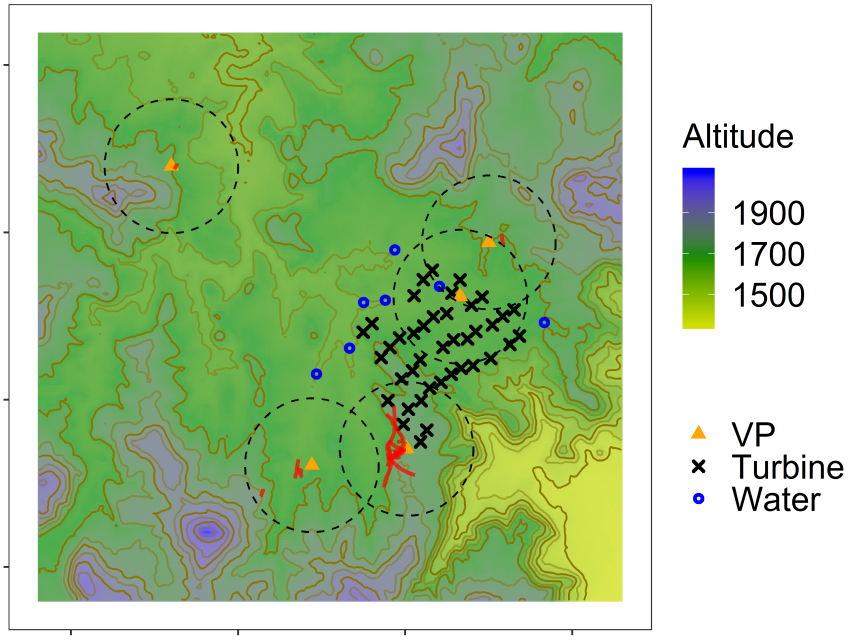


Figure 5.1: General topography of the study area and distribution of vantage points (orange triangles), wind turbines (black crosses), observed Verreaux's eagle's flights (red paths) and water bodies close to the wind farm (blue circles). A two kilometre buffer around the vantage points is represented by the dashed line circles and gives an estimation of the area observed from each vantage point. Contour lines correspond to levels of equal altitude and they are plotted every 50 meters.

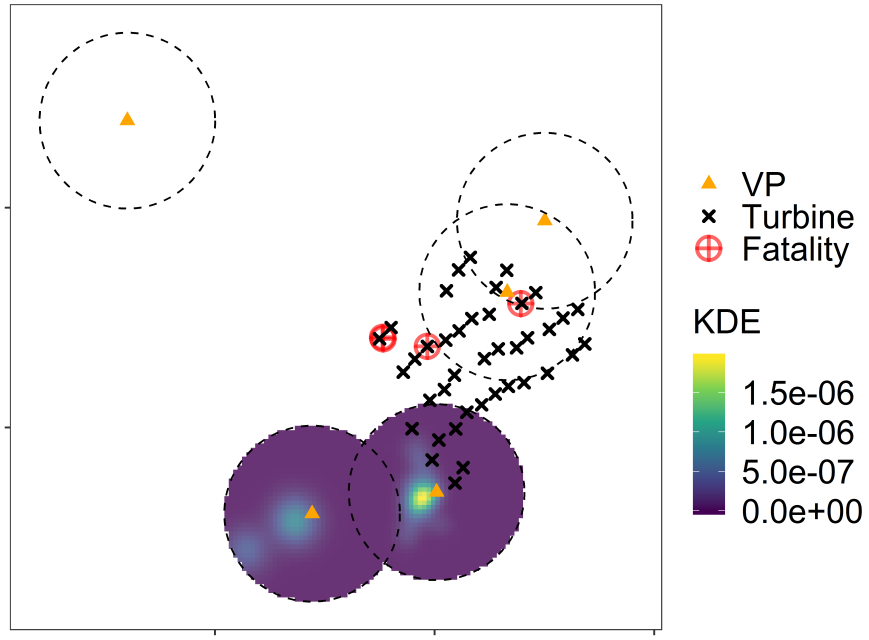


Figure 5.2: Long-term activity distribution based on kernel density estimates (blue to yellow gradient). Vantage points are represented by orange triangles. The areas observed from the vantage points are represented by the two kilometre buffer depicted by dashed-line circles. Only one of the wind turbines (black crosses) that recorded Verreaux's Eagle fatalities (represented by red icons) during the first year of operation fall within the view of a pre-construction vantage point. Kernel density estimates could only be computed for the two South-most vantage points, which were the only ones from which more than three flights were observed.

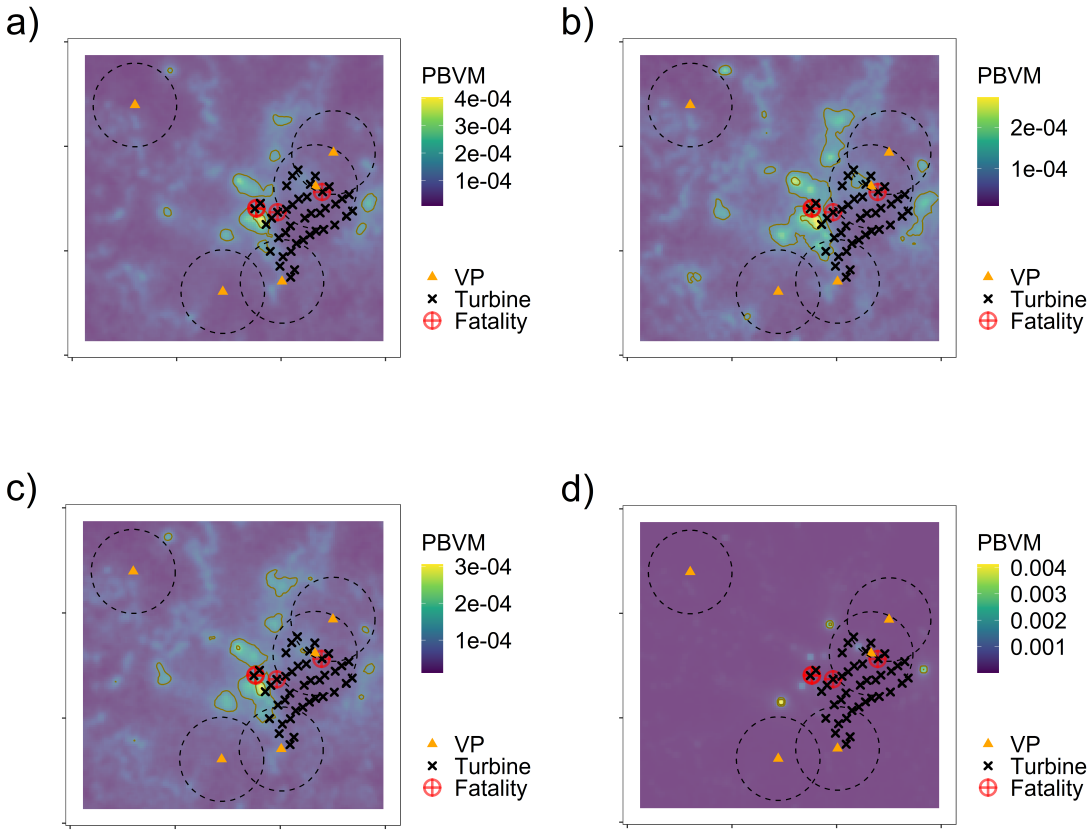


Figure 5.3: Long-term activity distribution based on simulations from potential-based velocity models (normalized counts of simulated locations per grid cell - PBVM, blue to yellow gradient). Vantage points are represented by orange triangles. Flights were simulated from: a) model fitted with vague priors; b) model fitted with informative priors; c) model fitted with informative prior to a set of mapping error realizations; d) predictions from informative priors. The areas observed from the vantage points are represented by the two kilometre buffer depicted by dashed-line circles. Red icons represent the fatalities observed during the first year of wind farm operation. The contours represent four equally spaced activity intensity levels.

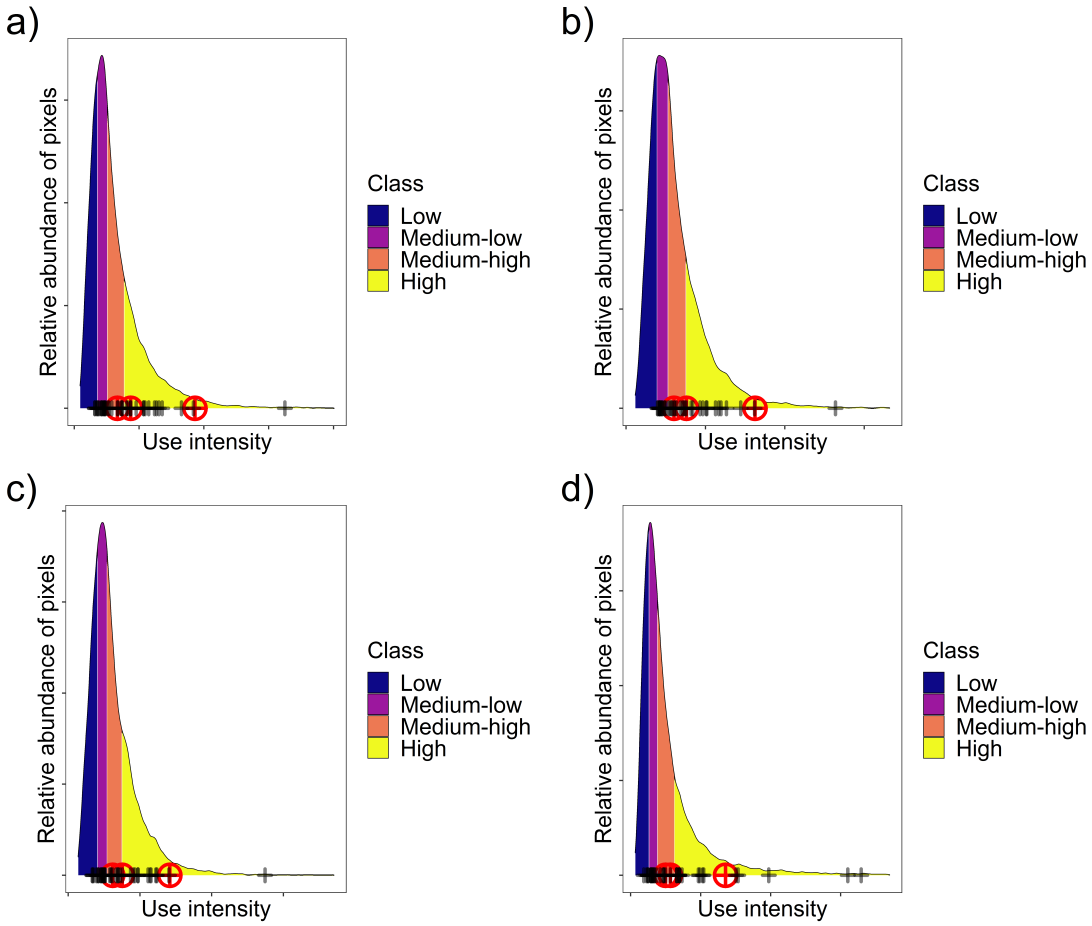


Figure 5.4: Relative abundance of grid cells classified in the different intensity classes of Verreaux's Eagle activity based on predictions by the potential-based velocity model. Flights were simulated from: a) model fitted with vague priors; b) model fitted with informative priors; c) model fitted with informative prior to a set of mapping error realizations; d) predictions from informative priors. The location of recorded eagle fatalities (red icons) in relation to the intensity classes is also plotted, as well as the locations of all other turbines (black crosses). Grid cells were classified into one of four categories: low, medium-low, medium-high and high. The same number of grid cells fall within each category.

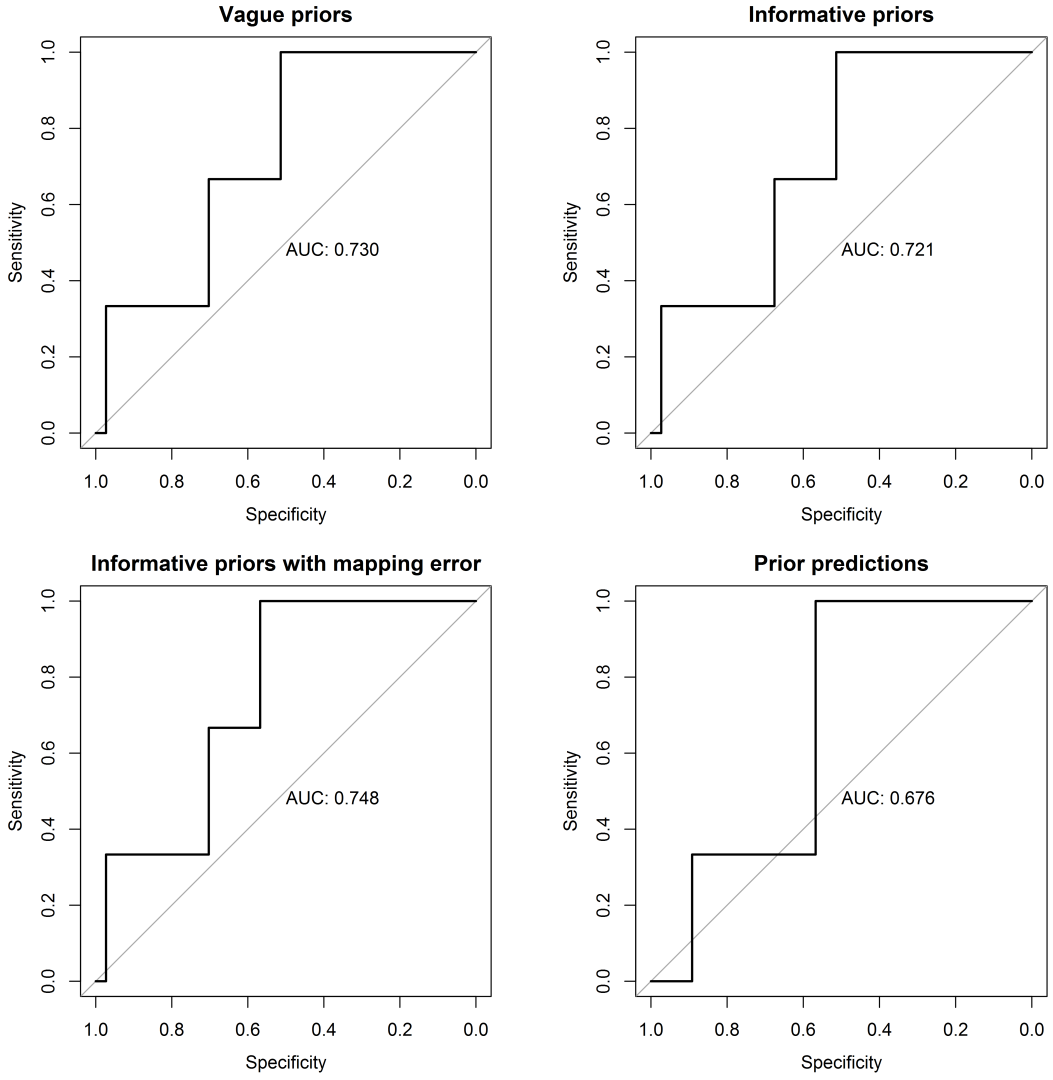


Figure 5.5: Receiver Operating Characteristic (ROC) curve representing the relationship between true positive rate (Sensitivity) and false positive rate (Specificity) when predicting fatalities based on activity intensity. Activity intensity was predicted using: models fitted with vague priors, models fitted with informative priors, models fitted with informative priors to a set of mapping error realization and finally, using only prior predictions. The area under the curve AUC was similar for all scenarios but using informative priors and accounting for mapping error was slightly better. Using only prior predictions performed the worst.

Table 5.4: Posterior distribution of population-level parameters of the potential-based velocity model fitted to the pre-contruction Verreaux's Eagle vantage point data. Posterior mean and standard deviation (in parenthesis). The column named *Inform.* correspond to values fitted to the raw data with informative priors. The column *Vague* presents the values fitted to the raw data using vague priors. The column *Obs. error* shows the values calculated based on the data augmented with 100 simulations of mapping error using informative priors. Effect estimated from vague priors are considerably larger than those estimated with informative priors. We can see that estimated parameters are very similar irrespective of whether simulated errors are included in the data when using informative priors.

	Inform.	Vague	Obs. error
Mean slope drift ($\mu_{\beta_{slp}}$)	0.017 (0.04)	0.069 (0.129)	0.011 (0.039)
Mean water body drift ($\mu_{\beta_{wb}}$)	0.027 (0.039)	0.089 (0.086)	0.024 (0.039)
Variation slope drift ($\sigma_{\beta_{slp}}$)	0.506 (0.145)	0.519 (0.154)	0.472 (0.23)
Variation water body drift ($\sigma_{\beta_{wb}}$)	0.239 (0.095)	0.232 (0.094)	0.189 (0.117)
Diffusion (σ)	4.305 (0.117)	4.313 (0.125)	4.33 (0.148)

5.4 Discussion

I estimated the relative collision risk of the Verreaux's Eagle at a wind farm in South Africa based on pre-construction vantage point data. The problem was approached by predicting where the eagles were likely to fly, based on flight trajectories observed from vantage points. To inform collision risk, I used two main pieces of information: i) the concentration of observed activity of the species in the different regions of the study area, and ii) the location and heading of the observed flight trajectories. Potential errors made by the observers when mapping the flights were explicitly modelled to separate this error from other sources of uncertainty, such as eagle behaviour. I then proceeded to validate the estimated distribution of collision risk based on pre-construction flight data against the actual fatalities observed during the first year of wind farm operation.

Firstly, the concentration of observed activity in different regions of the study area was estimated using kernel density estimation. While the method is appealing for its intuitive rationale, problems arise when the study area is not extensively covered by the surveys. In our wind farm, only two of the vantage points, covering seven turbines (out of 40), could be analysed using this method. Estimated kernel densities flagged certain regions, characterized by steep slopes to the west of the wind farm, as used more intensively than others. However, kernel density estimates do not take covariates into account and, therefore, cannot make use of these to predict outside of the observed areas (but see Millspaugh *et al.*, 2006). Furthermore, only one turbine where mortality was registered during the first year of monitoring was observed from a vantage point and, therefore, the location of the recorded fatalities could not have been predicted using kernel density estimates only.

Potential-based velocity models describe the mechanisms driving the heading of the flight trajectories and how it is influenced by landscape features. Once the mechanisms driving Verreaux's Eagle flights are understood, we may study the interaction of this species with, possibly unobserved, landscapes and the emerging activity patterns. These properties allowed me to predict activity intensity outside of the regions where flights were observed. In agreement with what kernel densities suggest, potential-based velocity models identified steep slopes as being used more often than other areas. Interestingly, in this study drift towards water features was estimated to be slightly larger than the drift towards steep slopes. These estimates support the original intuition of water features possibly being attractive to Verreaux's Eagles.

Augmenting the vantage point data set with simulated flights that incorporate different realizations of the mapping error allows us to investigate whether patterns in bird behaviour are consistent across error realizations. In this case study, parameters estimated using the set of error realization were similar to those estimated from the original data set. This suggests that behavioural patterns inferred from the data were not greatly affected by the mapping error introduced by the observers. Furthermore, predictions of eagle activity that were produced from models that incorporate mapping error had only slightly better predictive power than those that ignore it. It is important to note that the mapping error model considers erroneous trajectories to be centred at the true trajectory. Therefore, we should not expect large differences between the observed and simulated trajectories on average. In any case, the small reduction in the variation of drift parameters produced when accounting for mapping error had limited impact on the turbines that were flagged as presenting important risk for collision.

Only 14 Verreaux's Eagle flights were observed during pre-construction studies and most of them were recorded from a single vantage point. The limited amount of data and the relatively narrow range of conditions under which these flights were observed, challenged my confidence in the representativeness of the data and in the long-term activity predictions. For example, the long-term behaviour of the eagles might not be well represented by a single year of monitoring since no inter-annual variation is considered. In addition, the time gap between pre-construction studies and the first year of wind farm operation (ca. two years), disconnects the conditions under which flights and fatalities were recorded. Nevertheless, fatalities recorded during the first year of wind farm operation occurred in areas of high activity intensity identified by the predictions of potential-based velocity models. In fact, all model versions (i.e. with different type of priors and including, or not, mapping error) were successful at identifying the turbine that produced two collisions as likely to accumulate high levels of activity.

Conducting the analysis of the velocity in a Bayesian framework allowed me to use prior information about model parameters coming from pre-existing telemetry studies of the Verreaux's Eagle. Using informative priors takes some of the estimation workload away from the trajectories observed from the vantage points by constraining parameter estimates to be consistent with our knowledge about the ecology of the species. This regularization process is key to finding movement parameters that can be generalized beyond the conditions under which the study was conducted. Adding extra information to the analysis in data-poor situations is also important. There is a potential drawback of using priors obtained from models fitted to GPS data for the analysis of velocity

of trajectories observed from vantage points. Instantaneous speed usually cannot be recorded by observers in the field (at least not without supporting equipment, such as radar). Therefore, inference is limited to changes in movement direction, which may interfere in the comparison between a model fitted to GPS data (for which speed may be available at a higher resolution) and one fitted to vantage point data. However, the potential function formulation used in this study induces an effect on the direction of travel only, both for GPS and vantage point data, making them more comparable.

It came as a surprise that using informative priors to fit the movement model to the flight trajectories of the eagles, did not improve the correlation between the distribution of predicted activity intensity and the distribution of eagle fatalities. In fact, the long-term distribution of activity obtained by simulating flights from the priors directly (without fitting the model to data), had poor predictive accuracy compared to those simulated from models fitted to vantage point data. This speaks of the utility of vantage point data for providing local context to the general behavioural information contained in the priors. It is important to remember that the model fitted to the eagle's telemetry data was different to the one fitted to the vantage point data, and from this perspective modelling directly from the priors could lead to model misspecification. For example, we ignore the location of the nests of the eagles observed during the vantage points. However, we know from the model fitted to the GPS data that this is an important variable to consider. An additional discrepancy between models arises from nests typically being situated in steep slopes. Consequently, the effect estimated for steep slopes from the GPS data set is the direct effect of steep slopes, once the effect of nest has been accounted for (see Westreich & Greenland, 2013). However, in the vantage point data set the effect of nest is unaccounted for, and therefore the model estimates the total effect of steep slopes. This being said, the long-term distribution of flight activity simulated from models that used informative priors was very similar to that simulated from models that used vague priors. This serves as a prior sensitivity analysis and makes model predictions more compelling (see Gelman *et al.*, 2014).

My predictions of collision risk were solely based on the spatial distribution of eagle activity intensity levels, and I found a good correlation between the two, in contrast to de Lucas *et al.* (2008). This fact seems to support the idea that use intensity may be a good indicator for collision risk should the assessment be conducted for each turbine separately, rather than aggregated by wind farm or even vantage point (Ferrer *et al.*, 2012; de Lucas *et al.*, 2012; Heuck *et al.*, 2019). However, the spatial distribution of bird activity does not address the question of how many birds are likely to be killed by

the turbines, but only where they are most likely to be killed. A logical continuation for this research is to first estimate bird abundance in the general area, and then find the expected distribution of the activity using the methods described in this study. This approach would allow us to estimate the expected number of collisions at each turbine. It would also provide additional evidence to investigate whether bird abundance correlates with number of fatalities, once the spatial distribution of bird activity is taken into account. Further research on the use of potential-based velocity models could also be naturally extended to study avoidance of wind turbine positions by soaring birds.

In this case study, I have presented the collision risk assessment of a single wind farm, and I acknowledge that further validation is needed. However, my results suggest that: i) kernel density estimation is recommended for exploratory purposes, and ii) potential-based velocity models are promising tools to identify flight behaviours and habitat preferences that may put birds at risk of collision with wind turbine blades. However, vantage point surveys usually produce a limited amount of data and the drawing of flight trajectories is subject to observer interpretation. Based on the results presented here inference is relatively robust to the use of vague priors or ignoring mapping error; however, until further validation is conducted, I encourage the use of multiple sources of information during collision risk assessment, and avoid relying solely on vantage point observations. Bayesian analysis accommodates naturally and in a tractable way both the inclusion of different data sources and the incorporation of mapping error into the models. In the case analysed here, flight trajectories observed during pre-construction studies, and a spatially-explicit statistical analysis that accounts for known sources of uncertainty along the modelling process, identified the most problematic turbine locations in terms of collision risk. Furthermore, the fact that long-term activity predicted from vantage point data correlated better with observed fatalities than long-term activity predicted directly from the priors suggest that vantage point data offer some local context that priors obtained from other studies could miss.

Chapter 6

Conclusions

In this thesis, I develop an analytical framework to make statistical inference from vantage point surveys. The overarching objective is to understand where in the landscape soaring birds are more likely to fly, in order to optimize wind turbine placement so as to minimize collision risk.

Vantage point surveys are regularly used to predict collision risk of raptors with wind turbines, and to determine where collisions are likely to occur (Strickland *et al.*, 2011; Scottish Natural Heritage, 2014; Jenkins *et al.*, 2015). Traditionally, vantage point surveys are used to capture data on the abundance of flights over the wind farm, vantage point or turbine row (see Smallwood *et al.*, 2009). However, flight abundance does not always correlate with number of fatalities (de Lucas *et al.*, 2008; Ferrer *et al.*, 2012). Apparently, collisions are concentrated around certain turbines (de Lucas *et al.*, 2012), and risk situations are associated with particular environmental conditions (Barrios & Rodríguez, 2004; Smallwood *et al.*, 2009). To associate bird activity with certain habitats, or define the main fly ways in the area, observers map the bird flights they detect during vantage point surveys. However, these flight trajectories are usually reduced to some metric during the analysis of collision risk, such as closer distance to a turbine, or whether the bird visited certain habitat or not (see Barrios & Rodríguez, 2004; Smallwood *et al.*, 2009; Johnston *et al.*, 2014). This simplification results in the loss of valuable information.

In this thesis, I approach the problem estimating collision risk from an animal movement perspective, analysing the full trajectories mapped by the observers. Movement ecology offers an approach to studying wildlife that shifts the classical focus on population numbers to population redistribution (Nathan *et al.*, 2008). The information contained in the mapped flight trajectories about prevalent flight patterns helps us understand bird behaviour and predict the distribution of future activity.

In chapter 2, I estimate the long-term distribution of bird activity, using kernel density estimation. I develop a kernel density estimator that is designed to find the density of trajectories, rather than of independent locations - a requirement of classical kernel density estimators. I show that the estimator of trajectory density outperform methods that require independent locations at estimating the underlying distribution underlying simulated flights. However, as shown in the case study presented in chapter 5, the technique is not well suited to extrapolate to areas unobserved from vantage points. Therefore, its application leans towards the exploratory analysis and it works well to summarise potentially cluttered flight data. It provides a good alternative to methods that impose a grid over the study area and calculate abundance of flights at each grid cell (e.g. see Walker *et al.*, 2005). The main advantages of kernel densities over grid methods are: i) that there is no need to define the size of the grid cells arbitrarily, ii) kernel density provides a smooth density distribution, while grid-based densities may change sharply in contiguous cells, and iii) the kernel density estimator for trajectories takes into account the temporal dependence in the data.

Potential-based velocity models provide an alternative framework to understand habitat preferences of birds based on the systematic drift of flight trajectories in relation to landscape features. These models differ from kernel density estimators in that they describe the mechanisms driving the flight trajectories. In this thesis, I show how potential-based velocity models can be used to learn about movement behaviour of the Verraux's Eagle from vantage point observations. Movement parameters can then be used to estimate the long-term distribution of eagle activity by simulation. The spatial correlation between bird activity predictions calculated from pre-construction vantage point data using a potential-based velocity model, with fatalities observed during the first year of operation of a wind farm, showed good potential.

Conducting the analysis in a Bayesian framework, I incorporate existing knowledge about eagle behaviour. This information was obtained from telemetry studies conducted independently from the monitoring of the wind farm, and it was included as prior information in the models fitted to vantage point observations. One of the main advantages of this

approach is that information available on the species ecology has an explicit role in the inference drawn from the models. Another advantage is that part of the estimation workload is taken away from the (potentially limited) vantage point data available. Moreover, the simulation study conducted in chapter 3 suggests that model parameters may be substantially biased if vague priors are used to fit potential-based velocity models to vantage point observations; even in data-rich situations relative to what is expected from vantage point surveys. In the case study analysed in chapter 5, the effect of using vague priors in models that predict long-term activity from vantage point data, was not as important as expected, based on the simulations conducted in the previous chapter. However, because I have only analysed a single case study, it is still recommended to choose priors carefully when using potential-based velocity models for the analysis of vantage point surveys, and to conduct prior sensitivity analysis on the models. Furthermore, it is advised to conduct tracking studies using technologies that capture larger volumes of data - GPS for example - to better understand movement behaviour for those species for which this type of information might not be yet available. Even if the tracked individuals use an area different from the wind farm, general information about the movement behaviour of the species will be useful to make combined inference with vantage point observations.

Another common critique with regards to vantage point surveys is that observers may introduce errors when mapping the flight trajectories they observe. To my knowledge there are no published studies that investigate the mapping error made by vantage point observers, with the exception of errors in flight height estimation (Borkenhagen *et al.*, 2018). In chapter 4 I test several models to describe the magnitude and dynamics of the mapping error. I compare the flight trajectories traced by a drone (as a proxy for a bird) recorded by its on-board GPS with the same flights mapped by observers from vantage points. The model that best fitted the error (measured as the difference between GPS and visual trajectories) describes a damped harmonic motion in which the error oscillates around the error in location made when the drone was first detected. Under this model, the error is a stochastic process, with an expectation of zero, although its magnitude depends on the distance to the flying object, its height, the observer and characteristics of the landscape. A model formulation to estimate this error is provided in chapter 4.

After the flying object (e.g. bird) is first detected and visual tracking starts, the error evolves and may become larger over time. The standard deviation at the beginning of the flight may be in the vicinity of 150 meters, while after five minutes it may be as large as 500 meters, in any direction. The error oscillates around the initial error value, and

therefore, the shape of the flight might also be distorted. Thus, it is advised to avoid using raw observations to decide which turbine locations could be most problematic. Instead, it is recommended to take into account the properties of the error, and based on the observed flights, make a probabilistic assessment of where the true flights might have occurred. Furthermore, it is recommended to incorporate a model for the mapping error in any modelling of the movement characteristics of the species of interest. In the case study presented in chapter 5, I show how predictions of the distribution of fatalities during wind farm operation based on a potential-based velocity model fitted to Verreaux's Eagle flights observed during vantage point pre-construction surveys, improved substantially when accounting for the mapping error. This improvement was modest, perhaps because according to the mapping error model, different realizations of this error result in trajectories that are, on average, similar to the ones observed in the field. However, I believe that the magnitude of the error provided in this thesis is over-estimated, due to the experiments using a drone instead of a bird. The observers mapping the drone, which where all experienced in mapping bird flights from vantage points, pointed out that the device was more difficult to track than birds. Although the estimates provided in this thesis serve as a first approximation to the mappinge error (that possibly errs on the conservative side), over-estimation of the error may result on inflated uncertainty in movement parameters, when they are estimated from a set of mapping error realizations, as done in chapter 5. Less uncertain parameters would potentially lead to stronger patterns in the estimated activity intensity surface. Therefore, with higher parameter uncertainty it would be more difficult to separate turbines by the activity intensity they are expected to accumulate.

The results presented here show the value of considering the visual mapping of flight trajectories from vantage point as an alternative form of tracking. As such, techniques used within the animal movement framework have proved useful for the prediction of collision risk. I envisage that exploring the suite of techniques available for the analysis of animal movement further, may uncover new approaches for the analysis of vantage point observations. With a specific analysis in mind it is easier to design an appropriate protocol for data capture. In this sense, it is critical to adapt data capture during vantage point surveys to produce data that are amenable to be analysed within the animal movement paradigm (Nathan *et al.*, 2008). This and other aspects that deserve further study are discussed in the next section.

6.1 Future research

An issue that has not been tackled in this thesis is that of flight height estimation. Collision risk with wind turbines is considered to be intimately influenced by bird flight height (Schuster *et al.*, 2015). Bird collisions only occur when the birds fly at the height of the turbine rotor, which typically spans circular areas of 50 meters to 70 meters in radius. The typical clearance from the ground to the lower tip of the rotor is between 20 meters and 50 meters, and the maximum height up to 190 meters. These characteristics change for different turbine models, so these are just approximate numbers. Raptors may fly higher than 200 meters, and some species spend most of the time moving under 50 meters (e.g. Katzner *et al.*, 2012; Schaub *et al.*, 2019). Research suggests that collisions with wind turbines may occur when updraughts are too weak for the birds to gain height at a rate that allows them to fly over the rotor swept area (Barrios & Rodríguez, 2004). Therefore, identifying areas and times where birds tend to fly at the height of the rotor may help us make more precise predictions of collision risk.

Incorporating a third dimension into the models proposed in this thesis would not entail profound changes in their structure. I believe the limiting factor is to find a suitable way to capture height information and two-dimensional position simultaneously from the vantage points. I would suggest capturing measures of bearing, angle and distance to the bird with respect to the observer with some frequency, instead of plotting the flights on a map. In this way, we obtain a three-dimensional sequence of locations from which the trajectory can be recovered. Having specific target measures allows a more objective representation of the trajectories. We may also work towards improving specific measurements, for example using or developing different technology. The frequency with which these measurements should be taken needs further study. However, I consider one location every 30 seconds to be a good initial guess. The GPS data sets analysed in this thesis contained locations acquired every two-three minutes and still provided reasonable priors for models fitted to vantage point data. One location every 30 seconds is an even higher definition, and therefore, I believe this frequency should be suitable. However, vantage point data sets are typically small compared to GPS data sets and the flights mapped are relatively short (the flights used in chapter 5 were two minutes long on average). Therefore, acquiring locations at a rate lower than 30 seconds might result in overly simplified trajectories. The optimal frequency for data capture has not been investigated, and remains an open question. Until there is further detail on the subject, I suggests capturing trajectories with as much detail as possible during field work.

Avoidance of wind turbines by birds is another active research area (May, 2015). Birds might take evasive actions when approaching a turbine, and collision risk models typically incorporate a correction factor to account for these, when translating bird density numbers into predicted fatalities (Masden & Cook, 2016). This is what is called meso-avoidance (May, 2015). Micro-avoidance refers to last-second reactions to avoid collision (May, 2015). The term macro-avoidance is used to denote the avoidance of wind turbines or wind farms altogether (May, 2015). Despite reducing fatalities by collision, this effect represents an impact in itself. Birds might perceive areas altered by the installation of wind farms unattractive in terms of habitat quality. This effective habitat loss may produce the displacement of bird communities away from wind energy facilities.

Avoidance rates are usually calculated by comparing flight abundance estimated during the pre-construction phase of the projects with the abundance after construction (Cook *et al.*, 2018). However, as mentioned earlier, abundance might not correlate well with collision risk. A combined figure for micro- and meso-avoidance is, on occasions, calculated as the difference between observed and predicted mortality (Cook *et al.*, 2018). However, this number contains, in addition to avoidance, any misspecification of the collision risk model, as well as imprecisions in other inputs to the model (Band, 2012). The potential-based velocity models proposed in this thesis would naturally estimate drift of flight trajectories in relation to turbines. Studying reactions towards turbines we may explicitly estimate avoidance of these infrastructure at various scales. We may further account for this behaviour in the computation of long-term activity distribution, and hence in improve collision risk estimates.

The models presented in this thesis are not only applicable to the analysis of vantage point observations. Tracking technology is developing fast (Kays *et al.*, 2015), and the trajectories captured by both bird-borne devices, as well as by radar, are perfectly amenable to analysis using potential-based velocity models. In particular, avian radar now provides an excellent alternative to vantage point surveys, although at an extra cost (Katzner *et al.*, 2016a; May *et al.*, 2017; Jenkins *et al.*, 2018). Radar data provides bird locations at a high rate (few seconds) with a negligible mapping error compared to visually mapped trajectories, and it can monitor continuously for long periods of time capturing large volumes of data. Nevertheless, radar presents it own challenges in addition to its costs, such as not being able to associate species with recorded tracks. However, this technology is advancing fast and the application of the concepts and models presented in this thesis to radar data is an avenue worth exploring.

References

Arlot, S. & Celisse, A. (2010) A survey of cross-validation procedures for model selection. *Statistics Surveys* **4**, 40–79.

Arnett, E., Inkley, D., Johnson, D., Larkin, R., Manes, S., Manville, A., Mason, R., Morrison, M., Strickland, M. & Thresher, R. (2007) *Impacts of Wind Energy Facilities on Wildlife and Wildlife Habitat*. No. September in Impacts of Wind Energy on Wildlife and Wildlife Habitat, The Wildlife Society, Bethesda, Maryland.

Arnett, E.B. & May, R.F. (2016) Mitigating wind energy impacts on wildlife: Approaches for multiple taxa. *Human-Wildlife Interactions* **10**, 28–41.

Aschwanden, J., Stark, H., Peter, D., Steuri, T., Schmid, B. & Liechti, F. (2018) Bird collisions at wind turbines in a mountainous area related to bird movement intensities measured by radar. *Biological Conservation* **220**, 228–236.

Avgar, T., Potts, J.R., Lewis, M.A. & Boyce, M.S. (2016) Integrated step selection analysis: bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution* **7**, 619–630.

Baddeley, A., Rubak, E. & Turner, R. (2015) *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, London.

Band, B. (2012) Using a collision risk model to assess bird collision risks for offshore windfarms **02**, 62.

Barrientos, R., Martins, R.C., Ascensão, F., D'Amico, M., Moreira, F. & Borda-de Água, L. (2018) A review of searcher efficiency and carcass persistence in infrastructure-driven mortality assessment studies. *Biological Conservation* **222**, 146–153.

Barrios, L. & Rodríguez, A. (2004) Behavioural and environmental correlates of soaring-bird mortality at on-shore wind turbines. *Journal of Applied Ecology* **41**, 72–81.

Barton, P.S., Lentini, P.E., Alacs, E., Bau, S., Buckley, Y.M., Burns, E.L., Driscoll, D.A., Guja, L.K., Kujala, H., Lahoz-Monfort, J.J., Mortelliti, A., Nathan, R., Rowe, R. & Smith, A.L. (2015) Guidelines for Using Movement Science to Inform Biodiversity Policy. *Environmental Management* **56**, 791–801.

Benhamou, S. (2011) Dynamic approach to space and habitat use based on biased random bridges. *PloS one* **6**, e14592.

Benhamou, S. & Cornélis, D. (2010) Incorporating movement behavior and barriers to improve kernel home range space use estimates. *Journal of Wildlife Management* **74**, 1353–1360.

Bevanger, K., Berntsen, F., Clausen, S., Dahl, E., Flagstad, Ø., Follestad, A., Halley, D., Hanssen, F., Johnsen, L., Kvaløy, P., Lund-Hoel, P., May, R., Nygård, T., Pedersen, H., Reitan, O., Røskaft, E., Steinheim, Y., Stokke, B. & Vang, R. (2010) Pre- and post-construction studies of conflicts between birds and wind turbines in coastal Norway (BirdWind). Report on findings 2007-2010. Tech. rep., Norwegian Institute for Nature Research, Trondheim, Norway.

Bibby, C., Burgess, N., Hill, D. & Mustoe, S. (2000) *Bird Census Techniques*. Academic Press, second edn.

Biewener, A.A. & Patek, S.N. (2018) *Animal locomotion*. Oxford University Press, Oxford, UK, second edn.

BirdLife International (2016) Aquila verreauxii. The IUCN Red List of Threatened Species 2016: e.T22696067A95221980.

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

Borkenhagen, K., Corman, A.M. & Garthe, S. (2018) Estimating flight heights of seabirds using optical rangefinders and GPS data loggers: a methodological comparison. *Marine Biology* **165**, 1–7.

Bouten, W., Baaij, E.W., Shamoun-Baranes, J. & Camphuysen, K.C. (2013) A flexible GPS tracking system for studying bird behaviour at multiple scales. *Journal of Ornithology* **154**, 571–580.

Bowman, A.W.. (1984) An Alternative Method of Cross-Validation for the Smoothing of Density Estimates. *Biometrika* **71**, 353–360.

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

Cabrera-Cruz, S.A., Villegas-Patraca, R. & Thompson, D. (2016) Response of migrating raptors to an increasing number of wind farms. *Journal of Applied Ecology* **53**, 1667–1675.

Calenge, C. (2006) The package adehabitat for the r software: tool for the analysis of space and habitat use by animals. *Ecological Modelling* **197**, 1035.

Calenge, C., Dray, S. & Royer-Carenzi, M. (2009) The concept of animals' trajectories from a data analysis perspective. *Ecological Informatics* **4**, 34–41.

Carrete, M., Sánchez-Zapata, J.A., Benítez, J.R., Lobón, M. & Donázar, J.A. (2009) Large scale risk-assessment of wind-farms on population viability of a globally endangered long-lived raptor. *Biological Conservation* **142**, 2954–2961.

Carrete, M., Sánchez-Zapata, J.A., Benítez, J.R., Lobón, M., Montoya, F. & Donázar, J.A. (2012) Mortality at wind-farms is positively related to large-scale distribution and aggregation in griffon vultures. *Biological Conservation* **145**, 102–108.

Clemons, T.E. & Bradley, E.L. (2000) Nonparametric measure of the overlapping coefficient. *Computational Statistics and Data Analysis* **34**, 51–61.

Cook, A.S., Humphreys, E.M., Bennet, F., Masden, E.A. & Burton, N.H. (2018) Quantifying avian avoidance of offshore wind turbines: Current evidence and key knowledge gaps. *Marine Environmental Research* **140**, 278–288.

Cutting, J.E. & Vishton, P.M. (1995) Perceiving Layout and Knowing Distances.

Dahl, E.L., May, R., Hoel, P.L., Bevanger, K., Pedersen, H.C., Røskaft, E. & Stokke, B.G. (2013) White-tailed eagles (Haliaeetus albicilla) at the Smøla wind-power plant, central Norway, lack behavioral flight responses to wind turbines. *Wildlife Society Bulletin* **37**, 66–74.

Dai, K., Bergot, A., Liang, C., Xiang, W.N. & Huang, Z. (2015) Environmental issues associated with wind energy – A review. *Renewable Energy* **75**, 911–921.

de Lucas, M., Ferrer, M. & Janss, G.F.E. (2012) Using Wind Tunnels to Predict Bird Mortality in Wind Farms: The Case of Griffon Vultures. *PLoS ONE* **7**, 1–7.

de Lucas, M., Janss, G.F., Whitfield, D. & Ferrer, M. (2008) Collision fatality of raptors in wind farms does not depend on raptor abundance. *Journal of Applied Ecology* **45**, 1695–1703.

Desholm, M. (2009) Avian sensitivity to mortality: prioritising migratory bird species for assessment at proposed wind farms. *Journal of environmental management* **90**, 2672–9.

Dokter, A.M., Baptist, M.J., Ens, B.J., Krijgsveld, K.L. & van Loon, E.E. (2013) Bird Radar Validation in the Field by Time-Referencing Line-Transect Surveys. *PLoS ONE* **8**.

Douglas, D.J.T., Follestad, A., Langston, R.H.W. & Pearce-Higgins, J.W. (2012) Modelled sensitivity of avian collision rate at wind turbines varies with number of hours of flight activity input data. *Ibis* **154**, 858–861.

Drewitt, A.L. & Langston, R.H.W. (2006) Assessing the impacts of wind farms on birds. *Ibis* **148**, 29–42.

Duerr, A.E., Miller, T.A., Lanzone, M., Brandes, D., Cooper, J., O'Malley, K., Maisonneuve, C., Tremblay, J.A. & Katzner, T. (2015) Flight response of slope-soaring birds to seasonal variation in thermal generation. *Functional Ecology* **29**, 779–790.

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

Duong, T. (2007) ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. *Journal of Statistical Software* **21**, 1–16.

Duong, T. (2019) *ks: Kernel Smoothing*. R package version 1.11.5.

Eckstein, M.P. (2011) Visual search: A retrospective. *Journal of Vision* **11**, 14–14.

Efron, B. (1979) Bootstrap Methods : Another Look at the Jackknife. *The Annals of Statistics* **7**, 1–26.

Ellison, A.M. (2004) Bayesian inference in ecology. *Ecology Letters* **7**, 509–520.

Erickson, W.P., Johnson, G.D., Strickland, M.D., David P. Young, J., Sernka, K.J. & Good, R.E. (2001) Avian Collisions with Wind Turbines: A Summary of Existing

Studies and Comparisons to Other Sources of Avian Collision Mortality in the United States. Tech. rep., West Inc.,/NWCC, Washintong, D.C.

Erkelens, C.J. (2015) The perspective structure of visual space. *i-Perception* **6**, 1–13.

Ferrer, M., de Lucas, M., Janss, G.F.E., Casado, E., Muñoz, A.R., Bechard, M.J. & Calabuig, C.P. (2012) Weak relationship between risk assessment studies and recorded mortality in wind farms. *Journal of Applied Ecology* **49**, 38–46.

Fieberg, J., Matthiopoulos, J., Hebblewhite, M., Boyce, M.S. & Frair, J.L. (2010) Correlation and studies of habitat selection: problem, red herring or opportunity? *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 2233–2244.

Fleming, C., Fagan, W., Mueller, T., Olson, K., Leimgruber, P. & Calabrese, J. (2015) Rigorous home range estimation with movement data: a new autocorrelated kernel density estimator. *Ecology* **96**, 1182–1188.

Fleming, C.H., Sheldon, D., Fagan, W.F., Leimgruber, P., Mueller, T., Nandintsetseg, D., Noonan, M.J., Olson, K.A., Setyawan, E., Sianipar, A. & Calabrese, J.M. (2018) Correcting for missing and irregular data in home-range estimation. *Ecological Applications* **28**, 1003–1010.

Fortin, D., Beyer, H.L., Boyce, M.S., Smith, D.W., Duchesne, T. & Mao, J.S. (2005) Wolves influence elk movements: Behavior shapes a trophic cascade in Yellowstone National Park. *Ecology* **86**, 1320–1330.

Garvin, J.C., Jennelle, C.S., Drake, D. & Grodsky, S.M. (2011) Response of raptors to a windfarm. *Journal of Applied Ecology* **48**, 199–209.

Gasparatos, A., Doll, C.N.H., Esteban, M., Ahmed, A. & Olang, T.A. (2017) Renewable energy and biodiversity: Implications for transitioning to a Green Economy. *Renewable and Sustainable Energy Reviews* **70**, 161–184.

Geisler, W.S. (2008) Visual Perception and the Statistical Properties of Natural Scenes. *Annual Review of Psychology* **59**, 167–192.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A. & Rubin, D. (2014) *Bayesian Data Analysis*. CRC Press, Taylor and Francis Group, Boca Raton, FL, third edn.

Gitzen, R.A., Millspaugh, J.J. & Kernohan, B.J. (2006) Bandwidth Selection for Fixed-Kernel Analysis of Animal Utilization Distributions. *Journal of Wildlife Management* **70**, 1334–1344.

Green, R.E., Langston, R.H.W., McCluskie, A., Sutherland, R., Wilson, J.D. & Votier, S. (2016) Lack of sound science in assessing wind farm impacts on seabirds. *Journal of Applied Ecology* **53**, 1635–1641.

Gurarie, E., Fleming, C.H., Fagan, W.F., Laidre, K.L., Hernández-Pliego, J. & Ovaskainen, O. (2017) Correlated velocity models as a fundamental unit of animal movement: synthesis and applications. *Movement Ecology* **5**, 13.

Gürbüz, S.Z., Reynolds, D.R., Koistinen, J., Liechti, F., Leijnse, H., Shamoun-Baranes, J., Dokter, A.M., Kelly, J. & Chapman, J.W. (2015) Exploring the skies: Technological challenges in radar aeroecology. *IEEE National Radar Conference - Proceedings*, vol. 2015-June, pp. 817–822.

Hanley, J.A. & McNeil, B.J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29–36.

Hardey, J., Crick, H., Wernham, C., Riley, H., Etheridge, B. & Thompson, D. (2009) *Raptors: A Field Guide for Surveys and Monitoring*. The Stationary Office Limited, Edinburgh, second edi edn.

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

Hastie, T., Tibshirani, R. & Friedman, J. (2009) *The Elements of Statistical Learning*, vol. 1. Springer, second edi edn.

Hebblewhite, M. & Haydon, D.T. (2010) Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences* **365**, 2303–2312.

Heidenreich, N.B., Schindler, A. & Sperlich, S. (2013) Bandwidth selection for kernel density estimation: A review of fully automatic selectors. *AStA Advances in Statistical Analysis* **97**, 403–433.

Heuck, C., Herrmann, C., Levers, C., Leitão, P.J., Krone, O., Brandl, R. & Albrecht, J. (2019) Wind turbines in high quality habitat cause disproportionate increases in collision mortality of the white-tailed eagle. *Biological Conservation* **236**, 44–51.

Hodos, W. (2003) Minimization of Motion Smear : Reducing Avian Collisions with Wind Turbines.

Hooten, M.B. & Hobbs, N.T. (2015) A guide to Bayesian model selection for ecologists. *Ecological Monographs* **85**, 3–28.

Hooten, M.B., Johnson, D.S., McClintock, B.T. & Morales, J.M. (2017) *Animal movement : statistical models for telemetry data*. CRC Press, Boca Raton, FL.

Horne, J.S. & Garton, E.O. (2006) Likelihood cross-validation versus least squares cross-validation for choosing the smoothing parameter in kernel home-range analysis. *Journal of Wildlife Management* **70**, 641–648.

Horne, J.S., Garton, E.O., Krone, S.M. & Lewis, J.S. (2007) Analyzing animal movements using Brownian bridges. *Ecology* **88**, 2354–2363.

Hull, C.L. & Muir, S.C. (2013) Behavior and turbine avoidance rates of eagles at two wind farms in Tasmania, Australia. *Wildlife Society Bulletin* **37**, 49–58.

Jenkins, A.R., Reid, T., du Plessis, J., Colyn, R., Benn, G. & Millikin, R. (2018) Combining radar and direct observation to estimate pelican collision risk at a proposed wind farm on the Cape west coast, South Africa. *PLoS ONE* **13**, 1–21.

Jenkins, A.R., Rooyen, C.S.V., Smallie, J.J., Harrison, J., Diamond, M., Smit-Robinson, H. & Ralston, S. (2012) *Birds and Wind-Energy Best-Practice Guidelines: Best-Practice Guidelines for assessing and monitoring the impacts of wind-energy facilities on birds in southern Africa*. BirdLife South Africa and Endangered Wildlife Trust, 2nd edn.

Jenkins, A.R., Rooyen, C.S.V., Smallie, J.J., Harrison, J., Diamond, M., Smit-Robinson, H. & Ralston, S. (2015) *Birds and Wind-Energy Best-Practice Guidelines: Best-Practice Guidelines for assessing and monitoring the impacts of wind-energy facilities on birds in southern Africa*. November, BirdLife South Africa and Endangered Wildlife Trust, 3 edn.

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

Johnston, N.N., Bradley, J.E. & Otter, K.A. (2014) Increased Flight Altitudes among Migrating Golden Eagles Suggest Turbine Avoidance at a Rocky Mountain Wind Installation. *PloS one* **9**.

Kaldellis, J.K. & Zafirakis, D. (2011) The wind energy (r)evolution: A short review of a long history. *Renewable Energy* **36**, 1887–1901.

Kalman, R.E. (1960) A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering* **82**, 35.

Katzner, T., Benett, V., Miller, T., Duerr, A., Braham, M. & Hale, A. (2016a) Wind energy development : methods for assessing risks to birds and bats pre-construction. *Human–Wildlife Interactions* **10**, 42–52.

Katzner, T.E., Brandes, D., Miller, T., Lanzone, M., Maisonneuve, C., Tremblay, J.a., Mulvihill, R. & Merovich, G.T. (2012) Topography drives migratory flight altitude of golden eagles: Implications for on-shore wind energy development. *Journal of Applied Ecology* **49**, 1178–1186.

Katzner, T.E., Nelson, D.M., Braham, M.A., Doyle, J.M., Fernandez, N.B., Duerr, A.E., Bloom, P.H., Fitzpatrick, M.C., Miller, T.A., Culver, R.C.E., Braswell, L. & Dewoody, J.A. (2016b) Golden Eagle fatalities and the continental-scale consequences of local wind-energy generation. *Conservation Biology* **31**, 406–415.

Kays, R., Crofoot, M.C., Jetz, W. & Wikelski, M. (2015) Terrestrial animal tracking as an eye on life and planet. *Science* **348**, aaa2478.

Keating, K. & Cherry, S. (2009) Modeling utilization distributions in space and time. *Ecology* **90**, 1971–1980.

Kohavi, R. (1995) A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *International Joint Conference on Articial Intelligence (IJCAI)* pp. 1–7.

Kranstauber, B., Kays, R., Lapoint, S.D., Wikelski, M. & Safi, K. (2012) A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *Journal of Animal Ecology* **81**, 738–746.

Kulkarni, V.G. (2011) *Introduction to Modeling and Analysis of Stochastic Systems*, vol. 64 of *Springer Texts in Statistics*. Springer New York, New York, NY.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D. & Morales, J.M. (2012) Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology* **93**, 2336–2342.

Langston, R.H.W. & Pullan, J. (2003) Windfarms and Birds: an analysis of the effects of windfarms on birds, and guidance on environmental assessment criteria and site selection issues. *Convention on the Conservation of European Wildlife and Natural Habitats*, vol. Report T-PVS/Inf (2003) 12, by BirdLife International to the Council

of Europe, Bern Convention on the Conservation of European Wildlilfe and Natural Habitats, pp. 1–58, RSPB/BirdLife, UK.

Laver, P.N. & Kelly, M.J. (2008) A critical review of home range studies. *Journal of Wildlife Management* **72**, 290–298.

Levin, S.A., Carpenter, S.R., Godfray, H.C.J., Levin, S.A.A. & Carpenter, S.R.R. (2009) *The Princeton Guide to Ecology*. Princeton University Press, Oxfordshire, UK.

Liechti, F., Guélat, J. & Komenda-Zehnder, S. (2013) Modelling the spatial concentrations of bird migration to assess conflicts with wind turbines. *Biological Conservation* **162**, 24–32.

Loss, S.R., Dorning, M.A. & Diffendorfer, J.E. (2019) Biases in the Literature on Direct Wildlife Mortality from Energy Development. *BioScience* .

Loss, S.R., Will, T. & Marra, P.P. (2015) Direct Mortality of Birds from Anthropogenic Causes. *Annu. Rev. Ecol. Evol. Syst* **46**, 99–120.

Marques, A.T., Batalha, H., Rodrigues, S., Costa, H., Pereira, M.J.R., Fonseca, C., Mascarenhas, M. & Bernardino, J. (2014) Understanding bird collisions at wind farms: An updated review on the causes and possible mitigation strategies. *Biological Conservation* **179**, 40–52.

Martínez-Abraín, A., Tavecchia, G., Regan, H.M., Jiménez, J., Surroca, M. & Oro, D. (2012) Effects of wind farms and food scarcity on a large scavenging bird species following an epidemic of bovine spongiform encephalopathy. *Journal of Applied Ecology* **49**, 109–117.

Masden, E.A. & Cook, A.S.C.P. (2016) Avian collision risk models for wind energy impact assessments. *Environmental Impact Assessment Review* **56**, 43–49.

Masden, E.A., Reeve, R., Desholm, M., Fox, A.D., Furness, R.W. & Haydon, D.T. (2012) Assessing the impact of marine wind farms on birds through movement modelling. *Journal of the Royal Society, Interface / the Royal Society* **9**, 2120–30.

Matsuoka, S.M., Mahon, C.L., Handel, C.M., Sólymos, P., Bayne, E.M., Fontaine, P.C. & Ralph, C.J. (2014) Reviving common standards in point-count surveys for broad inference across studies. *The Condor* **116**, 599–608.

May, R., Masden, E.A., Bennet, F. & Perron, M. (2019) Considerations for upscaling

individual effects of wind energy development towards population-level impacts on wildlife. *Journal of Environmental Management* **230**, 84–93.

May, R., Steinheim, Y., Kvaløy, P., Vang, R. & Hanssen, F. (2017) Performance test and verification of an off-the-shelf automated avian radar tracking system. *Ecology and Evolution* **7**, 5930–5938.

May, R.F. (2015) A unifying framework for the underlying mechanisms of avian avoidance of wind turbines. *Biological Conservation* **190**, 179–187.

McClintock, B.T., Johnson, D.S., Hooten, M.B., Ver Hoef, J.M. & Morales, J.M. (2014) When to be discrete: The importance of time formulation in understanding animal movement. *Movement Ecology* **2**, 1–14.

McElreath, R. (2019) *Statistical Rethinking. A Bayesian Course with Examples in R and Stan.* Chapman and Hall / CRC Press, 8th march, edn.

Michelot, T. & Blackwell, P.G. (2019) State-switching continuous-time correlated random walks. *Methods in Ecology and Evolution* **10**, 637–649.

Millspaugh, J.J., Nielson, R.M., McDonald, L., Marzluff, J.M., Gitzen, R.A., Rittenhouse, C.D., Hubbard, M.W. & Sheriff, S.L. (2006) Analysis of resource selection using utilization distributions. *Journal of Wildlife Management* **70**, 384–395.

Mucina, L. & Rutherford, M. (eds.) (2010) *The vegatation of South Africa, Lesotho and Swaziland. Strelitzia 19.* South African National Biodiversity Institute, Pretoria.

Murgatroyd, M., Avery, G., Underhill, L.G. & Amar, A. (2016a) Adaptability of a specialist predator: the effects of land use on diet diversification and breeding performance of Verreaux's eagles. *Journal of Avian Biology* **47**, 834–845.

Murgatroyd, M., Photopoulou, T., Underhill, L.G., Bouten, W. & Amar, A. (2018) Where eagles soar: Fine-resolution tracking reveals the spatiotemporal use of differential soaring modes in a large raptor. *Ecology and Evolution* **8**, 6788–6799.

Murgatroyd, M., Underhill, L.G., Bouten, W. & Amar, A. (2016b) Ranging behaviour of Verreaux's eagles during the pre-breeding period determined through the use of high temporal resolution tracking. *PLoS ONE* **11**, 1–18.

NASA JPL (2013) NASA Shuttle Radar Topography Mission Global 1 arc second [Data set].

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse,

P.E. (2008) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences* **105**, 19052–19059.

Neal, R. (2011) MCMC Using Hamiltonian Dynamics. *Handbook of Markov Chain Monte Carlo* (eds. S. Brooks, A. Gelman, G. Jones & X.L. Meng), chap. 5, pp. 116–62, Chapman and Hall/CRC, Boca Raton, FL.

Nelson, E. (2001) *Dynamical Theories of Brownian Motion*. Princeton University Press, second edn.

Neumann, M.H. & Polzehl, J. (1998) Simultaneous bootstrap confidence bands in nonparametric regression. *Journal of Nonparametric Statistics* **9**, 307–333.

New, L., Bjerre, E., Millsap, B., Otto, M.C. & Runge, M.C. (2015) A Collision Risk Model to Predict Avian Fatalities at Wind Facilities: An Example Using Golden Eagles, Aquila chrysaetos. *Plos One* **10**.

Noonan, M.J., Tucker, M.A., Fleming, C.H., Akre, T.S., Alberts, S.C., Ali, A.H., Altmann, J., Antunes, P.C., Belant, J.L., Beyer, D., Blaum, N., Böhning-Gaese, K., Cullen, L., de Paula, R.C., Dekker, J., Drescher-Lehman, J., Farwig, N., Fichtel, C., Fischer, C., Ford, A.T., Goheen, J.R., Janssen, R., Jeltsch, F., Kauffman, M., Kappeler, P.M., Koch, F., LaPoint, S., Markham, A.C., Medici, E.P., Morato, R.G., Nathan, R., Oliveira-Santos, L.G.R., Olson, K.A., Patterson, B.D., Paviolo, A., Ramalho, E.E., Rösner, S., Schabo, D.G., Selva, N., Sergiel, A., Xavier da Silva, M., Spiegel, O., Thompson, P., Ullmann, W., Zięba, F., Zwijacz-Kozica, T., Fagan, W.F., Mueller, T. & Calabrese, J.M. (2019) A comprehensive analysis of autocorrelation and bias in home range estimation. *Ecological Monographs* **89**, 1–21.

Northrup, J.M. & Wittemyer, G. (2013) Characterising the impacts of emerging energy development on wildlife, with an eye towards mitigation. *Ecology letters* **16**, 112–25.

Ovaskainen, O., de Knegt, H.J. & Delgado, M.d.M. (2016) *Quantitative Ecology and Evolutionary Biology: Integrating Models with Data*. Oxford University Press, Oxford, UK, first edit edn.

Parton, A. & Blackwell, P.G. (2017) Bayesian Inference for Multistate 'Step and Turn' Animal Movement in Continuous Time. *Journal of Agricultural, Biological, and Environmental Statistics* **22**, 373–392.

Patterson, T.A., Parton, A., Langrock, R., Blackwell, P.G., Thomas, L. & King, R. (2017) Statistical modelling of individual animal movement: an overview of key meth-

ods and a discussion of practical challenges. *AStA Advances in Statistical Analysis* **101**, 399–438.

Pebesma, E. (2018) Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* **10**, 439–446.

Perold, V., Ralston-Paton, S. & Ryan, P. (2020) On a collision course? The large diversity of birds killed by wind turbines in South Africa. *Ostrich* **91**, 228–239.

Péron, G. (2019) Modified home range kernel density estimators that take environmental interactions into account. *Movement Ecology* **7**, 1–8.

Poessel, S.A., Brandt, J., Mendenhall, L., Braham, M.A., Lanzone, M.J., McGann, A.J. & Katzner, T.E. (2018) Flight response to spatial and temporal correlates informs risk from wind turbines to the California Condor. *The Condor* **120**, 330–342.

Preisler, H.K., Ager, A.A., Johnson, B.K. & Kie, J.G. (2004) Modeling animal movements using stochastic differential equations. *Environmetrics* **15**, 643–657.

Preisler, H.K., Ager, A.A. & Wisdom, M.J. (2013) Analyzing animal movement patterns using potential functions. *Ecosphere* **4**, 1–13.

Proffitt, D.R. (2006) Distance Perception. *Current Directions in Psychological Science* **15**, 131–135.

QGIS Development Team (2017) Qgis geographic information system .

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

R Core Team (2019) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ralston-Paton, S. (2017) *Verreaux's Eagle and Wind Farms: Guidelines for impact assessment, monitoring and mitigation*. BirdLife South Africa occasional report series, BirdLife South Africa, Hoboken, NJ, USA.

Ralston-Paton, S., J., S., Pearson A. & R., R. (2017) Wind energy's impacts on birds in South Africa: A preliminary review of the results of operational monitoring at the first wind farms of the Renewable Energy Independent Power Producer Procurement Programme in South Africa. *BirdLife South Africa Occasional Report Series No. 2.* pp. 1–30.

Reid, T., Kruger, S., Whitfield, P.D. & Amar, A. (2015) Using Spatial analysis of bearded vulture movement in southern Africa to inform wind turbine placement. *Journal of Applied Ecology* **52**.

REN21 (2018) Renewables 2018 global status report. Tech. rep., REN21 Secretariat, Paris.

REN21 (2019) Renewables 2019 global status report. Tech. rep., REN21 Secretariat, Paris.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F. & Dormann, C.F. (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **40**, 913–929.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C. & Müller, M. (2011) proc: an open-source package for r and s+ to analyze and compare roc curves. *BMC Bioinformatics* **12**, 77.

Russell, J.C., Hanks, E.M., Haran, M. & Hughes, D. (2018) A spatially varying stochastic differential equation model for animal movement. *Annals of Applied Statistics* **12**, 1312–1331.

Sappington, J.M., Longshore, K.M. & Thompson, D.B. (2007) Quantifying Landscape Ruggedness for Animal Habitat Analysis : A Case Study Using Bighorn Sheep in the Mojave Desert. *Journal of Wildlife Management* **71**, 1419–1426.

Särkkä, S. & Solin, A. (2019) *Applied Stochastic Differential Equations*. Cambridge University Press.

Schach, S. (1971) Weak Convergence Results for a Class of Multivariate Markov Processes. *The Annals of Mathematical Statistics* **42**, 451–465.

Schaub, T., Klaassen, R.H., Bouten, W., Schlaich, A.E. & Koks, B.J. (2019) Collision risk of Montagu's Harriers Circus pygargus with wind turbines derived from high resolution GPS tracking. *Ibis* .

Schmaljohann, H., Liechti, F., Bächler, E., Steuri, T. & Bruderer, B. (2008) Quantification of bird migration by radar - A detection probability problem. *Ibis* **150**, 342–355.

Schuster, E., Bulling, L. & Köppel, J. (2015) Consolidating the State of Knowledge: A

Synoptical Review of Wind Energy's Wildlife Effects. *Environmental Management* **56**, 300–331.

Schuster, E.F. & Gregory, G.G. (1981) On the Nonconsistency of Maximum Likelihood Nonparametric Density Estimators. *Computer Science and Statistics: Proceedings of the 13th Symposium on the Interface*, pp. 295–298, Springer US, New York, NY.

Scott, D.W. (1992) *Multivariate density estimation: theory, practice, and visualization.* John Wiley & Sons, Inc., New York.

Scottish Natural Heritage (2014) Recommended bird survey methods to inform impact assessment of onshore wind farms. Tech. Rep. May, Scottish Natural Heritage.

Shepard, E.L.C., Wilson, R.P., Rees, W.G., Grundy, E., Lambertucci, S.a. & Vosper, S.B. (2013) Energy landscapes shape animal movement ecology. *The American naturalist* **182**, 298–312.

Silverman, B. (1986a) Density estimation for statistics and data analysis. *Chapman and Hall*, vol. 37, pp. 1–22.

Silverman, B.W. (1986b) *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London; New York.

Simmons, R.E. (2005) Verreaux's eagle. *Roberts' Birds of Southern Africa* (eds. P.A.R. Hockey, W.R.J. Dean & P.G. Ryan), John Voelcker Bird Book Fund, Cape Town, 7th edition edn.

Sinai, M.J. & He, Z.J. (1998) Terrain influences the accurate judgement of distance. *Nature* **395**, 497–500.

Smallwood, K.S., Rugge, L. & Morrison, M.L. (2009) Influence of Behavior on Bird Mortality in Wind Energy Developments. *Journal of Wildlife Management* **73**, 1082–1098.

Stan Development Team (2018a) RStan: the R interface to Stan. R package version 2.18.2.

Stan Development Team (2018b) RStan: the R interface to Stan. R package version 2.18.2.

Stan Development Team (2018c) RStan: the R interface to Stan. R package version 2.18.2.

Stone, M. (1977) An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Journal of the Royal Society, Series B* **39**, 44–47.

Strickland, D., Arnett, E., Erickson, W., Johnson, D., Johnson, G., Morrison, M., Shaffer, J. & Warren-Hicks, W. (2011) *Comprehensive Guide to Studying Wind Energy/Wildlife Interactions*. June, Washintong, D.C.

Sur, M., Belthoff, J.R., Bjerre, E.R., Millsap, B.A. & Katzner, T. (2018) The utility of point count surveys to predict wildlife interactions with wind energy facilities: An example focused on golden eagles. *Ecological Indicators* **88**, 126–133.

Tarjan, L.M. & Tinker, M.T. (2016) Permissible Home Range Estimation (PHRE) in Restricted Habitats: A New Algorithm and an Evaluation for Sea Otters. *PLOS ONE* **11**, e0150547.

Taylor, M.R., Peacok, F. & Wanless, R. (eds.) (2015) *The 2015 Eskom Red Data Book of Birds of Southern Africa, Lesotho and Swaziland*. BirdLife South Africa, Johannesburg.

Thaxter, C.B., Buchanan, G.M., Carr, J., Butchart, S.H.M., Newbold, T., Green, R.E., Tobias, J.A., Foden, W.B., O'Brien, S. & Pearce-Higgins, J.W. (2017) Bird and bat species' global vulnerability to collision mortality at wind farms revealed through a trait-based assessment. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20170829.

Todd, J.T., Oomes, A.H.J., Koenderink, J.J. & Kappers, A.M.L. (2001) On the affine structure of perceptual space. *Psychological science : a journal of the American Psychological Society / APS* **12**, 191–196.

Turchin, P. (1998) *Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants*. 1, Beresta Books, Middletown, DE, 2015 edition edn.

Turchin, P. (2015) *Quantitative Analysis of Movement: measuring and modelling population redistribution in animals and plants*. Beresta Books, USA.

Uhlenbeck, G.E. & Ornstein, L.S. (1930) On the Theory of the Brownian Motion. *Physical Review* **36**, 823–841.

UNFCCC (2015) Adoption of the Paris Agreement. *FCCC/CP/2015/10/Add.1*.

Van Winkle, W. (1975) Comparison of Several Probabilistic Home-Range Models. *The Journal of Wildlife Management* **39**, 118–123.

Wagner, M. (1985) The metric of visual space. *Perception & Psychophysics* **38**, 483–495.

Walker, D., McGrady, M., McCluskie, A., Madders, M. & McLeod, D. (2005) Resident Golden Eagle ranging behaviour before and after construction of a windfarm in Argyll. *Scottish Bird* **25**, 24–40.

Walls, R., Pendlebury, C., Budgey, R., Brookes, K. & Paul, P. (2009) Revised best practice guidance for the use of remote techniques for ornithological monitoring at offshore windfarms.

Watson, R.T., Kolar, P.S., Ferrer, M., Nygård, T., Johnston, N., Hunt, W.G., Smit-Robinson, H.A., Farmer, C.J., Huso, M. & Katzner, T.E. (2018) Raptor Interactions With Wind Energy: Case Studies From Around the World. *Journal of Raptor Research* **52**, 1–18.

Weitzman, M. (1970) Measure of the overlap of income distribution of white and Negro families in the United States.

Westreich, D. & Greenland, S. (2013) The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients. *American Journal of Epidemiology* **177**, 292–298.

Wickham, H. (2017) *tidyverse: Easily Install and Load the 'Tidyverse'*. R package version 1.2.1.

Witt, J.K., Stefanucci, J.K., Riener, C.R. & Proffitt, D.R. (2007) Seeing beyond the target: Environmental context affects distance perception. *Perception* **36**, 1752–1768.

Wood, S. (2017) *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2 edn.

Worton, B. (1987) A review of models of home range for animal movement. *Ecological Modelling* **38**, 277–298.

Worton, B.J. (1989) Kernel methods for estimating the utilization distribution in home-range studies. *Ecology* **70**, 164–168.

Yang, Z. & Purves, D. (2003) A statistical explanation of visual space. *Nature Neuroscience* **6**, 632–40.

Young, D.J., Erickson, W., Strickland, M., Good, R. & Sernka, K. (2003) Comparison of Avian Responses to UV-Light-Reflective Paint on Wind Turbines Comparison of Avian Responses to UV-Light-Reflective Paint on Wind Turbines.

Zhang, Y. & Yang, Y. (2015) Cross-validation for selecting a model selection procedure. *Journal of Econometrics* **187**, 95–112.

Appendix A

Transitional density of the potential-based velocity model

The model used for describing the bird location at time t is the following:

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t \mathbf{V}_s \mathrm{d}s \tag{A.1}$$

$$\mathrm{d}\mathbf{V}_t = \beta(\gamma(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t)\mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t \tag{A.2}$$

where $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ is the preferred velocity at time t and is represented by a vector in the direction of the nearest point of attraction at time t and of magnitude $|\mathbf{V}_t|$. The coefficient β, modulates the drift towards (if positive) or away from (if negative) a target location. The term $\sigma \mathrm{d}\mathbf{B}_t$, for $\sigma > 0$, incorporates random forces that act on the bird, changing its velocity.

$$\mathrm{d}\mathbf{B}_t \sim N(\mathbf{0}, \mathrm{d}t\mathbf{I}) \tag{A.3}$$

In case there are K target locations at time t, then,

$$\mathrm{d}\mathbf{V}_t = \sum_{k=1}^{K} \beta_k(\gamma_k(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t)\mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t \tag{A.4}$$

To derive the transitional density and for clarity, I will use a model with 2 target locations and I simplify the notation from $\gamma_k(\mathbf{X}_t, \mathbf{V}_t)$ to γ_k.

$$\mathrm{d}\mathbf{V}_t = \beta_1(\gamma_1 - \mathbf{V}_t)\mathrm{d}t + \beta_2(\gamma_2 - \mathbf{V}_t)\mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t \tag{A.5}$$

At short time horizons the target locations will most likely not change, so I will ignore that γ changes over time. This will be helpful for simulating flights, producing "steps" at short intervals, although not so much for calculating an equilibrium solution.

We can first re-write the equation as:

$$\frac{\mathrm{d}\mathbf{V}_t}{\mathrm{d}t} + (\beta_1 + \beta_2)\mathbf{V}_t = \beta_1\gamma_1 + \beta_2\gamma_2 + \sigma \mathrm{d}\mathbf{B}_t \tag{A.6}$$

$$\mathrm{d}\mathbf{V}_t + (\beta_1 + \beta_2)\mathbf{V}_t\mathrm{d}t = (\beta_1\gamma_1 + \beta_2\gamma_2)\mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t \tag{A.7}$$

This expression suggest the following integrating factor to solve the ordinary differential equation:

$$IF = e^{\int(\beta_1+\beta_2)\mathrm{d}t} = e^{(\beta_1+\beta_2)t} \tag{A.8}$$

I will let $\kappa = \beta_1 + \beta_2$ for a more compact notation. Multiplying equation (6) by IF and integrating,

$$e^{\kappa t}\mathrm{d}\mathbf{V}_t + e^{\kappa t}\kappa\mathbf{V}_t\mathrm{d}t = e^{\kappa t}(\beta_1\gamma_1 + \beta_2\gamma_2)\mathrm{d}t + e^{\kappa t}\sigma \mathrm{d}\mathbf{B}_t \tag{A.9}$$

Considering that the Itô differential for a function $f(x_t, t)$ is:

$$\mathrm{d}[f(x_t, t)] = \frac{\partial}{\partial t}f(x_t, t)\mathrm{d}t + \frac{\partial}{\partial x}f(x_t, t)\mathrm{d}x + \frac{1}{2}\frac{\partial^2}{\partial x^2}f(x_t, t)\mathrm{d}x^2 \tag{A.10}$$

We recognize that:

$$\mathrm{d}[e^{\kappa t}\mathbf{V}_t] = e^{\kappa t}\kappa\mathbf{V}_t\mathrm{d}t + e^{\kappa t}\mathrm{d}\mathbf{V}_t \tag{A.11}$$

The right hand side of A.11 is the left hand side of A.9. Substituting A.11 into A.9 and integrating both sides from zero to t we obtain:

$$\int_0^t e^{\kappa s}\mathbf{V}_s\mathrm{d}s = \int_0^t e^{\kappa s}(\beta_1\gamma_{1s} + \beta_2\gamma_{2s})\mathrm{d}s + \int_0^t e^{\kappa s}\sigma\mathrm{d}\mathbf{B}_s \tag{A.12}$$

Here, I assume that at short time horizons, γ will most likely not change. Then, I can write:

$$e^{\kappa t}\mathbf{V}_t - \mathbf{V}_0 = (\beta_1\gamma_1 + \beta_2\gamma_2)\int_0^t e^{\kappa s}\mathrm{d}s + \sigma\int_0^t e^{\kappa s}\mathrm{d}\mathbf{B}_s = \tag{A.13}$$

$$= \frac{(\beta_1\gamma_1 + \beta_2\gamma_2)}{\kappa}\Big[e^{\kappa t} - 1\Big] + \sigma\int_0^t e^{\kappa s}\mathrm{d}\mathbf{B}_s \tag{A.14}$$

And finally,

$$\mathbf{V}_t = \mathbf{V}_0 e^{-\kappa t} + \frac{(\beta_1\gamma_1 + \beta_2\gamma_2)}{\kappa}\Big[1 - e^{-\kappa t}\Big] + \sigma\int_0^t e^{-\kappa(t-s)}\mathrm{d}\mathbf{B}_s \tag{A.15}$$

This equation derives in a transitional density, such that:

$$E[\mathbf{V}_t] = \mathbf{V}_0 e^{-\kappa t} + \frac{(\beta_1\gamma_1 + \beta_2\gamma_2)}{\kappa}\Big[1 - e^{-\kappa t}\Big] \tag{A.16}$$

$$Var[\mathbf{V}_t] = \sigma^2\int_0^t e^{-2\kappa(t-s)}\mathrm{d}_s = \frac{\sigma^2}{2\kappa}\Big[1 - e^{-2\kappa t}\Big] \tag{A.17}$$

Therefore, the transitional density for short time horizons, can be written as:

$$\mathbf{V}_t \sim N\Big(\mathbf{V}_0 e^{-\kappa t} + \frac{(\beta_1\gamma_1 + \beta_2\gamma_2)}{\kappa}\Big[1 - e^{-\kappa t}\Big], \frac{\sigma^2}{2\kappa}\Big[1 - e^{-2\kappa t}\Big]\Big) \tag{A.18}$$

Appendix B

Transitional density of the stochastic damped harmonic motion

The model selected as best describing the evolution of the observation error with time, resembles a damped stochastic harmonic oscillator:

$$\begin{aligned} \mathrm{d}\mathbf{D}_t &= \mathbf{V}_t \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= -\alpha(\mathbf{D}_t - \mathbf{D}_0)\mathrm{d}t - \beta \mathbf{V}_t \mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t \\ \mathrm{d}\mathbf{B}_t &\sim N(0, \mathrm{d}t) \end{aligned} \tag{B.1}$$

To start solving the system of equations, I define the variable $\mathbf{S}_t = \mathbf{D}_t - \mathbf{D}_0$. It can be proven that the rate of change of $\mathbf{S}_t$ is the same as the rate of chance in $\mathbf{D}_t$:

$$\mathrm{d}\mathbf{S}_t = \mathrm{d}(\mathbf{D}_t - \mathbf{D}_0) = \mathrm{d}\mathbf{D}_t - 0 = \mathrm{d}\mathbf{D}_t \tag{B.2}$$

Then, I can express the system in the following vector form (Särkkä & Solin, 2019):

$$\mathrm{d}\mathbf{X}_t = \mathbf{F}\mathbf{X}_t \mathrm{d}t + \mathbf{L}\mathrm{d}\mathbf{B}_t \tag{B.3}$$

Where,

$$\mathbf{X}_t = \begin{bmatrix} \mathbf{S}_t \\ \mathbf{V}_t \end{bmatrix}; \mathbf{F} = \begin{bmatrix} 0 & 1 \\ -\alpha & -\beta \end{bmatrix}; \mathbf{L} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{B.4}$$

$\mathbf{B}_t$ is a Brownian motion with diagonal diffusion matrix $\mathbf{Q} = \sigma^2\mathbf{I}$.

We can solve the above vector equation by rearranging and multiplying both sides by the integrating factor corresponding to the matrix exponential $e^{-\mathbf{F}t}$:

$$e^{-\mathbf{F}t}\mathrm{d}\mathbf{X}_t - e^{-\mathbf{F}t}\mathbf{F}\mathbf{X}_t\mathrm{d}t = e^{-\mathbf{F}t}\mathbf{L}\mathrm{d}\mathbf{B}_t \tag{B.5}$$

Considering that the Itô differential for a function $f(x_t, t)$ is:

$$\mathrm{d}[f(x_t,t)] = \frac{\partial}{\partial t}f(x_t,t)\mathrm{d}t + \frac{\partial}{\partial x}f(x_t,t)\mathrm{d}x + \frac{1}{2}\frac{\partial^2}{\partial x^2}f(x_t,t)\mathrm{d}x^2 \tag{B.6}$$

We recognize that:

$$\mathrm{d}[e^{-\mathbf{F}t}\mathbf{X}_t] = -e^{-\mathbf{F}t}\mathbf{F}\mathbf{X}_t\mathrm{d}t + e^{-\mathbf{F}t}\mathrm{d}\mathbf{X}_t \tag{B.7}$$

The right hand side of B.7 is the left hand side of B.5. Substituting B.7 into B.5 and integrating both sides from zero to t we obtain:

$$\begin{aligned} e^{-\mathbf{F}t}\mathbf{X}_t - \mathbf{X}_0 &= \int_0^t e^{-\mathbf{F}s}\mathbf{L}\mathrm{d}\mathbf{B}_s \\ \mathbf{X}_t &= e^{\mathbf{F}t}\mathbf{X}_0 + \int_0^t e^{\mathbf{F}(t-s)}\mathbf{L}\mathrm{d}\mathbf{B}_s \end{aligned} \tag{B.8}$$

This is the formulation of the stochastic process $\mathbf{X}_t$. Now, we can find its expected value and variance:

$$\mathbb{E}[\mathbf{X}_t] = \mathbb{E}[e^{\mathbf{F}t}\mathbf{X}_0] + \mathbb{E}[\int_0^t e^{\mathbf{F}(t-s)}\mathbf{L}\mathrm{d}\mathbf{B}_s] = e^{\mathbf{F}t}\mathbf{X}_0 \tag{B.9}$$

Since, $\mathbb{E}[\mathbf{B}_t] = 0$.

$$Var[\mathbf{X}_t] = \mathbb{E}[(\mathbf{X}_t - \mathbb{E}[\mathbf{X}_t])^2] = \mathbb{E}[(\int_0^t e^{\mathbf{F}(t-s)}\mathbf{L}\mathrm{d}\mathbf{B}_s)^2] = \int_0^t e^{\mathbf{F}(t-s)}\mathbf{L}\mathbf{Q}\mathbf{L}^T(e^{\mathbf{F}(t-s)})^T\mathrm{d}s \tag{B.10}$$

Where I made use of the fact that $\mathbb{E}[\mathbf{B}_t\mathbf{B}_s^T] = (t-s)\mathbf{Q}$ for $t > s$ (Särkkä & Solin, 2019).

The equations that solve the second order differential equation described by B.1 involve finding the roots of the characteristic polynomial of the second order equation and these are characterized by $r = (-\beta \pm \sqrt{\beta^2 - 4\alpha})/2$. Depending on the particular values of α and β, the radical will produce real roots (if $\beta^2 - 4\alpha > 0$) or complex roots (if $\beta^2 - 4\alpha < 0$). To simplify the derivation of the solution and provided that in this chapter, the quantity $\beta^2 - 4\alpha < 0$, I will focus on this case, only.

To simplify notation, it is convenient to let $\omega = (\sqrt{4\alpha - \beta^2})/2$ and $b = \beta/2$. Then, we can write an expression for the matrix $e^{\mathbf{F}t}$ ($\mathbf{F}$ was defined in B.4) as:

$$e^{\mathbf{F}t} = \begin{bmatrix} e^{-bt}(\frac{b}{\omega}sin(\omega t) + cos(\omega t)) & \frac{e^{-bt}}{\omega}sin(\omega t) \\ -e^{-bt}\frac{\alpha}{\omega}sin(\omega t) & e^{-bt}(cos(\omega t) - \frac{b}{\omega}sin(\omega t)) \end{bmatrix} \tag{B.11}$$

Then, the transition density of $\mathbf{X}_t = (\mathbf{S}_t, \mathbf{V}_t)$ - the probability density at a point $\mathbf{X}_t$ at time t, given location $\mathbf{X}_s$ at time s - will be Gaussian:

$$\mathbf{X}_t|\mathbf{X}_0 \sim N(\mu_t, \mathbf{P}_t) \tag{B.12}$$

The parameters μ_t and $\mathbf{P}_t$ are a short hand for $\mu_t|\mathbf{X}_0$ and $\mathbf{P}_t|\mathbf{X}_0$ respectively and are computed substituting B.11 and B.4 in B.9 and B.10.

$$\mu_t = \begin{bmatrix} \frac{e^{-bt}}{\omega}[(\mathbf{S}_0 b + \mathbf{V}_0)sin(\omega t) + \mathbf{S}_0\omega cos(\omega t)] \\ -\frac{e^{-bt}}{\omega}[(\mathbf{S}_0\alpha + \mathbf{V}_0 b)sin(\omega t) - \mathbf{V}_0\omega cos(\omega t)] \end{bmatrix} \tag{B.13}$$

The covariance matrix $\mathbf{P}_t$ results in the integral:

$$\mathbf{P}_t = \int_0^t \begin{bmatrix} \left(\frac{\sigma e^{-bs}}{\omega} sin(\omega s)\right)^2 & \frac{\sigma^2 e^{-2bs}}{2\omega^2}(\omega sin(2\omega s) - 2bsin^2(\omega s)) \\ \frac{\sigma^2 e^{-2bs}}{2\omega^2}(\omega sin(2\omega s) - 2bsin^2(\omega s)) & \left(\sigma e^{-bs}(cos(\omega s) - \frac{b}{\omega} sin(\omega s))\right)^2 \end{bmatrix} ds \tag{B.14}$$

Each entry of the matrix can be calculated by taking the integral of the respective element in the matrix. These integrals result in complicated expressions that I write down for completeness.

$$\begin{aligned} P_{11} &= \frac{\sigma^2}{2\alpha\beta\omega^2}(\omega^2 + e^{-2bt}(b^2 cos(2\omega t) - \omega b sin(2\omega t) - \alpha)) \\ P_{12} = P_{21} &= \frac{\sigma^2 e^{-2bt}}{2\omega^2} sin^2(\omega t) \\ P_{22} &= \frac{\sigma^2}{2\beta\omega^2}(\omega^2 + e^{-2bt}(b^2 cos(2\omega t) + \omega b sin(2\omega t) - \alpha)) \end{aligned} \tag{B.15}$$

Appendix C

Stan model to fit potential-based velocity model to flight trajectories observed from vantage points

In this chapter I present the stan code to run a basic potential-based velocity model. This code was prepared to fit the model to flight trajectories observed from vantage points. The code can also be used for fitting the model to telemetry data; however, in if working with such data, it might be worth exploring the use of algorithms such as the Kalman filter (Kalman, 1960; Johnson *et al.*, 2008).

Lets first remember the form of the model:

$$\begin{aligned} \mathrm{d}\mathbf{X}_t &= \mathbf{V}_t \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= \beta(\gamma(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t)\mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t. \end{aligned} \tag{C.1}$$

where $\mathbf{X}_t$ is the state of the process (e.g. location of the animal), $\mathbf{V}_t$ is the velocity at which the process is changing (e.g. velocity of the moving animal), $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ is the preferred velocity that changes with the state and velocity of the process and $\sigma \mathrm{d}\mathbf{B}_t$ is a Brownian motion with variance scale σ. The parameter β modulates the rate at which

the velocity of the process drifts back to the preferred velocity $\gamma(\mathbf{X}_t, \mathbf{V}_t)$.

If there are K target locations (attractive or repulsive) at time t, then,

$$\mathrm{d}\mathbf{V}_t = \sum_{k=1}^{K} \beta_k(\gamma_k(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t)\mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t. \tag{C.2}$$

Recall further that the preferred velocity $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ imposed by the potential field is given by the negative of the gradient of the potential function at the current state and by the current travelling speed:

$$\gamma(\mathbf{X}_t, \mathbf{V}_t) = -\nabla_H |\mathbf{V}_t|, \tag{C.3}$$

where $\mathbf{X}_t$ is the location of the bird at time t, $|\mathbf{V}_t|$ is its travelling speed and ∇_H is the gradient of the potential function H.

The drift coefficients β can change for different flights. Therefore, for each type of target location k and flight i,

$$\begin{aligned} \beta_{ki} &= \mu_{\beta_k} + \eta_i \\ \eta_i &\sim \mathrm{N}(0, \sigma_{\beta_k}) \end{aligned} \tag{C.4}$$

This formulation specifies a random effect η_i on each flight's drift towards target location k.

We may now prepare the data as if we were to fit a linear model, such that:

$$\begin{aligned} y_t &= \beta^\top \mathbf{U}_t \delta t + \sigma \epsilon_t \\ \epsilon &\sim N(0, \delta t) \end{aligned} \tag{C.5}$$

Therefore, we set $d\mathbf{V}_t = y_t$, $\mathbf{U}_t = \gamma(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t$ and $d\mathbf{B}_t = \epsilon_t$. Where $\mathbf{U}_t$ has k columns, each corresponding to a different type of target location (i.e. steep slopes, water bodies, etc.) and N rows, where N is the total number of locations recorded along all flights.

Therefore, we need a variable that indexes each observed location with its respective flight. In the dataset, this variable is called group (because it groups locations into groups).

The model in stan is as follows:

```
// Input data

data {
    // Total number of locations
    int<lower = 0> N;
    // Number of flights
    int<lower = 0> J;
    // Number of types of target locations (slope and water bodies, K = 2)
    int<lower = 0> K;

    // Vector of changes in velocity
    vector[N] dv;
    // Vector of groups
    int group[N];
    // Matrix of drift terms
    vector[K] u[N];
    // Vector of times
    vector[N] dt;

    // Prior mean (pmean) and sig (psig) for mean drift (mubeta)
    vector[K] pmu_mubeta;
    vector<lower = 0>[K] psig_mubeta;

    // Prior mean (pmean) and sig (psig) for variation in drift (sigbeta)
    vector[K] pmu_sigbeta;
    vector<lower = 0>[K] psig_sigbeta;

    // Prior mean (pmu) for diffusion parameter (sigma)
    real pmu_sigma;
    // Prior sigma (psig) for diffusion parameter (sigma)
    real psig_sigma;
}
```

```
// Model parameters

parameters {
    // Mean drifts
    vector[K] mu_beta;

    // Random effect of flight on drift
    vector<lower = 0>[K] eta;

    // Individual effects of flights on drift
    vector[K] delta_group[J];

    // Diffusion parameter
    real<lower=0> sigma;
}

transformed parameters{
    // Drift parameters of individual flights
    vector[K] beta[J];

    // Drift at each location
    vector[N] drift;

    // Diffusion at each location
    vector[N] diff;

    // Non-centred parameterization of drift parameters
    for(j in 1:J){
        for(k in 1:K){
            beta[j,k] = mu_beta[k] + eta[k] * delta_group[j,k];
        }
    }

    // Drift and diffusion at each location
```

```
    for(i in 1:N){
        drift[i] = dot_product(u[i], beta[group[i]]) * dt[i];
        diff[i] = sigma * sqrt(dt[i]);
    }

}



// Model specification

model {
    // Mean drift for each type of target location
    mu_beta ~ normal(pmu_mubeta, psig_mubeta);

    // Standardized effects of individual flights on drift
    for(j in 1:J){
        delta_group[j] ~ normal(0, 1);
    }

    // Scale of effects of individual flights on drift
    eta ~ cauchy(pmu_sigbeta, psig_sigbeta);



    // Diffusion parameter
    sigma ~ cauchy(pmu_sigma, psig_sigma);

    // Likelihood
    dv ~ normal(drift, diff);

}
```

Appendix D

Simulation code for potential-based velocity models

This chapter serves two purposes. Firstly, I present the R code necessary to simulate the motion imposed by potential-based velocity models. Secondly, I present potential-based velocity models graphically, to provide some insight about how these models operate. There are different ways to approach the simulation of trajectories from a stochastic differential equation. Here I present a method based on forward difference. This algorithm is appealing due to its simplicity, but requires a simulating small time step δt at a time. Otherwise the simulations become unstable. An alternative way of simulating trajectories is making use of the transitional density of the stochastic process, which I present in a appendix A. By using the transitional density we gain in stability but we need to be prepared to assume that the gradient of the potential stays constant during the time step δt. This is also explained in the same appendix where I derive the transitional density of the process. Finally, note that the code could certainly be optimized further, but I favoured clarity over performance.

Potential-based velocity model

Let's remember the form of the potential-based velocity model used in chapter 3. The basic structure of this model was:

$$\begin{aligned} \mathrm{d}\mathbf{X}_t &= \mathbf{V}_t \mathrm{d}t \\ \mathrm{d}\mathbf{V}_t &= \beta(\gamma(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t)\mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t. \end{aligned} \tag{D.1}$$

where $\mathbf{X}_t$ is the state of the process (e.g. location of the animal), $\mathbf{V}_t$ is the velocity at which the process is changing (e.g. velocity of the moving animal), $\gamma(\mathbf{X}_t, \mathbf{V}_t)$ is the preferred velocity that changes with the state and velocity of the process and $\sigma \mathrm{d}\mathbf{B}_t$ is a Brownian motion with variance scale σ. The parameter β modulates the rate at which the velocity of the process drifts back to the preferred velocity $\gamma(\cdot)$.

If there are K target locations (attractive or repulsive) at time t, then,

$$\mathrm{d}\mathbf{V}_t = \sum_{k=1}^{K} \beta_k(\gamma_k(\mathbf{X}_t, \mathbf{V}_t) - \mathbf{V}_t)\mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t. \tag{D.2}$$

Note that this model can be reformulated as:

$$\mathrm{d}\mathbf{V}_t = \sum_{k=1}^{K} \beta_k \gamma_k(\mathbf{X}_t, \mathbf{V}_t)\mathrm{d}t - (\sum_{k=1}^{K} \beta_k)\mathbf{V}_t \mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t. \tag{D.3}$$

By making $(\sum_{k=1}^{K} \beta_k) = \kappa$, we may re-write to

$$\mathrm{d}\mathbf{V}_t = \sum_{k=1}^{K} \beta_k \gamma_k(\mathbf{X}_t, \mathbf{V}_t)\mathrm{d}t - \kappa \mathbf{V}_t \mathrm{d}t + \sigma \mathrm{d}\mathbf{B}_t. \tag{D.4}$$

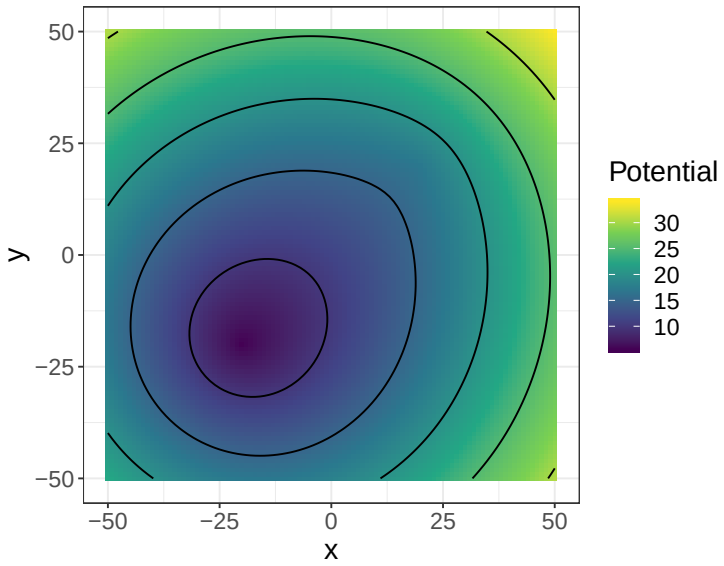
The first term on the right hand side controls the preferred velocity the process drifts towards, whereas the second term is the rate at which the current velocity decreases to zero. This notation is convenient because the first term on the right hand side can be visualised as the effective velocity field driving the system.

Simulations

Lets start by defining the shape of the potential field. I will use the same form of potential that I used in chapter 3. In that formulation, the potential was equal to the distance from target locations. For simplicity, I define the potential with only two reference target points $\mu = (\mu_x, \mu_y)$ and $\eta = (\eta_x, \eta_y)$.

$$H(x,y) = d_\mu + d_\eta = \sqrt{(x-\mu_x)^2 + (y-\mu_y)^2} + \sqrt{(x-\eta_x)^2 + (y-\eta_y)^2} \tag{D.5}$$

```
# Clean workspace
rm(list = ls())

# Load libraries
library(tidyverse)
library(viridis)
# Prepare a plotting theme
mytheme <- list(coord_equal(),
                theme_bw(),
                theme(text = element_text(size = 16)))

# Create spatial domain and specify a potential function with
# target points at mu = (20,20) and eta = (-20,-20)
mu = c(20, 20); eta = c(-20,-20)

# Define beta coefficients (one for each target location)
beta <- c(0.1, 0.3)
kappa <- sum(beta)

# Spatial domain and potential
pixsize <- 1
omega <- expand.grid(x = seq(-50, 50, pixsize),
                     y = seq(-50, 50, pixsize)) %>%
    mutate(dmu = beta[1] * sqrt((x - mu[1])^2 + (y - mu[2])^2),
           deta = beta[2] * sqrt((x - eta[1])^2 + (y -eta[2])^2),
           pot = dmu + deta)

# Plot potential
ggplot(omega) +
    geom_raster(aes(x = x, y = y, fill = pot)) +
    scale_fill_viridis_c(name = "Potential") +
    geom_contour(aes(x, y, z = pot), col = "black") +
    mytheme
```



The preferred velocity imposed by the potential field is given by the negative gradient of the potential function. More precisely the preferred velocity in chapter 3 was defined as:

$$\gamma(\mathbf{X}_t, \mathbf{V}_t) = -|\mathbf{V}_t| \nabla_H, \tag{D.6}$$

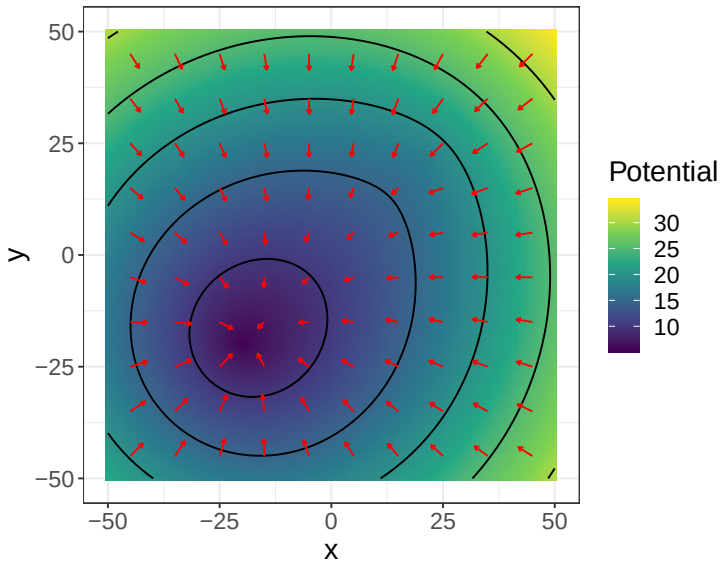
where $\mathbf{X}_t$ is the location of the bird at time t, $|\mathbf{V}_t|$ is its travelling speed and ∇_H is the gradient of the potential function H.

We first calculate the gradient of H by computing its first partial derivatives:

$$\nabla_H = \begin{bmatrix} \frac{\partial H}{\partial x} \\ \frac{\partial H}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{x_t - \mu_x}{d_\mu} + \frac{x_t - \eta_x}{d_\eta} \\ \frac{y_t - \mu_y}{d_\mu} + \frac{y_t - \eta_y}{d_\eta} \end{bmatrix}, \tag{D.7}$$

```
# function to calculate the gradient of Euclidean distance function
eucGrad<- function(x, y, mu){

    d <- sqrt((x - mu[1])^2 + (y - mu[2])^2)

    matrix(c((x - mu[1])/d, (y - mu[2])/d), ncol = 2, byrow = FALSE)

}

# Add partial derivatives to the main data frame
```

```
omega <- omega %>%
    mutate(dfdx = (beta[1] * eucGrad(x, y, mu) +
                       beta[2] * eucGrad(x, y, eta))[,1],
           dfdy = (beta[1] * eucGrad(x, y, mu) +
                       beta[2] * eucGrad(x, y, eta))[,2])

# Auxiliary gradient data frame to create arrows at some points
gradpix <- 10
grad_scale <- 10 # to plot larger arrows

grad <- expand.grid(x = seq(-45, 45, gradpix),
                    y = seq(-45, 45, gradpix)) %>%
    mutate(dfdx = (beta[1] * eucGrad(x, y, mu) +
                       beta[2] * eucGrad(x, y, eta))[,1],
           dfdy = (beta[1] * eucGrad(x, y, mu) +
                       beta[2] * eucGrad(x, y, eta))[,2])

ggplot(omega) +
    geom_raster(aes(x = x, y = y, fill = pot)) +
    scale_fill_viridis_c(name = "Potential") +
    geom_contour(aes(x, y, z = pot), col = "black") +
    geom_segment(data = grad,
                 aes(x = x, y = y,
                     xend = x - dfdx * grad_scale,
                     yend = y - dfdy * grad_scale),
                 arrow = arrow(length = unit(0.1, "cm")),
                 col = "red") +
    mytheme
```

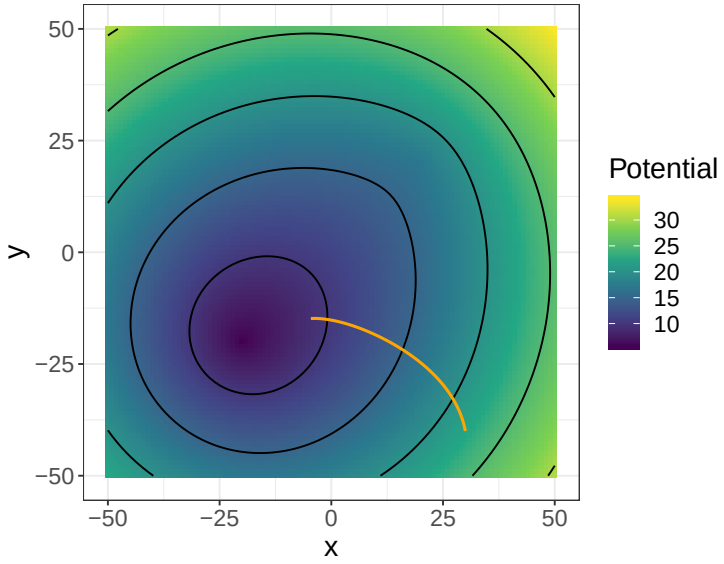


We may now simulate the deterministic dynamics of a bird moving within this field:

```
# SIMULATE 100 SECONDS OF BIRD MOTION

# Total time
Time <- 100

# Initial location
r0 = c(30, -40)

# Set an initial velocity
v0 = c(-1, 5)

# Time step
dt <- 0.1

# steps
n <- Time/dt + 1 # plus the initial time

# Data frame to store location and velocity at each time step
traj <- data.frame(x = numeric(length = n),
                   y = numeric(length = n),
                   vx = numeric(length = n),
                   vy = numeric(length = n),
                   dvx = numeric(length = n),
                   dvy = numeric(length = n))
```

```
# Initial velocity drift:
# - speed times neg gradient
gamma <- -sqrt(sum(v0^2)) * (beta[1] * eucGrad(r0[1], r0[2], mu) +
                                beta[2] * eucGrad(r0[1], r0[2], eta))

# - drift
dv <- (gamma - kappa * v0) * dt


# fill in first row of the data frame
traj[1, ] <-  c(r0, v0, dv)


for(i in 2:n){

    # Update position
    r <- r0 + v0 * dt

    # update velocity
    v <- v0 + dv

    # update velocity drift
    # - speed times neg gradient
    gamma <- -sqrt(sum(v^2)) * (beta[1] * eucGrad(r[1], r[2], mu) +
                                    beta[2] * eucGrad(r[1], r[2], eta))

    # - drift
    dv <- (gamma - kappa * v) * dt

    # fill in new row of the data frame
    traj[i, ] <-  c(r, v, dv)

    # iterate
    r0 <- r; v0 <- v

}

ggplot(omega) +
    geom_raster(aes(x = x, y = y, fill = pot)) +
    scale_fill_viridis_c(name = "Potential") +
    geom_contour(aes(x, y, z = pot), col = "black") +
    geom_path(data = traj, aes(x = x, y = y),
```

```
              col = "orange", size = 0.8) +
    mytheme
```

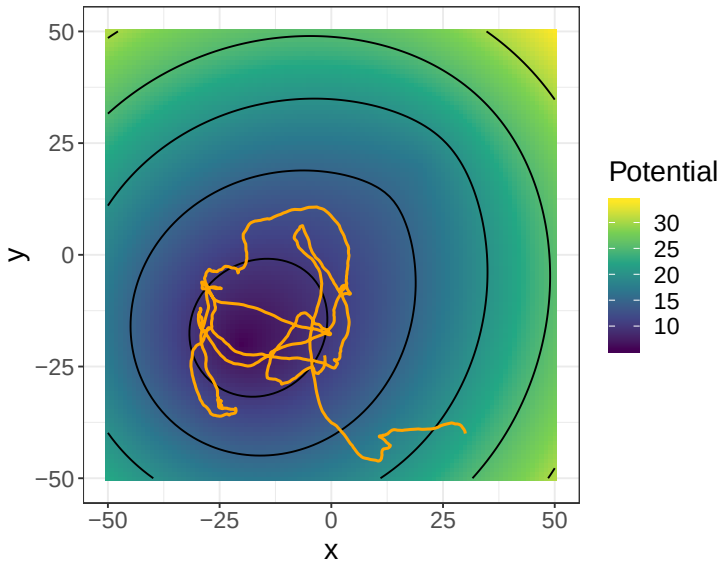


Lastly, we may include the stochastic change in velocity into the simulation, in the form of a Brownian motion and simulate a trajectory within the potential field

```
# SIMULATE 100 SECONDS OF BIRD MOTION

set.seed(34763)

# Set scale of the Brownian motion variance
sigma <- 3

# Total time
Time <- 100

# Initial location
r0 = c(30,-40)

# Set an initial velocity
v0 = c(-1, 5)

# Time step
dt <- 0.1

# steps
n <- Time/dt + 1 # plus the initial time
```

```
# Data frame to store location and velocity at each time step
traj <- data.frame(x = numeric(length = n),
                   y = numeric(length = n),
                   vx = numeric(length = n),
                   vy = numeric(length = n),
                   dvx = numeric(length = n),
                   dvy = numeric(length = n))


# Initial velocity drift
# - speed times neg gradient
gamma <- -sqrt(sum(v0^2)) * (beta[1] * eucGrad(r0[1], r0[2], mu) +
                               beta[2] * eucGrad(r0[1], r0[2], eta))

# - drift
drift <- (gamma - kappa * v0) * dt

# Initial velocity diffusion
diff <- rnorm(2, 0, sigma * sqrt(dt))

# Initial change in velocity
dv <- drift + diff

# fill in first row of the data frame
traj[1, ] <-  c(r0, v0, dv)


for(i in 2:n){

    # Update position
    r <- r0 + v0 * dt

    # update velocity
    v <- v0 + dv

    # update velocity drift
    # - speed times neg gradient
    gamma <- -sqrt(sum(v^2)) * (beta[1] * eucGrad(r[1], r[2], mu) +
                                  beta[2] * eucGrad(r[1], r[2], eta))

    # - drift
    drift <- (gamma - kappa * v) * dt
```

```
    # update velocity diffusion
    diff <- rnorm(2, 0, sigma * sqrt(dt))

    # update change in velocity
    dv <- drift + diff

    # fill in new row of the data frame
    traj[i, ] <-  c(r, v, dv)

    # iterate
    r0 <- r; v0 <- v

}

ggplot(omega) +
    geom_raster(aes(x = x, y = y, fill = pot)) +
    scale_fill_viridis_c(name = "Potential") +
    geom_contour(aes(x, y, z = pot), col = "black") +
    geom_path(data = traj, aes(x = x, y = y),
              col = "orange", size = 0.8) +
    mytheme
```



Note that the trajectory that incorporates the diffusion term is much longer than its deterministic counterpart, despite them being of the same duration. Under this model, the diffusion term not only perturbs the motion imposed by the potential field, but it is also the main generator of movement. For example, in chapter 3, I show how larger

diffusion parameters capture an increase in eagle activity towards midday.

Appendix E

Relative bias of potential-based velocity model parameters fitted to simulations

Mean relative bias (RB, in the vertical axis and colour gradient) of the posterior distribution of the parameters fitted to the simulated vantage point observations, and effect of number of vantage points and flight duration per vantage point, as predicted by the GAM. An $RB = 0$ means that the posterior is unbiased. An $RB = 1$ means that bias is as large as the parameter value that generated the data. An $RB = 2$, the bias is twice as large as the data-generating parameter, etc. Posterior RB is compared for different priors: uninformative (left), informative (centre - these priors come from parameters of a different eagle than the one use to generate the data), and data-generating distributions (right column - these are the distributions that were used to generate the data). The surface shows the effect of increasing the number of vantage points and the observed fight time per vantage point, on the relative bias.

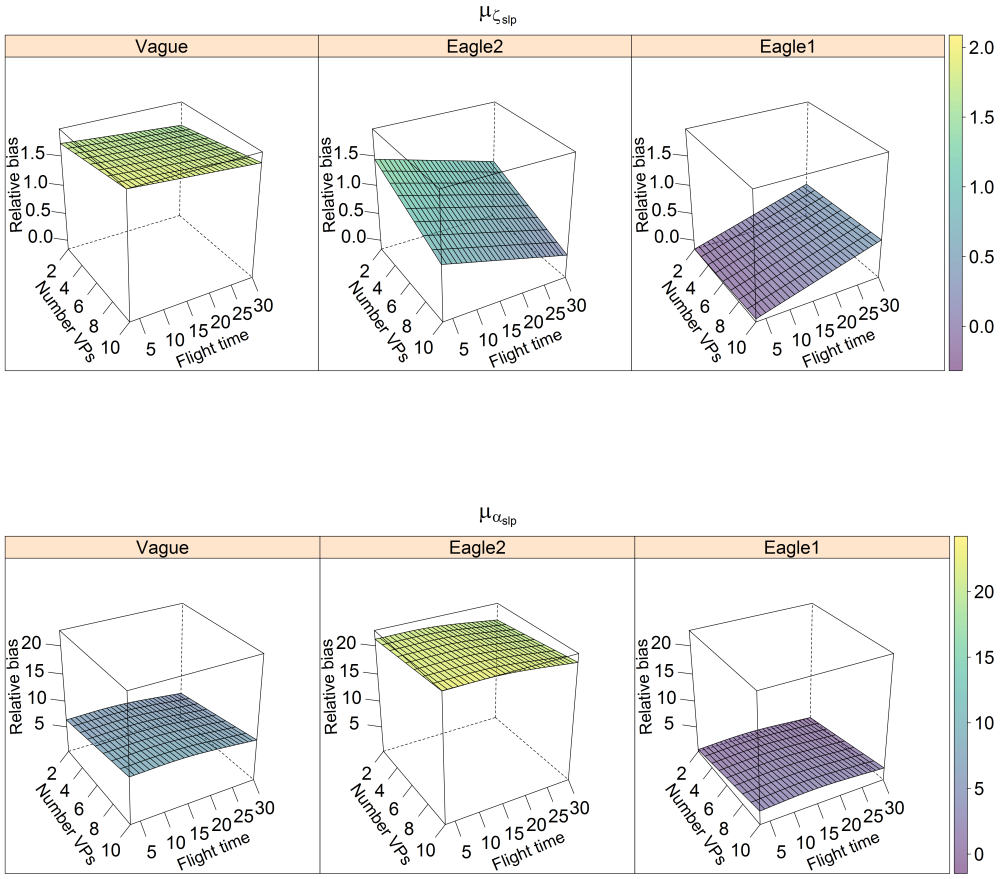


Figure E.1: $\mu_{\zeta_{slp}}$ is the mean drift towards slopes at noon across different days. $\mu_{\alpha_{slp}}$ is the mean change in drift with time of day, across different days.

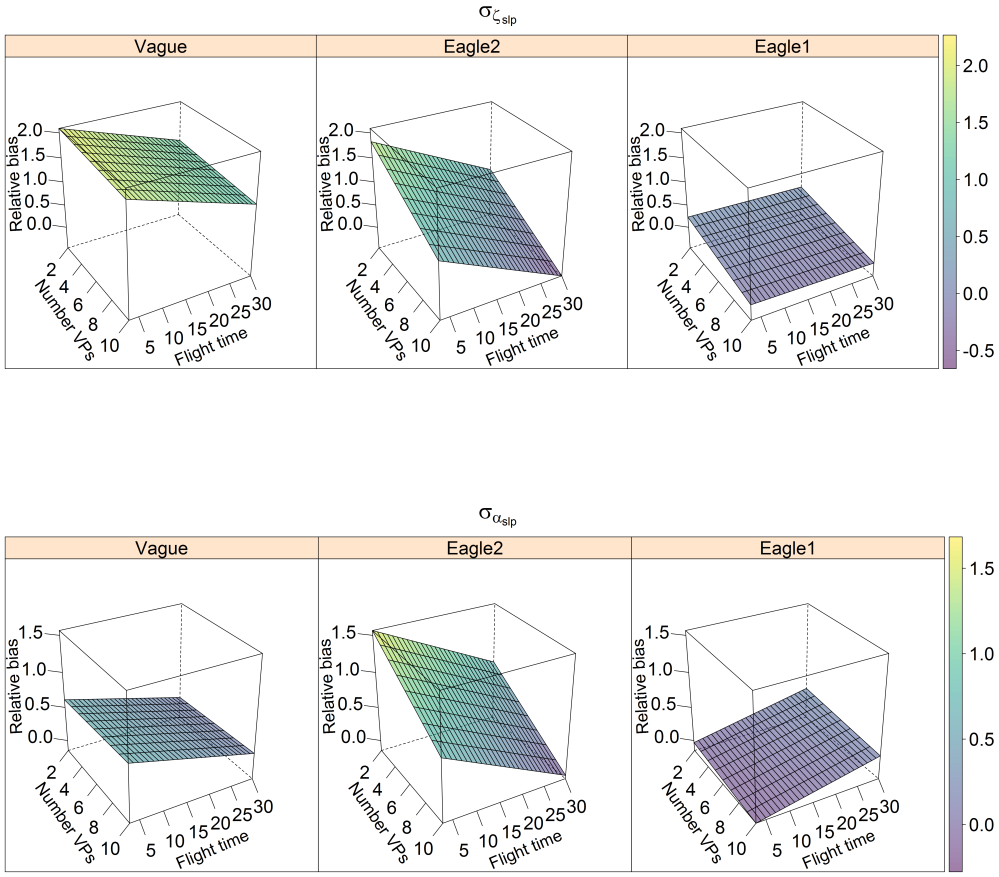


Figure E.2: $\sigma_{\zeta_{slp}}$ is the standard deviation of the change in drift towards slopes with time of day across different days. $\sigma_{\alpha_{slp}}$ is the standard deviation of the change in drift with time of day, across different days.

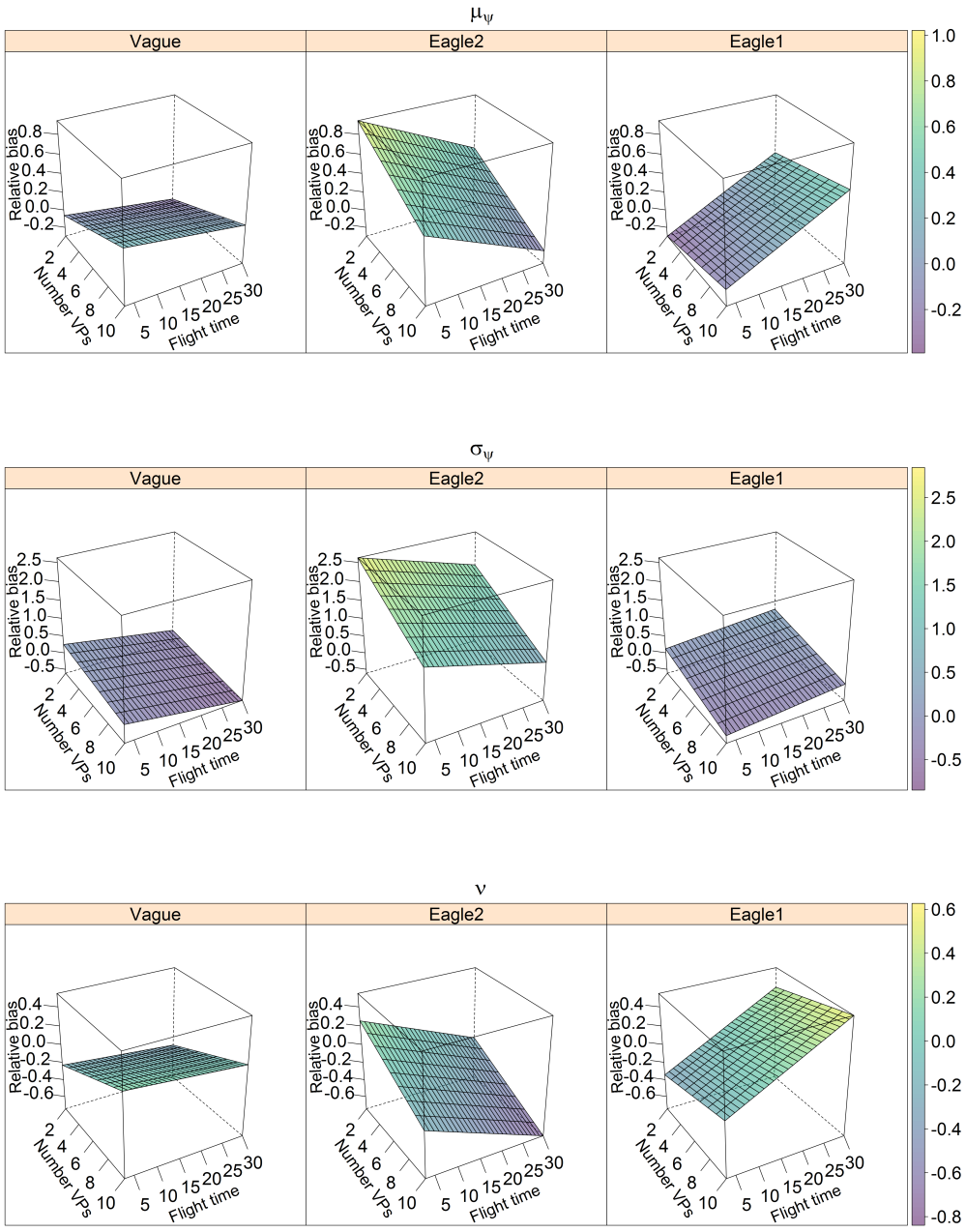


Figure E.2: μ_ψ corresponds to the mean change in diffusion with time of day across different days, whereas σ_ψ is the standard deviation of the change in diffusion with time of day across different days. ν corresponds to the maximum diffusion any day.

Appendix F

Overlapping coefficients of potential-based velocity model parameters fitted to simulations

Mean overlapping coefficient (OVL, in the vertical axis and colour gradient) of the posterior distributions of the potential-based velocity model parameters fitted to the simulated vantage point observations, as predicted by the GAM (see chapter 3). An OVL = 0 means that the posterior and the data-generating distributions are completely disjoint. An OVL = 1 that the posterior and the data-generating distributions agree exactly. Overlapping coefficient is compared for different priors: uninformative (left), informative (centre - these priors come from parameters of a different eagle than the one use to generate the data), and data-generating distributions (right column - these are the distributions that were used to generate the data). The surface shows the effect of increasing the number of vantage points and the observed flight time per vantage point, on the relative bias.

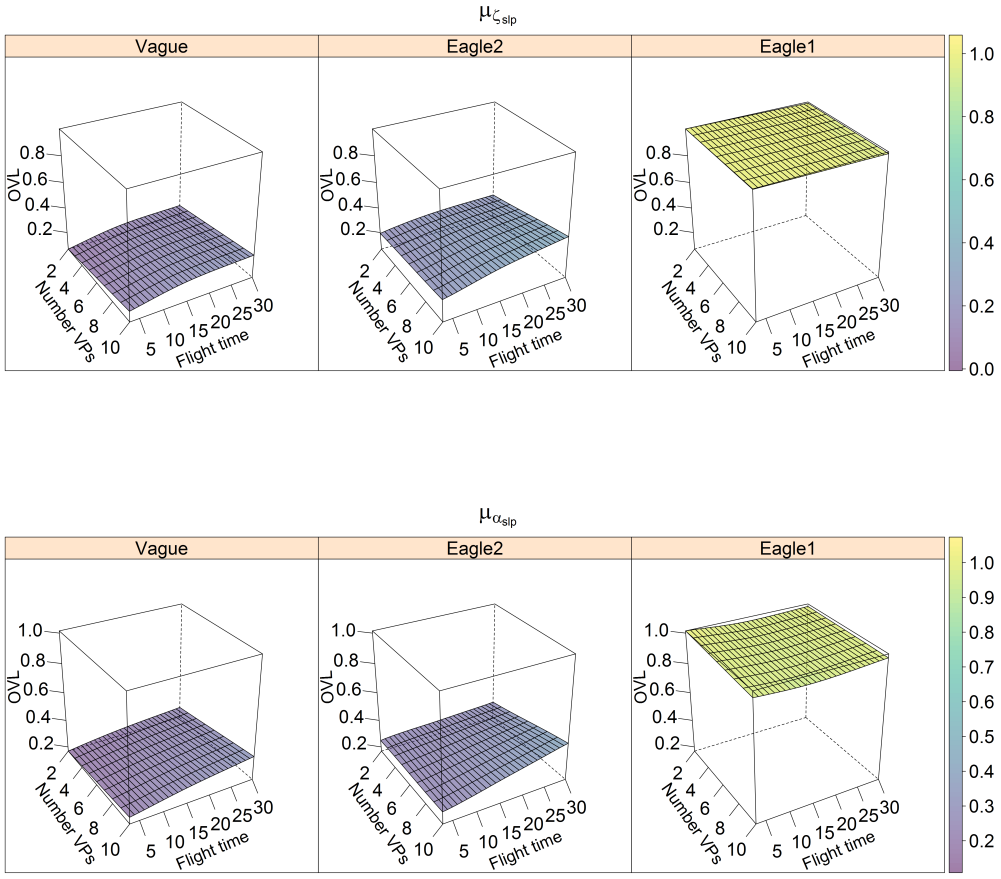


Figure F.1: $\mu_{\zeta_{slp}}$ is the mean drift towards slopes at noon across different days. $\mu_{\alpha_{slp}}$ is the mean change in drift with time of day, across different days.

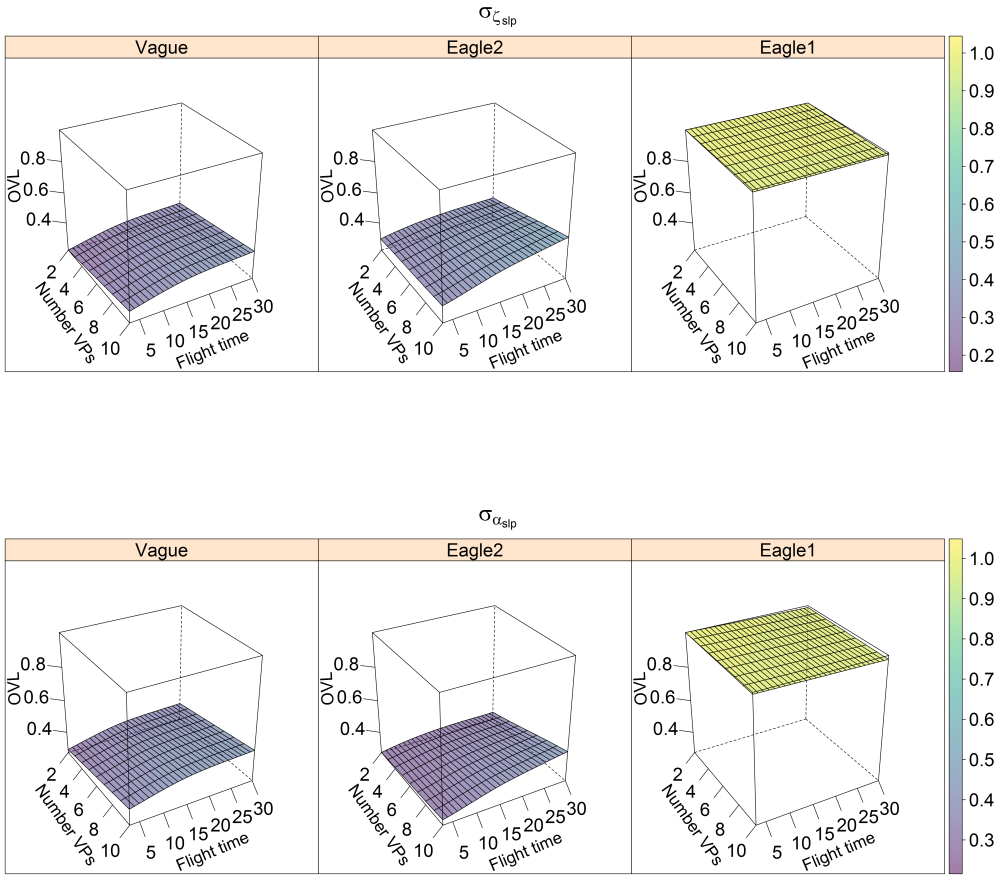


Figure F.2: $\sigma_{\zeta_{slp}}$ is the standard deviation of the change in drift towards slopes with time of day across different days. $\sigma_{\alpha_{slp}}$ is the standard deviation of the change in drift with time of day, across different days.

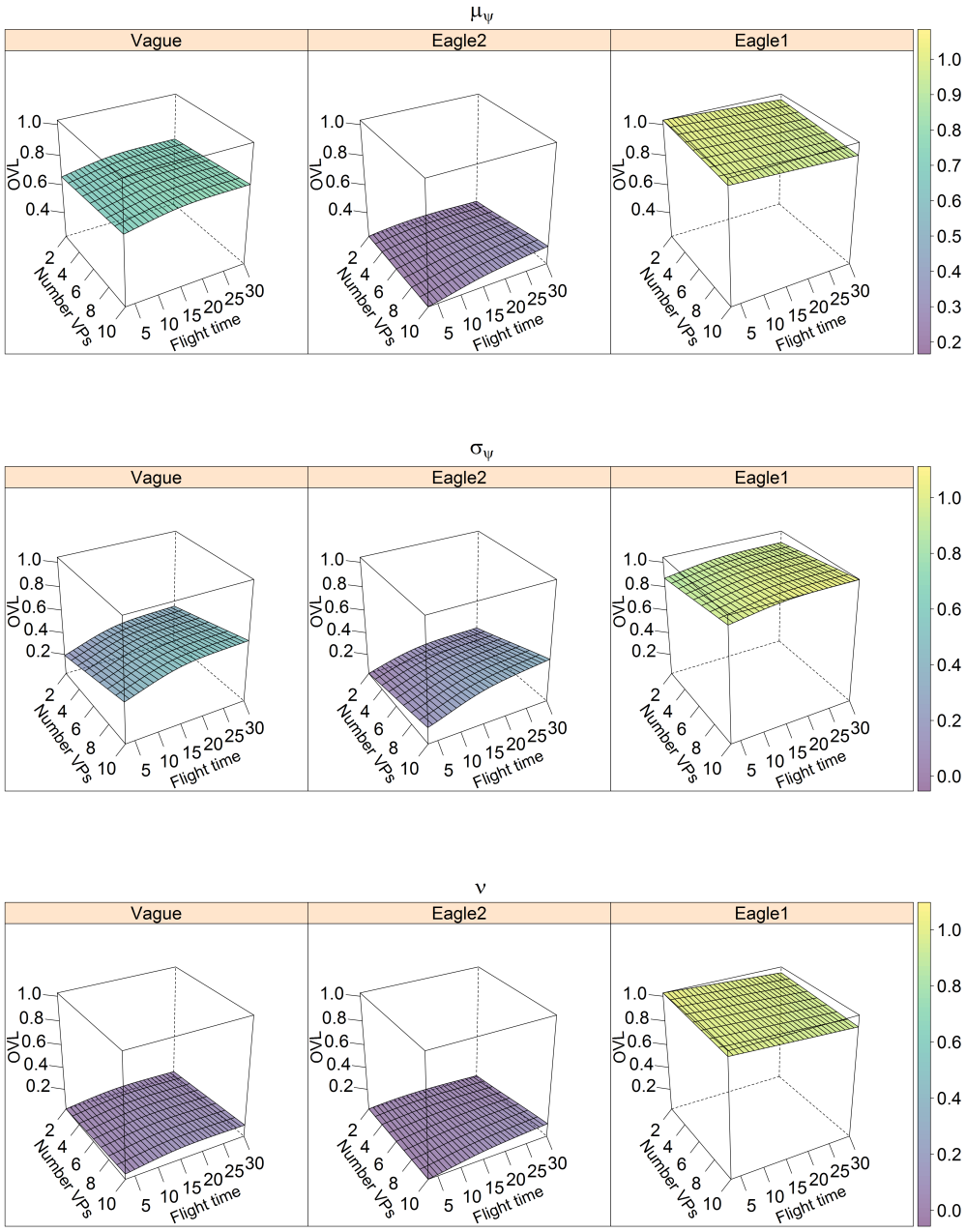


Figure F.2: μ_ψ corresponds to the mean change in diffusion with time of day across different days, whereas σ_ψ is the standard deviation of the change in diffusion with time of day across different days. ν corresponds to the maximum diffusion any day.