

# Evaluating, Implementing, and Applying Methods for Analyzing Animal Biotelemetry Data

DISSERTATION

zur Erlangung des Doktorgrades  
der Fakultät für Forstwissenschaften und Waldökologie  
der Georg-August-Universität Göttingen

vorgelegt von  
Johannes Michael Signer  
geboren in Arlesheim, Schweiz

Göttingen, 2015

1. Gutachter: Prof. Dr. Niko Balkenhol
2. Gutachterin: Prof. Dr. Kerstin Wiegand

Tag der mündlichen Prüfung: 29.01.2016

# Contents

<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Conceptual considerations . . . . .	2
1.2 Analytical considerations . . . . .	3
1.3 Technical considerations . . . . .	5
1.4 Aim of this thesis . . . . .	8
1.5 Structure of this thesis . . . . .	8
1.6 References . . . . .	9
<b>2 Error detection and handling in GPS data</b>	<b>13</b>
2.1 Abstract . . . . .	14
2.2 Introduction . . . . .	14
2.3 Methods . . . . .	15
2.4 Results . . . . .	18
2.5 Discussion . . . . .	18
2.6 References . . . . .	22
2.7 Appendix: A practical illustration . . . . .	23
<b>3 Reproducible Home Ranges (rhr)</b>	<b>27</b>
3.1 Abstract . . . . .	28
3.2 Introduction . . . . .	28
3.3 Methods . . . . .	29
3.4 Discussion . . . . .	33
3.5 Acknowledgements . . . . .	36
3.6 References . . . . .	36
3.7 Appendix 1: An brief introduction to the <code>rhr</code> package . . . . .	38
<b>4 Home range core areas: use and estimation</b>	<b>45</b>
4.1 Abstract . . . . .	46
4.2 Introduction . . . . .	46
4.3 Methods . . . . .	47
4.4 Results . . . . .	49
4.5 Discussion . . . . .	52

4.6	References . . . . .	52
<b>5</b>	<b>Does estimator choice influence our ability to detect changes in home range size?</b>	<b>55</b>
5.1	Abstract . . . . .	56
5.2	Background . . . . .	56
5.3	Methods . . . . .	58
5.4	Results . . . . .	61
5.5	Discussion . . . . .	66
5.6	Acknowledgements . . . . .	67
5.7	References . . . . .	67
5.8	Appendix 1: Additional Figures . . . . .	69
<b>6</b>	<b>Detecting influence of environment on animal movement</b>	<b>79</b>
6.1	Abstract . . . . .	80
6.2	Introduction . . . . .	80
6.3	Methods . . . . .	81
6.4	Results . . . . .	87
6.5	Discussion . . . . .	91
6.6	Acknowledgments . . . . .	92
6.7	References . . . . .	93
6.8	Appendix 1: Additional Figures and Tables . . . . .	94
<b>7</b>	<b>Conclusions</b>	<b>99</b>
7.1	Findings of this thesis . . . . .	99
7.2	Other aspects of animal biotelemetry . . . . .	100
7.3	Outlook . . . . .	101
7.4	References . . . . .	102
<b>8</b>	<b>Curriculum vitae</b>	<b>105</b>

# List of Figures

1.1	Movement ecology paradigm . . . . .	3
2.1	Conceptual illustration of how to regularize movement trajectories . . . .	17
2.2	Distribution of the times of day when relocations were recorded by year .	19
2.3	Distribution of the time when fixes were taken . . . . .	20
2.4	Illustration of data preparing process for one animal . . . . .	21
2.A1	Track split into bursts . . . . .	25
2.A2	Interpolated track . . . . .	25
3.1	Screen shots of the graphical user interface that we developed as part of a package for the statistical software package R . . . . .	31
3.A1	GUI for the <code>rhr</code> package to load data . . . . .	40
3.A2	Mapping data . . . . .	40
3.A3	Selecting analyses. . . . .	41
3.A4	An example of the automatically generated report for one animal. . . . .	42
3.A5	Plot of MCP home range estimate . . . . .	42
3.A6	Weekly home range size per animal and grouped by study area (different color) with a smoothed mean for each study area. . . . .	44
4.1	One realization for each simulation scenario . . . . .	48
4.2	Theoretical results of the area-independent method after Seaman and Powell (1990) to estimate core areas . . . . .	50
4.3	Results of different core area estimates . . . . .	51
5.1	Results of simulation scenario 1a . . . . .	62
5.2	Variability of home range size estimates (scenario 1a) . . . . .	63
5.3	Variability of home range estimators with respect to sample size . . . . .	64
5.4	Power to detect a significant effect of the amount of favorable habitat on home range size . . . . .	64
5.5	Results from empirical data . . . . .	65
5.A1	Underlying pattern of simulation . . . . .	69
5.A2	Realizations of simulated landscapes . . . . .	70
5.A3	Fraction of relocations located in favorable habitat as a function of the amount of favorable habitat in the landscape. . . . .	71
5.A4	One realization of simulation scenario 1a divided by month . . . . .	72
5.A5	Averaged home range sizes over time (scenario 1a) . . . . .	73
5.A6	Averaged changes of home range size over time (scenario 1a) . . . . .	74

*List of Figures*

5.A7 Results of simulation scenario 1b . . . . .	75
5.A8 Avergaed home range size over time (scenario 1b) . . . . .	76
5.A9 Averaged changes of home range size over time (scenario 1b) . . . . .	77
6.1 Illustrations of the four different random walk simulation scenarios . . . .	83
6.2 Semivariogram and the dependency of lag-specific diffusion rates on envi- ronmental covariates . . . . .	84
6.3 Power to detect influence of environmental covariates on animal movement	89
6.4 Lag-specific diffusion rates of a female red deer for altitude, temperature and rain . . . . .	90
6.A1 Power to detect an effect of environmental covariates on animal movement for the full data set . . . . .	95

# List of Tables

3.1	Analytical steps implemented in the <code>rhr</code> package . . . . .	30
3.2	A non-exhaustive overview of some software packages available to analyze animal movement data . . . . .	35
5.1	Parameterization of the simulation study . . . . .	61
6.1	Alternative hypotheses tested with the simulation study . . . . .	85
6.2	Results based on applying the statistical framework on movement data of a red deer individual . . . . .	88
6.A1	Error rates of different simulation scenarios and maximum time lags for the semivariogram . . . . .	94
6.A2	Error rates of different simulation scenarios and maximum time lags for the semivariogram . . . . .	96
6.A3	Results based on applying the statistical framework on movement data of a red deer with different amounts of missing data. . . . .	96





# Acknowledgements

This thesis would not have been possible without the many inspiring ideas, help, discussion, and continuous support from many other people. Especially, I would like to thank:

- Prof. Dr. Niko Balkenhol – for a very supportive and encouraging supervision and for giving me the freedom to explore new ideas.
- Prof. Dr. John Fieberg and Prof. Dr. Otso Ovaskainen – for showing me new perspectives on how to address scientific problems, supervising parts of this thesis, and their support.
- Prof. Dr. Kerstin Wiegand – for kindly serving as a co-referee.
- All coauthors – for providing new insights, challenging questions, and help with manuscript writing.
- Horst Reinecke, Flavia Koch, Marcus Meißner, Horst Leitner – for helping me to put the biology back to my at times a bit technical view on movement ecology.
- Hendrik Edelhoff and Femke Pflüger – for being in the office since the very beginning.
- All members of the Department of Wildlife Sciences Göttingen and Fieberg Lab – for making either side of the Atlantic such a pleasurable place to work.
- Sebastian Schoneberg – for teaming up on countless statistical challenges and Prof. Dr. Thomas Kneib for providing them.
- Katrin Bolbeth, Femke Pflüger, Sarah Fritts, and Maria Hendrich – for proofreading.
- R core Team – for R.
- Felix Kreier, Martin zur Nedden, and Björn Reineking – for sparking my interest in science as general and, more particularly, in ecological modeling.
- Everyone who is not mentioned here, but who contributed to my thesis in some way or other.
- My parents, family, and friends – for their encouragement and support.
- Maria – for her love, patience, and support.



# Summary

Movement ecology is an emerging discipline within ecology. Researchers addressing basic and applied questions within the movement ecology framework often rely on animal biotelemetry data. Biotelemetry sensors can collect different types of data (from tracking an animal's position to measuring its heart rate). Rapid technological advances in satellite based navigation and tracking devices enable researchers to track animals with smaller devices, leading to steadily increasing sampling rates. To characterize space requirements of tracked animals, the concept of a home range is often used.

This thesis starts with a general introduction (chapter 1), that connects the following chapters to the wider conceptual and analytical picture. The main parts of this thesis focus on the analysis of animal tracking data using home range analyses (chapters 2 through 5) and how to statistically test for the influence of (environmental) covariates on animal movement (chapter 6). Starting at the level of data management, chapter 2 highlights the need for a data model when working with tracking data. Next, a new package for program R is introduced that implements the previously discussed data model and provides functionality for the analysis of animal tracking data (chapter 3). Within the `rhr` (reproducible home ranges) package, the most commonly used estimators for home-range analyses are implemented. Further, functionalities for the automatic reporting of results and a graphical user interface are provided. After these technical aspects of tracking data, different methods for the estimation of home-range core areas (chapter 4) and home ranges as such (chapter 5) are discussed. Moving beyond home ranges, chapter 6 introduces a method to test which tests whether an animal's movement track is influenced by environmental covariates. Finally, this thesis concludes that: (i) amounts of data (tracking data and auxiliary environmental data) will continue to increase in the future; (ii) tools to handle, manage, and analyze them are of great importance (chapter 2 and 3); (iii) researchers should not get lost in methods and lose sight of the wider biological picture, rather, they should use established as well as new methods to answer interesting biological questions (chapter 4 through 6).

**Keywords:** home range; movement ecology; tracking data; biotelemetry; R



# Zusammenfassung

Bewegungsökologie ist ein junges und aufstrebendes Teilgebiet der Ökologie. Daten zum Bearbeiten grundlegender und angewandter Fragestellungen in der Bewegungsökologie werden oft mittels Biotelemetrie gewonnen. Die Biotelemetrie bietet die Möglichkeit unterschiedlichste Daten aufzuzeichnen (z.B. die Position von Tieren oder deren Herzfrequenz). Der schnelle technologische Fortschritt in der Satellitennavigation und bei der Entwicklung neuer Sensoren hat zur Folge, dass immer mehr und feiner aufgelöste Daten zur Verfügung stehen. Für Fragestellungen die die räumlichen Ansprüche von Tieren untersuchen werden oft so genannte Streifgebiete angewendet.

In der Einleitung (Kapitel 1) werden die Hauptteile dieser Arbeit in den größeren Kontext der Bewegungsökologie eingeordnet. Die folgenden Kapitel beschäftigen sich mit der Analyse von Telemetriedaten anhand von Streifgebieten (Kapitel 2 bis 5) und damit, wie der Einfluss von (Umwelt-)Kovariaten auf die Bewegung von Tieren statistisch getestet werden kann (Kapitel 6). Ein Datenmodell stellt einen unabdingbaren Teil jeder Telemetriestudie dar. Grundlegend für jede Telemetriestudie ist es, die Positionsdaten in einer zugänglichen Form verfügbar zu machen. So ist es möglich, diese Daten für weiterführende Analysen zur Verfügung zu stellen. Im 2. Kapitel dieser Arbeit, wird die Notwendigkeit eines geeigneten Datenmodelles diskutiert und ein solches vorgestellt. Im Folgenden wird das Paket `rhr` (reproducible home ranges) für die statistische Programmierumgebung R vorgestellt, welches das zuvor besprochene Datenmodell implementiert und Methoden zur Analyse von Telemetriedaten zur Verfügung stellt (Kapitel 3). Zusätzlich werden Funktionen zum automatischen Erstellen eines Ergebnisberichtes und eine grafische Benutzeroberfläche durch das Paket `rhr` vorgestellt. Nach diesen eher technischen Aspekten zur Analyse von Telemetriedaten werden verschiedene Methoden zur Schätzung von Kernarealen von Streifgebieten (Kapitel 4) und zur Schätzung von Streifgebieten an sich (Kapitel 5) verglichen. Im abschließenden 6. Kapitel dieser Arbeit wird eine neue Methode vorgestellt, die den Einfluss von (Umwelt-)Kovariaten auf die Bewegung von Tieren testet. Zusammenfassend kann festgestellt werden, dass (i) die in der Bewegungsökologie verfügbaren Datenmengen weiterhin zunehmen werden; (ii) Werkzeuge zur Verwaltung, Verarbeitung und Analyse dieser Datenmengen von großer Bedeutung sind (Kapitel 2 und 3) und (iii) die Entwicklung neuer analytischer Methoden schnell vorangeht. Trotzdem ist es wichtig, dass biologische Zusammenhänge nicht in einer Methodendiskussion untergehen und dass bewährte sowie neue Methoden eingesetzt werden um biologisch relevante Fragen zu beantworten (Kapitel 4 bis 6).

**Schlagwörter:** Streifgebiete; Bewegungsökologie; Telemetriedaten; Biotelemetrie; R



# 1

## Introduction

### Contents

---

<b>1.1</b>	<b>Conceptual considerations . . . . .</b>	<b>2</b>
<b>1.2</b>	<b>Analytical considerations . . . . .</b>	<b>3</b>
1.2.1	Mechanistic models of animal movement . . . . .	4
1.2.2	Analytical models of animal movement . . . . .	4
1.2.3	Statistical models of animal movement . . . . .	4
1.2.4	Home ranges in the context of movement ecology . . . . .	5
<b>1.3</b>	<b>Technical considerations . . . . .</b>	<b>5</b>
1.3.1	Collecting animal relocation data . . . . .	6
1.3.2	Managing relocation data . . . . .	6
1.3.3	Analyzing data . . . . .	7
<b>1.4</b>	<b>Aim of this thesis . . . . .</b>	<b>8</b>
<b>1.5</b>	<b>Structure of this thesis . . . . .</b>	<b>8</b>
<b>1.6</b>	<b>References . . . . .</b>	<b>9</b>

---

The ecology of wild animals is often difficult to study through direct observations in the field. Researchers have made use of technological advances such as telemetry (i.e., following animals remotely) or camera trapping, to better study wild animals in their natural habitats. Telemetry has become the method of choice for many questions a wildlife biologist may ask (e.g., What is the preferred habitat of a given population? Does a species migrate, and if so – how far? How do changing environmental conditions influence habitat selection of the species of interest?). Broadly speaking, most telemetry studies are comprised of three parts: (i) a *technical part* that deals with the collection and management of tracking data, (ii) an *analytical part* that deals with the analyses of the tracking data (i.e, which statistical models can be used), and (iii) finally, in the *conceptual part*, findings of the study are connected to the wider ecological theory.

The thesis presented here starts with describing the stage in which data have been obtained from tracking devices and the researcher is challenged with the questions on how to store, organize and analyze these data. In the introduction, I will briefly discuss prior

steps (i.e., how to obtain data) and how these findings connect to the more conceptual models of movement ecology and ecology in general. In the remaining chapters of this thesis, I will focus on the organization of movement data (chapter 2) and tools to analyze movement data (chapters 3 through 6). Finally, in the discussion I elaborate on the conclusions made from the findings of this thesis and attempt to develop an outlook of what lies ahead of us in the field of animal telemetry.

## **1.1 Conceptual considerations**

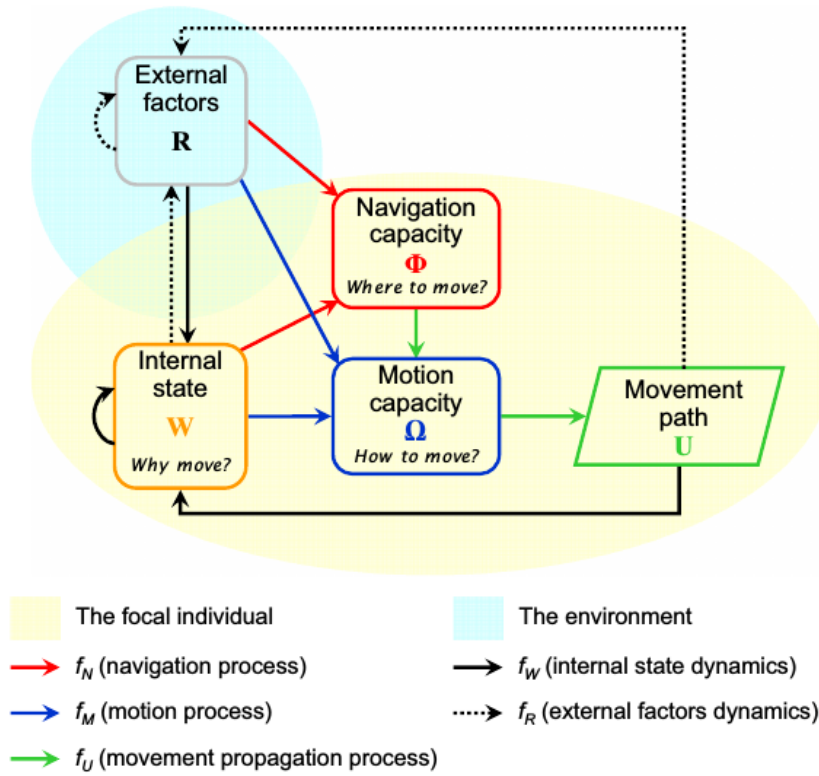
Naturalists have long recognized that animals move in a non-random fashion (Seton 1909; Burt 1943), and exhibit site fidelity (i.e., movement is restricted to a smaller area than would be expected if an animal's movement were governed by a random walk; Spencer et al. 1990; Powell 2000). Restricted animal movement is often theoretically conceptualized using the home range concept (Burt 1943; Spencer 2012; Börger et al. 2008).

Burt (1943, p. 351) defined home ranges as: "The area traversed by an individual in its normal activities of food gathering, mating, and caring for young. Occasional sallies outside the area, perhaps exploratory in nature, should not be considered as in part of the home range". Or one may think of an animal's home range in terms of different cognitive maps that summarize various spatio-temporal utility surfaces representing different resources for the animal's survival (e.g., food resources or shelter).

Burt's (1943) definition of home ranges is rather descriptive, focuses on the biological concepts, and does not explicitly take the underlying mechanisms into consideration. This often makes it challenging to generate mechanistic and testable hypotheses. For example, Powell (2000) discusses that maintaining a home range is associated with costs. The animal has to invest into remembering and developing a home range. These costs must be outweighed by the benefits of possessing a home range (Powell 2000). Burt's (1943) definition assumes the presence of a home range and does not allow to ask under which circumstances we would expect an animal to possess a home range (i.e. if the benefits of maintaining a home range, exceed the costs of a home range). Such questions can be accommodated within the movement ecology paradigm introduced by Nathan et al. (2008). The movement ecology paradigm considers animal movement as a fundamental process in the biology of animals, having profound influence on their population dynamics and distribution (Morales et al. 2010), interactions with other individuals (Delgado et al. 2014), habitat selection (Van Moorter et al. 2015), genetic structure (Cushman and Lewis 2010), and biodiversity as a whole (Jeltsch et al. 2013). The realized movement path of an animal results from its movement capacity, navigational capacity, internal state, and external factors (Nathan et al. 2008, Fig. 1.1). Although no work has explicitly investigated the conceptual relationship between movement ecology and home ranges, in theory, the movement paradigm can accommodate the home range concept.

In order to test conceptual hypotheses and predictions of movement ecology, one has to contrast them with real data. Analytical models are required to model observed data and confront those data with theoretical predictions. In the next chapter, I will discuss





**Figure 1.1:** The movement ecology paradigm as described by Nathan et al. (2008, p. 19054). The realized movement path ( $U$ ) of an animal is modeled as a function of an animal's current position, movement capacity, navigational capacity, internal states, and external factors (e.g., environment conditions). Copyright (2008) National Academy of Sciences, U.S.A.

different categories of analytical models that have been used to model movement data.

## 1.2 Analytical considerations

Analytical methods for understanding animal movement data, and home ranges in particular, have been subject of active research over the last few decades. Different schools of thought have emerged on how to model home ranges and how to think about them (Börger et al. 2008; Moorcroft 2012). Broadly, these approaches can be classified into three categories: mechanistic models, that attempt to model the underlying biological process; (ii) stochastic models, that model the observed movement path as a realization of a stochastic process; and (iii) statistical smoothing techniques, that characterize and summarize the observed pattern of a movement path.

### 1.2.1 Mechanistic models of animal movement

This group of analytical approaches considers the home range of an animal as the emergent property of various underlying internal (e.g., the navigation capacity of an animal) and external processes (e.g., seasonal weather dynamics). Typically, individual based models are used to model such processes (e.g., Buchmann et al. 2011; Anadón et al. 2012) and movement data are then used, among other data from the literature or other field studies, to parameterize such models and test different competing hypotheses on the observed patterns. Mechanistic models can produce very realistic patterns but are often difficult to parameterize due to a large number of required parameters. An alternative approach to mechanistic models are home range models that combine resource selection analysis and simple movement rules (e.g., synoptic home ranges; Horne et al. 2008).

### 1.2.2 Analytical models of animal movement

Different random walk models, that often originated in physics, have been used in ecology for decades to model animal movement (e.g., Skellam 1951). One of the difficulties of random walk models is the formulation of stochastic processes that have home-ranging behavior as an emerging property of the process with a realistic internal structure of space use. Early approaches used attractions to a focal point, often realized through the so-called Ornstein-Uhlenbeck process (Dunn and Gipson 1977), that can be conveniently fitted to real data. The main drawback of an Ornstein-Uhlenbeck process is its stationary distribution, which follows a bivariate normal distribution. Hence, the assumptions on the internal structure of the home range are unrealistic. More recent generalizations and extensions to the Ornstein-Uhlenbeck have been developed (e.g., Blackwell 1997; Fleming et al. 2014b; Fleming et al. 2014a).

### 1.2.3 Statistical models of animal movement

There are different statistical and geometrical approaches for obtaining estimates of a home range from observed movement data. These methods are typically either relatively simple bivariate smoothing techniques (such as kernel density estimation) or the calculation of geometric summaries of the observed point cloud. The kernel density estimation (Worton 1989) and the minimum convex polygon method are the most prominent. Geometric methods (such as minimum convex polygon) are strictly speaking not statistical, but have been included in this group of estimators because they are usually used for the same kind of studies as kernel density estimation. Although simple, these methods are useful for mapping, describing, and comparing observed patterns of an animal movement data but do not touch on the underlying process.

Statistical home range models allow estimation of an utilization distribution and of core areas from observed movement data. The utilization distribution is a bivariate (non-)parametric probability density function that describes the intensity of use in space. The utilization distribution is often used to find areas of high importance (i.e., core areas) and to determine the size of home ranges. Home range sizes are obtained by cutting the

utilization distribution at a given level (this is often referred to as the isopleth level) and only considering areas with a greater probability of use. Typically, a 95% level is used to estimate the home range size (meaning, the 5% of the utilization distribution with the lowest probability of use are disregarded) and a 50% level is used for core areas (see chapter 4 for more details).

There are some hybrid approaches that combine traditional statistical models with stochastic or mechanistic approaches. For example, the (dynamic) Brownian Bridge Movement Model incorporates a stochastic process (Brownian motion) to model the position of an animal between two relocations (Horne et al. 2007; Kranstauber et al. 2012). The synoptic home range model is an attempt to combine movement data with environmental data and estimate home ranges and habitat selection simultaneously (Horne et al. 2008).

Except for chapter 6, where the theory of stochastic processes is used, the methods of this thesis belong to the class of statistical models.

#### 1.2.4 Home ranges in the context of movement ecology

The different approaches to home range estimation have been developed largely in isolation from each other (Börger et al. 2008). Statistical modeling techniques are historically rooted in behavioral ecology and natural history, stochastic approaches often originate from more mathematically inclined researchers, and mechanistic models were developed with optimal foraging theory as a theoretical justification (Börger et al. 2008). The recently proposed movement paradigm (Nathan et al. 2008) has the potential to unify different views on home ranges by regarding home ranges as an emergent property of the movement process (i.e., the results of animals internal state, movement ability, navigational capacity, and external covariates). This is conceptually appealing and should make it relatively easy, at least conceptually, to think about costs and benefits of animals for maintaining a home range (Powell 2000).

### 1.3 Technical considerations

Once the conceptual and analytical framework for the analysis of telemetry data is in place, one can start to think about how to practically collect, store, manage, and analyze telemetry data. Such technical considerations are of particular importance for applied scientists and wildlife managers that wish to answer practical questions and provide scientifically sound recommendations to policy makers. First and foremost the question arises on how data are collected. Next, solutions for data handling need to be developed (e.g., how to store data in a structured and normalized way so that data can be accessed as needed). Finally, choices concerning the computational analysis tools and software need to be made. A final important step is reporting results of the analysis and making the findings available to the community in a reproducible way.

### 1.3.1 Collecting animal relocation data

In the early days of telemetry, the method was seen as a way to increase the "observational power" (Craighead and Craighead 1965) of a researcher. Telemetry was used to find an animal more easily so behavioral observations could be made. However, the technology was soon used for estimating the home range (e.g., Craighead 1976; Laundre and Keller 1981) and habitat selection of animals. The early telemetry systems used very high frequency (VHF) transmitters to track the position of animals. These VHF transmitters required a lot of manpower in the field to follow animals. VHF transmitters only provide sparse information of an animal's position, often accompanied by considerable errors. The development and availability of satellite navigation helped telemetry studies to enter a new era. The use of Global Position Systems (GPS) allowed researchers to follow animals at regular time intervals without having to be in the field (a detailed review of the development of GPS sensors is given by Tomkiewicz et al. 2010). Another major step forward was achieved through the wide availability of mobile communication. The wide spread availability of the Global System for Mobile communication (GSM) network and Short Messaging Service (SMS) made it possible for researchers to track animals via GPS and receive their position in near real time by SMS (Dettki et al. 2004). Such tracking systems have been used to implement decision support systems for adaptive management. For example, Wall et al. (2014) used a near real time system to monitor elephants in Kenya and Wal et al. (2015) used a near real time system to monitor red kites in Scotland and detect excursions.

### 1.3.2 Managing relocation data

With the increasing use of GPS tracking technologies in animal telemetry studies, the discipline of movement ecology entered the age of big data (Urbano et al. 2010; Kays et al. 2015). The number of relocations that are collected per individual has increased by one to three orders of magnitude. Telemetry studies with several thousands to tens of thousands of relocations per animal have become the norm, rather than the exception. Additionally, auxiliary data are becoming commonly available in increasingly finer spatial and temporal resolution (Kuenzer et al. 2014; Neumann et al. 2015). There are initiatives that automatically extract environmental data to corresponding movement paths (e.g., Dodge et al. 2013). The immense increase of relocation data and the availability of matching environmental data requires appropriate data models and software tools to handle, manage, and analyze such data effectively (Urbano et al. 2010; Calenge et al. 2009, and also chapter 2 of this thesis). There are different solutions to manage such data: from highly customizable but rather technical solutions (Urbano and Cagnacci 2014, e.g., the EuroDeer system) to web based frameworks (e.g., Dwyer et al. 2015) that make movement data readily available to the end user.

### 1.3.3 Analyzing data

The analysis of movement data can be tricky with respect to the underlying analytical model, the amount of data, and the assumptions of the methods of choice. From now on, I will only consider the statistical models for the analysis of animal movement data (as discussed above). The formulation of the underlying model has little to no importance when movement data are analyzed within the statistical framework. The most commonly used method for home range estimation (i.e., kernel density estimation) only requires data points to be independent and identically distributed (iid). The assumption of temporal independence has caused a lot of discussion and confusion in the home range community (Swihart and Slade 1985; Solla et al. 1999; Fieberg 2007). Data points originating from GPS telemetry studies are usually temporally autocorrelated in space and in time. Empirical studies and simulation studies (e.g., Fieberg 2007; Solla et al. 1999) have suggested that a violation of the independence assumption may be acceptable. Modern methods, such as autocorrelated kernel density estimation (Fleming et al. 2015), can account for the autocorrelation in the data but become analytically more sophisticated and already fall into the group of stochastic models.

The availability of large data sets led researchers to question if traditional home range estimators, that have been developed some time ago, are still valid. Kie et al. (2010) concluded that traditional estimators are likely to be used in future. Due to the wealth of available data, these estimators can be used for smaller time intervals (i.e., estimation of daily, weekly, or monthly ranges). Such estimates can then be analyzed through other modeling approaches (e.g., linear mixed models) to address new interesting questions (Fieberg and Börger 2012, this thesis chapter 5).

The inherently spatial notion of movement data leads to a tight coupling of Geographic Information Systems (GIS) and home range analyses. In the past, estimates of home ranges were often obtained from extensions of GISs. Such extensions exist for ArcGIS (e.g., Rodgers et al. 2007), QIGS, or OpenJump (Steiniger and Hunter 2012).

Alternatively to extensions of GISs, many high level statistical programming languages (e.g., R, python, Julia, matlab) provide an implementation of the statistical estimation methods used to estimate home ranges (e.g., kernel density estimation). There are specialized packages for home range estimation, for example for Program R (R Core Team 2015). `adehabitatHR` (Calenge 2006), `move` (Kranstauber and Smolla 2015), `rhr` (Signer and Balkenhol 2015, also chapter 3 of this thesis), are examples for packages that extend program R.

### Reproducible research

Reproducibility is one of the core tenets of science. Traditionally, the method section of any scientific manuscript provided sufficient information to reproduce the entire experiment and ideally allow another scientist to obtain the same results as obtained in the original experiment. However, increasingly complex settings – animal telemetry studies being a case in point – and inherently complex workflows for data analysis, have made reproducibility more challenging. Laver and Kelly (2008), for example highlighted that

many home range studies did not provide the software they used for data analyses. Outside the home range community there have been several calls for more reproducibility in general (McNutt 2014) and more specifically, to make computer codes available (Barnes 2010), and publish data in repositories (e.g., <http://datadryad.org/>). Because home range studies are inherently difficult to reproduce as a whole, it would be desirable to at least try to increase reproducibility at the level of analyses. More reproducible and standardized results are more credible and allow the reuse of data in meta studies (Laver and Kelly 2008; Gula and Theuerkauf 2013).

One way to implement reproducible research is literate programming (Knuth 1984). With literate programming, one no longer separates the writing process from the data analysis. Results from the data analysis are generated on the fly. Different programming environments have implementations of literate programming. For the analysis of movement data, literal programming initiatives for Program R (R Core Team 2015) are of interest. R `Sweave` (Leisch 2002), and more recently `knitr` (Xie 2013), provide functionality to mix  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  with R code. Recent developments towards more web driven data presentation, better interchangeability, and ease of use fostered the emergence of different markdown dialects. Markdown is a very lightweight markup language that allows quick and versatile documentation and literate programming. R has its own dialect markdown dialect, R-markdown (Baumer and Udwin 2015; Allaire et al. 2015), which together with `knitr` package, forms a powerful environment for reproducible research.

## 1.4 Aim of this thesis

The overall aim of this thesis is to evaluate existing methods for the estimation of home ranges, making methods available to researchers through implementation in user friendly software, and to evaluate the performance of methods under different real world scenarios. I hope that the results of this thesis will help applied researchers choose the correct method for the situation and free some of their research time to think about interesting and biological meaningful questions.

## 1.5 Structure of this thesis

After the introduction and before the conclusion and outlook in chapter 7, this thesis is structured into five chapters that discuss some of the issues commented on in the introduction in much greater detail.

**Chapter 2** highlights in more detail the need for a sound data model that can accommodate movement data and environmental data. Desirable properties for such a data model are discussed and its requirements are compared to readily available implementations of data models. Finally, I propose a new data model for medium scaled home range studies, describe its implementation, and illustrate it by applying it to red deer (*Cervus elaphus*) tracking data.

**Chapter 3** discusses available software methods for the estimation of home ranges and presents a new package to perform home range analyses for the statistical software R. The `rhr` package (reproducible home ranges) provides access to the most commonly used home range estimators. In addition, a graphical user interface and an automatic generation of a report summarizing results are implemented.

**Chapter 4** compares four different methods to estimate the core area of a home range. Core areas are areas that are used with a higher than expected intensity. I used simple simulations to generate data for three different scenarios and applied four different core area estimators.

**Chapter 5** presents a new approach to compare different home range estimators. While comparisons of home range estimators in the past often focused on type 1 and type 2 error rates of the estimators, I argue in this chapter that estimators should be compared with regard to their ability to detect an effect or the variability of estimates. In addition, this chapter investigates how sensitive results are to estimator choice, when the aim of the study is to detect changes in home range size.

**Chapter 6** presents an extension to the semi-variance approach for the analysis of movement data (Fleming et al. 2014b). It tests if environmental covariates influence movement rates of animals.

## 1.6 References

- Allaire, J., J. Cheng, Y. Xie, J. McPherson, W. Chang, J. Allen, H. Wickham, A. Atkins, and R. Hyndman (2015). *rmarkdown: Dynamic Documents for R*. R package version 0.8.1 (cit. on p. 8).
- Anadón, J., T. Wiegand, and A. Giménez (2012). “Individual-based movement models reveals sex-biased effects of landscape fragmentation on animal movement”. In: *Ecosphere* 3.7, art64 (cit. on p. 4).
- Barnes, N. (2010). “Publish your computer code: it is good enough”. In: *Nature* 467.7317, pp. 753–753 (cit. on p. 8).
- Baumer, B. and D. Udwin (2015). “R Markdown”. In: *Wiley Interdisciplinary Reviews: Computational Statistics* 7.3, pp. 167–177 (cit. on p. 8).
- Blackwell, P. (1997). “Random diffusion models for animal movement”. In: *Ecological Modelling* 100.1, pp. 87–102 (cit. on p. 4).
- Börger, L., B. D. Dalziel, and J. M. Fryxell (2008). “Are there general mechanisms of animal home range behaviour? A review and prospects for future research”. In: *Ecology letters* 11.6, pp. 637–650 (cit. on pp. 2, 3, 5).
- Buchmann, C. M., F. M. Schurr, R. Nathan, and F. Jeltsch (2011). “An allometric model of home range formation explains the structuring of animal communities exploiting heterogeneous resources”. In: *Oikos* 120.1, pp. 106–118 (cit. on p. 4).
- Burt, W. H. (1943). “Territoriality and home range concepts as applied to mammals”. In: *Journal of Mammalogy* 24.3, pp. 346–352 (cit. on p. 2).
- Calenge, C. (2006). “The package adehabitat for the R software: a tool for the analysis of space and habitat use by animals”. In: *Ecological Modelling* 197.3, pp. 516–519 (cit. on p. 7).
- Calenge, C., S. Dray, and M. Royer-Carenzi (2009). “The concept of animals’ trajectories from a data analysis perspective”. In: *Ecological Informatics* 4.1, pp. 34–41 (cit. on p. 6).
- Craighead, F. C. (1976). “Grizzly bear ranges and movement as determined by radiotracking”. In: *Bears: Their Biology and Management*, pp. 97–109 (cit. on p. 6).

## Chapter 1 Introduction

- Craighead, F. and J. Craighead (1965). “Tracking grizzly bears”. In: *BioScience* 15.2, pp. 88–92 (cit. on p. 6).
- Cushman, S. A. and J. S. Lewis (2010). “Movement behavior explains genetic differentiation in American black bears”. In: *Landscape ecology* 25.10, pp. 1613–1625 (cit. on p. 2).
- Delgado, M. d. M., V. Penteriani, J. M. Morales, E. Gurarie, and O. Ovaskainen (2014). “A statistical framework for inferring the influence of conspecifics on movement behaviour”. In: *Methods in Ecology and Evolution* 5.2, pp. 183–189 (cit. on p. 2).
- Dettki, H., G. Ericsson, and L. Edenius (2004). “Real-time moose tracking: an internet based mapping application using GPS/GSM-collars in Sweden”. In: *Alces* 40, pp. 13–21 (cit. on p. 6).
- Dodge, S., G. Bohrer, R. Weinzierl, S. C. Davidson, R. Kays, D. Douglas, S. Cruz, J. Han, D. Brandes, and M. Wikelski (2013). “The environmental-data automated track annotation (Env-DATA) system: linking animal tracks with environmental data”. In: *Movement Ecology* 1.3, pp. 1–14 (cit. on p. 6).
- Dunn, J. E. and P. S. Gipson (1977). “Analysis of radio telemetry data in studies of home range”. In: *Biometrics* 33.1, pp. 85–101 (cit. on p. 4).
- Dwyer, R. G., C. Brooking, W. Brimblecombe, H. A. Campbell, J. Hunter, M. Watts, and C. E. Franklin (2015). “An open Web-based system for the analysis and sharing and analysis of animal tracking data”. In: *Animal Biotelemetry* 3.1, pp. 1–11 (cit. on p. 6).
- Fieberg, J. (2007). “Kernel density estimators of home range: smoothing and the autocorrelation red herring”. In: *Ecology* 88.4, pp. 1059–1066 (cit. on p. 7).
- Fieberg, J. and L. Börger (2012). “Could you please phrase home range as a question?” In: *Journal of Mammalogy* 93.4, pp. 890–902 (cit. on p. 7).
- Fleming, C. H., J. M. Calabrese, T. Mueller, K. A. Olson, P. Leimgruber, and W. F. Fagan (2014a). “From fine-scale foraging to home ranges: A semi-variance approach to identifying movement modes across spatiotemporal scales”. In: *The American Naturalist* 183.5, E154–E167 (cit. on p. 4).
- Fleming, C. H., W. F. Fagan, T. Mueller, K. A. Olson, P. Leimgruber, and J. M. Calabrese (2015). “Rigorous home range estimation with movement data: a new autocorrelated kernel density estimator”. In: *Ecology* 96.5, pp. 1182–1188 (cit. on p. 7).
- Fleming, C. H., J. M. Calabrese, T. Mueller, K. A. Olson, P. Leimgruber, and W. F. Fagan (2014b). “Non-Markovian maximum likelihood estimation of autocorrelated movement processes”. In: *Methods in Ecology and Evolution* 5.5, pp. 462–472 (cit. on pp. 4, 9).
- Gula, R. and J. Theuerkauf (2013). “The need for standardization in wildlife science: home range estimators as an example”. In: *European Journal of Wildlife Research* 59.5, pp. 713–718 (cit. on p. 8).
- Horne, J. S., E. O. Garton, S. M. Krone, and J. S. Lewis (2007). “Analyzing animal movements using Brownian bridges”. In: *Ecology* 88.9, pp. 2354–2363 (cit. on p. 5).
- Horne, J. S., E. O. Garton, and J. L. Rachlow (2008). “A synoptic model of animal space use: simultaneous estimation of home range, habitat selection, and inter/intra-specific relationships”. In: *Ecological Modelling* 214.2, pp. 338–348 (cit. on pp. 4, 5).
- Jeltsch, F., D. Bonte, G. Pe’er, B. Reineking, P. Leimgruber, N. Balkenhol, B. Schröder, C. M. Buchmann, T. Mueller, N. Blaum, et al. (2013). “Integrating movement ecology with biodiversity research-exploring new avenues to address spatiotemporal biodiversity dynamics”. In: *Movement Ecology* 1.6 (cit. on p. 2).
- Kays, R., M. C. Crofoot, W. Jetz, and M. Wikelski (2015). “Terrestrial animal tracking as an eye on life and planet”. In: *Science* 348.6240, aaa2478 (cit. on p. 6).
- Kie, J. G., J. Matthiopoulos, J. Fieberg, R. A. Powell, F. Cagnacci, M. S. Mitchell, J.-M. Gaillard, and P. R. Moorcroft (2010). “The home-range concept: are traditional estimators still relevant with modern telemetry technology?” In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365.1550, pp. 2221–2231 (cit. on p. 7).
- Knuth, D. E. (1984). “Literate programming”. In: *The Computer Journal* 27.2, pp. 97–111 (cit. on p. 8).
- Kranstauber, B., R. Kays, S. D. LaPoint, M. Wikelski, and K. Safi (2012). “A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement”. In: *Journal of Animal Ecology* 81.4, pp. 738–746 (cit. on p. 5).
- Kranstauber, B. and M. Smolla (2015). *move: Visualizing and Analyzing Animal Track Data*. R package version 1.5.514 (cit. on p. 7).
- Kuenzer, C., M. Ottinger, M. Wegmann, H. Guo, C. Wang, J. Zhang, S. Dech, and M. Wikelski (2014). “Earth observation satellite sensors for biodiversity monitoring: potentials and bottle-



- necks". In: *International Journal of Remote Sensing* 35.18, pp. 6599–6647 (cit. on p. 6).
- Laundre, J. W. and B. L. Keller (1981). "Home-range use by coyotes in Idaho". In: *Animal Behaviour* 29.2, pp. 449–461 (cit. on p. 6).
- Laver, P. N. and M. J. Kelly (2008). "A critical review of home range studies". In: *The Journal of Wildlife Management* 72.1, pp. 290–298 (cit. on pp. 7, 8).
- Leisch, F. (2002). "Sweave: Dynamic generation of statistical reports using literate data analysis". In: *Compstat*. Springer, pp. 575–580 (cit. on p. 8).
- McNutt, M. (2014). "Journals unite for reproducibility". In: *Science* 346.6210, pp. 679–679 (cit. on p. 8).
- Moorcroft, P. R. (2012). "Mechanistic approaches to understanding and predicting mammalian space use: recent advances, future directions". In: *Journal of Mammalogy* 93.4, pp. 903–916 (cit. on p. 3).
- Morales, J. M., P. R. Moorcroft, J. Matthiopoulos, J. L. Frair, J. G. Kie, R. A. Powell, E. H. Merrill, and D. T. Haydon (2010). "Building the bridge between animal movement and population dynamics". In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365.1550, pp. 2289–2301 (cit. on p. 2).
- Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse (2008). "A movement ecology paradigm for unifying organismal movement research". In: *Proceedings of the National Academy of Sciences* 105.49, pp. 19052–19059 (cit. on pp. 2, 3, 5).
- Neumann, W., S. Martinuzzi, A. B. Estes, A. M. Pidgeon, H. Dettki, G. Ericsson, and V. C. Radeloff (2015). "Opportunities for the application of advanced remotely-sensed data in ecological studies of terrestrial animal movement". In: *Movement ecology* 3.1 (cit. on p. 6).
- Powell, R. A. (2000). "Animal home ranges and territories and home range estimators". In: *Research techniques in animal ecology: controversies and consequences* 442 (cit. on pp. 2, 5).
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria (cit. on pp. 7, 8).
- Rodgers, A. R., A. Carr, H. Beyer, L. Smith, and J. Kie (2007). *HRT: home range tools for ArcGIS* (cit. on p. 7).
- Seton, E. T. (1909). *Life-histories of northern animals: an account of the mammals of Manitoba*. Vol. 1. Charles Scribner's Sons, New York (cit. on p. 2).
- Signer, J. and N. Balkenhol (2015). "Reproducible home ranges (rhr): A new, user-friendly R package for analyses of wildlife telemetry data". In: *Wildlife Society Bulletin* 39.2, pp. 358–363 (cit. on p. 7).
- Skellam, J. G. (1951). "Random dispersal in theoretical populations". In: *Biometrika* 38.1-2, pp. 196–218 (cit. on p. 4).
- Solla, S. R. D., R. Bonduriansky, and R. J. Brooks (1999). "Eliminating autocorrelation reduces biological relevance of home range estimates". In: *Journal of Animal Ecology* 68.2, pp. 221–234 (cit. on p. 7).
- Spencer, S. R., G. N. Cameron, and R. K. Swihart (1990). "Operationally defining home range: temporal dependence exhibited by hispid cotton rats". In: *Ecology* 71.5, pp. 1817–1822 (cit. on p. 2).
- Spencer, W. D. (2012). "Home ranges and the value of spatial information". In: *Journal of Mammalogy* 93.4, pp. 929–947 (cit. on p. 2).
- Steiniger, S. and A. J. Hunter (2012). "OpenJUMP HoRAE-A free GIS and toolbox for home-range analysis". In: *Wildlife Society Bulletin* 36.3, pp. 600–608 (cit. on p. 7).
- Swihart, R. K. and N. A. Slade (1985). "Testing for independence of observations in animal movements". In: *Ecology* 66.4, pp. 1176–1184 (cit. on p. 7).
- Tomkiewicz, S. M., M. R. Fuller, J. G. Kie, and K. K. Bates (2010). "Global positioning system and associated technologies in animal behaviour and ecological research". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1550, pp. 2163–2176 (cit. on p. 6).
- Urbano, F. and F. Cagnacci (2014). *Spatial database for GPS wildlife tracking data: a practical guide to creating a data management system with PostgreSQL/PostGIS and R*. Springer Science & Business Media (cit. on p. 6).
- Urbano, F., F. Cagnacci, C. Calenge, H. Dettki, A. Cameron, and M. Neteler (2010). "Wildlife tracking data management: a new vision". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1550, pp. 2177–2185 (cit. on p. 6).
- Van Moorter, B., C. M. Rolandsen, M. Basille, and J.-M. Gaillard (2015). "Movement is the glue connecting home ranges and habitat selection". In: *Journal of Animal Ecology*, in press (cit. on p. 2).

## Chapter 1 Introduction

- Wal, R. van der, C. Zeng, D. Heptinstall, K. Ponnampereuma, C. Mellish, S. Ben, and A. Sidharthan (2015). “Automated data analysis to rapidly derive and communicate ecological insights from satellite-tag data: A case study of reintroduced red kites”. In: *Ambio* 44.4, pp. 612–623 (cit. on p. 6).
- Wall, J., G. Wittemyer, B. Klinkenberg, and I. Douglas-Hamilton (2014). “Novel opportunities for wildlife conservation and research with real-time monitoring”. In: *Ecological Applications* 24.4, pp. 593–601 (cit. on p. 6).
- Worton, B. J. (1989). “Kernel methods for estimating the utilization distribution in home-range studies”. In: *Ecology* 70.1, pp. 164–168 (cit. on p. 4).
- Xie, Y. (2013). *Dynamic Documents with R and knitr*. Vol. 29. CRC Press (cit. on p. 8).

# 2

## Error detection and handling in GPS data

This chapter is accepted as: **Signer, J.** & Balkenhol, N. Error detection and handling in GPS data *Jahrestagungsband der Gesellschaft für Wildtier- und Jagdforschung (GWJF)*.

**Authors contributions:** JS and NB conceived the conceptual framework, JS implemented the code, and JS and NB wrote the manuscript.

### Contents

---

<b>2.1</b>	<b>Abstract</b> . . . . .	<b>14</b>
<b>2.2</b>	<b>Introduction</b> . . . . .	<b>14</b>
<b>2.3</b>	<b>Methods</b> . . . . .	<b>15</b>
2.3.1	Data Model and Implementation . . . . .	15
2.3.2	Burstifying trajectories . . . . .	16
2.3.3	Data Quality . . . . .	16
2.3.4	Case Study . . . . .	18
<b>2.4</b>	<b>Results</b> . . . . .	<b>18</b>
2.4.1	Data model and implementation . . . . .	18
2.4.2	Case study . . . . .	18
<b>2.5</b>	<b>Discussion</b> . . . . .	<b>18</b>
<b>2.6</b>	<b>References</b> . . . . .	<b>22</b>
<b>2.7</b>	<b>Appendix: A practical illustration</b> . . . . .	<b>23</b>

---

## 2.1 Abstract

Wild animals are by their nature often difficult to observe and study. Hence, wildlife biologists often rely on remote data collecting devices such as Global Positioning Systems (GPS). Location data from GPS-collars have become popular for studying the behavior, resource or space use of wild animals. Many analytical methods (e.g., habitat selection, home range analysis) that are used to answer interesting biological questions and guide management decisions heavily depend on GPS location data. In order to obtain correct results, it is fundamental to screen GPS data prior to any analysis for potential errors. We give a short conceptual overview of the importance of a data model and the kind of errors that potentially occur when working with GPS data. We then discuss how wildlife professionals can handle these errors to improve the accuracy of location data and illustrate this with a data set from a red deer (*Cervus elaphus*) population from northern Germany.

## 2.2 Introduction

Data collected with GPS (Global Positioning System) telemetry are widely used to study wild animals and provide insights into their behavior (Cagnacci et al. 2010; Hebblewhite and Haydon 2010). Questions regarding habitat selection of animals (e.g., Thurfjell et al. 2014), behavioral states (e.g., Gurarie et al. 2009; Gurarie et al. 2015), space use (e.g., Laver and Kelly 2008), inter- and intraspecific interactions (e.g., Long and Nelson 2013) and movement in general (e.g., Avgar et al. 2013) can be addressed with telemetry data. With technological advances researchers are increasingly faced with larger data sets that open opportunities to investigate new questions, but are also accompanied by challenges of handling data adequately.

Dealing with GPS relocation data can be overwhelming and tools for managing data are needed (Urbano et al. 2010). Initiatives to harmonize and store data are available through data providers, such as EuroDeer or Movebank. Such database systems organize GPS telemetry data and can perform outlier detection. However, often GPS telemetry data are not stored in such databases and are only available through delimiter separated text files. Here, we focused on methods for situations where no database system with a sophisticated data model is available. Several other studies have introduced data models. For example, Calenge et al. (2009) introduced two different trajectory types (distinguishing trajectories of ordered data with and without time stamps). Urbano et al. (2010) developed a powerful database system based on PostGIS and PostgreSQL that is implemented with EuroDeer. Kranstauber et al. (2011) developed a data model for Movebank that is also tightly coupled to the `move` package for package R (R Core Team 2015; Kranstauber and Smolla 2015). Finally, Pebesma (2012) introduced a set of data models in R for handling spatio-temporal data implemented in the `spacetime` package. Ideally, a data model for animal tracking data would (1) integrate different types of trajectories (sensu Calenge et al. 2009); (2) accommodate attribute data of relocations, such as the habitat or time of the day when a relocation was recorded; (3)

provide methods to manage tracking data and interact with other (environmental) covariates and (4) be implemented in a widely used and freely available software solution. Once an appropriate data model is applied to tracking data, it is widely recognized that it is important to check data quality (Frair et al. 2010; Urbano et al. 2010; Bjørneraas et al. 2010) and detect outliers. Ideally, the analytical method accounts for erroneous observations (Patterson et al. 2008) and no cleaning of the data is necessary prior to the analysis. However, statistical methods that are able to account for observation errors (i.e., Bayesian state space models) are often non trivial to fit and usually require custom-written code. Since we often use analytical approaches that do not account for an observation model (e.g., non state space models), preprocessing of the data is required. This involves removing erroneous relocations that are beyond the study region, where the GPS failed, that do not fit a given sampling regime, do not have sufficient accuracy (often measured through the number of satellites used to obtain a relocation) or lie too far apart (i.e., distances that the animal was not able to move). In this article we start with introducing a data model for animal tracking data that builds on previously defined models. After a detailed description of the data model we show how this data model can be used to manage tracking data and to detect erroneous GPS locations based on space and time. Finally, we illustrate the implemented methods with relocation data from a red deer (*Cervus elaphus*) population from northern Germany and discuss the implementation, constraints and further plans for extensions.

## 2.3 Methods

### 2.3.1 Data Model and Implementation

A trajectory is characterized by a set of observations where the animal of interest was directly or indirectly observed. Each observation is characterized by an x and y coordinate that uniquely defines its position in space at a given point in time, and an ordering attribute (often a time stamp). In addition each relocation can have 0 to many additional attributes (e.g., habitat type, temperature, number of satellites used to obtain an observation). We distinguish three types of trajectories (following Calenge et al. (2009)) based on whether time is known and on the regularity of the spacing that are represented in three different classes. The simplest trajectory type consists only of ordered relocations, but no time stamp is available. The second trajectory type consists of coordinates with associated time stamps. Finally, the third trajectory type consists of coordinates with regular time stamps (i.e., two relocations are always separated by exactly the same amount of time). This type of trajectory can in most cases only be obtained through methods that regularize the trajectory. We distinguish two spatial components for trajectories: the relocations (the points where an animal was observed) and segments (the segments between two consecutive observations as linear interpolations between the start and end points). Each component can optionally have attribute data. By default a set of attributes are calculated for the segment attributes (e.g., time difference, length or turning angle; see also Calenge et al. (2009)).

We implemented the highlighted data model in Program R (R Core Team 2015) and the `rhr` package (Signer and Balkenhol 2015). Spatial positions of animals are implemented using spatial classes for R that are available through the `sp` package (Pebesma and Bivand 2005). Spatio-temporal data are represented using the `spacetime` package in R (Pebesma 2012). The `spacetime` package represents spatial data using `sp` classes and time using the `xts` package (Ryan and Ulrich 2014). Beside the implementation of the classes themselves, we provide methods to access and assign spatial components and their attributes of trajectories using R's standard methods for manipulating objects. We also provide a set of methods that make already existing methods available to interact with other spatial data (i.e., raster layers with environmental information). Further methods to split a trajectory, calculate basic summary statistics (e.g., number of relocations, time span, bounding box) and mean squared displacement are available.

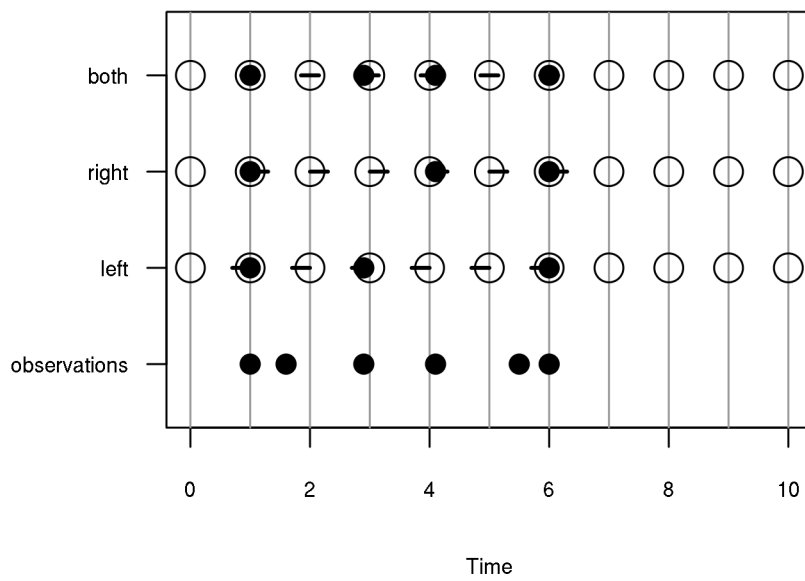
### 2.3.2 Burstifying trajectories

When tracking animals there are often periods of relocations followed by gaps with no observations. Such periods of continuous observations from the same animal are often referred to as bursts. In other words, a burst splits a trajectory of an animal in one or more complete sub-trajectories (i.e., there are no gaps). We deliberately did not implement bursts as part of our data model, but think they are useful in the sense of sub-trajectories. Instead we always work on trajectory objects. R provides lists as a very flexible data structure that can accommodate various subsets of trajectories. We provide methods to (1) regularize a trajectory (Fig. 2.1) and (2) burstify or split trajectories which results in a list of trajectories.

To regularize a trajectory, observations of an existing trajectory are aligned, within a time window, to a new trajectory (Fig. 2.1). New observations that are too far apart from any existing observation become empty observations, in the sense that they only contain time stamps and no relocations or segments. When regularizing trajectories the user can choose whether observations before, after or to both sides of a given new observation are considered (Fig. 2.1). The old observation that is closest in time to the new observation is then chosen. Once a trajectory is regular, we can apply the concept of bursts. Either the trajectory is split by some covariate (e.g., by day or year) or the specially designed method to burstify the trajectory is applied. Burstify will split a trajectory into sub-trajectories after a prespecified number of missing observation in a regular trajectory. This two step approach of regularizing, and splitting or bursting a trajectory will achieve similar results to building bursts into the data model, but retains more flexibility.

### 2.3.3 Data Quality

When working with GPS relocation data, errors can occur with regard to the actual measurement of the relocation (Graves and Waller 2006; Bjørneraas et al. 2010) and the study design. In the first case, missing or erroneous spatial data are recorded. In the second case, spatial data are recorded but outside the planned deployment period



**Figure 2.1:** Conceptual illustration of how a trajectory is regularized. The bottom row represents the actual observations. The top three rows illustrate the new trajectories (open circles) at regular time intervals with difference one. Black bars indicate the search radius that is either before, after or to both sides of the empirical observations. If a new empirical observations is within the search radius, it is considered in the new trajectory (black points).

of the sensor (e.g., a sensor was not on the animal, or remained on the animal after the study terminated). Relocations with missing spatial data are relatively easy to detect and eliminate. Relocations with erroneous GPS data can be detected through attribute data of fixes (e.g., number of satellites used to record a fix or the delution of precision [DOP] of the relocation recording). With an appropriate data model (like the one we suggested above) it should also be easier to filter relocations for certain periods (i.e., the exact duration a collar was deployed) or specific times/areas of interest.

### **2.3.4 Case Study**

To demonstrate the data, we used 107,468 relocations from a telemetry study from northern Germany (data are described in Reinecke et al. 2014). For this illustration, we prepared the data set in such a way to use the year with the most relocations available and to have two relocations per day (one at midnight and one at noon).

## **2.4 Results**

### **2.4.1 Data model and implementation**

We implemented the data model within the package `rhr` (Signer and Balkenhol 2015) for Program R (R Core Team 2015). The data model is recognized by all functions within the `rhr` package. Hence it is easy to prepare data (e.g., regularize or burstify trajectories) prior to the actual analysis (example code is available from the package website: [rhr.spamwell.net](http://rhr.spamwell.net)).

### **2.4.2 Case study**

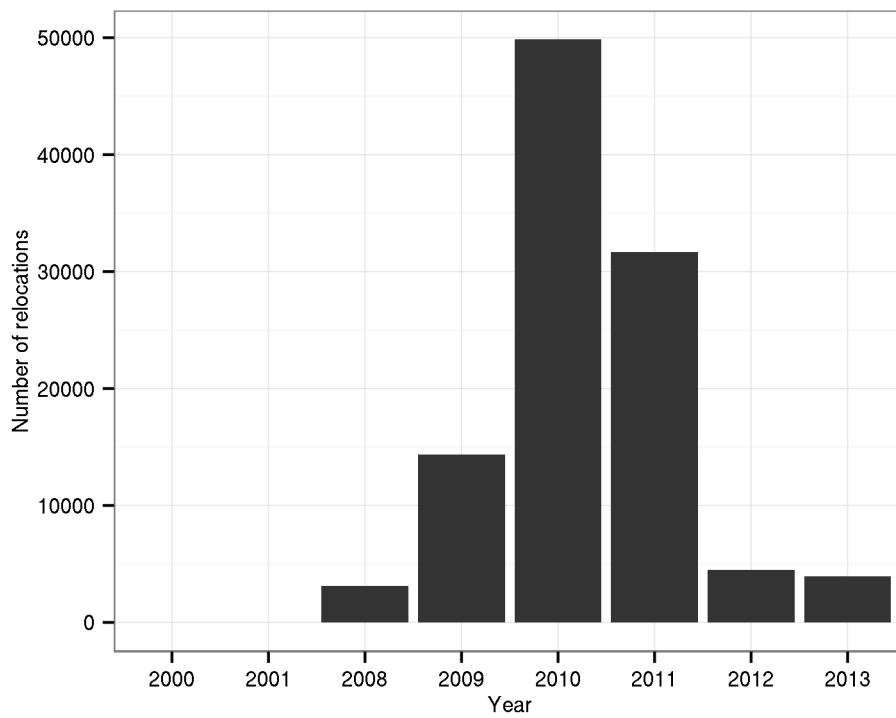
We read the data from a separator delimited text file. In the first step, we created a trajectory with space and time. We then visually determined that for the year 2010 most relocations were available (Fig. 2.2). Inspecting the distribution of the time of the day when relocations were recorded revealed that the intended interval of 6 hours was detectable, but significant noise was present (Fig. 2.3). Finally, we regularized the trajectory to only a maximum of two relocations per day: the relocations closest to noon and midnight (Fig. 2.4).

## **2.5 Discussion**

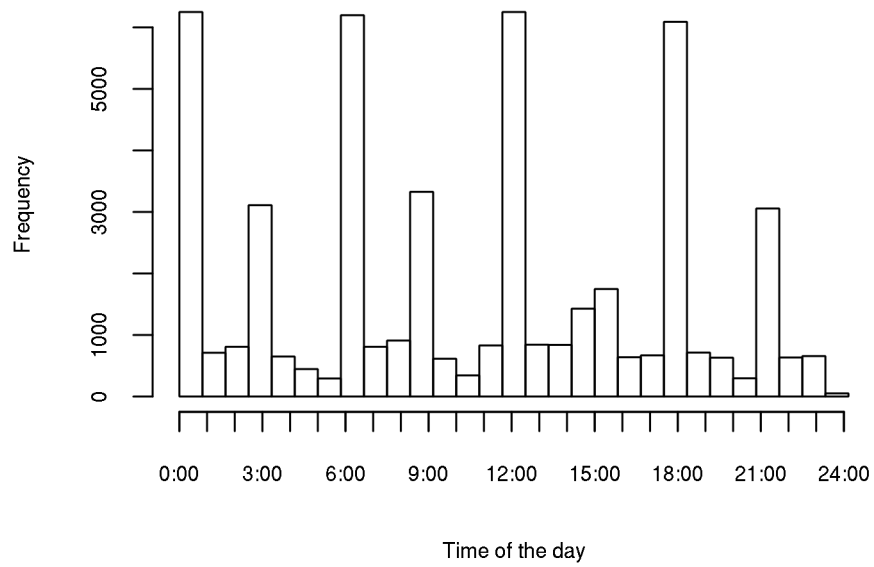
When working with GPS telemetry data, a solid data model is essential. A good data model can greatly improve efficiency and help to avoid errors during the analysis and detect errors within the data. We have extended previous works and implemented a data model for telemetry data for the `rhr` package for the analysis of telemetry data within Program R (R Core Team 2015).

The data model we implemented has the capabilities to represent GPS telemetry data (optionally also other telemetry data) and provides methods to query and manage

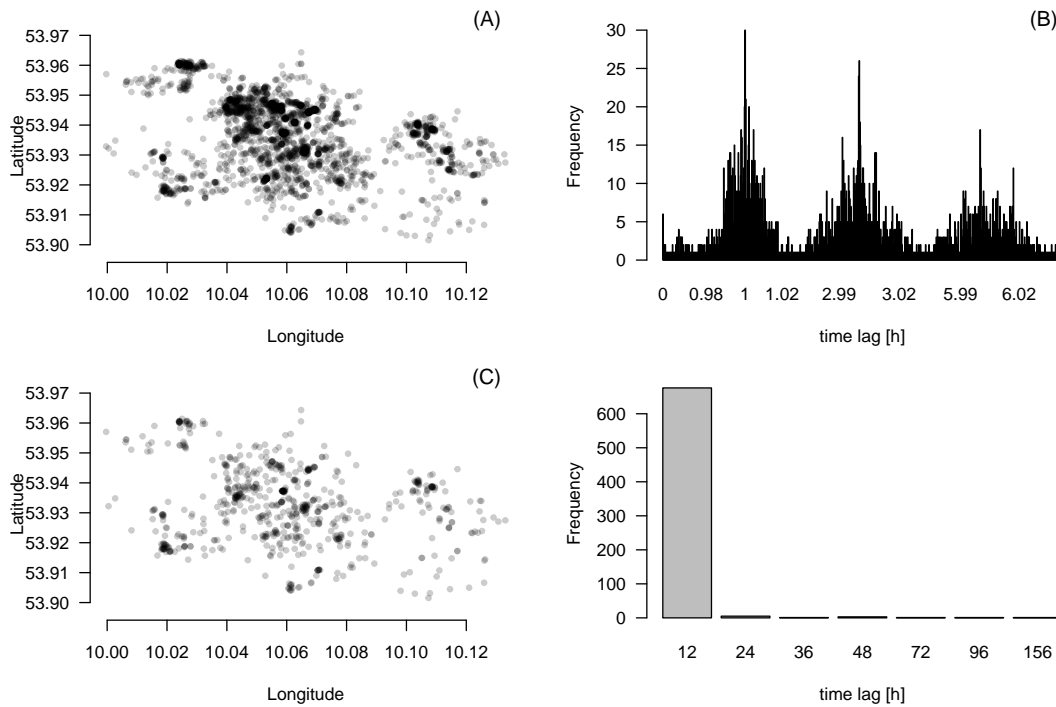




**Figure 2.2:** Distribution of the times of day when relocations were recorded by year. Relocations are unevenly distributed across years.



**Figure 2.3:** Distribution of the time when fixes were taken. While all GPS collars were programmed to take a fix every 6 hours, substantial noise occurred.



**Figure 2.4:** Illustration of data preparing process for one animal. Starting with the full track for 2010 (panel A) and the distribution of time lags between two relocations (panel B; note that lags above 7 hours occur but are omitted here). We regularized the track to consider only two relocations a day (panel C). We considered relocations closest to midnight and noon with a search window of 5 hours before and after the actual relocation. Some days with missing data persist (panel D).

telemetry data. Among others, methods are available to regularize a path (i.e., ensure that the time interval between two relocations is always identical), split a trajectory into two or more sub-trajectories based on some criteria or into bursts, if gaps between relocation periods exist. Furthermore, the data model provides an infrastructure to save attribute information for the relocations themselves, and also for the segments between two relocations.

We demonstrate the usefulness of such a data model with a data set of a red deer population from northern Germany. We were interested in correcting the trajectory to obtain two relocations per day, one at midnight and one at noon. Using the newly implemented methods, it was a simple task to create such a trajectory.

Further extensions of this data model could include methods to detect interactions between trajectories (animals), more sophisticated error detection mechanisms (e.g., routines suggested by Bjørneraas et al. 2010) and to move to the next step of analysis telemetry data using path segmentation and/or step selection function.

## 2.6 References

- Avgar, T., A. Mosser, G. S. Brown, and J. M. Fryxell (2013). “Environmental and individual drivers of animal movement patterns across a wide geographical gradient”. In: *Journal of Animal Ecology* 82.1, pp. 96–106 (cit. on p. 14).
- Bjørneraas, K., B. Van Moorter, C. M. Rolandsen, and I. Herfindal (2010). “Screening global positioning system location data for errors using animal movement characteristics”. In: *Journal of Wildlife Management* 74.6, pp. 1361–1366 (cit. on pp. 15, 16, 21).
- Cagnacci, F., L. Boitani, R. A. Powell, and M. S. Boyce (2010). “Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1550, pp. 2157–2162 (cit. on p. 14).
- Calenge, C., S. Dray, and M. Royer-Carenzi (2009). “The concept of animals’ trajectories from a data analysis perspective”. In: *Ecological informatics* 4.1, pp. 34–41 (cit. on pp. 14, 15).
- Frair, J. L., J. Fieberg, M. Hebblewhite, F. Cagnacci, N. J. DeCesare, and L. Pedrotti (2010). “Resolving issues of imprecise and habitat-biased locations in ecological analyses using GPS telemetry data”. In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365.1550, pp. 2187–2200 (cit. on p. 15).
- Graves, T. A. and J. S. Waller (2006). “Understanding the causes of missed global positioning system telemetry fixes”. In: *Journal of Wildlife Management* 70.3, pp. 844–851 (cit. on p. 16).
- Gurarie, E., R. D. Andrews, and K. L. Laidre (2009). “A novel method for identifying behavioural changes in animal movement data”. In: *Ecology Letters* 12.5, pp. 395–408 (cit. on p. 14).
- Gurarie, E., C. Bracis, M. Delgado, T. D. Meckley, I. Kojola, and C. M. Wagner (2015). “What is the animal doing? Tools for exploring behavioral structure in animal movements”. In: *Journal of Animal Ecology*, in press (cit. on p. 14).
- Hebblewhite, M. and D. T. Haydon (2010). “Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology”. In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365.1550, pp. 2303–2312 (cit. on p. 14).
- Kranstauber, B., A. Cameron, R. Weinzerl, T. Fountain, S. Tilak, M. Wikelski, and R. Kays (2011). “The Movebank data model for animal tracking”. In: *Environmental Modelling & Software* 26.6, pp. 834–835 (cit. on p. 14).
- Kranstauber, B. and M. Smolla (2015). *move: Visualizing and Analyzing Animal Track Data*. R package version 1.5.514 (cit. on p. 14).
- Laver, P. N. and M. J. Kelly (2008). “A critical review of home range studies”. In: *The Journal of Wildlife Management* 72.1, pp. 290–298 (cit. on p. 14).
- Long, J. A. and T. A. Nelson (2013). “Measuring dynamic interaction in movement data”. In: *Transactions in GIS* 17.1, pp. 62–77 (cit. on p. 14).
- Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos (2008). “State-space models of individual animal movement”. In: *Trends in ecology & evolution* 23.2, pp. 87–94 (cit. on p. 15).
- Pebesma, E. (2012). “spacetime: Spatio-temporal data in r”. In: *Journal of Statistical Software* 51.7, pp. 1–30 (cit. on pp. 14, 16).
- Pebesma, E. J. and R. S. Bivand (2005). “Classes and methods for spatial data in R”. In: *R News* 5.2, pp. 9–13 (cit. on p. 16).
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria (cit. on pp. 14, 16, 18).
- Reinecke, H., L. Leinen, I. Thißen, M. Meißner, S. Herzog, S. Schütz, and C. Kiffner (2014). “Home range size estimates of red deer in Germany: environmental, individual and methodological correlates”. In: *European journal of wildlife research* 60.2, pp. 237–247 (cit. on p. 18).
- Ryan, J. A. and J. M. Ulrich (2014). *xts: eXtensible Time Series*. R package version 0.9-7 (cit. on p. 16).
- Signer, J. and N. Balkenhol (2015). “Reproducible home ranges (rhr): A new, user-friendly R package for analyses of wildlife telemetry data”. In: *Wildlife Society Bulletin* 39.2, pp. 358–363 (cit. on pp. 16, 18).
- Thurfjell, H., S. Ciuti, and M. S. Boyce (2014). “Applications of step-selection functions in ecology and conservation”. In: *Movement Ecology* 2.4 (cit. on p. 14).
- Urbano, F., F. Cagnacci, C. Calenge, H. Dettki, A. Cameron, and M. Neteler (2010). “Wildlife tracking data management: a new vision”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1550, pp. 2177–2185 (cit. on pp. 14, 15).

## 2.7 Appendix: A practical illustration

### Data generation

To illustrate the data model with a small example, we consider a simple simulated data set. We assume that an animal was tracked for one year and its position was recorded every 12 hours. We randomly removed 50% of the data. The aim of this exercise is to regularize the data to one relocation at noon each day. This can be achieved in two ways: (i) splitting the data into bursts with the same sampling rate, or (ii) linearly interpolate missing locations.

```
library(lubridate)
library(rhr)
set.seed(123)

dat_true <- data.frame(
  x = cumsum(rnorm(720)),
  y = cumsum(rnorm(720)),
  time = ymd_hm("2015-01-01 12:00") + hours(0:719 * 12) +
    minutes(round(rnorm(720, 0, 5))) # add random error for time
)
dat <- dat_true[-(sample(1:nrow(dat_true), 360)), ]
```

### Tracks

With the simulated data we can create an object of class `RhrTrack` using the constructor `rhrTrack`.

```
coordinates(dat) <- ~x+y
track1 <- rhrTrack(dat, time = dat$time)
```

In the next step we will regularize the track so that relocations are aligned to a new and regular sequence of time steps. Once the time stamps are regular, we can either linearly interpolate missing steps or burstify the trajectory (i.e., select subsets of the trajectory without missing values).

### Regularize tracks

To regularize a trajectory we have to create a new sequence of dates and specify a direction and time window within which steps are selected:

```
new_time <- ymd_hm("2015-01-01 12:00") + hours(0:360 * 24)
track2 <- rhrRegularize(track1, new_time, "both", 30 * 60)
```

We could also be more restrictive on the time window (e.g., only a 10 min window):

```
new_time <- ymd_hm("2015-01-01 12:00") + hours(0:360 * 24)
track3 <- rhrRegularize(track1, new_time, "both", 10 * 60)
```

In both cases the track is aligned to the new sequence, but has missing observations.

```
is.regular(track2)
```

```
## [1] FALSE
```

```
is.regular(track3)
```

```
## [1] FALSE
```

This is because missing values were introduced in the data generation process.

### Burstify

Burstifying splits the track into different subtracks, each of which is regular (i.e., points are separated by equidistant time):

```
track4 <- rhrBurstify(track2)
plot(track4)
```

We can now check that all partial tracks are regular:

```
all(sapply(track4, is.regular))
```

```
## [1] TRUE
```

### Interpolate

An alternative method is to linearly interpolate missing points:

```
track5 <- rhrInterpolate(track2)
plot(track2)
points(rhrPoints(track5), col = "blue")
points(rhrPoints(track2), col = "blue")
```

Again, the track is now regular.

```
is.regular(track5)
```

```
## [1] TRUE
```

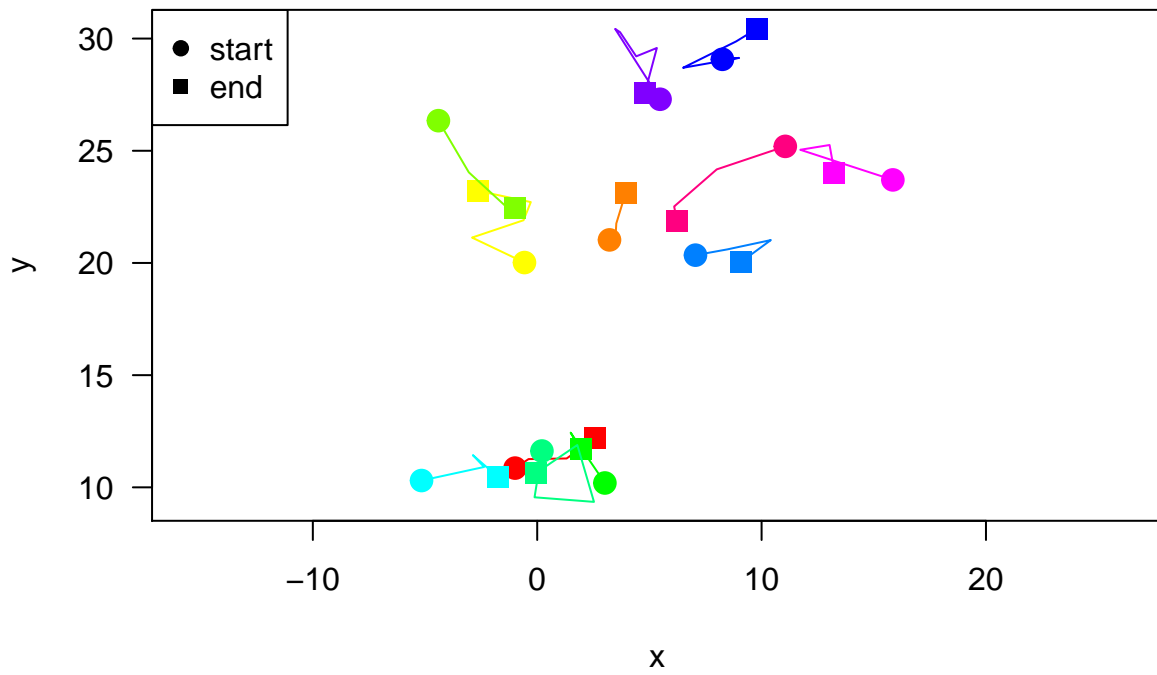


Figure 2.A1: Track split into burst. Only small parts of the trajectory are regular.

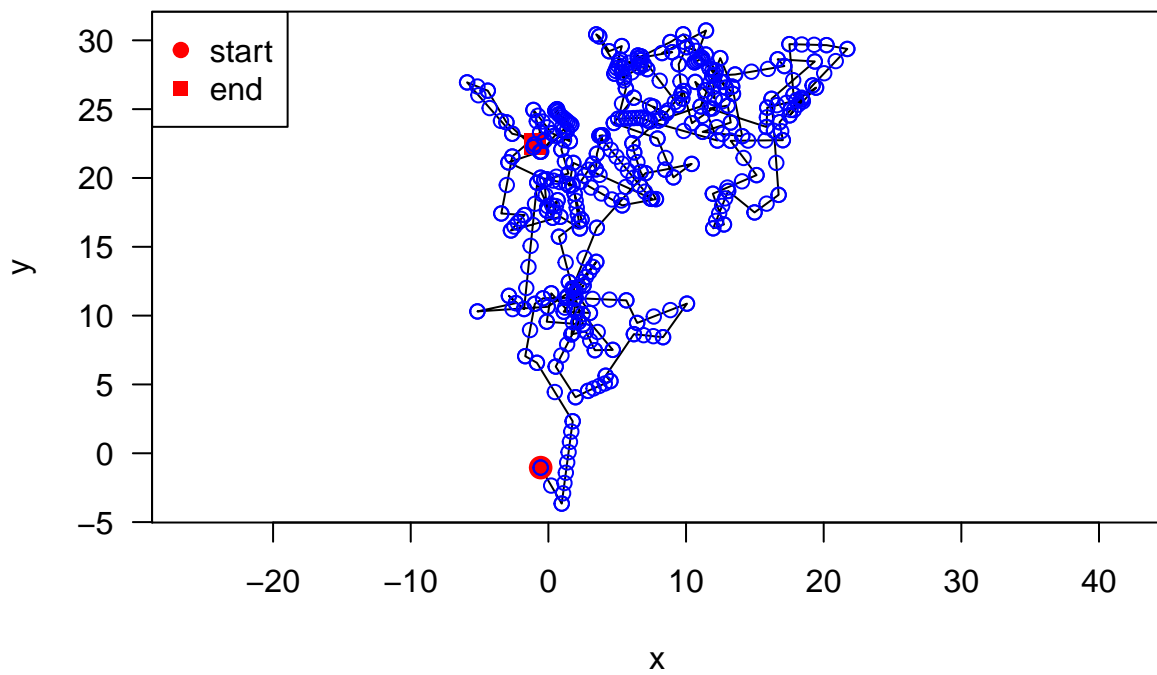


Figure 2.A2: Linear interpolation of missing relocations.

## **Outlook**

This short illustration presented the data model discussed in the main text of this chapter and how it may be used to regularize a trajectory. Reasons why a regular trajectory is required can be different. For example, if several animals are tracked, often tracking schedules are often different and for the analysis it may be desirable to have a standardized sampling rate.

Whether interpolating (Fig. 2.A2) or burstifying (Fig. 2.A1) a trajectory is the method of choice will depend on the question asked. Burstifying has the advantage that no 'new' data are generated and if there are only small gaps, it may be a suitable strategy to split the path in continuous chunks. Interpolating has the advantage that the whole path is regularly sampled. However, it might be desirable to have more sophisticated interpolation methods than linear interpolation.



# 3

## Reproducible Home Ranges (rhr): A New, User-Friendly R Package for Analyses of Wildlife Telemetry Data

This chapter is published as: **Signer, J. & Balkenhol, N.** (2015). Reproducible home ranges (rhr): A new, user-friendly R package for analyses of wildlife telemetry data. *Wildlife Society Bulletin..* DOI: <http://dx.doi.org/10.1002/wsb.539>.

**Authors contributions:** NB provided the initial idea, JS wrote most of the code, and JS and NB wrote the manuscript.

### Contents

---

<b>3.1</b>	<b>Abstract</b> . . . . .	<b>28</b>
<b>3.2</b>	<b>Introduction</b> . . . . .	<b>28</b>
<b>3.3</b>	<b>Methods</b> . . . . .	<b>29</b>
	3.3.1 Preliminary Data Analysis . . . . .	29
	3.3.2 Home Range Estimation . . . . .	31
	3.3.3 How to Use the Package . . . . .	33
<b>3.4</b>	<b>Discussion</b> . . . . .	<b>33</b>
<b>3.5</b>	<b>Acknowledgements</b> . . . . .	<b>36</b>
<b>3.6</b>	<b>References</b> . . . . .	<b>36</b>
<b>3.7</b>	<b>Appendix 1: An brief introduction to the rhr package</b> . . . .	<b>38</b>

---

### 3.1 Abstract

Analyzing wildlife tracking data frequently involves the estimation of home ranges. However, home range studies frequently lack important analytical steps, or only insufficiently report results. This makes it difficult for other researchers to evaluate, compare, and reproduce results from published home range studies. To facilitate more thorough home range analyses and reporting of analytical details, we developed a package for the statistical software package R that offers a user-friendly platform for comprehensive home range analyses. Importantly, the package automatically generates a summary report that contains all analytical parameters used during analyses, and lists the main findings. To improve usability of the package, we also provide a graphical user interface that can be called from R without any programming skills. We currently implemented the calculation of site fidelity, time to statistical independence, minimum convex polygon, kernel density estimation, Brownian Bridge Movement Model, Jennrich–Turner Ellipses, local convex hull, estimation of home range asymptote, and area-independent core-area estimation.

### 3.2 Introduction

Radiotracking studies are essential for understanding movement ecology, behavior, and habitat relationships of wild animals in space and time (Cagnacci et al. 2010). Relocation data obtained via tracking devices are often analyzed by estimating animal home ranges—the area traversed by an individual during its normal activities of food gathering, mating, and caring for its offspring (Burt 1943; Powell and Mitchell 2012). Even with new technological developments such as Global Positioning System telemetry or acceleration-based activity sensors, the estimation of home ranges remains one of the most useful ways for analyzing radiotracking data (Kie et al. 2010). Recent conceptual advances for home range estimation highlight the importance of precisely defining research questions and testable hypotheses before collaring wildlife and conducting home range analyses (Fieberg and Börger 2012). Once relocation data have been collected within this question-driven framework, home ranges can be analyzed with a myriad of different methods (Mohr 1947; Swihart and Slade 1985; Worton 1989; Seaman and Powell 1990; Getz et al. 2007; Laver and Kelly 2008). These methods have different advantages and limitations (Downs and Horner 2009; Lichti and Swihart 2011; Cumming and Cornélis 2012), so that no standardized method of home range analysis exists (Gula and Theuerkauf 2013).

Unfortunately, the variety of methods available for home range estimation also makes it difficult to objectively evaluate published results of many home range studies. This is because results and parameter values of home range analyses are often not reported adequately, and important analytical steps are often missing (Laver and Kelly 2008). Consequently, Laver and Kelly (2008) urged researchers to conduct certain analytical steps before actual home range analyses, and requested minimum editorial standards for reporting home range analyses (see Table 3.1). The critique of Laver and Kelly

(2008) fits into a broader debate in current science on reproducible research. In order to guarantee reproducibility—one of the central tenants of science (Fomel and Claerbout 2009)—several authors encouraged use of Free and Open Source Software for scientific computing (Rocchini and Neteler 2012), use of a scripted work flow (Peng 2011), and making workflows accessible to the entire scientific community (Barnes 2010). Several software products and extensions for Geographic Information Systems (GIS) are available to estimate home ranges with various methods (see Table 3.2). However, current software solutions are often closed-source and require commercial licenses (for the home range software itself or for the hosting GIS); for example, the Home Range Extension for ESRI ArcGIS by Rodgers et al. (2007), or software Ranges by Kenward and Hodder (1996). Other packages do not possess automatic reporting facilities (e.g., OpenHoRAE; Steiniger and Hunter 2012), or require programming skills, which not every wildlife manager or student has (e.g., the adehabitatHR package for R Calenge 2006). To our knowledge, no free and open-source package exists that implements all steps suggested by Laver and Kelly (2008) and allows a reproducible work flow for home range analyses. To improve the current situation and to provide a software platform implementing the recommendations of Laver and Kelly (2008), we present a new R package, `rhr`, that enables users to perform home range analyses using the most common estimators and keep track of all analytical steps, parameter values, and results. Importantly, all analytical steps can be called through a convenient graphical user interface (Fig. 3.1). We adhere to all principles of scientific reproducibility by using the freely available and open-source framework R (R Core Team 2015) for the package, and by introducing an automatic and standardized report that will facilitate reproducibility and meta-analysis of future home range studies. We encourage researchers to submit this report as an electronic supplement to any home range study they publish.

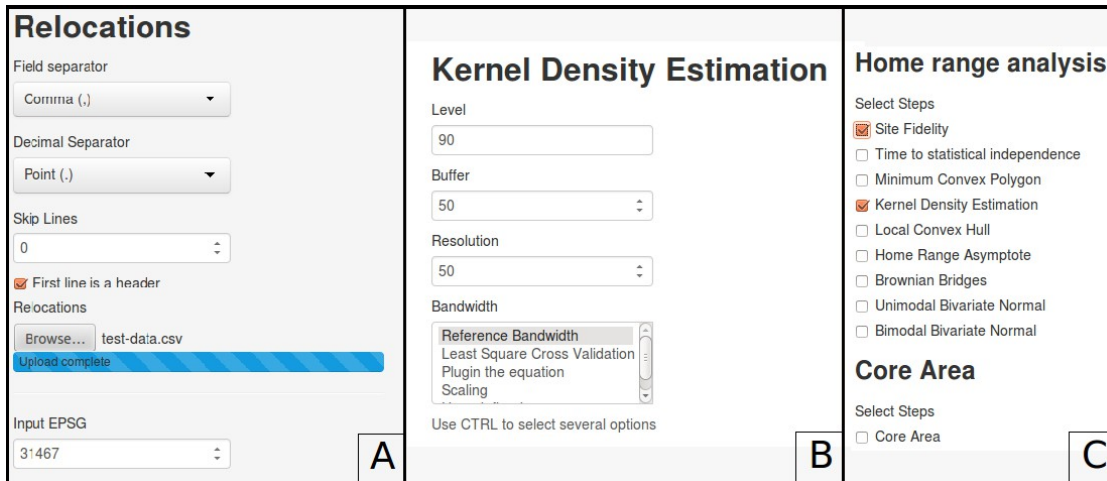
## 3.3 Methods

### 3.3.1 Preliminary Data Analysis

When estimating home ranges from animal tracking data, several preliminary analytical steps can be considered. First, before actual home ranges are estimated, site fidelity and time to statistical independence (TTSI) can be determined for each animal (Laver and Kelly 2008). Absence of site fidelity may be found if the animal is purely nomadic or was tracked only during a migratory phase. Estimating home ranges in the absence of site fidelity may contain little information (Munger 1984); home range analysis may not be the optimal analytical tool for such a data set, and the researcher may choose a different method, such as the Brownian Bridge Movement Model (Horne et al. 2008). We implemented tests for site fidelity as suggested by Spencer et al. (1990). Time to statistical independence finds the critical time interval, after which 2 subsequent relocations are statistically independent, which is important information for accurate interpretation of home range studies (Solla et al. 1999). To determine TTSI, Schoener's  $V$  is calculated for increasing time intervals until the expected value under independence

**Table 3.1:** The analytical steps implemented in the *rhr* package that we developed as part of a package for the statistical software package R that offers a user-friendly platform for comprehensive home range analyses. We closely followed the suggestions of Laver and Kelly (2008).

Step	Purpose	Implementation
Site fidelity	Animals may be monitored in non-stationary phases of their life cycle or over short time spans. Both cases can lead to the absence of site fidelity. In such cases, home range analysis may not be the appropriate analytical method, whereas a movement model such as Brownian Bridge methods may be more appropriate	Random trajectories are generated from the actual data. An index of linearity and the mean squared distance from the center of activity is compared between the true and simulated trajectories
Time to statistical independence (TTSI)	Some authors suggested (e.g., Laver and Kelly 2008) that TTSI should be calculated	This is the time interval that is required until 2 consecutive relocations are statistically independent. We implemented the approach of Swihart and Slade (1985)
Home range estimation	Estimating the actual home range is a main step for many telemetry studies. Different estimators are suitable to answer different research questions related to animal space use, so we provide several of the most commonly used parametric and nonparametric methods	Minimum convex polygon, kernel density estimation with various approaches to calculate the bandwidth (reference bandwidth, plugin the equation and least-square cross validation), parametric home range estimators (e.g., uni- and bimodal bivariate normal distribution) and local convex polygons are implemented
Core area estimation	To objectively identify whether certain areas of the home range are used more intensively than others, core areas are estimated	We implemented the method of Seaman and Powell (1990) to estimate the core area. This method determines the core area based on intensity of use, rather than an arbitrary isopleth
Report results and parameter values used	In order to be able to refer back to the analysis at a later point in time, it is important that all parameter values are stored	All parameters values used during analyses and the main findings are automatically summarized in a report if the package is used through the graphical user interface. All results are saved and available for further processing, with a GIS or R facilitating further exploration through summary statistic (e.g., mean and variance of home range size for different age classes) and more complex models



**Figure 3.1:** Screen shots of the graphical user interface that we developed as part of a package for the statistical software package R that offers a user-friendly platform for comprehensive home range analyses. Panel A shows how data can be uploaded from a text delimited file. All analytical steps have a set of properties that can be adjusted. As an example, properties for kernel density estimation are shown in panel B. Each of the analytical steps can be chosen individually or all together (as shown in panel C).

(i.e., 2) is reached within a given confidence interval (Swihart and Slade 1985). Time to statistical independence has been the subject of much debate in the past and it was argued that tracking data are inherently autocorrelated (Solla et al. 1999; Börger et al. 2006; Fieberg 2007). We fully agree with the problems that arise with TTSI and do not encourage researchers to subsample their data based on TTSI. However, we felt it was important to provide this functionality, because it was one of the key steps suggested by Laver and Kelly (2008).

### 3.3.2 Home Range Estimation

To estimate actual home ranges, we currently implemented 5 of the most commonly applied methods: Minimum Convex Polygon, Kernel Density Estimation, Local Convex Hull, Brownian Bridge Movement Model, and parametric home ranges. Minimum Convex Polygon is the simplest home range estimator and calculates a convex polygon of a certain percentage of relocations closest to the centroid of the relocation cloud. Minimum Convex Polygon is still commonly used, but several studies have discouraged its use for intraspecific studies (Nilsen et al. 2008). Kernel Density Estimation fits a kernel density surface to the relocations (Worton 1989) using a bivariate Kernel function (e.g., a Gaussian Kernel). Kernel Density Estimation is the most frequently used method in recent studies, and has been evaluated in numerous cases (Downs and Horner 2009; Lichti and Swihart 2011; Cumming and Cornélias 2012). One of the main challenges

when using Kernel Density Estimation is the determination of the required bandwidth parameter. Several methods exist to estimate the bandwidth from the data. Reference bandwidth estimation is a simple and fast method, but it has been shown to overestimate home ranges if data do not originate from unimodal distribution (Gitzen and Millsbaugh 2003). On the other hand, least-square cross validation often fails to find a solution (i.e., it does not converge; Seaman et al. 1999). Alternatively the bandwidth can be estimated with the Plug in the Equation method (Sheather and Jones 1991). Gitzen et al. (2006) found that there is no optimal method to estimate the bandwidth for Kernel Density Estimation. The choices of the bandwidth selection should depend on the questions asked. For example for single and tight clusters of relocations, the reference bandwidth may be appropriate; on the other hand, if relatively smooth outer contours are desired, the Plug in the Equation method is more suitable. Finally the least-square cross validation method has its strength in situations where tight clumps are to be identified (but see Gitzen et al. 2006, for a full comparison). No optimal method for determining the bandwidth exists; therefore, all 3 options for setting this parameter are available in the *rhr* package. For serially autocorrelated data, the Brownian Bridge Movement Model (Horne et al. 2007; Wells et al. 2014) is made available through a wrapper to the implementation in *adehabitatHR* (Calenge 2006). The Local Convex Hull method was more recently introduced by Getz and Wilmers (2004) and refined by Getz et al. (2007). Convex polygons are calculated for each point with a given number of neighbors and consequently merged. Three different methods exist and are implemented for determining the number of neighbors to be used (k nearest neighbors; all neighbors within a radius r; all neighbors that can be reached within a distance a, starting with the closest one). Local hulls are then ordered by size in decreasing order and merged at given levels. In addition to nonparametric methods such as Kernel Density Estimation or Brownian Bridge Movement Model, we also implemented 2 parametric methods: Jennrich–Turner Ellipses (Jennrich and Turner 1969) and their generalization for the bimodal case (Don and Rennolls 1983). These methods can be used for small, spatially uncorrelated sample sizes. One advantage of parametric home ranges is that they can be put in an information theoretic model-selection framework (Horne and Garton 2006), which is also implemented in the package. To assess reliability of estimated home ranges, the minimum number of relocations that are necessary for home range areas to reach an asymptote can be estimated. We currently implemented estimation of home range asymptotes for Minimum Convex Polygon and Kernel Density Estimation. As a cutoff value for the asymptote, we followed the suggestions of Laver and Kelly (2008): an asymptote is reached if a predefined confidence interval (e.g., 95%) of bootstrapped home range sizes falls within a specified range (e.g., 5%) of the home range area using all relocations for at least n consecutive estimates (Laver and Kelly 2008). We highlight that a lack of asymptotic home-range estimates does necessarily mean that an insufficient number of relocations has been gathered, because an asymptote may not be reached for various other reasons, such as seasonal home range shifts, or because an incorrect home range model is applied (see Gautestad and Mysterud 1993; Gautestad and Mysterud 1995). Finally, core areas are commonly estimated by a certain percentage home range

level (e.g., 25% or 50%). However, Seaman and Powell (1990), and Laver and Kelly (2008) argued for an area-independent core-area estimation. Thus, we implemented the method of Seaman and Powell (1990) where the threshold for defining the core area is the inflection point when plotting fraction of home range used (this is the fraction of the home range with a higher relative frequency than the pixel under investigation) against the fraction of maximum relative frequency (e.g., relative density of use in decreasing order).

### 3.3.3 How to Use the Package

The package requires R (Version 3.1 or higher) and a modern internet browser (Firefox 23 or higher is recommended and tested). Once the package is installed and loaded into R (detailed instruction are available from the package website <http://rhr.spamwell.net/>), the graphical user interface can be called with a single command. Data can be loaded from delimiter-separated text files through the graphical user interface (e.g., a file with comma-separated values or with tab-separated values) or directly from R during start-up, so that all R-supported file formats can be used. In addition to the previously described analytical functionality, the package also includes several functions to facilitate the handling of large telemetry data sets (e.g., by selecting and analyzing relocation data only for certain areas, time periods, or individuals). Once analyses are finished, results are presented within the user interface and a report is automatically generated that contains all results and parameter values used. In addition, all results are saved into a directory, including all figures (e.g., home range graphics). Results are saved in appropriate formats for further processing in R and GIS. A description on how to use the package and more details on the implemented methods is available from the package website. Additionally, we provide a mailing list (<https://listserv.gwdg.de/mailman/listinfo/rhr-discussion>) where questions, bug reports, and feature requests for the package can be addressed and discussed.

## 3.4 Discussion

The main motivation for writing this package was to increase reproducibility and enhance usability of home range analysis with the statistical software package R, with a specific focus on practitioners and applied researchers. We do not attempt to replace existing options for home range analyses (see Table 3.2), but hope to encourage researchers to conduct more complete home range analyses by using our R package. Currently, most methods implemented in the package are based on recommendations by Laver and Kelly (2008). However, we stress that we do not give preference to any specific method, and that our package still requires sound scientific thinking and analytical decision-making when estimating and interpreting home ranges. We agree with Gula and Theuerkauf (2013) that it would be convenient to have a standardized way to conduct home range analysis. However, because of the complexity of home range analyses and the necessity to carefully match research questions with appropriate methods (Fieberg and Börger 2012),

### *Chapter 3 Reproducible Home Ranges (rhr)*

we doubt that a cookbook approach for home range analysis is possible or desirable. However, we do believe that it is possible to report results of performed analyses in a uniform and exhaustive way. Thus, we hope that the `rhr` package is a first step toward more reproducible home range analyses. In upcoming versions of the package, we envisage to include methods considering environmental covariates (e.g., synoptic home ranges; Horne et al. 2008), static and dynamic animal interactions (Long et al. 2014), and path-level approaches, which find increasing consideration in wildlife research (Cushman 2010).



**Table 3.2:** A non-exhaustive overview of some software packages available to analyze animal movement data (CLI, command line interface; i.e., some level of programming are required by the user; GUI, graphical user interface).

Name	Platform	Interface	License <sup>1</sup>	References
adehabitathR	Any (runs within R)	GUI	GPL-V3	Calenge (2006)
Ranges	Windows and Macintosh	GUI	No information available (commercial)	Kenward and Hodder (1996)
Animove for Sextante	Any (runs within QGIS 1.8)	GUI	GPL-V3	Quantum (2012)
OpenHORA	Any (runs within Open-Jump)	GUI	GPL-V2	Steiniger and Hunter (2012)
Home Range Extension	Windows (runs within ArcView)	GUI	No information available (free to download)	Rodgers and Carr (1998)
Home Range Tools	Windows (runs within ArcGIS)	GUI	No information available (free to download)	Rodgers et al. (2007)
Animal Movement Extension	Windows (runs within ArcView)	GUI	No information available (free to download)	Hooge and Eichenlaub (1997)
Adobe	Windows (runs within ArcGIS)	GUI	No information available (free to download)	Laver (2005)
GME	Windows (runs within ArcGIS)	GUI	Individual license (free to download)	Beyer (2012)
rhr	Any (runs within R)	CLI & GUI	GPL-V3	This paper

<sup>1</sup>GPL: stands for GNU General Public License, which is a widely used software license that enables the end user to study and modify the source code of the software (see <http://www.gnu.org/licenses/gpl.html> for more information).

## 3.5 Acknowledgements

We thank D. Johnson, E. O. Garton, B. S. Cade, and one anonymous reviewer for comments and suggestions on a previous draft. M. Meißner and H. Reinecke generously provided us with a sample data set. Discussion with H. Reinecke and H. Edelhoff improved the development on the package. Additionally, the development of this package was eased and inspired partly from many other packages (e.g., *adehabitatHR* for R).

## 3.6 References

- Barnes, N. (2010). “Publish your computer code: it is good enough”. In: *Nature* 467.7317, pp. 753–753 (cit. on p. 29).
- Beyer, H. L. (2012). “Geospatial Modelling Environment (Version 0.7. 2.1)”. In: *URL: <http://www.spatialecology.com/gme>* (cit. on p. 35).
- Börger, L., N. Franconi, F. Ferretti, F. Meschi, G. De Michele, A. Gantz, and T. Coulson (2006). “An integrated approach to identify spatiotemporal and individual-level determinants of animal home range size”. In: *The American Naturalist* 168.4, pp. 471–485 (cit. on p. 31).
- Burt, W. H. (1943). “Territoriality and home range concepts as applied to mammals”. In: *Journal of Mammalogy* 24.3, pp. 346–352 (cit. on p. 28).
- Cagnacci, F., L. Boitani, R. A. Powell, and M. S. Boyce (2010). “Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1550, pp. 2157–2162 (cit. on p. 28).
- Calenge, C. (2006). “The package *adehabitat* for the R software: a tool for the analysis of space and habitat use by animals”. In: *Ecological Modelling* 197.3, pp. 516–519 (cit. on pp. 29, 32, 35).
- Cumming, G. S. and D. Cornéllis (2012). “Quantitative comparison and selection of home range metrics for telemetry data”. In: *Diversity and Distributions* 18.11, pp. 1057–1065 (cit. on pp. 28, 31).
- Cushman, S. A. (2010). “Animal movement data: GPS telemetry, autocorrelation and the need for path-level analysis”. In: *Spatial complexity, informatics, and wildlife conservation*. Springer, pp. 131–149 (cit. on p. 34).
- Don, B. and K. Rennolls (1983). “A home range model incorporating biological attraction points”. In: *The Journal of Animal Ecology*, pp. 69–81 (cit. on p. 32).
- Downs, J. A. and M. W. Horner (2009). “A Characteristic-Hull Based Method for Home Range Estimation”. In: *Transactions in GIS* 13.5-6, pp. 527–537 (cit. on pp. 28, 31).
- Fieberg, J. (2007). “Kernel density estimators of home range: smoothing and the autocorrelation red herring”. In: *Ecology* 88.4, pp. 1059–1066 (cit. on p. 31).
- Fieberg, J. and L. Börger (2012). “Could you please phrase home range as a question?”. In: *Journal of Mammalogy* 93.4, pp. 890–902 (cit. on pp. 28, 33).
- Fomel, S. and J. F. Claerbout (2009). “Reproducible research”. In: *Computing in Science & Engineering* 11.1, pp. 5–7 (cit. on p. 29).
- Gaustestad, A. O. and I. Mysterud (1993). “Physical and biological mechanisms in animal movement processes”. In: *Journal of Applied Ecology* 30.3, pp. 523–535 (cit. on p. 32).
- (1995). “The home range ghost”. In: *Oikos* 74.2s, pp. 195–204 (cit. on p. 32).
- Getz, W. M., S. Fortmann-Roe, P. C. Cross, A. J. Lyons, S. J. Ryan, and C. C. Wilmers (2007). “LoCoH: nonparameteric kernel methods for constructing home ranges and utilization distributions”. In: *PloS one* 2.2, e207 (cit. on pp. 28, 32).
- Getz, W. M. and C. C. Wilmers (2004). “A local nearest-neighbor convex-hull construction of home ranges and utilization distributions”. In: *Ecography* 27.4, pp. 489–505 (cit. on p. 32).
- Gitzen, R. A. and J. J. Millspaugh (2003). “Comparison of least-squares cross-validation bandwidth options for kernel home-range estimation”. In: *Wildlife Society Bulletin* 31.3, pp. 823–831 (cit. on p. 32).
- Gitzen, R. A., J. J. Millspaugh, and B. J. Kernohan (2006). “Bandwidth selection for fixed-kernel analysis of animal utilization distributions”. In: *Journal of Wildlife Management* 70.5, pp. 1334–1344 (cit. on p. 32).
- Gula, R. and J. Theuerkauf (2013). “The need for standardization in wildlife science: home range es-

- timators as an example". In: *European Journal of Wildlife Research* 59.5, pp. 713–718 (cit. on pp. 28, 33).
- Hooge, P. and B. Eichenlaub (1997). "Animal movement extension to arcview. ver. 1.1. Alaska Science Center-Biological Science Office". In: *US Geological Survey, Anchorage, AK, USA* (cit. on p. 35).
- Horne, J. S. and E. O. Garton (2006). "Selecting the best home range model: an information-theoretic approach". In: *Ecology* 87.5, pp. 1146–1152 (cit. on p. 32).
- Horne, J. S., E. O. Garton, S. M. Krone, and J. S. Lewis (2007). "Analyzing animal movements using Brownian bridges". In: *Ecology* 88.9, pp. 2354–2363 (cit. on p. 32).
- Horne, J. S., E. O. Garton, and J. L. Rachlow (2008). "A synoptic model of animal space use: simultaneous estimation of home range, habitat selection, and inter/intra-specific relationships". In: *Ecological Modelling* 214.2, pp. 338–348 (cit. on pp. 29, 34).
- Jennrich, R. and F. Turner (1969). "Measurement of non-circular home range". In: *Journal of Theoretical Biology* 22.2, pp. 227–237 (cit. on p. 32).
- Kenward, R. and K. Hodder (1996). *RANGES V: an analysis system for biological location data*. Natural Environment Research Council (cit. on pp. 29, 35).
- Kie, J. G., J. Matthiopoulos, J. Fieberg, R. A. Powell, F. Cagnacci, M. S. Mitchell, J.-M. Gaillard, and P. R. Moorcroft (2010). "The home-range concept: are traditional estimators still relevant with modern telemetry technology?" In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365.1550, pp. 2221–2231 (cit. on p. 28).
- Laver, P. (2005). "ABODE: Kernel home range estimation for ArcGIS, using VBA and ArcObjects". In: *User manual, Beta version*, pp. 1–62 (cit. on p. 35).
- Laver, P. N. and M. J. Kelly (2008). "A critical review of home range studies". In: *The Journal of Wildlife Management* 72.1, pp. 290–298 (cit. on pp. 28–33).
- Lichti, N. I. and R. K. Swihart (2011). "Estimating utilization distributions with kernel versus local convex hull methods". In: *The Journal of Wildlife Management* 75.2, pp. 413–422 (cit. on pp. 28, 31).
- Long, J. A., T. A. Nelson, S. L. Webb, and K. L. Gee (2014). "A critical examination of indices of dynamic interaction for wildlife telemetry studies". In: *Journal of Animal Ecology* 83.5, pp. 1216–1233 (cit. on p. 34).
- Mohr, C. O. (1947). "Table of equivalent populations of North American small mammals". In: *American Midland Naturalist* 37.1, pp. 223–249 (cit. on p. 28).
- Munger, J. C. (1984). "Home ranges of horned lizards (*Phrynosoma*): circumscribed and exclusive?" In: *Oecologia* 62.3, pp. 351–360 (cit. on p. 29).
- Nilsen, E. B., S. Pedersen, and J. D. Linnell (2008). "Can minimum convex polygon home ranges be used to draw biologically meaningful conclusions?" In: *Ecological Research* 23.3, pp. 635–639 (cit. on p. 31).
- Peng, R. D. (2011). "Reproducible research in computational science". In: *Science* 334.6060, p. 1226 (cit. on p. 29).
- Powell, R. A. and M. S. Mitchell (2012). "What is a home range?" In: *Journal of Mammalogy* 93.4, pp. 948–958 (cit. on p. 28).
- Quantum, G. (2012). "Development Team, 2012". In: *Quantum GIS geographic information system* (cit. on p. 35).
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria (cit. on p. 29).
- Rocchini, D. and M. Neteler (2012). "Let the four freedoms paradigm apply to ecology". In: *Trends in Ecology & Evolution* 27.6, pp. 310–311 (cit. on p. 29).
- Rodgers, A. R. and A. P. Carr (1998). "HRE: the home range extension for ArcView". In: *Ontario Ministry of Natural Resources, Centre for Northern Forest Ecosystem Research, Thunder Bay, Ontario, Canada* (cit. on p. 35).
- Rodgers, A. R., A. Carr, H. Beyer, L. Smith, and J. Kie (2007). *HRT: home range tools for ArcGIS* (cit. on pp. 29, 35).
- Seaman, D. E., J. J. Millsbaugh, B. J. Kernohan, G. C. Brundige, K. J. Raedeke, and R. A. Gitzen (1999). "Effects of sample size on kernel home range estimates". In: *The Journal of Wildlife Management* 63.2, pp. 739–747 (cit. on p. 32).
- Seaman, D. E. and R. A. Powell (1990). "Identifying patterns and intensity of home range use". In: *Bears: their biology and management*, pp. 243–249 (cit. on pp. 28, 30, 33).
- Sheather, S. J. and M. C. Jones (1991). "A reliable data-based bandwidth selection method for kernel density estimation". In: *Journal of the Royal Sta-*

## Chapter 3 Reproducible Home Ranges (*rhr*)

- tistical Society. Series B (Methodological)* 53.3, pp. 683–690 (cit. on p. 32).
- Solla, S. R. D., R. Bonduriansky, and R. J. Brooks (1999). “Eliminating autocorrelation reduces biological relevance of home range estimates”. In: *Journal of Animal Ecology* 68.2, pp. 221–234 (cit. on pp. 29, 31).
- Spencer, S. R., G. N. Cameron, and R. K. Swihart (1990). “Operationally defining home range: temporal dependence exhibited by hispid cotton rats”. In: *Ecology* 71.5, pp. 1817–1822 (cit. on p. 29).
- Steiniger, S. and A. J. Hunter (2012). “OpenJUMP HoRAE-A free GIS and toolbox for home-range analysis”. In: *Wildlife Society Bulletin* 36.3, pp. 600–608 (cit. on pp. 29, 35).
- Swihart, R. K. and N. A. Slade (1985). “Testing for independence of observations in animal movements”. In: *Ecology* 66.4, pp. 1176–1184 (cit. on pp. 28, 30, 31).
- Wells, A. G., C. C. Blair, E. O. Garton, C. G. Rice, J. S. Horne, J. L. Rachlow, and D. O. Wallin (2014). “The Brownian bridge synoptic model of habitat selection and space use for animals using GPS telemetry data”. In: *Ecological Modelling* 273, pp. 242–250 (cit. on p. 32).
- Worton, B. J. (1989). “Kernel methods for estimating the utilization distribution in home-range studies”. In: *Ecology* 70.1, pp. 164–168 (cit. on pp. 28, 31).

### 3.7 Appendix 1: An brief introduction to the *rhr* package

#### Obtaining the *rhr* package

The *rhr* package requires R (>3.2) and can be installed by:

```
install.packages(  
  "rhr",  
  repos=c("http://78.47.85.98/R", "http://cran.rstudio.com/"),  
  dep = TRUE)
```

#### Data generation and set up

We consider a very simple scenario, where movement data are simulated for 100 animals living in two different landscapes for 1 year. We sample the animals every 12 hours and thus have 720 relocations per animal.

```
library(magrittr)  
library(dplyr)  
library(rhr)  
library(lubridate)  
library(ggplot2)  
  
set.seed(123)  
  
dat <- data.frame(  
  id = rep(paste0("A", 1:100), each = 720),  
  region = rep(c("A", "B"), each = 720 * 50),  
  x = c(unlist(replicate(50, cumsum(rnorm(720))), FALSE)),
```

```

      unlist(replicate(50, cumsum(rnorm(720, 0, 3)), FALSE))),
y = c(unlist(replicate(50, cumsum(rnorm(720)), FALSE)),
      unlist(replicate(50, cumsum(rnorm(720, 0, 3)), FALSE))),
time = rep(ymd_hm("2015-01-01 12:00") + hours(0:719 * 12), 100) +
      minutes(round(rnorm(720 * 100, 0, 5))))
)

```

The data are stored in an ordinary `data.frame`.

```

str(dat)

## 'data.frame':    72000 obs. of  5 variables:
## $ id      : Factor w/ 100 levels "A1","A10","A100",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ region: Factor w/ 2 levels "A","B": 1 1 1 1 1 1 1 1 1 1 ...
## $ x      : num  -0.56 -0.791 0.768 0.839 0.968 ...
## $ y      : num  -1.211 -2.212 -1.977 -0.943 -1.335 ...
## $ time   : POSIXct, format: "2015-01-01 11:59:00" "2015-01-01 23:50:00" ...

```

### Using the GUI (Graphical User Interface)

To work with the GUI, first the data need to be saved as a text file (here a temporary directory is used).

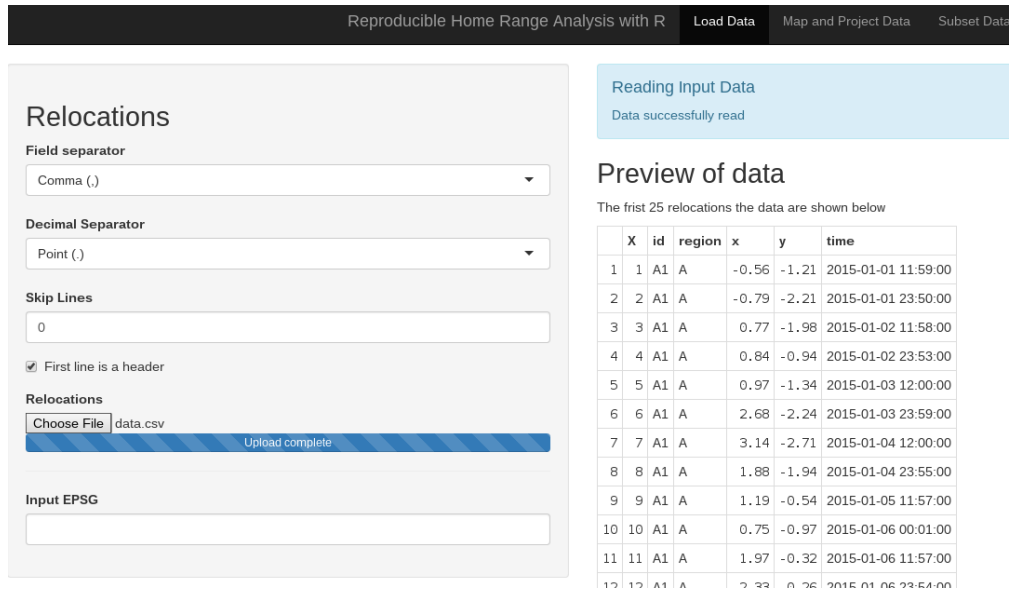
```
write.csv(dat, file.path(tempdir(), "data.csv"))
```

Next we can start the GUI with:

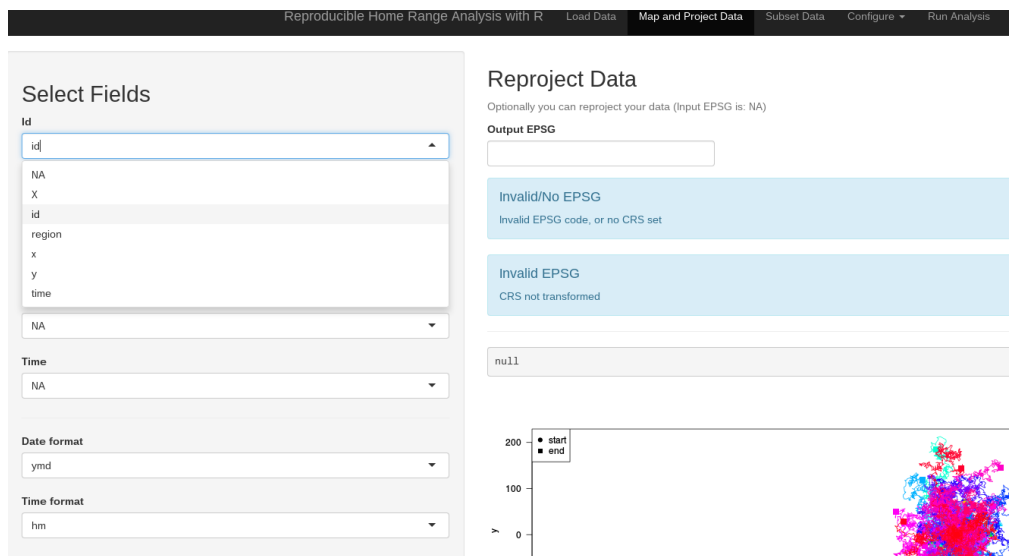
```
rhrGUI()
```

After typing the command `rhrGUI` a new tab in the web browser will open, that allows the user to upload data from the most commonly used text files (Fig. 3.A1). The next step consists of defining which columns in the text file contain which information. For example in the simulated data set the columns `x` and `y` contain the coordinates, and `id` the id of the animal (Fig. 3.A2).

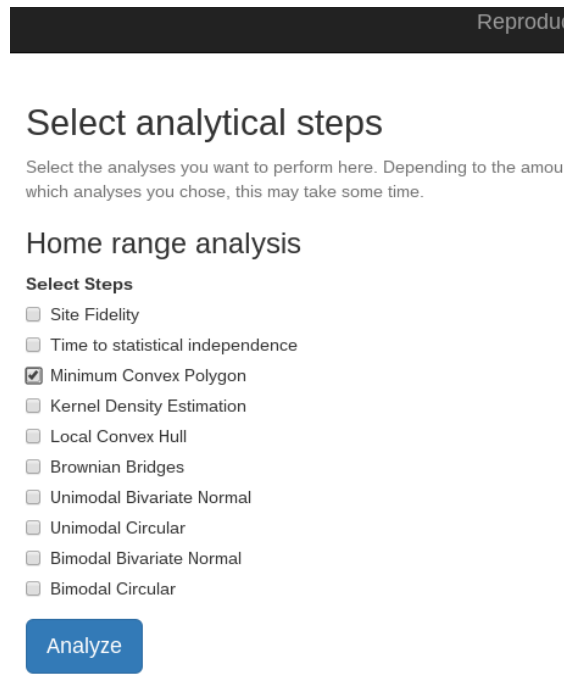
Further tools for data management are available, but not shown here. For example the user can transform the coordinates reference system of the data, select a subset of animals, or subset points based on space and/or time. Many different methods can be chosen to analyze movement data. Here only Minimum Convex Polygon (Fig. 3.A3) is used. Once the analysis finished, all results (`Rds`-files of the result objects and spatial data as ESRI Shapefile or GeoTIFF rasters) are written to a temporary directory. Further, a HTML report is generated and displayed. (Fig. 3.A4).



**Figure 3.A1:** GUI for the *rhr* package to load data. The GUI supports most commonly used delimiter separated text formats, such as \*.csv



**Figure 3.A2:** Mapping data. This step lets the application know which information is saved in which column of the text file.



**Figure 3.A3:** Selecting analyses.

## Command line

### Replicating the results of the GUI

It is relatively easy to reproduce the same results using the command line.

```
dat <- read.csv(file.path(tempdir(), "data.csv"))
res1 <- lapply(split(dat, dat$id), function(x) rhrMCP(x[, c("x", "y")]))
```

We can also obtain a plots, similar to Fig. 3.A4 for each animal by using `plot` (Fig. 3.A5).

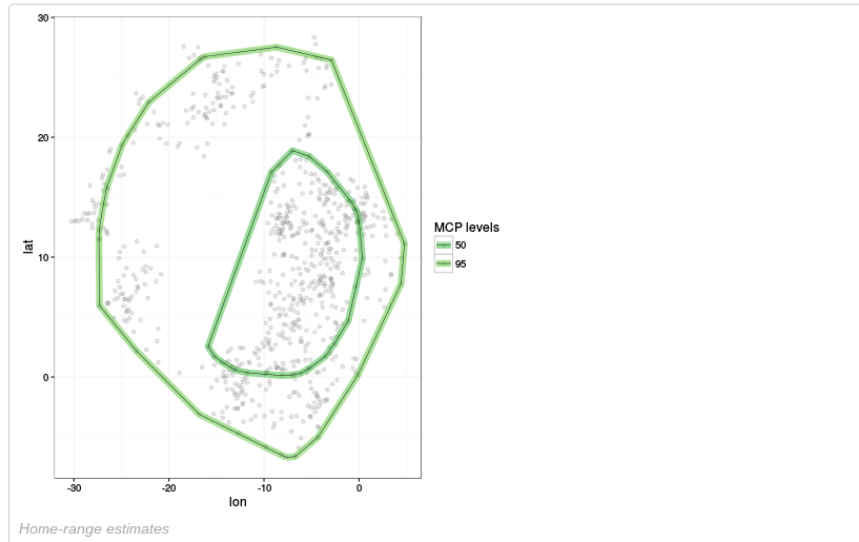
```
plot(res1[[3]])
```

Similar to the `plot` function, there is also a generic function to obtain the area of a home range estimate. `rhrArea` works for all home range estimators that are implemented in the `rhr` package.

```
rhrArea(res1[[3]])
```

```
##      level      area
## 95      95 1327.096
```

Results for A30



Home Range Area

Level	Area
50	200.7105
95	791.2149

Home range areas

Figure 3.A4: An example of the automatically generated report for one animal.

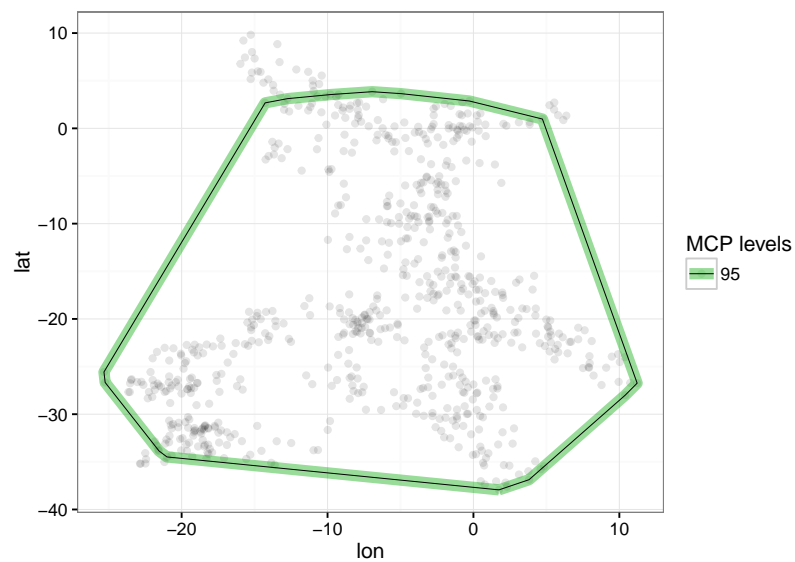


Figure 3.A5: Plot of MCP home range estimate using the `plot` function from the `rhr` package.



### Using *RhrTracks*

A researcher may be interested in only one relocation per week, this can be easily achieved using tracks (see also the appendix of chapter 4).

```
dat <- read.csv(file.path(tempdir(), "data.csv"))
dat$time <- ymd_hms(dat$time)
coordinates(dat) <- ~x+y
track <- rhRTracks(dat, dat$time, dat$id)

new_track <- seq(rhRTracksStart(track), length.out = 50, by = "1 week")
track1 <- rhRRegularize(track, new_track, side = "both", windw = 60 * 60 * 5)

res2 <- lapply(track1, rhRMCP)
```

We can now check if the home ranges size differ for the full data set (12h sampling rate) and for the reduced data set with one relocation per week.

```
quantile(sapply(res1, function(x) rhRArea(x)$area), probs = c(0.025, 0.975))
```

```
##      2.5%      97.5%
## 421.5368 12559.1657
```

```
quantile(sapply(res2, function(x) rhRArea(x)$area), probs = c(0.025, 0.975))
```

```
##      2.5%      97.5%
## 292.301 9992.173
```

This does not seem to be the case.

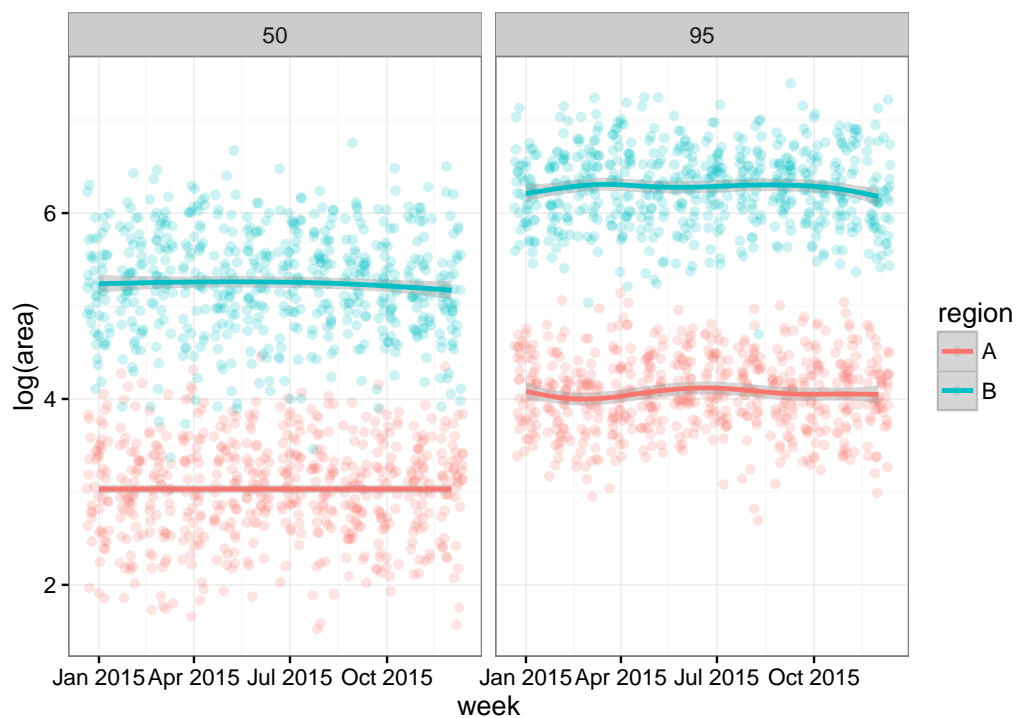
### Chaining operations

The *rhR* package supports in many aspects the use of pipelines (e.g., chaining operations). This can be useful for addressing slightly more complex task, such as plotting the home range size per week or grouping the animals by the area. We make use of the R packages *dplyr*<sup>2</sup>, *magrittr*<sup>3</sup> and *ggplot2*.

```
dat <- read.csv(file.path(tempdir(), "data.csv"))
dat %>% mutate(week = floor_date(ymd_hms(time), "month")) %>%
  group_by(region, id, week) %>%
  do(rhRArea(rhRMCP(data.frame(. $x, . $y), level = c(50, 95)))) %>%
```

<sup>2</sup>*dplyr* provides a fast, consistent and convenient way to work with `data.frame` and alike objects in R.

<sup>3</sup>*magrittr* provides pipelines in R to chain operations.



**Figure 3.A6:** Weekly home range size per animal and grouped by study area (different color) with a smoothed mean for each study area.

```
ggplot(., aes(x = week, y = log(area), color = region)) +  
  geom_jitter(alpha = 0.2) +  
  facet_grid(. ~ level) +  
  geom_smooth(aes(x = week, y = log(area), color = region)) +  
  theme_bw()
```

Note, that only the first four lines are required to do calculations and the remainder is only for plotting.

### More information on the *rhr* package

The package is documented as most other packages. The documentation can be accessed through `help("rhrMCP")` or `?rhrMCP` to read the documentation for calculating Minimum Convex Polygons.

Slightly more verbose documentations are available at the package website: <http://jmsigner.github.io/rhrman/>.

# 4

## Home range core areas: use and estimation

This chapter is accepted as: **Signer, J.** & Balkenhol, N. Home range core areas: use and estimation *Tagungsband der Vereinigung der Wildbiologen und Jagdwissenschaftler*.

**Authors contributions:** JS and NB conceived the study, JS performed the simulation and analyses, and JS and NB wrote the manuscript.

### Contents

---

<b>4.1</b>	<b>Abstract</b>	<b>46</b>
<b>4.2</b>	<b>Introduction</b>	<b>46</b>
<b>4.3</b>	<b>Methods</b>	<b>47</b>
4.3.1	Data generation	47
4.3.2	Core area estimation	47
4.3.3	Analysis	49
<b>4.4</b>	<b>Results</b>	<b>49</b>
<b>4.5</b>	<b>Discussion</b>	<b>52</b>
<b>4.6</b>	<b>References</b>	<b>52</b>

---

## 4.1 Abstract

To understand the space use of animals, researchers often use home range estimators as the analytical tool of choice. Since the very early days of home ranges, it was recognized that the space use within a home range is heterogeneous. Core areas are commonly used to characterize the internal structure of a home range. Knowledge of the structure of home ranges and thus the delineation of core areas is of great importance for interpreting results of home range studies accurately. Further, knowledge of core areas are essential for successful management and for conservation recommendations based on home range data. While there has been some progress on the theory of core areas and methods on how to estimate core areas, these advances are largely ignored in practice. We attempt to find a definition for core areas and compare four different methods of estimating core areas using a simple simulation approach. Results suggest that a recent Bayesian approach gives the most reliable estimates for the size of core area size.

## 4.2 Introduction

Understanding the space use of animals in order to investigate basic and applied research questions is of great interest for wildlife researchers. Recent advances in tracking technology enable researchers to collect position data of animals on unprecedented scales (e.g., Cagnacci et al. 2010). For example, GPS (Global Position System) devices allow researchers to record the position of animals at regular intervals (often, several relocations per day or even per hour are recorded). Data gathered through such devices are often analyzed using home ranges (Kie et al. 2010), a widely used concept in wildlife biology to analyze the space use of animals. A home range is defined as, that area traversed by an individual during its normal activities of food gathering, mating, and caring for their young. Occasional sallies outside the area, perhaps exploratory in nature, should not be considered part of the home range (Burt 1943). However, from the beginning of range analysis, it was recognized that areas within a home range are not used homogeneously (Hayne 1949). By separating areas of greater use intensity within a home range from areas with lower usage intensity, the concept of core areas emerged. While there is a very clear and widely accepted definition of home range (as given by Burt 1943, see above), core areas lack such a precise and widely recognized definition. Kaufmann (1961) introduced the term core area for describing areas that encounter a greater intensity of use than other areas within an animal's home range. Further definitions of core areas appeared with the discussion of new methods to estimate core areas. Samuel et al. (1985) defined core areas as areas of intensive use and Seaman and Powell (1990) defined core areas as areas of concentrated use within an animal's home range. More recently, Wilson et al. (2010) used the definition that a home range can be divided into several parts, with each part having a different but homogeneous density of use. These definitions require an estimate of the density of use, such as the utilization distribution, which is often obtained from probabilistic home range estimators (e.g., the kernel density estimator). Based on Wilson et al. (2010), we define a core area as an

area with a homogeneously greater density of use than other areas within an animal's home range. Thus, it follows that we should be able to separate a home range into two or more disjoint parts with different densities of use. We then classify the part of a home range with the greatest density of use as the core area and the remaining part of the home range as the non-core area. For simplicity, we consider only cases in which the home range is split into two disjoint areas (i.e., the core area and the non-core area). Each part of the home range needs to meet the requirement of a homogeneous density use within this part. The problem that arises now is: How do we separate a home range into two parts? Or, phrased differently: Which isopleth will divide a home range into two parts in such a way that within each part the density of use is homogeneous? There have been several methodological advances and developments of theory around the estimation of core areas in recent years, but they are often ignored by home range studies (Laver and Kelly 2008). The aim of this study is to compare different methods to estimate core areas. We compared four different methods to divide home ranges in core areas and non-core areas using data from simple simulations. Specifically, we focused on 1.) an arbitrary isopleth (i.e., the 50% isopleth of the utilization distribution), area independent methods after 2.) Seaman and Powell (1990), and 3.) Vander Wal and Rodgers (2012), and finally 4.) a Bayesian approach after Wilson et al. (2010). Finally, we discuss the results of applying previously mentioned methods to the simulated data and suggest alternative methods.

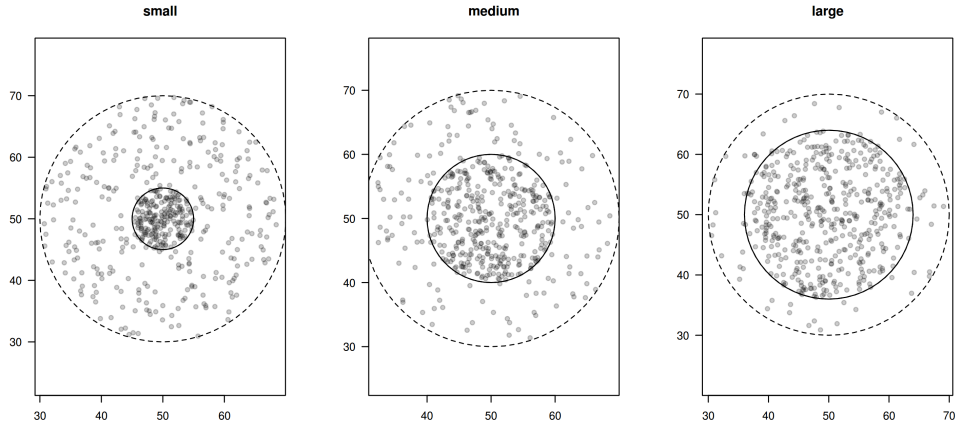
## 4.3 Methods

### 4.3.1 Data generation

We simulated relocation data in a neutral landscape within a circle centered at 50, 50 and with a diameter of 40 units. True core areas were simulated as circles with the same center, but smaller diameters of 10, 20 and 40 units respectively. We simulated relocation data from a Poisson point process with an intensity of 10 per unit area for core areas and 1 per unit area for non-core areas. We randomly sampled 500 relocations without replacement for the analysis (Fig. 4.1). We replicated simulations for each scenario 150 times to obtain a sampling distribution for estimates of the core area size. Simulations were conducted with Program R (R Core Team 2015) and the package `spatstat` (Baddeley and Turner 2004).

### 4.3.2 Core area estimation

Home range literature has proposed different methods to estimate core areas. Here, we compare a subset of four different methods that were either recently developed or are commonly used. The simplest method to distinguish between core area and non-core area parts of a home range is to use an arbitrary isopleth (e.g., 50%). This method is widely used, even though it has been criticized for its lack of accounting for intra- and interspecific differences in space use patterns of animals (Laver and Kelly 2008). We included this method as comparison and because it is likely the most widely used



**Figure 4.1:** One realization for each simulation scenario. Relocations were simulated as a Poisson point process with greater intensities for the core areas (enclosed by the solid line) than for the non-core areas of a home range (enclosed by the dashed line).

method. We delineated core areas at the 50% isopleth of the utilization distribution of the kernel density estimates.

Seaman and Powell (1990) suggested an area-independent method for the delineation of core areas which takes into account the relative intensity of use. To obtain the isopleth for delineating the core areas, one plots the percentage of the maximum relative frequency on the y-axis (e.g., if for a given grid cell the relative frequency is 0.02 and the maximum relative frequency in the whole home range is 0.35, the cell under investigation would obtain a value of  $0.02 / 0.35 = 0.057$ ) and the percentage of the home range on the x-axis (i.e., the number of pixels) which have a higher usage density. In the resulting graph (Fig. 4.2), the value on the y-axis of the point on the y-axis that is furthest away from the line with slope -1 and intercept 1 is subtracted from 1 and used as a cut off value for the utilization distribution. In theory, three distinct patterns of space use can be distinguished with this method: random, regular and clumped habitat use (Fig. 4.2, panel A-C). Random habitat use leads to a pattern where the resulting line remains close to the straight line, regular space use results in a line that is always above the straight line, and clumped use leads to a line that is always below the straight line (Fig. 4.2, panel D-F).

To obtain the separating isopleth according to Vander Wal and Rodgers (2012), the isopleth volume can be plotted on the x-axis against the home range size at a given isopleth level as a percentage of the home range size at the 100% isopleth. The delineating isopleth is obtained in two steps. First, an exponential regression line of the form  $\ln(PA) = \ln(b_0) + b_1 * IV$  is fitted to the data. Where,  $PA$  is the percentage of the maximum home range area (y-axis),  $IV$  is the isopleth volume (x-axis),  $b_0$  is the intercept, and  $b_1$  the slope. In the second step, the partial derivative with respect to

$IV$  is taken and set to one (to obtain the value of  $IV$  where the proportion of home range increases at a greater rate than the probability of use in the corresponding area). Finally, this equation is solved for  $IV$  to obtain the desired isopleth (see also Vander Wal and Rodgers 2012).

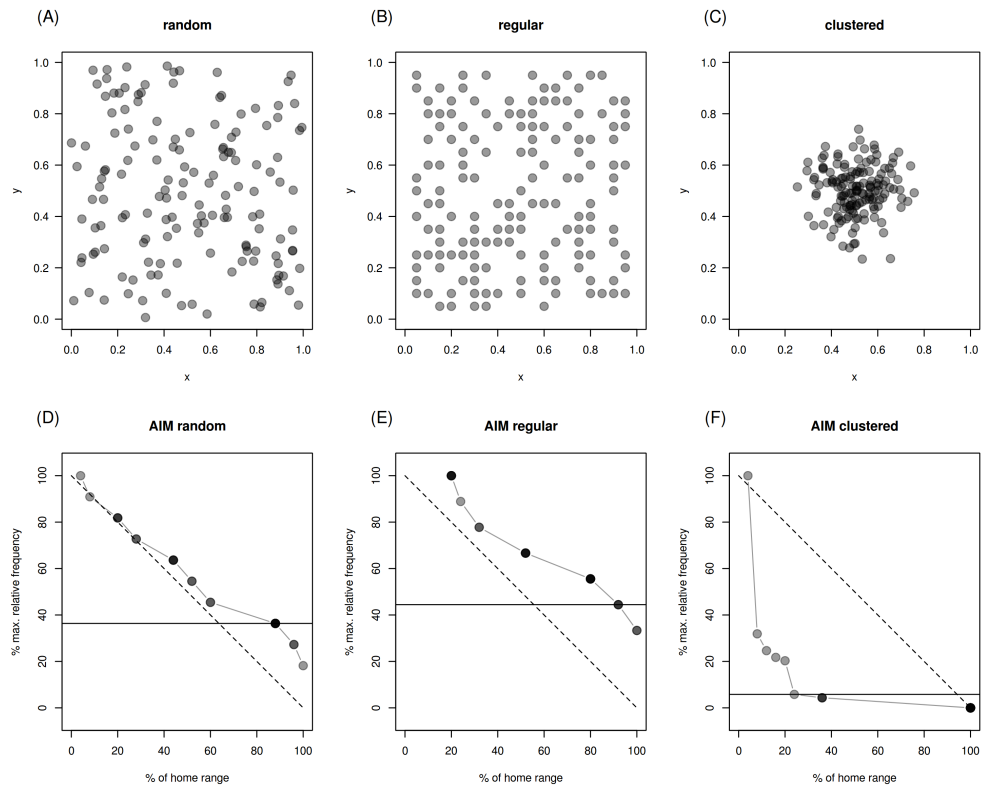
Wilson et al. (2010) developed a method that separates the home range of an animal into two or more parts, based on complete spatial randomness. In a nutshell, we assume that within a home range there are two distinct point patterns in complete spatial randomness and that an isopleth exists which can separate these two areas. Furthermore, we assume that this isopleth is not fixed, but random, which results in a Bayesian approach. In order to obtain an estimate of the isopleth, we divide the study area into an equal number of bins, which can be modeled as a multinomial distribution. As a prior, we used a Dirichlet distribution with hyperpriors of 0.1. This indicates weak prior knowledge that the isopleth is unlikely to be located too near at either boundary. Because this model is non-conjugate, Wilson et al. (2010) developed a Metropolis-Hastings sampler to obtain samples from the posterior. Wilson et al. (2010) also provide a solution for dividing the home range into more than three parts, which we do not consider here. We used a slightly adapted Metropolis-Hastings sampler to obtain the posterior distribution, generated 3000 samples per estimation, and used a burn-in phase of 600 with a thinning of 6. We used the posterior mode to obtain a point estimate for the isopleth that separates the core area from the non-core area.

### 4.3.3 Analysis

We used kernel density estimation with reference bandwidth to obtain estimates of the utilization distribution (UD) of home ranges. Home ranges were estimated with program R and the `rhr` package (Signer and Balkenhol 2015). For each simulation, we applied each of the previously described methods for the estimation of the isopleth value to delineate core areas. We then calculated the area of the core area for each replication to obtain a distribution of core areas for each method and compared this distribution to the true known area of the core area. We also calculated the mean of the estimated core area sizes and compared these means to the true known size of the core area by calculating a factor by which the true core area size is over- or underestimated. To obtain this factor, we divided the estimated core area size by the true known core area size.

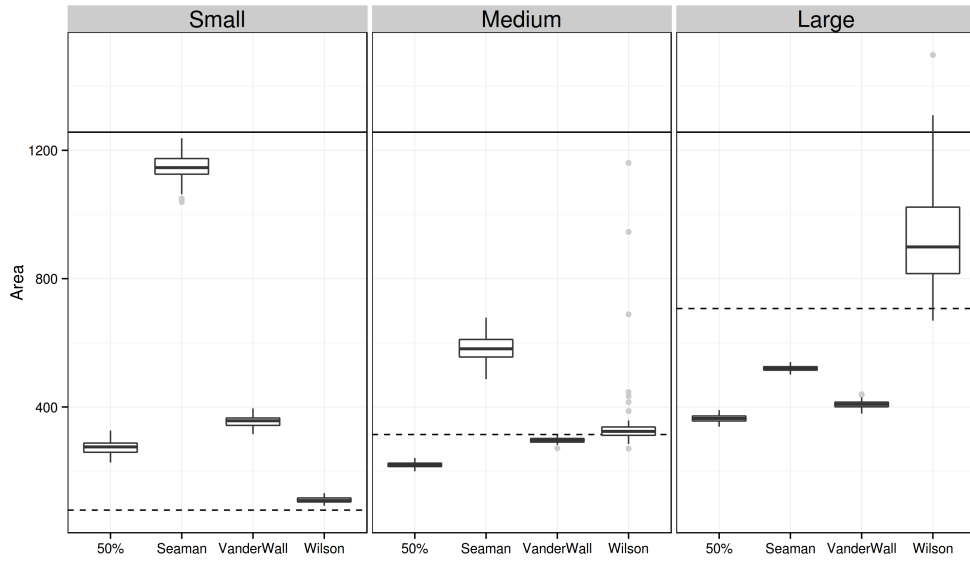
## 4.4 Results

Magnitude of core area size differed between methods and from the true core area size. Differences showed variable patterns (i.e., no method overestimated or underestimated the core area size for all three simulation scenarios consistently; Fig. 4.3). For small true core areas, all four core area estimators overestimated the core area size. This most pronounced for the Seaman and Powell Method (the true core area size was overestimated by a factor of 14.61). The arbitrary isopleth and Vander Wal Method overestimated the true core area size by a factor of 3.50 and 4.52 respectively. Finally, the Willson Method



**Figure 4.2:** Theoretical results of the area-independent method after Seaman and Powell (1990) to estimate core areas. Different space use patterns (Panels A-C) lead to different patterns when plotting the percentage of home range used vs. the percentage of maximum relative frequencies. The isopleth to divide between core area and non-core area is obtained by 100 – the value of the maximum relative frequency that is farthest away from the line (horizontal solid line, Panel D-F) with slope -1 and intercept 100 (black dashed line; Panel D-F).





**Figure 4.3:** Methods to estimate core areas differed with regard to precision (i.e., how close the estimates were to the known truth) and variability. The arbitrary isopleth and the Seaman and Powell Method (1990) never accurately estimated the true core area size and, moreover, the arbitrary isopleth showed less variability. For intermediate true core area sizes, the method after Vander Wal and Rodgers (2012) performed well with regard to precision. The Vander Wal Method showed little variability for all three scenarios. The method after Wilson et al. (2010) performed best with regard to precision, although variability was greatest for large true core areas. True core area sizes are indicated by the horizontal dashed line and true home range sizes are indicated by the horizontal solid line.

overestimated the true core area size by a factor of only 1.39. For intermediate true core area sizes, all four methods were closer to the true core area size. While the arbitrary isopleths underestimated the core area size by a factor of 0.79 and the Seaman and Powell Method overestimated the true core area size by a factor of 1.84, the Vander Wal Method and Wilson Method estimated the core area size as expected (i.e., the width of the 95% percentile of the sampling distribution overlapped the true core area size). The Wilson Method showed greater variability (i.e., more spread of the sampling distributions). Finally, for large true core areas, the arbitrary isopleth, the Seaman and Powell Method and the Vander Wal Method underestimated the core area size by a factor of 0.51, 0.73 and 0.57, respectively; and only the Wilson Method overestimated it by a factor of 1.30. However, the Wilson Method was the only method which covered the true core area by the 95% percentile interval of the sampling distribution.

Variability of estimates differed between estimators and true core area size. The

variability was consistently small using the arbitrary isopleth and Vander Wal Method, decreased with increasing core area size using the Seaman and Powell method and increased with increasing true core area using the Wilson Method.

## 4.5 Discussion

In the introduction we attempted to define the widely applied concept of a core area and settled with a definition that closely followed Wilson et al. (2010): core areas are areas with a homogeneously greater density of use than other areas within a home range. According to this definition, we used simple simulations to simulate relocation data with different true core area sizes and applied four methods to estimate the core area again. None of the methods we compared was able to recover the true core areas in a satisfactory way for all three modeling scenarios (small, medium and large true core areas). The simulations were relatively simple and disregarded individual differences. Simulations could potentially mimic three different species or environments. According to the literature, the core area size should not have varied greatly using the arbitrary isopleth. For the remaining three estimators, we expected the estimates of the core area size to follow the known true core area size. The arbitrary isopleth did not increase in the different scenarios. This was also true for using the Vander Wal Method. Contrarily, the Seaman and Powell method resulted in larger estimates of core areas with smaller true core areas (Fig. 4.3). Only the Wilson Method resulted in similar increases in estimated and true core area size.

Finally, the question of the biological relevance of core areas arises. Wray et al. (1992) already questioned the existence of core areas and the meaningfulness of the core area concept in general. From a biological point of view, it can not necessarily be said that core areas are of greater importance than non-core areas. It cannot be assumed that because non core areas are used less frequently that they are biologically less important (e.g., non-core areas could contain vital resources that are only occasionally visited, such as mating areas, important migratory routes, or water resources in dry areas). With ever increasing amounts of data available through GPS telemetry and more advanced analytical methods [e.g., step selection functions (Thurfjell et al. 2014) or path segmentation (Gurarie et al. 2009)], traditional core area estimators may be supplemented in future with additional insights to gain a better biological understanding of core areas. For now, we must conclude that none of the methods we used provides reliable core area estimates in all cases. While our simulations suggest that the Bayesian approach of Wilson et al. (2010) may be suitable for most data sets, from a biological standpoint, absolute values of core area sizes, as well as the spatial delineation of these core areas, need to be interpreted with caution.

## 4.6 References

Baddeley, A. J. and R. Turner (2004). *Spatstat: An R Package for Analyzing Spatial Point Patterns* (cit. on p. 47).

- Burt, W. H. (1943). "Territoriality and home range concepts as applied to mammals". In: *Journal of Mammalogy* 24.3, pp. 346–352 (cit. on p. 46).
- Cagnacci, F., L. Boitani, R. A. Powell, and M. S. Boyce (2010). "Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1550, pp. 2157–2162 (cit. on p. 46).
- Gurarie, E., R. D. Andrews, and K. L. Laidre (2009). "A novel method for identifying behavioural changes in animal movement data". In: *Ecology Letters* 12.5, pp. 395–408 (cit. on p. 52).
- Hayne, D. W. (1949). "Calculation of size of home range". In: *Journal of Mammalogy* 30.1, pp. 1–18 (cit. on p. 46).
- Kaufmann, J. H. (1961). *The ecology and social behavior of the coati, Nasua narica, of Barro Colorado Island, Panama*. University of California, Berkeley (cit. on p. 46).
- Kie, J. G., J. Matthiopoulos, J. Fieberg, R. A. Powell, F. Cagnacci, M. S. Mitchell, J.-M. Gaillard, and P. R. Moorcroft (2010). "The home-range concept: are traditional estimators still relevant with modern telemetry technology?" In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365.1550, pp. 2221–2231 (cit. on p. 46).
- Laver, P. N. and M. J. Kelly (2008). "A critical review of home range studies". In: *The Journal of Wildlife Management* 72.1, pp. 290–298 (cit. on p. 47).
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria (cit. on p. 47).
- Samuel, M. D., D. Pierce, and E. O. Garton (1985). "Identifying areas of concentrated use within the home range". In: *The Journal of Animal Ecology* 54.3, pp. 711–719 (cit. on p. 46).
- Seaman, D. E. and R. A. Powell (1990). "Identifying patterns and intensity of home range use". In: *Bears: their biology and management*, pp. 243–249 (cit. on pp. 46–48, 50).
- Signer, J. and N. Balkenhol (2015). "Reproducible home ranges (rhr): A new, user-friendly R package for analyses of wildlife telemetry data". In: *Wildlife Society Bulletin* 39.2, pp. 358–363 (cit. on p. 49).
- Thurfjell, H., S. Ciuti, and M. S. Boyce (2014). "Applications of step-selection functions in ecology and conservation". In: *Movement Ecology* 2.4 (cit. on p. 52).
- Vander Wal, E. and A. Rodgers (2012). "An individual-based quantitative approach for delineating core areas of animal space use". In: *Ecological Modelling* 224.1, pp. 48–53 (cit. on pp. 47–49, 51).
- Wilson, R. R., M. B. Hooten, B. N. Strobel, and J. A. Shivik (2010). "Accounting for individuals, uncertainty, and multiscale clustering in core area estimation". In: *The Journal of Wildlife Management* 74.6, pp. 1343–1352 (cit. on pp. 46, 47, 49, 51, 52).
- Wray, S., W. Cresswell, P. White, and S. Harris (1992). "What, if anything, is a core area? An analysis of the problems of describing internal range configurations". In: *Wildlife telemetry: remote monitoring and tracking of animals*, pp. 256–271 (cit. on p. 52).



# 5

## Does estimator choice influence our ability to detect changes in home range size?

This chapter is published as: **Signer, J.**, Balkenhol, N., Ditmer, M. & Fieberg, J. (2015) Does estimator choice influence our ability to detect changes in home range size? *Animal Biotelemetry*. DOI: <http://dx.doi.org/10.1186/s40317-015-0051-x>.

**Authors contributions:** JF provided the initial idea. JS performed the modeling and performed the statistical analyses together with JF. The bear movement data was provided by MD. All authors contributed to the interpretation of results and the writing of the manuscript.

### Contents

---

<b>5.1</b>	<b>Abstract</b> . . . . .	<b>56</b>
<b>5.2</b>	<b>Background</b> . . . . .	<b>56</b>
<b>5.3</b>	<b>Methods</b> . . . . .	<b>58</b>
5.3.1	Simulation model . . . . .	58
5.3.2	Scenarios . . . . .	58
5.3.3	Analysis . . . . .	59
5.3.4	Bear data . . . . .	60
<b>5.4</b>	<b>Results</b> . . . . .	<b>61</b>
5.4.1	Scenario 1a/b: Modeling temporally changing movement patterns	61
5.4.2	Scenario 2: Modeling effects of habitat fragmentation on movement patterns . . . . .	62
5.4.3	Bear data . . . . .	65
<b>5.5</b>	<b>Discussion</b> . . . . .	<b>66</b>
<b>5.6</b>	<b>Acknowledgements</b> . . . . .	<b>67</b>
<b>5.7</b>	<b>References</b> . . . . .	<b>67</b>
<b>5.8</b>	<b>Appendix 1: Additional Figures</b> . . . . .	<b>69</b>

---

## 5.1 Abstract

Estimates of home range size are frequently used to compare areal requirements of animals over time or space. Comparative studies of home range estimators have highlighted extreme differences among general classes of methods (e.g., polygon-based and kernel density-based estimators) and sensitivity to the choice of various tuning parameters (e.g., amount of smoothing). These studies, however, have largely failed to consider how estimates of home range size are typically used in applied research. We illustrate simulation-based methods for comparing estimators, which focus on relative differences in home range size (over time or space), rather than their absolute magnitude. We also consider GPS location data from a black bear (*Ursus americanus*) from northwestern Minnesota, USA to illustrate the relevance to real-world data applications.

In our examples, estimates of home range size often differed considerably in absolute magnitude. Yet, for relative differences, the choice of home range estimator was often negligible. Furthermore, choosing the right estimator was less important than other aspects of study design (e.g., number of animals followed).

Many questions in ecology focus on changes in space-use patterns (over space or time). For these types of questions, home range estimators should be evaluated in terms of their ability to detect these spatial and temporal patterns. More importantly, home range estimation should be seen as a means to an end – i.e., estimators provide indices useful for addressing interesting biological questions or hypotheses – rather than as an end to itself.

## 5.2 Background

Animals interact with conspecifics and their environment, leading to non-random patterns of space-use (Börger et al. 2008). Several different analytical methods have been proposed for quantifying these patterns, including home range estimation (e.g., Anderson et al. 2005; Börger et al. 2006b), habitat and step selection models (e.g., Manly et al. 2007; Thurfjell et al. 2014), and Bayesian State-Space Models that fit a mixture of random walks to movement data (e.g., Patterson et al. 2008; McClintock et al. 2012). Whereas the latter two approaches often require custom written code and fine tuning to fit a specific data set, a variety of off-the-shelf home range estimators can be easily implemented in multiple software platforms (R, ArcGIS, etc.).

Because of their accessibility, home range estimators are frequently used to compare space-use patterns for animals living in different landscapes (e.g., Walter et al. 2009; Reinecke et al. 2014), or along spatial gradients (e.g., Morellet et al. 2013)). With the increase of fine-scale spatio-temporal data afforded by Global Positioning Technology (GPS), short-term (weekly, monthly) estimates of home range size are now also commonly used to explore changes in space-use patterns over time (e.g., Börger et al. 2006b; Knight et al. 2009; Van Beest et al. 2011).

When using home range estimators to summarize space-use patterns, it is important not to conflate these statistics with the biological concept of an animal's home range

(Fieberg and Börger 2012; Kie et al. 2010; Powell and Mitchell 2012), defined by Burt (Burt 1943), as "That area traversed by an individual in its normal activities of food gathering, mating, and caring for young. Occasional sallies outside the area, perhaps exploratory in nature, should not be considered part of the home range". An animal's home range is arguably best thought of in terms of cognitive maps, a complex summary of various spatio-temporal utility surfaces representing different resources (e.g., food, shelter, escape routes or mating and breeding areas), which the animal actively updates as it interacts with its environment (Nathan et al. 2008; Powell and Mitchell 2012). Home range estimators, by contrast, provide simple, often static summaries of space-use in terms of a boundary on a map, an area, or a probability distribution of the relative time spent in different areas. An estimate of home range size is best viewed as an index of space-use or movement cost to meet an individual's needs, a response measure that can be related to other measured covariates in order to gain insights into how animals interact with their environment or other organisms (e.g., habitat types and configurations, waterways, urban areas or other GPS-tagged individuals).

An advantage of viewing home range estimates as indices is that it emphasizes the need to start with an interesting and meaningful biological question or hypothesis, rather than viewing home range estimation as an end in itself (Fieberg and Börger 2012). It also suggests that we need to think differently about comparative studies of home range estimators. We should not only compare the ability of estimators to capture (or exclude) particular areas of space used (or not used) by an animal, but also increase emphasis on whether the estimator does a good job of capturing an important biological signal in the data that is relevant to an underlying question or hypothesis of interest. A biased estimator for home range size may still prove useful if it does a good job of detecting *changes* in home range size. For many research questions, this will mean that home range size is often expected to be proportional to movement cost.

The aim of this manuscript is to illustrate these points by investigating a property of common home range estimators, the area, through a simulation study. We consider two different simulation scenarios capturing two broadly defined research questions. In the first scenario, we are interested in detecting changes in the amount of space used by an individual over time. In the second scenario, we are interested in comparing the amount of space used by different individuals living in landscapes with different amounts of favorable habitat. In both cases, we simulate animal movement using discrete time steps and calculate home ranges using different estimators, sampling rates and analysis intervals (scenario 1 only; movement behavior did not change over time for scenario 2). In both cases it is natural to ask, 'What home range estimator is most appropriate for detecting differences in the amount of space used (changes over time in scenario 1 or differences among landscapes with different amounts of habitat in scenario 2)'. In addition, we ask 'What analysis interval is most appropriate for exploring temporal space-use patterns?' (scenario 1 only). We explore these questions using simulated data, but we also consider GPS location data from a black bear (*Ursus americanus*) inhabiting northwestern Minnesota, USA to illustrate the relevance of our simulation results, and method of estimator comparison to real-world data applications.

## 5.3 Methods

### 5.3.1 Simulation model

We use a spatially-explicit, individual-based model to simulate animal movement. The simulated animal moves in discrete time steps,  $t_1, \dots, t_n$ , in homogeneous (scenario 1) or heterogeneous (scenario 2) space. We denote the position of the animal at time step  $t$  as vector  $\mathbf{x}_t = (x_t, y_t)$  containing the  $x$  and  $y$  coordinates at time  $t$ . At each time step, the animal selects 20 candidate locations indexed with the superscript  $j$  ( $\mathbf{x}_{t+1}^1, \dots, \mathbf{x}_{t+1}^{20}$ ) for the position at the next time step ( $t + 1$ ). The coordinates for each candidate location ( $\mathbf{x}_{t+1}^j$ ) are determined by a turning angle  $\phi_j$  and step length  $s_j$  from the current position ( $\mathbf{x}_t$ ). The turning angle is drawn from a wrapped Cauchy distribution with concentration parameter  $\rho_t$ . The step length is drawn from an exponential distribution with rate  $\lambda_t$ . Next the animal assesses the quality of each candidate location,  $\mathbf{x}_{t+1}^j$  in terms of habitat quality at  $\mathbf{x}_{t+1}^j$  and the euclidean distance between  $\mathbf{x}_{t+1}^j$  and  $\mathbf{x}_0$ —the home range center—resulting in a vector of weights,  $\boldsymbol{\Omega}_j$ , for each candidate location. Weights for habitat quality and distance to home range center are then summed, resulting in a single weight,  $\omega_{t+1}^j$ , for each candidate location for the position of the animal at the time  $t + 1$ . The animal then chooses one of the candidate locations randomly, with probability proportional to the weights  $\omega_1, \dots, \omega_{20}$ .

Model parameterization for the different simulation scenarios are shown in Table 5.1. We used time steps of 5 minutes and modeled the individual for one year, so that 104,832 movement steps were available for each simulated individual.

### 5.3.2 Scenarios

We investigate two different scenarios that reflect two different biological questions. In the first scenario an animal is tracked for one year, and the researcher uses estimates of short term "home range" size as an index to quantify how movement costs change over time. In the second scenario, multiple animals in different landscapes are followed, and the researcher uses estimates of home range size to detect differences in movement costs between landscapes that vary with respect to the proportions of favorable habitat.

In scenario 1a, we assume that the tendency of the animal to move away from its home range center changes over the course of a year. To model the underlying pattern, we use a rescaled Gaussian density function centered on the middle of the year (Panel A in Fig. 5.A1 in the appendix). For scenario 1b, we add additional sinusoidal noise to the underlying Gaussian density function, to represent changes in movement that occur on smaller time scales (Panel B in Fig. 5.A1 in the appendix).

For the second scenario, movement parameters do not vary temporally, but the landscape is heterogeneous. We simulated landscapes with different amounts of favorable habitat, using a modified version of the random cluster algorithm (Saura and Martinez-Millan 2000) as implemented in Efford (2015). The random cluster algorithm generates landscapes based on spatial autocorrelation of habitat patches controlled through a fragmentation parameter  $p$ , and the expected amount of habitat  $A$ . We simulated landscapes



with different amounts of favorable habitat,  $A$ , by drawing this value from an uniform distribution between 0.1 and 0.9 (i.e. 10 to 90% of the landscape consists of favorable habitat), while holding the degree of fragmentation constant ( $p = 0.3$ ). We allowed turning angles to depend on the habitat type (favorable or unfavorable), and we weighted proposed locations according to their habitat type and the distance of each candidate location to the home range center (Table 5.1). In favorable habitat, chosen turning angles were less concentrated (this resembles within-patch foraging) than in unfavorable habitat (resembling inter-patch movement).

For both scenarios, we assumed the home range center to be at  $x = 50$  and  $y = 50$ , the center of the landscape. If this placed the home range center in unfavorable habitat, we assigned the home range center to the closest pixel with favorable habitat. All simulations were programmed and run using the program R (R Core Team 2015).

### 5.3.3 Analysis

We estimated home ranges for different sampling rates and analysis intervals (only scenario 1). We sampled each realization of the movement path at three different sampling rates (2, 4 and 24 observations per day). We chose these sampling rates to resemble commonly used sampling schedules of real GPS collars. We then split the sampled data into either weekly or monthly analysis intervals and calculated home ranges for each combination of sampling rate and analysis interval. We set a threshold of 20 for the minimum number of relocations required for any combination of sampling rate and analysis interval. Hence, we did not consider weekly home ranges for a sampling rate of 2 observations per day.

We estimated home ranges with the minimum convex polygon (MCP), kernel density estimator (KDE) and local convex hull (LoCoH). Although a variety of more sophisticated home range estimators have been developed in recent years, MCPs and KDEs are the simplest and most widely applied approaches (Laver and Kelly 2008; Kie et al. 2010). We also included LoCoH because previous studies (e.g., Scull et al. 2012; Reinecke et al. 2014) found that home range estimates from LoCoH may be more useful for wildlife managers, because of their ability to better detect sharp boundaries. For KDE, we used bivariate normal kernels on a 100 by 100 grid with a resolution of 1 unit. We did not rescale the data before calculating bandwidth and we used the same bandwidth for both dimensions. We used 4 different methods for choosing the bandwidth: 1) the reference (href) method (Worton 1989), 2) least-squares cross validation (hls cv; Seaman and Powell 1996), 3) plug-in-the-equation (hpi; Sheather and Jones 1991), and 4) an *ad hoc* scaling of the reference bandwidth (hscaled; Kie 2013)) This last approach attempts to find the smallest bandwidth that results in a contiguous home range area. For least-squares cross validation and *ad hoc* scaling, we performed a grid search with values ranging from  $0.1 \times$  reference bandwidth to  $2 \times$  reference bandwidth. For least-squares cross validation, we searched for a global minimum. If the algorithm did not converge (this only happened once for scenario 1b), we excluded the simulation run from the analysis. With the LoCoH approach, we used the type  $k$  method, with  $k$  determined by the square root of the number of fixes. In all cases, we estimated the area associated

with the 95 % percent isopleth. All home range analyses were implemented using Program R (Version 3.1.1) (R Core Team 2015) and the *rhr* (1.0.010) package (Signer and Balkenhol 2015).

For scenario 1, we explored, graphically, temporal patterns in home range estimates and their variability across multiple simulated data sets for each unique combination of: sampling rate, home range estimator, and analysis interval. Since we were particularly interested how home ranges can be used as a proxy to quantify movement costs, we also looked at differences in home range size, i.e. the numerical first derivative:  $\frac{\text{home range size}_t - \text{home range size}_{t-1}}{\Delta t}$  versus  $t$ . We calculated  $\frac{\Delta \text{home range size}}{\Delta t}$  for each realization and then averaged these values at each time point. In order to investigate the variance of different home range estimators, we calculated the difference between the 97.5 percentile and 2.5 percentile of the home range size and  $\frac{\Delta \text{home range size}}{\Delta t}$  sampling distribution, and used the width of this interval as a measure of estimator precision at each time point.

For scenario 2, we sampled  $n = (5, 10, \dots, 100)$  individuals from the population and fit linear models relating  $\log(\text{home range size})$  to the amount of favorable habitat,  $A$ ,  $\log(\text{home range size}) = \beta_0 + \beta_1 A$ . Then, we repeated this process for  $R = 100$  times for each home range estimator, sampling rate and sampling size  $n$ . We then calculated the interquartile range and the difference of the 97.5 percentile and 2.5 percentile of the sampling distribution for each combination of sample size, sampling rate and estimator, and again used the widths of these intervals as a measure of estimator precision. For  $\beta_1$ , the slope, we calculated for each estimate a 95% confidence interval and looked at the number of times we obtained significant results (i.e. the confidence interval did not overlap 0).

### 5.3.4 Bear data

During May - July, 2007 – 2011, we captured American black bears living in northwestern Minnesota, U.S.A using baited barrel traps as part of a larger study investigating how bears utilize an agriculturally dominated and fragmented landscape (Ditmer 2014). We immobilized bears using a combination of ketamine hydrochloride and xylazine, or premixed tiletamine hydrochloride and zolazepam (Telazol®), Elkins-Sinn, Cherry Hill, NJ, USA). Bears were fit with either collars with store-on-board GPS devices or GPS collars using the Iridium satellite system as a data link. During annual winter den visits we adjusted collar fit and downloaded data from collars. Bear handling was approved by the University of Minnesota’s Institutional Animal Care and Use Committees (permit no. 1002A77516). For the analyses presented, we chose to utilize data from a representative 8-year old female bear without cubs of the year. This individual was outfitted with a GPS collar from Vectronic Aerospace (Berlin, Germany) throughout 2012. We chose this bear because of the complete year of data she provided (den exit to den entrance), frequent fix schedule (fixes attempted every 2 h), and high level of home range fidelity (based on 5 years of GPS-collar data). We used 1,711 successful GPS-collar locations (mean time step = 2.54 h, SD = 1.17 h) from 30 April 2012 to 28 October 2012. We

**Table 5.1:** Parameter values at time step,  $t$ , for finding one of 20 candidate locations (upper half) and for weighting the candidate locations based on habitat and distance to the home range center.  $\mu$  and  $\sigma$  are constant with values of 182 and 36.4 respectively. Step length,  $s$ , habitat weight,  $\omega_{hab}$ , and concentration parameter,  $\rho$  are constant for scenario 1 and vary for scenario 2. In contrast  $\lambda_t$  varies over time following a rescaled Gaussian density function for scenario 1 and is constant for scenario 2.

Parameter	Scenario 1a	Scenario 1b	Scenario 2
<i>Determining candidate locations</i>			
$\rho_t$	0.5	0.5	$\begin{cases} 0.01 & \mathbf{x}_t \text{ in patch} \\ 0.8 & \mathbf{x}_t \text{ in matrix} \end{cases}$
$\lambda_t^j$	$-e^{\frac{(t-\mu)^2}{2\sigma^2}} + 2$	$-e^{\frac{(t-\mu)^2}{2\sigma^2}} + 2 - 0.04 \sin\left(\frac{0.2t}{2\pi}\right)$	0.01
<i>Weighting candidate locations</i>			
$\omega_{hab}$	1	1	$\begin{cases} 1 & \mathbf{x}_t \text{ in patch} \\ 0.1 & \mathbf{x}_t \text{ in matrix} \end{cases}$
$\theta_t$	$-0.4e^{\frac{(t-\mu)^2}{2\sigma^2}} + 0.5$	$-0.4e^{\frac{(t-\mu)^2}{2\sigma^2}} + 0.5 - 0.1 \sin\left(\frac{0.2t}{2\pi}\right)$	0.01

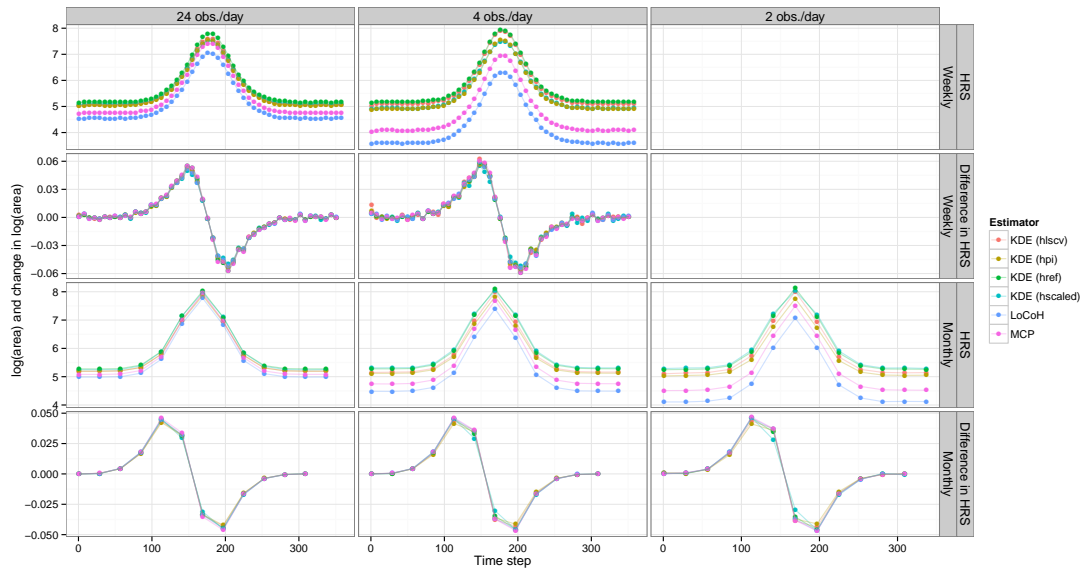
used 1, 2, 4 and 8-week analysis intervals and estimated the home range size using the same home range estimators as were used in the scenario analyses (see above).

## 5.4 Results

### 5.4.1 Scenario 1a/b: Modeling temporally changing movement patterns

For scenario 1a, absolute magnitude of home range size differed (Fig. 5.1), with KDE methods producing larger home range size than MCP and LoCoH. Yet, all estimators captured the underlying temporal pattern. Further, estimates of changes in home range size,  $\left(\frac{\Delta_{\text{home range size}}}{\Delta t}\right)$ , were nearly indistinguishable (Fig. 5.1), although the magnitude of changes depended somewhat on the analysis interval (Fig. 5.1), with longer analysis intervals resulting in estimated rates of change that were less variable and smaller in absolute magnitude. Results for scenario 1b were similar (to scenario 1a), but only short analysis intervals (weekly) were able to recover the signal of the short term sinusoidal variation of changes of space requirements (Fig. 5.A7 in the appendix).

The variability of home range size estimates was dependent on the estimator and the time of the year (Fig. 5.2). For weekly analysis intervals only two estimators, KDE (hscaled) and LoCoH, had relatively constant variability (i.e. the width of the 95 percentile interval of the sampling distribution did not depend on the time). For all other estimators and analysis intervals, the width of the 95 percentile of the sampling distri-

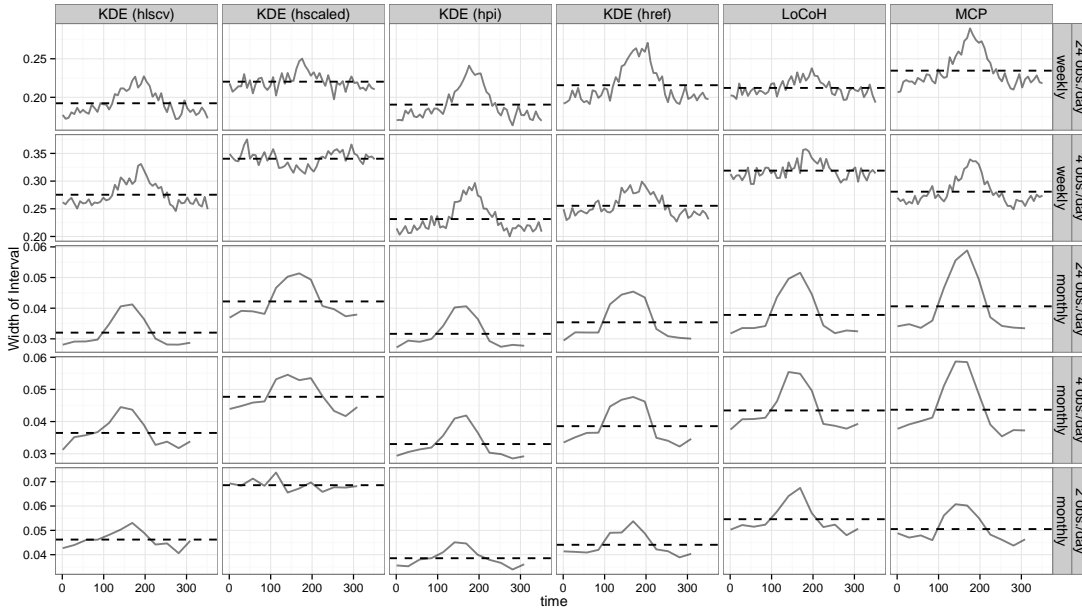


**Figure 5.1:** Results of simulation scenario 1a. Average estimates of  $\log(\text{home range size})$  (first and third row) and changes of  $\log(\text{home range size})$  (second and fourth row) from 1000 realization of animals moving according to simulation model 1a at different analysis intervals (rows) and sampling rates (rows). We present results for all 6 estimators (different colors), two analysis intervals and three sampling rates. We required at least 20 relocations for any combination of sampling rate and analysis interval. Hence we did not consider weekly home ranges for two observation per day.

bution varied with time, with estimates exhibiting more variability in the middle of the year when animals moved more extensively. Across the range of simulation scenarios (and averaging over time), the variability was smallest for the KDE (hpi) and largest for the KDE (hscaled). Estimates of home range size (and also rates of change) were less variable when using longer time intervals and when sampling more frequently (Fig. 5.2 and supplementary figures).

#### 5.4.2 Scenario 2: Modeling effects of habitat fragmentation on movement patterns

Intercepts associated with the linear model relating  $\log(\text{home range size})$  to the amount of favorable habitat differed in magnitude, again suggesting absolute estimates of home range size varied with estimator choice (Fig. 5.3). For small sample sizes ( $n = 5, 10$ ), the sampling distributions of intercepts ( $\beta_0$ ) overlapped, but differences among estimators became more pronounced as more individuals were followed. Intercepts associated with KDE (hlscv) and KDE (hpi) were generally smaller than MCP, KDE (href), and KDE (hscaled), with LoCoH falling in between (Fig. 5.3). Estimates of  $\beta_1$ , which summarizes

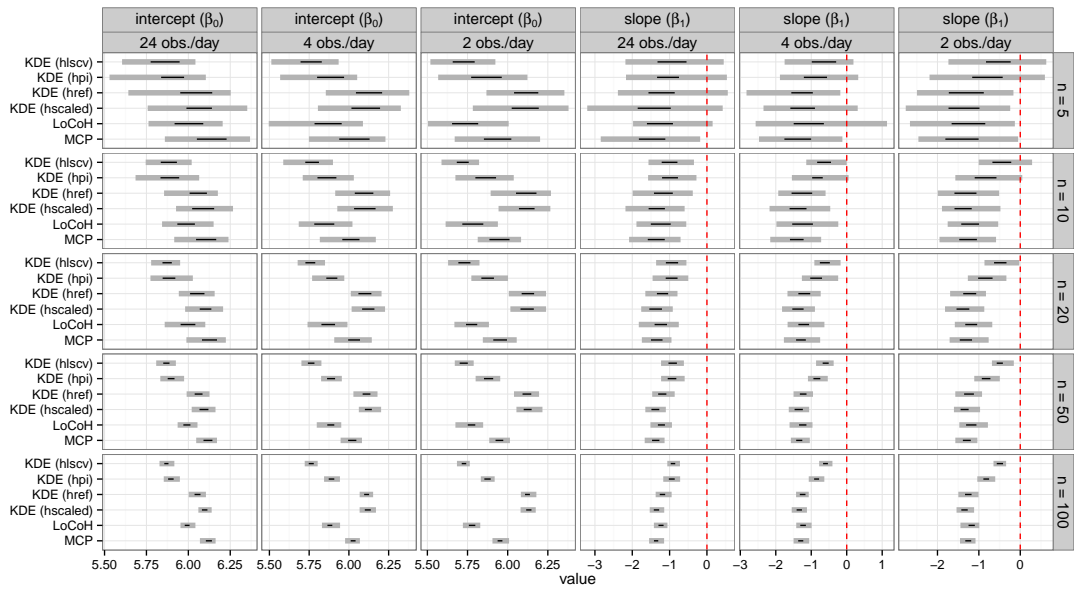


**Figure 5.2:** Results of simulation scenario 1a. Difference between the 97.5 percentile and 2.5 percentile of the sampling distributions of home range size as a measure of the variability of estimates of home range size. We present variability estimates from 1000 realizations of animals moving according to simulation scenario 1a for all six estimators (columns), three analysis intervals and two sampling rates (rows), averaged at each time stamp (solid line) and averages over all time steps (dashed line).

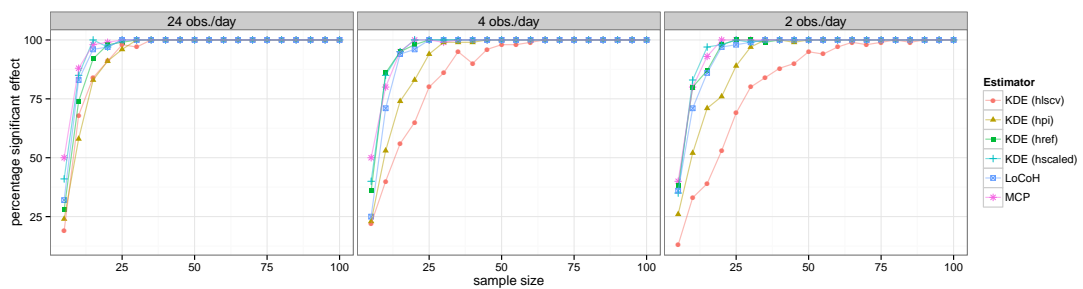
the influence of habitat on home range size, were closer to 0 for KDE (lscv) and KDE (hpi); estimates of  $\beta_1$  were similar in magnitude (and larger in absolute value) for all other estimators (Fig. 5.3 and supplementary figures).

Precision, and thus the power to detect an influence of the amount of favorable habitat on movement costs, depended mainly on the number of animals followed, but was also influenced to a lesser extent by the sampling rate (Fig. 5.3 and Fig. 5.4).

At the highest sampling rates (24 relocations per day) all estimators had a high power to detect an effect of habitat on home range size (Fig. 5.3). With lower sampling rates (e.g., 2 relocations per day), however, many more animals were necessary to detect significant effects, and choice of estimator become more important. In particular KDE (lscv, hpi) had lower power to detect an effect of habitat than other methods (Fig. 5.4), largely because the sampling distributions of these two estimators were centered on values closer to 0 (Fig. 5.3).



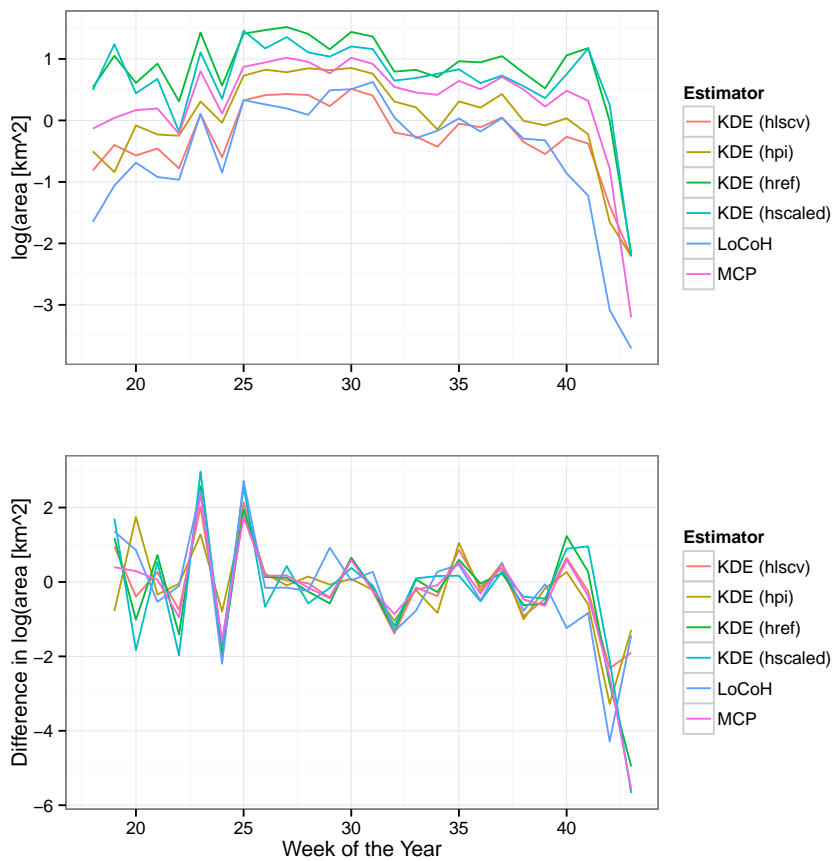
**Figure 5.3:** Variability of home range estimators with respect to sample size. Difference between the 97.5 percentile and 2.5 percentile (horizontal gray bars), and the interquartile range (horizontal solid black lines) of the sampling distributions of coefficients associated with a linear model relating  $\log(\text{home range size})$  to the amount of favorable habitat of animals moving according to simulation model 2. We present results for six home range estimators for three different sampling rates (columns) and five different sample sizes (i.e., the number of animals that were tracked). Absence of an effect of amount of favorable habitat on home range size is indicated if the difference between the 97.5 percentile and 2.5 percentile of the sampling distribution of the slope overlaps 0 (vertical dashed red line).



**Figure 5.4:** Power to detect a significant effect of the amount of favorable habitat on home range size. Power is plotted as a function of sample size (i.e. the number of animals tracked) for six different home range estimators (colors) and sampling rates (different panels).

### 5.4.3 Bear data

We observed a similar, although noisier pattern, with real data originating from a female bear. Absolute home range size was small in the beginning and the end of the year (this is when bears exit (spring) and prepare to enter (fall) the den). During summer, energy demands increase and bears expand their range to meet their daily needs. All of the estimators captured this trend, but the estimates were of different magnitudes. LoCoH estimates were the most conservative (i.e., smallest) and KDE (href) estimated the largest home ranges (but see also the top panel of Fig. 5.5). Similar to the simulated data from scenario 1a, differences between estimators became negligible when we evaluated *changes* in the amount of space used over time (bottom panel of Fig. 5.5).



**Figure 5.5:** Results from empirical data. Estimates of home range size (top panel) and changes of home range size (bottom panel) with six different home range estimators over time of a female black bear from northwestern Minnesota.

## 5.5 Discussion

In the introduction, we suggested that it could be advantageous, from a statistical and conceptual point of view, to think about home range estimators as an index of an animal's areal requirement or movement costs. In such cases, the absolute magnitude of home range size is often less important than how estimates of home range size change over space and/or time. Importantly, we found that although estimates of home range size often differed considerably in absolute magnitude, for relative differences, the choice of home range estimator was often negligible. Furthermore, results of simulation scenario 2 suggest that choosing the right estimator is often less important than other aspects of study design (e.g., number of animals followed; see also Börger et al. (2006a).

Other studies have shown that results may be robust to choice of home range estimator (e.g., Nilsen et al. 2008; Van Beest et al. 2011). Van Beest et al. (2011) found that different home range estimators (KDE and LoCoH) provided qualitatively similar results (i.e. the direction of the relationship did not depend on the home range estimator) regarding the most important factors influencing variation in home range size across different spatio-temporal scales for moose (*Alces alces*) in Norway. Also Nilsen et al. (2008) were able to draw the same biological conclusions relating home range size from MCP- and KDE-based estimators to body mass for interspecific comparative studies. In congruence with Börger et al. (2006a) our results suggest that efforts should be directed towards collaring more individuals at the expense of sampling rate; results from scenario 2 highlight this point.

Although, comparative studies of home range estimators are common (e.g., Seaman and Powell 1996; Lichti and Swihart 2011; Schuler et al. 2014) most simulation studies tend to focus on bias associated with estimates of home range size, which requires that one defines a true home range, a concept that can be difficult to quantify using only (real or simulated) animal locations (Fieberg 2007; Powell and Mitchell 2012). For many questions, we suggest that a focus on variability of home range estimates under different circumstances will lead to more informative comparisons. The simulation conducted under scenario 1 would lead us to KDE (hpi) because it was the most precise estimator. By contrast, the power analysis aimed at detecting the influence of habitat on resource needs would lead us to suggest any method *except* KDE (hpi or hlscv). These conflicting results highlight the need to begin with clearly defined research questions and goals prior to choosing an appropriate home range estimation method (Fieberg and Börger 2012).

Our results suggest that for questions like: do male individuals require more space than females? does latitude or elevation influence the space-use of a species?, or, do individuals of a given species need more space in summer than in winter?, the choice of the estimator will likely be less important than emphasized in the literature while other decisions a researcher might make (e.g., sample size, sample rate or analysis interval) are at least as important as estimator choice. Yet, because different estimators give different absolute estimates of home range size, it is still important, whenever possible, to use the same home range estimator to analyze all location data. Further, given the differences in implementation of home range estimators (e.g., variability in how KDE



(lscv) is implemented (Laver and Kelly 2008)), it is also important to use the same software whenever possible. As a corollary, we strongly encourage researchers to archive raw location data (e.g., in archives like Movebank or dryad) to facilitate meta-analytic hypotheses testing of broad-scale ecological patterns.

Although some research question may require absolute estimates of home range size, most often, home range estimates are used to explore changes in space-use patterns over time or space. In these cases, home range estimators should be evaluated in terms of their ability to detect these patterns. More importantly, home range estimation should be seen as a means to an end, i.e., estimators provide indices useful for addressing interesting biological questions or hypotheses – rather than as an end to itself (Fieberg and Börger 2012).

## 5.6 Acknowledgements

We thank R. Powell, and 1 anonymous reviewer for helpful comments on the manuscript. We also thank the wildlife discussion group at Göttingen University for valuable feedback on the manuscript.

## 5.7 References

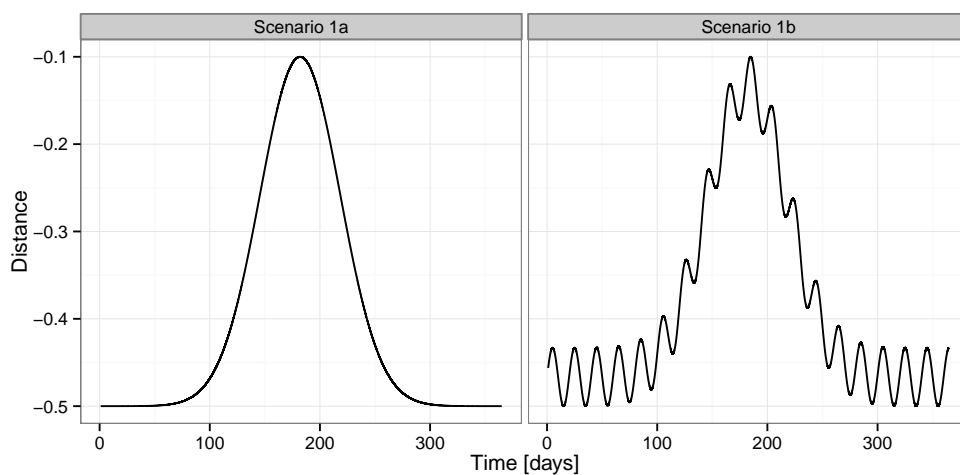
- Anderson, D. P., J. D. Forester, M. G. Turner, J. L. Frair, E. H. Merrill, D. Fortin, J. S. Mao, and M. S. Boyce (2005). “Factors influencing female home range sizes in elk (*Cervus elaphus*) in North American landscapes”. In: *Landscape Ecology* 20.3, pp. 257–271 (cit. on p. 56).
- Börger, L., B. D. Dalziel, and J. M. Fryxell (2008). “Are there general mechanisms of animal home range behaviour? A review and prospects for future research”. In: *Ecology letters* 11.6, pp. 637–650 (cit. on p. 56).
- Börger, L., N. Franconi, G. De Michele, A. Gantz, F. Meschi, A. Manica, S. Lovari, and T. Coulson (2006a). “Effects of sampling regime on the mean and variance of home range size estimates”. In: *Journal of Animal Ecology* 75.6, pp. 1393–1405 (cit. on p. 66).
- Börger, L., N. Franconi, F. Ferretti, F. Meschi, G. De Michele, A. Gantz, and T. Coulson (2006b). “An integrated approach to identify spatiotemporal and individual-level determinants of animal home range size”. In: *The American Naturalist* 168.4, pp. 471–485 (cit. on p. 56).
- Burt, W. H. (1943). “Territoriality and home range concepts as applied to mammals”. In: *Journal of Mammalogy* 24.3, pp. 346–352 (cit. on p. 57).
- Ditmer, M. A. (2014). “American black bears: strategies for living in a fragmented, agricultural landscape”. PhD thesis. Saint Paul, Minnesota: University of Minnesota (cit. on p. 60).
- Efford, M. (2015). *secr: Spatially explicit capture-recapture models*. R package version 2.9.5 (cit. on p. 58).
- Fieberg, J. (2007). “Kernel density estimators of home range: smoothing and the autocorrelation red herring”. In: *Ecology* 88.4, pp. 1059–1066 (cit. on p. 66).
- Fieberg, J. and L. Börger (2012). “Could you please phrase home range as a question?” In: *Journal of Mammalogy* 93.4, pp. 890–902 (cit. on pp. 57, 66, 67).
- Kie, J. G. (2013). “A rule-based ad hoc method for selecting a bandwidth in kernel home-range analyses”. In: *Animal Biotelemetry* 1.1, pp. 1–12 (cit. on p. 59).
- Kie, J. G., J. Matthiopoulos, J. Fieberg, R. A. Powell, F. Cagnacci, M. S. Mitchell, J.-M. Gaillard, and P. R. Moorcroft (2010). “The home-range concept: are traditional estimators still relevant with modern telemetry technology?” In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 365.1550, pp. 2221–2231 (cit. on pp. 57, 59).
- Knight, C. M., R. E. Kenward, R. E. Gozlan, K. H. Hodder, S. S. Walls, and M. C. Lucas (2009). “Home-range estimation within

- complex restricted environments: importance of method selection in detecting seasonal change". In: *Wildlife Research* 36.3, pp. 213–224 (cit. on p. 56).
- Laver, P. N. and M. J. Kelly (2008). "A critical review of home range studies". In: *The Journal of Wildlife Management* 72.1, pp. 290–298 (cit. on pp. 59, 67).
- Lichti, N. I. and R. K. Swihart (2011). "Estimating utilization distributions with kernel versus local convex hull methods". In: *The Journal of Wildlife Management* 75.2, pp. 413–422 (cit. on p. 66).
- Manly, B., L. McDonald, D. Thomas, T. L. McDonald, and W. P. Erickson (2007). *Resource selection by animals: statistical design and analysis for field studies*. Springer Science & Business Media (cit. on p. 56).
- McClintock, B. T., R. King, L. Thomas, J. Matthiopoulos, B. J. McConnell, and J. M. Morales (2012). "A general discrete-time modeling framework for animal movement using multi-state random walks". In: *Ecological Monographs* 82.3, pp. 335–349 (cit. on p. 56).
- Morellet, N., C. Bonenfant, L. Börger, F. Ossi, F. Cagnacci, M. Heurich, P. Kjellander, J. D. Linnell, S. Nicoloso, P. Sustr, et al. (2013). "Seasonality, weather and climate affect home range size in roe deer across a wide latitudinal gradient within Europe". In: *Journal of animal ecology* 82.6, pp. 1326–1339 (cit. on p. 56).
- Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse (2008). "A movement ecology paradigm for unifying organismal movement research". In: *Proceedings of the National Academy of Sciences* 105.49, pp. 19052–19059 (cit. on p. 57).
- Nilsen, E. B., S. Pedersen, and J. D. Linnell (2008). "Can minimum convex polygon home ranges be used to draw biologically meaningful conclusions?" In: *Ecological Research* 23.3, pp. 635–639 (cit. on p. 66).
- Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos (2008). "State-space models of individual animal movement". In: *Trends in ecology & evolution* 23.2, pp. 87–94 (cit. on p. 56).
- Powell, R. A. and M. S. Mitchell (2012). "What is a home range?" In: *Journal of Mammalogy* 93.4, pp. 948–958 (cit. on pp. 57, 66).
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria (cit. on pp. 59, 60).
- Reinecke, H., L. Leinen, I. Thißen, M. Meißner, S. Herzog, S. Schütz, and C. Kiffner (2014). "Home range size estimates of red deer in Germany: environmental, individual and methodological correlates". In: *European journal of wildlife research* 60.2, pp. 237–247 (cit. on pp. 56, 59).
- Saura, S. and J. Martinez-Millan (2000). "Landscape patterns simulation with a modified random clusters method". In: *Landscape ecology* 15.7, pp. 661–678 (cit. on p. 58).
- Schuler, K. L., G. M. Schroeder, J. A. Jenks, and J. G. Kie (2014). "Ad hoc smoothing parameter performance in kernel estimates of GPS-derived home ranges". In: *Wildlife Biology* 20.5, pp. 259–266 (cit. on p. 66).
- Scull, P., M. Palmer, F. Frey, and E. Kraly (2012). "A comparison of two home range modeling methods using Ugandan mountain gorilla data". In: *International Journal of Geographical Information Science* 26.11, pp. 2111–2121 (cit. on p. 59).
- Seaman, D. E. and R. A. Powell (1996). "An evaluation of the accuracy of kernel density estimators for home range analysis". In: *Ecology* 77.7, pp. 2075–2085 (cit. on pp. 59, 66).
- Sheather, S. J. and M. C. Jones (1991). "A reliable data-based bandwidth selection method for kernel density estimation". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 53.3, pp. 683–690 (cit. on p. 59).
- Signer, J. and N. Balkenhol (2015). "Reproducible home ranges (rhr): A new, user-friendly R package for analyses of wildlife telemetry data". In: *Wildlife Society Bulletin* 39.2, pp. 358–363 (cit. on p. 60).
- Thurfjell, H., S. Ciuti, and M. S. Boyce (2014). "Applications of step-selection functions in ecology and conservation". In: *Movement Ecology* 2.4 (cit. on p. 56).
- Van Beest, F. M., I. M. Rivrud, L. E. Loe, J. M. Milner, and A. Mysterud (2011). "What determines variation in home range size across spatiotemporal scales in a large browsing herbivore?" In: *Journal of Animal Ecology* 80.4, pp. 771–785 (cit. on pp. 56, 66).
- Walter, W. D., K. C. VerCauteren, H. Campa III, W. R. Clark, J. W. Fischer, S. E. Hygnstrom, N. E. Mathews, C. K. Nielsen, E. M. Schaubert, T. R. Van Deelen, et al. (2009). "Regional assessment on influence of landscape configuration and connectivity on range size of white-tailed deer". In: *Landscape Ecology* 24.10, pp. 1405–1420 (cit. on p. 56).

Worton, B. J. (1989). "Kernel methods for estimating the utilization distribution in home-range studies". In: *Ecology* 70.1, pp. 164–168 (cit. on p. 59).

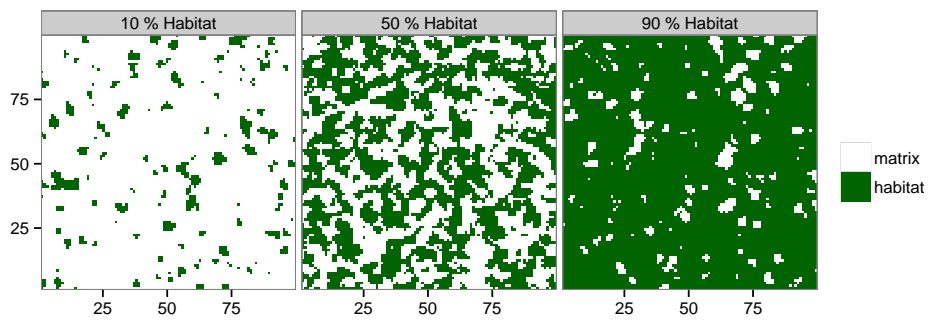
## 5.8 Appendix 1: Additional Figures

### Underlying pattern Scenario 1a and 1b

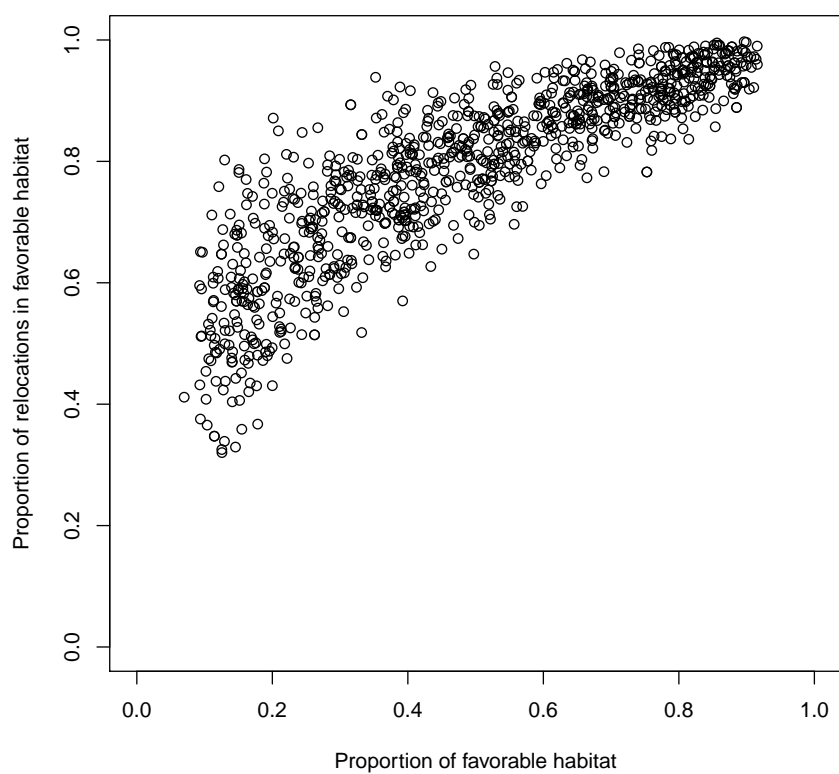


**Figure 5.A1:** To simulate varying movement costs over the course of one year we used a rescaled Gaussian density function. For 1b scenario we added additional sinusoidal noise.

## Landscapes Scenario 2

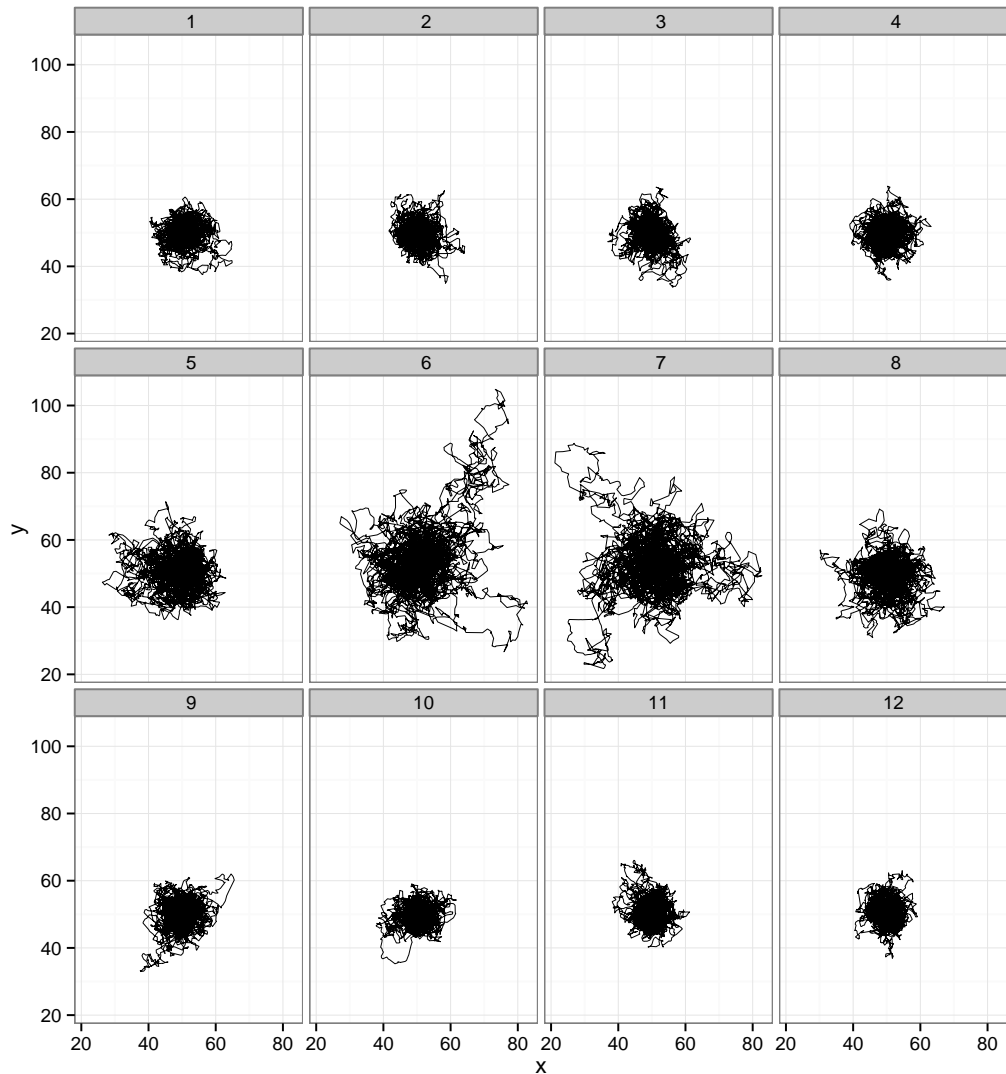


**Figure 5.A2:** Three examples of simulated landscapes with different amounts of favorable habitat, that were used for simulation scenario 2. Animals selected with a higher probability habitat cells and walked faster (larger step lengths) in the matrix (unfavorable habitat).



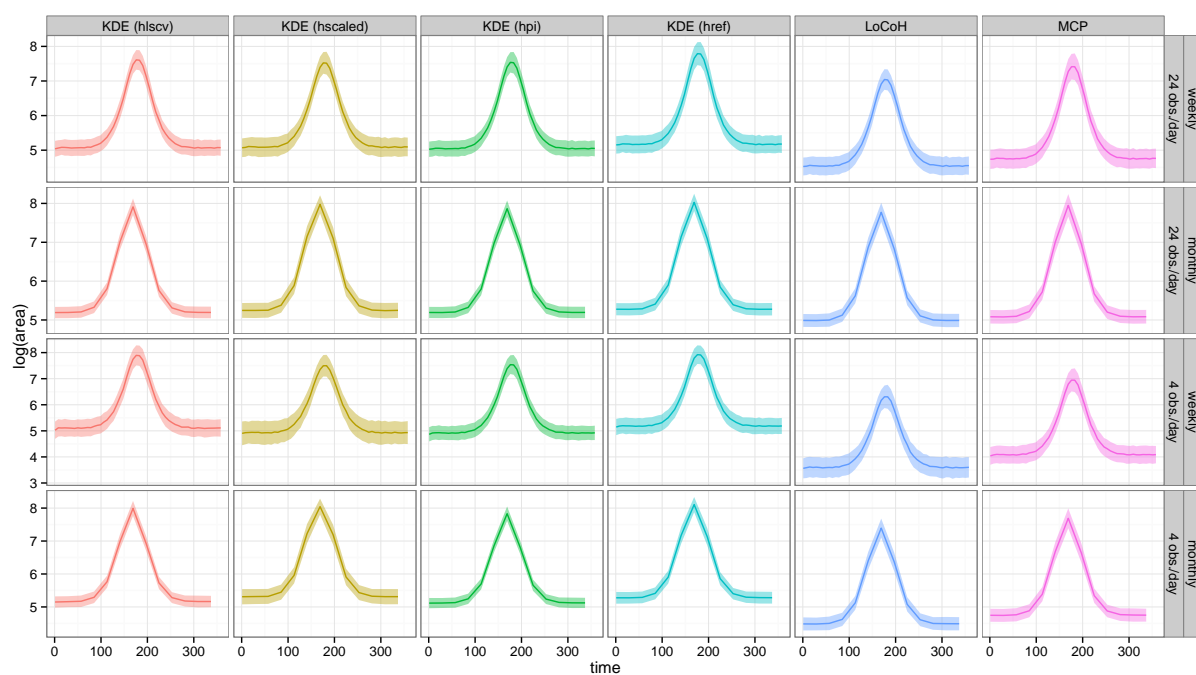
**Figure 5.A3:** Fraction of relocations located in favorable habitat as a function of the amount of favorable habitat in the landscape.

### Example Simulation Scenario 1

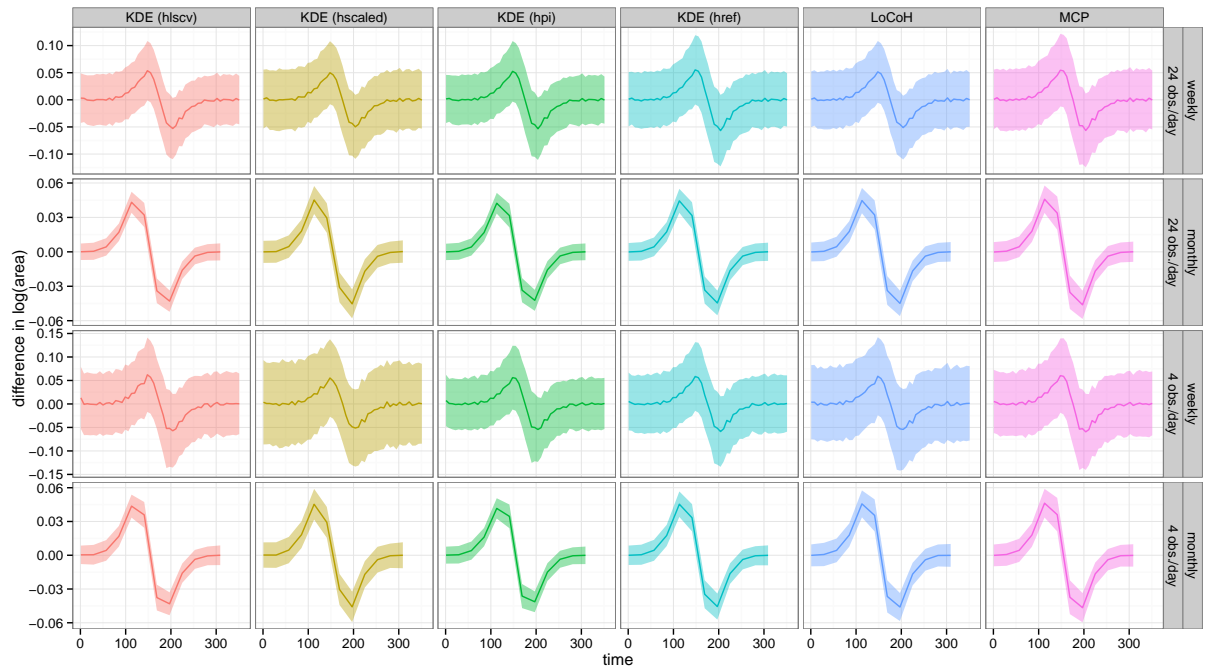


**Figure 5.A4:** One realization of simulation scenario 1a divided by month. Different panels show the movement path for different months. The pullback mechanism towards the home range center gets weaker in the middle of the year (months 5-8) and the animal is able to expand its range to account for higher movement costs.

## Results scenario 1a



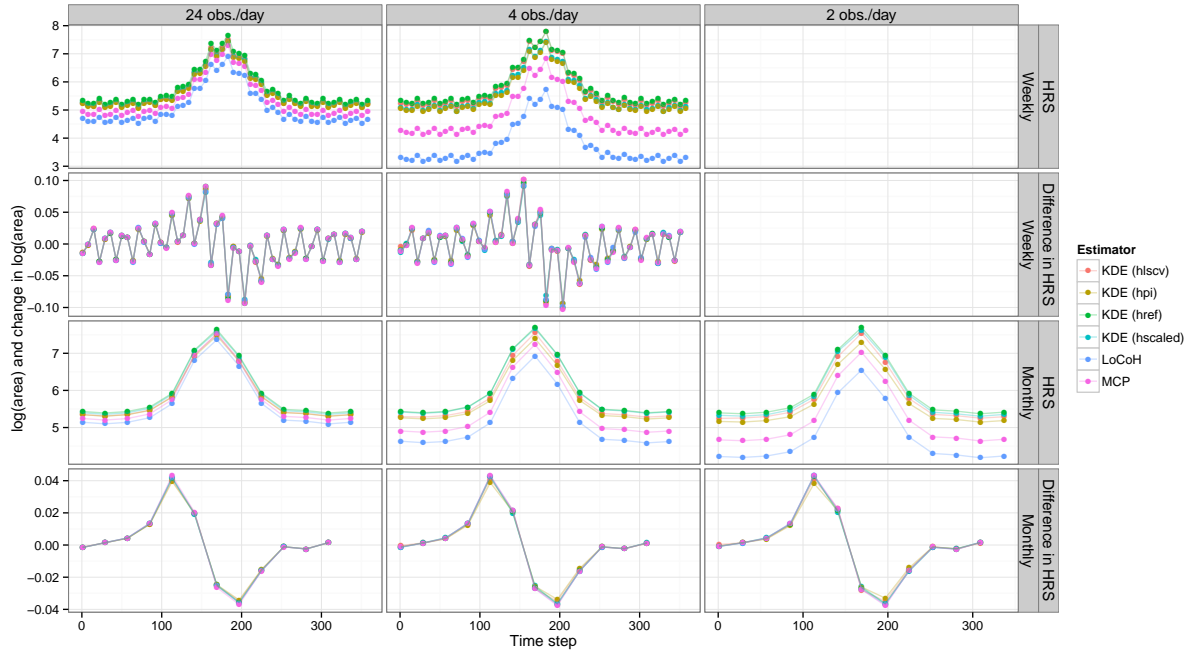
**Figure 5.A5:** Averaged  $\log(\text{HRS})$  ( $\pm$  standard deviation) of 1000 realizations of animals moving according to simulation scenario 1a for six different home range estimators (columns) and different analysis intervals and sampling rates (rows).



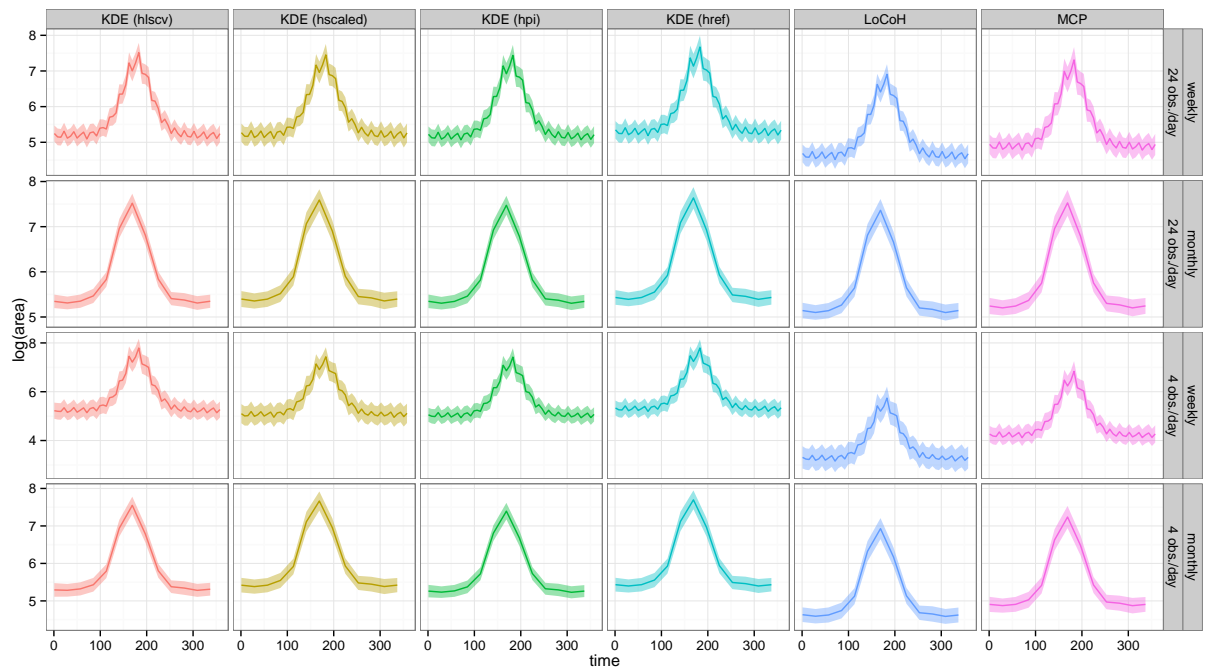
**Figure 5.A6:** Averaged changes of  $\log(\text{HRS})$  ( $\pm$  standard deviation) from 1000 realizations of animals moving according to simulation scenario 1a for six different home range estimators (columns) and different analysis intervals and sampling rates (rows).



Results scenario 1b

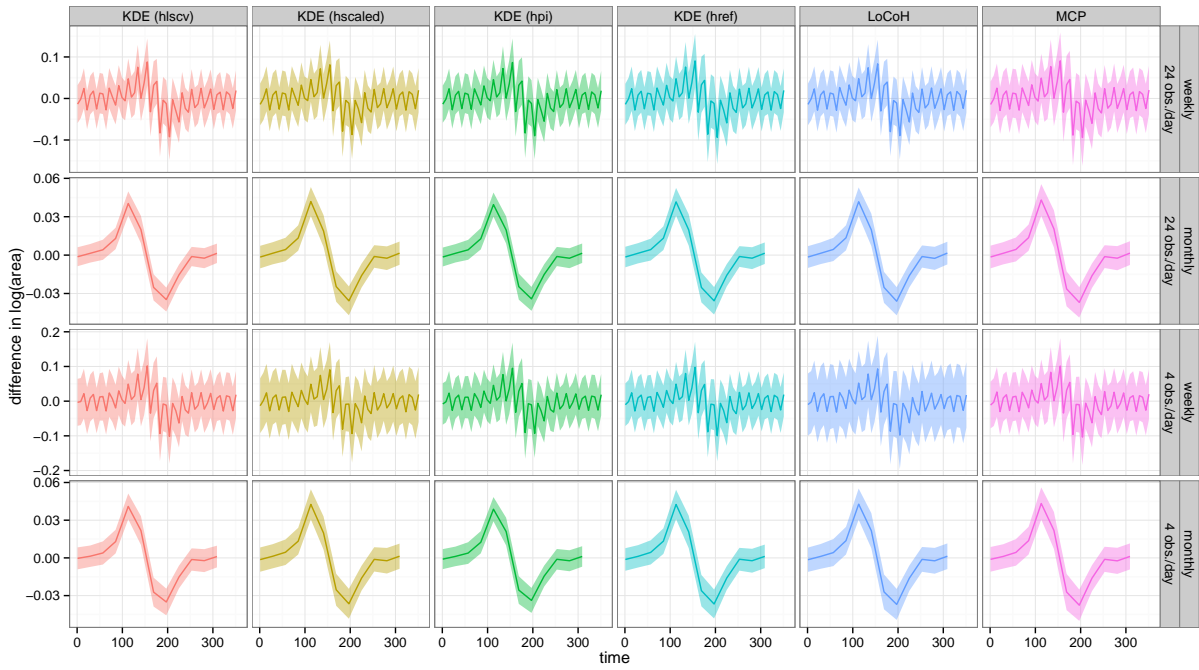


**Figure 5.A7:** Average estimates of  $\log(\text{HRS})$  (first and third row) and changes of  $\log(\text{HRS})$  (second and fourth row) from 1000 realization of animals moving according to simulation model 1b at different analysis intervals (rows) and sampling rates (rows). We present results for all 6 estimators (different colors), two analysis intervals and three sampling rates. We required at least 20 relocations for any combination of sampling rate and analysis interval. Hence we did not considered weekly home ranges for two observation per day



**Figure 5.A8:** Averaged  $\log(\text{HRS})$  ( $\pm$  standard deviation) of 1000 realizations of animals moving according to simulation scenario 1b for six different home range estimators (columns) and different analysis intervals and sampling rates (rows).

5.8 Appendix 1: Additional Figures



**Figure 5.A9:** Averaged changes of  $\log(\text{HRS})$  ( $\pm$  standard deviation) from 1000 realizations of animals moving according to simulation scenario 1b for six different home range estimators (columns) and different analysis intervals and sampling rates (rows).



# 6

## Detecting the influence of environmental covariates on animal movement: a semi-variance approach

An improved version of this chapter was submitted for publication to *Methods in Ecology and Evolution* as: **Signer, J** & Ovaskainen, O. Detecting the influence of environmental covariates on animal movement: a semivariance approach.

**Authors contributions:** OO provided the idea and the statistical framework. JS implemented the framework, performed the simulations, and applied the framework to the case study. JS and OO contributed to the interpretation of the results and writing of the manuscript.

### Contents

---

<b>6.1</b>	<b>Abstract</b>	<b>80</b>
<b>6.2</b>	<b>Introduction</b>	<b>80</b>
<b>6.3</b>	<b>Methods</b>	<b>81</b>
6.3.1	A motivating example	81
6.3.2	Extending the semi-variance framework to account for environmental covariates	82
6.3.3	Examining the performance of the statistical framework with a simulation study	86
6.3.4	Case study with red deer movement	87
<b>6.4</b>	<b>Results</b>	<b>87</b>
6.4.1	Simulation study	87
6.4.2	The red deer case study	88
<b>6.5</b>	<b>Discussion</b>	<b>91</b>
<b>6.6</b>	<b>Acknowledgments</b>	<b>92</b>
<b>6.7</b>	<b>References</b>	<b>93</b>
<b>6.8</b>	<b>Appendix 1: Additional Figures and Tables</b>	<b>94</b>

---

## 6.1 Abstract

Movements of organisms are almost always affected by environmental covariates. For example, the movement rates of animals in unsuitable habitats are often different from movement rates of animals foraging in high quality habitats. Different statistical methods to detect such effects exist, but they often rely on complete and regularly sampled data, or require the formulation of an explicit movement model.

We propose an extension to a recently introduced semi-variance framework to identify the effects of environmental or other kinds of covariates on animal movement. Our extension also applies for sparse and irregularly sampled data, and it does not require the formulation of an explicit movement model.

Within this extension of the semi-variance framework, the observed movement rates at different time lags are modeled as a linear regression of the covariates. To account for the inherent autocorrelation in semi-variance data, we test for the statistical significance of the influence of covariates using a permutation approach.

Our approach correctly identified covariates that influenced or did not influence movement rates in a simulation study. In a case study based on tracking data of a single female red deer (*Cervus elaphus*) individual from Southern Austria, an application of the method showed that movements rates peak during periods with intermediate temperature, and at intermediate altitudes, but they do not co-vary with precipitation.

## 6.2 Introduction

Many research questions in movement ecology are addressed by analyzing data on observed movement tracks of freely moving animals rather than data originating from controlled experiments. Deriving inferences from such data is not straightforward, as realized movement tracks result from a combination of many factors, such as the internal state of the individual, its interactions with con- and heterospecifics, and spatio-temporal variation in abiotic environmental conditions (Nathan et al. 2008). Data on movements and environmental features are becoming available at an increasing rate and resolution (Urbano et al. 2010; Dodge et al. 2013; Kays et al. 2015), opening a wide range of possibilities to address new questions in movement ecology (Kuenzer et al. 2014). Like with any statistical exercise with a large number of potential explanatory variables, identifying the causal explanatory variables and isolating influences from other confounding effects can be challenging. This is in particular the case with the analysis of movement tracks because many of the potential explanatory variables are typically not available, and because of the inherently auto-correlated nature of movement tracks.

A multitude of methods have been developed to identify different behavioral states from movement data, and to model the influence of environmental covariates on the movement of animals. These include, e.g. the use of Bayesian state-space models (e.g., Morales et al. 2004; Forester et al. 2007; Patterson et al. 2008), heterogeneous-space diffusion models (e.g., Ovaskainen 2004; Ovaskainen et al. 2008), agent based models (e.g., Avgar et al. 2013b; Bohrer et al. 2014), mixed effects models (Avgar et al. 2013b),

and step selection functions (e.g., Fortin et al. 2005; Thurfjell et al. 2014). Schick et al. (2008) stressed that analyses of movement data should account for variation in behavioral modes, variation in environmental conditions, and to the data sampling process (e.g., variation in sampling rate, missing observations due to low satellite coverage or technical failures, and so on). To overcome the problem of irregular sampling rates and incomplete data, Fleming et al. (2014b) and Fleming et al. (2014a) proposed to use a semi-variance framework as a general method for analyzing movement data. By applying a semi-variance function (SVF) to relocation data, one can use information at all time lags, not only at the sampling rate. Consequently, the SVF approach allows for detecting behavioral modes or influence of covariates that operate on different time scales than the sampling rate (Fleming et al. 2014a).

In this study, we extend the semi-variance approach to develop a simple but general framework that enables testing of influence of environmental covariates on the movement of animals. We first illustrate through simulations what kind of signature environmental heterogeneity (e.g. spatial variation in habitat structure or temporal variation in weather conditions) leaves to the semi-variance function. We then develop a permutation-based test for examining whether the influence of environmental covariates can be judged to be statistically significant, and demonstrate the validity of the test using simulated data. Finally, we demonstrate the practical relevance of our approach by applying it to relocation data on a single female red deer from Southern Austria. We close by discussing how our framework can be extended and integrated within more mechanistic approaches to movement modelling.

## 6.3 Methods

### 6.3.1 A motivating example

Consider two movement tracks (Fig. 6.1AC), one resulting from a homogeneous space random walk (Brownian motion, panel A) and the other one from a heterogeneous space random walk, in which step lengths are shorter in the patches than in the matrix (Fig. 6.1C). Fig. 6.2AB show variograms for these data, i.e., the squared displacement  $R_{ij}^2$  between locations for time steps  $i$  and  $j$ , as a function of the time lag  $\Delta t_{ij}$ . The time lags  $\Delta t_{ij}$  can be arbitrary, but if the data are collected at a regular time step  $\Delta$  like is the case here, it holds that  $\Delta t_{ij} = (j - i)\Delta$ . In these panels, the red tiles are the individual data points, whereas the dashed lines are the empirical means  $E[R_{ij}^2]$  based on averaging the data for each time lag. As it is well known (Gurarie and Ovaskainen 2011; Fleming et al. 2014a), the variogram for random walk increases linearly with the time lag (Fig. 6.2AB),

$$E[R_{ij}^2] = 4D\Delta t_{ij}, \quad (6.1)$$

where  $D$  is the diffusion parameter that can be derived from the parameterization of the random walk. For the heterogeneous space random walk the variogram shows a linear increase for the pooled data (black line in Fig. 6.2B). If dividing the data based

on whether the individual was originally in a patch or in the matrix, the rate of the linear increase is lower for those cases where the individual was initially in a patch (Fig. 6.2B; green line for patch and blue line for matrix). Figs. 6.2CD show the same data, but now translated to the diffusion coefficient  $D_{ij}$  that corresponds to the displacement from time step  $i$  to time step  $j$ , defined by

$$D_{ij} = \frac{R_{ij}^2}{4\Delta t_{ij}}. \quad (6.2)$$

Based on Eq. 6.1, for homogeneous space random walk the expectation of  $D_{ij}$  is a constant for any time lag  $\Delta t_{ij}$  (Fig. 6.2C), so that  $E[D_{ij}] = D$ . For the heterogeneous space random walk,  $E[D_{ij}]$  is constant for the pooled data, but for small time lags it shows two different values if the data are split according to the original location of the individual (Fig. 6.2D).

Fleming et al. (2014a) illustrated how the variogram approach can be used as a data exploration tool to separate different movement types, such as pure random walk and random walk in the presence of home range behavior (e.g., the Ornstein-Uhlenbeck process; Dunn and Gipson 1977). In this paper, we extend the semi-variance framework to obtain insight on covariates that may influence movements. Our starting point is the observation that the variogram can include information about differential movements in different habitat types (Fig. 6.2A-D), or more generally, about the influence of any covariates on movements.

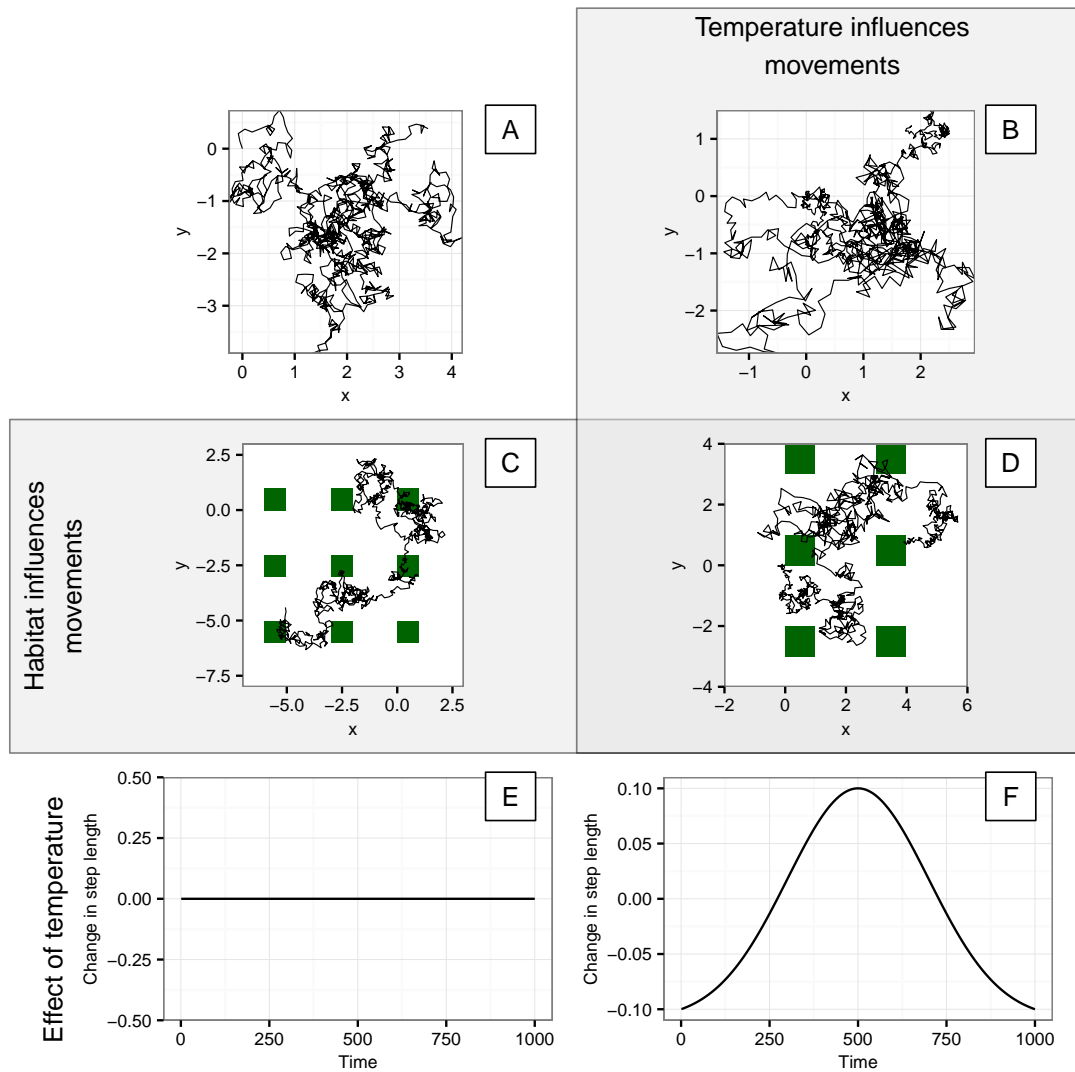
### 6.3.2 Extending the semi-variance framework to account for environmental covariates

Our approach stems from the idea that displacement-specific diffusion rates can be modeled as a function of covariates. As an example, Fig. 6.2E shows how the observed diffusion rate depends on the underlying habitat type. Instead of the habitat type in which the individual is initially located (Figs. 6.2BD), in Fig. 6.2E we use as the covariate the fraction of locations between the original and final location (including the endpoints) which belong to the matrix. Fig. 6.2E illustrates that movements restricted solely within a patch show a lower diffusion rate than movements restricted solely within the matrix, as expected from the longer step size in the matrix. For movement tracks that include sections both within the matrix and the patch, the mean diffusion rate increases with the fraction of locations which are within the matrix. For the log-transformed diffusion rate, the increase is approximately linear, suggesting that the linear model

$$\log(D_{ij}) = \alpha + \beta z_{ij} + \epsilon_{ij} \quad (6.3)$$

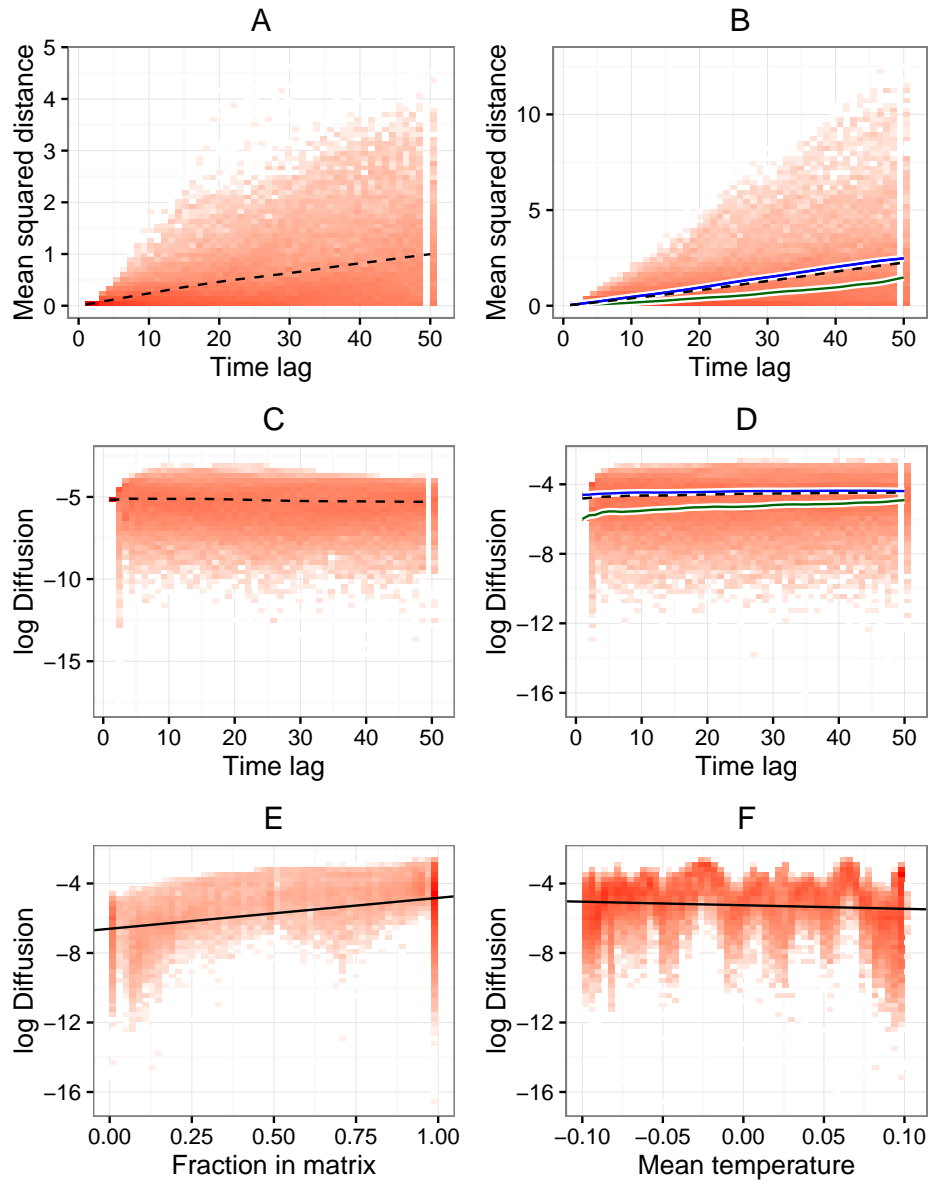
is a good statistical description of the data. Here the covariate  $z_{ij}$  is the fraction of locations from time step  $i$  to time step  $j$  that fall into matrix. Thus the log-transformed diffusion rate within a patch is estimated by the intercept  $\alpha$ , whereas the log-transformed diffusion rate within the matrix is estimated by  $\alpha + \beta$ .





**Figure 6.1:** Illustrations of the four different random walk simulation scenarios.

Panels ABCD show movement paths simulated for 1000 time steps with random walk models and panels EF show the effect of temperature on step lengths. In panel A step lengths were assumed to be constant ( $L = 0.15$ ), in panel C they were assumed to be shorter in the patches ( $L = 0.1$ ; green) than in the matrix ( $L = 0.2$ ; white), and in panel B step lengths were additively influenced by temperature as shown in panel F. The influence of temperature peaked at  $t = 500$  ( $L = 0.25$ ). In panel D step lengths were influenced by habitat and temperature jointly.



**Figure 6.2:** The semivariogram and the dependency of lag-specific diffusion rates on environmental covariates. Panels A and B show semivariograms derived from the data of panels A and C in Fig. 6.1. The lines show the mean values for the pooled data (black dashed line), for steps that start in habitat (green line), and for steps that start in the matrix (blue line). Panels C and D show the same data, but we have converted the mean squared displacement into log-transformed diffusion rate by Eq. 6.2. Panels E and F show the log-transformed diffusion coefficients (computed by Eq. 6.2) as a function of fraction of steps within matrix between the two relocations (panel E; for movement path Fig. 6.1C), and as a function of temperature average over the relocations (panel F).

**Table 6.1:** Alternative hypotheses tested with the simulation study (Fig. 6.3 and Table 6.A1 in the Appendix). Each hypothesis was tested by fitting a specific model to the semi-variance data, and computing the probability  $P$  that the regression coefficient related to the hypothesis (shown in bold) is smaller (in absolute value) than the null distribution. The null-distribution was obtained by permuting the variable whose significance is to be tested.

Hypothesis	Test statistic	Model to fit	Permutation of
Habitat influences movement	$P(H)$	$\log(D) = \alpha + \beta_H H$	habitat
Habitat influences movement, if controlling for temperature	$P(H T)$	$\log(D) = \alpha + \beta_H H + \beta_T T$	habitat
Temperature influences movement	$P(T)$	$\log(D) = \alpha + \beta_T T$	temperature
Temperature influences movement, if controlling for habitat	$P(T H)$	$\log(D) = \alpha + \beta_H H + \beta_T T$	temperature

More generally, we can extend the linear model of Eq. 6.3 to include an arbitrary number of covariates, yielding

$$\log(D_{ij}) = \alpha + \sum_{k=1}^m \beta_k z_{k,ij} + \epsilon_{ij} \quad (6.4)$$

where  $z_{k,ij}$  is the covariate  $k$  measured for the displacement from time step  $i$  to time step  $j$ . In the case of simulated data to be described below, we consider two covariates: the fraction of the locations that fall into matrix ( $z_{1,ij}$ ), and the mean temperature between time steps  $i$  and  $j$  ( $z_{2,ij}$ ). In Fig. 6.2EF, we show the diffusion rate as a function of these two covariates for data generated by simulations in which habitat type influenced the movement rate but temperature not. As expected, the slope is non-zero for  $z_{1,ij}$  (Fig. 6.2E) but essentially zero for  $z_{2,ij}$  (Fig. 6.2F).

In a nutshell, the statistical approach we propose is based on selecting the set of environmental variables that are hypothesized to influence movements, fitting the Eq. 6.4 to data, and examining if the parameters  $\beta_k$  are non-zero.

Table 6.1 illustrates the collection of models for the case of two covariates (termed habitat,  $H$ , and temperature,  $T$ ) that include or exclude the influences of the two factors, and which can thus be used to disentangle the marginal and joint influences of those factors on the movement. Let us start from the model  $\alpha + \beta_H H$  that only includes the effect of habitat, which model can be used to ask whether the movement rate co-varies with habitat. The reason why applying Eq. 6.4 for this question is not straightforward is that the residuals are far from independent. To see this, assume that there are  $n$

locations ( $i = 1, \dots, n$ ). While the number of movements steps and thus the degrees of freedom is  $n - 1$ , the number of data points is  $n(n - 1)/2$ , as this is the number of ordered pairs of locations  $ij$  with  $j > i$ . The data points are correlated both within and among time lags. For example, for time lag  $\Delta t_{ij} = 5$ , one data point is given by the movement track from time 1 to time 6 (assuming regularly sampled data with time step  $\Delta t = 1$ ), and another data point is given by the movement track from time 2 to time 7. These two data points share the same track from time 2 to time 6. Similarly, the track from 1 to 5 yields a data point for time lag  $\Delta t = 4$ , which data point is correlated with that from 1 to 6 (with time lag  $\Delta t = 5$ ). As the variogram data points are dependent, the statistical significance of the regression slope  $\beta_H H$  being non-zero cannot be examined from the usual p-value of the linear regression. To overcome this problem, one possibility would be to explicitly model the autocorrelation structure. However, as a simpler alternative, we apply a permutation approach, in which we permute the covariate of interest to break the connection between the covariates and the movement track. To construct data conforming with the null hypothesis of habitat not having an effect, we permute the column associated to habitat in the data matrix. The distribution of effect sizes  $\beta_H$  derived from the permuted data matrices yield the distribution of  $\beta_H$  values corresponding to the null hypothesis.

In Table 6.1 we denote by  $P(H)$  the fraction of cases for which the observed effect size of the model  $\alpha + \beta_H H$  is greater (in absolute value) than values of the null distribution obtained by the permutation approach.  $P(H)$  measures the evidence in the data that movements co-vary with habitat type, if not controlling for any other variables. To account for confounding effects, we measure the evidence by  $P(H|T)$  in the data that movements co-vary with habitat type, after controlling for the effect of temperature.  $P(H|T)$  can be computed by fitting a model with both habitat and temperature ( $\alpha + \beta_H H + \beta_T T$ ; second row in Table 6.1), and comparing the effect of habitat ( $\beta_H$ ) to that obtained for null models where habitat has been permuted but temperature not. Analogously,  $P(T)$  examines if movement rate co-varies with temperature, and  $P(T|H)$  examines if movement rate co-varies with temperature after controlling for the effect of habitat (Table 6.1).

### 6.3.3 Examining the performance of the statistical framework with a simulation study

We used simulated data to test the performance of the proposed extension to the semi-variance framework. Simulations were conducted using a continuous space and discrete time model. We considered a fully balanced study design in which the movements could be influenced by two factors, named ‘habitat’ and ‘temperature’, thus resulting in four different scenarios. Habitat was assumed to consist of patches surrounded by matrix (Fig. 6.1C), whereas temperature was assumed to vary in a seasonal fashion and to peak in the middle of the study period (Fig. 6.1F).

For the null model (no covariates influencing movement), the animal performed a random walk with constant step length ( $L = 0.15$ ) and uniformly distributed turning angles (from 0 to  $2\pi$ ). In cases where habitat influenced movement, we mimicked for-

aging movement in suitable habitats and traversal movements in unsuitable habitats by assuming the step lengths  $L = 0.10$  for patches and  $L = 0.20$  for matrix. In cases where temperature influenced movements, we assumed that the step length increased with increasing temperature (Fig. 6.1F).

We simulated movement tracks for 1000 time steps (say days) for each of the four scenarios (with and without habitat and temperature influencing movement). To assess the influence of sample size, we subsampled the movement tracks by selecting every 10th data point, yielding 100 locations, and conducted analyses both for the full and subsampled data. We performed the four tests of Table 6.1 (with 200 permutations) to each of the simulated data sets to test whether there was evidence of the two covariates shaping the movement data. To explore the influence of the maximal time lag included, we repeated the analyses with four different maximum time lags (25, 50, 100 days). We replicated each simulation scenario 100 times to assess the type 1 and type 2 errors.

### 6.3.4 Case study with red deer movement

We used relocation data of one female red deer (*Cervus elaphus*) tracked in an alpine habitat in Southern Austria. For this analysis we used one daily position (recorded at noon) for one year from April 2014 to April 2015, thus in total there are 358 locations (for 7 days the location was missing). The animal was equipped with a 'Vectronics GPS Plus' GPS collar (permission to capture deer was granted by the Amt der Kärntner Landesregierung Abt. 10). We were interested in testing if (i) altitude, (ii) temperature and (iii) rain had a significant influence on the movement of the red deer. We calculated the covariates as the mean altitude and the mean temperature between the first and last location, and the portion of time without rainfall between the first and last location. For altitude and temperature we included also squared values of the covariates to test whether movement activity is maximized at intermediate values. To explore the ability of the framework to deal with missing data, we repeated the analyses and randomly removed 5, 10, 25 and 50% locations 100 times to assess the match with the full data set. The weather data was obtained from the Austrian Weather service. We used an  $\alpha$ -value of 0.05 to determine significance.

All analyses (both for simulated and case study) were implemented in Program R (R Core Team 2015). The code used to generate the simulated data, as well the red deer data, are available in the online Appendix.

## 6.4 Results

### 6.4.1 Simulation study

Generally, the influence of covariates on movement rates was correctly detected by the proposed statistical framework (Fig. 6.3). The ability of the method to identify a significant effect of a covariate was highest for an intermediate maximal time lag (Fig. 6.3). Further, the results were more accurate for the full data (see 6.A1 in the Appendix) than the sub-sampled data (Fig. 6.3), but the difference between these two was not large.

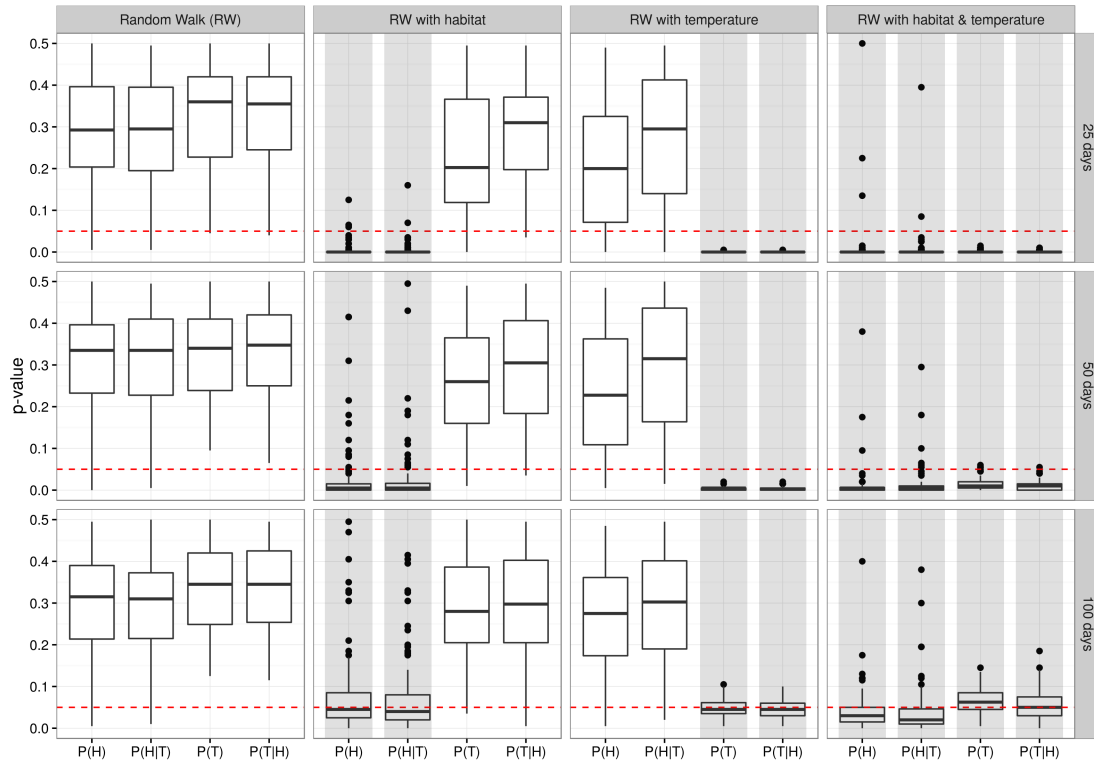
For the majority of simulation scenarios, the type 1 error rates were lower than 0.05 indicating that on average our tests of statistical significance are somewhat conservative (Fig. 6.3; Appendix). The maximal type 1 error rate over the scenarios was 0.095.

### 6.4.2 The red deer case study

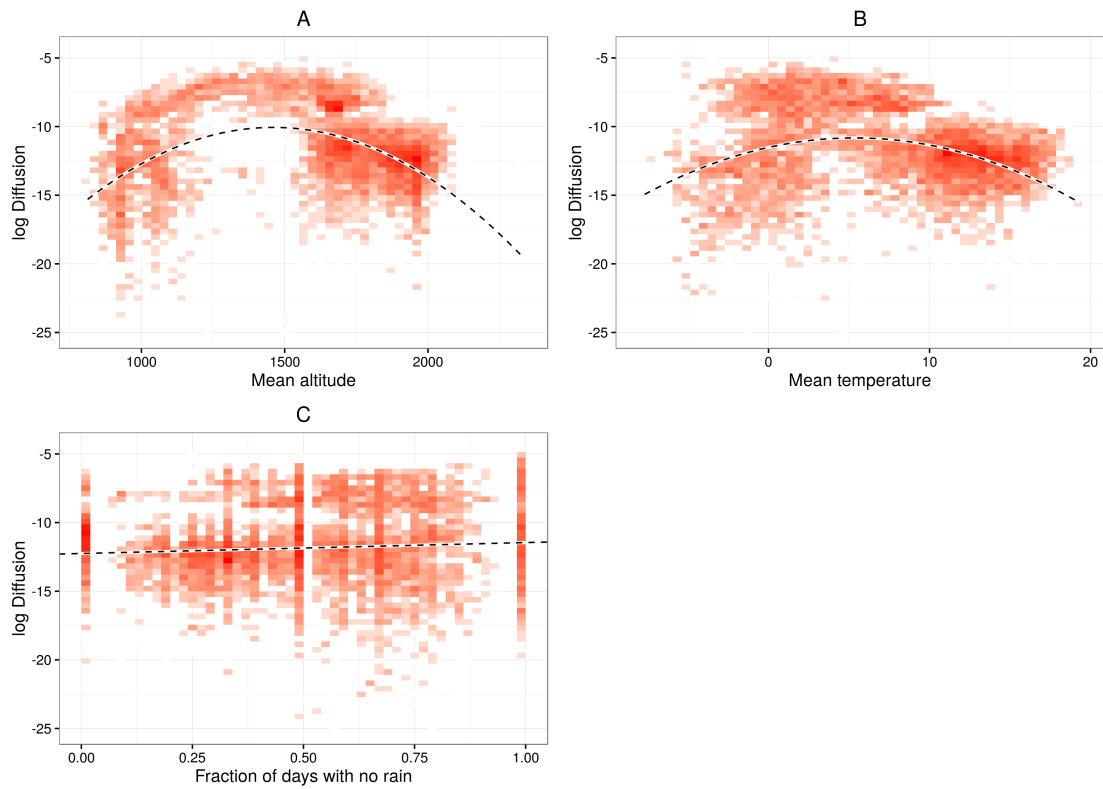
Movements of the female red deer were significantly influenced by altitude and temperature but not by rainfall (Table 6.2). The movement rates peaked at intermediate altitudes and temperatures (Fig. 6.4). Results did not change after randomly removing a moderate amount (10-20%) of the locations (Table 6.A3), suggesting robustness of the framework to missing locations.

**Table 6.2:** Results based on applying the statistical framework on movement data of a red deer individual (see Methods). The numbers show the p-values from permutation tests examining the influences of altitude (A), temperature (T) and rainfall (R), with the other factors not being controlled for. For altitude and temperature, second order terms were included to test the hypothesis that movement rates peak at an intermediate altitude or temperature.

Hypothesis	p-value
P(A)	0.015
P(A+A <sup>2</sup> )	0.000
P(A T+T <sup>2</sup> )	0.000
P(A+A <sup>2</sup>  T+T <sup>2</sup> )	0.000
P(A R)	0.005
P(A+A <sup>2</sup>  R)	0.000
P(A R+T+T <sup>2</sup> )	0.000
P(A+A <sup>2</sup>  R+T+T <sup>2</sup> )	0.000
P(T)	0.035
P(T+T <sup>2</sup> )	0.000
P(T A+A <sup>2</sup> )	0.000
P(T+T <sup>2</sup>  A+A <sup>2</sup> )	0.000
P(T R)	0.065
P(T+T <sup>2</sup>  R)	0.000
P(T A+A <sup>2</sup> +R)	0.000
P(T+T <sup>2</sup>  A+A <sup>2</sup> +R)	0.000
P(R)	0.315
P(R A+A <sup>2</sup> )	0.240
P(R T+T <sup>2</sup> )	0.490
P(R A+A <sup>2</sup> +T+T <sup>2</sup> )	0.345



**Figure 6.3:** Power to detect influence of environmental covariates on animal movement. The box plots show the distributions of p-values from different hypothesis tests (Table 6.1) for different simulation scenarios (columns) and different values of the maximum lag included in the semivariogram (rows). Results below the red dashed line are considered significant ( $\alpha = 0.05$ ). Tests for which we expected significant results are highlighted with gray vertical bars. The box plots show the median (horizontal line), 50% quantile (box) and 95% quantiles (vertical lines) of the 100 p-values obtained by replicating the simulations. Black dots indicate outliers. For the same data in numerical format, see Table 6.A1 in the Appendix. The results are shown for data subsampled to 100 data points. For the corresponding figure based on the full data, see Fig. 6.A1 and Table 6.A2 in the Appendix.



**Figure 6.4:** Lag-specific diffusion rates of a female red deer for altitude, temperature and rain. Lag-specific diffusion rates are displayed as a function of the mean altitude of all points within a lag (including end points; A), mean temperature between the first and the last relocation (B), and the fraction of days with no rain between the first and last relocation (including end points; C). Model predictions (Eq. 6.4) include a square term for altitude and temperature (black dashed line).



## 6.5 Discussion

We have presented a statistical framework to test for the effect of environmental covariates extending the semi-variance approach for the analysis of movement data (Fleming et al. 2014a), that allows addressing questions such as ‘Is an individual’s movement influenced by a given environmental covariate and in which direction is the effect?’. Both a simulation study and a study that applied our framework to real data demonstrated its ability to detect signals of environmental covariates in the movement track. Contrary to other approaches (e.g., McClintock et al. 2012; Avgar et al. 2013a), our framework is free of assumptions regarding the underlying movement process and tests for the influence of covariates on movements by permutations. By modelling the movement rates as a function of covariates, our framework is related to a recent extension of the Brownian Bridge Movement Model (Horne et al. 2007), that maps velocities. Buchin et al. (2015) developed and used this approach to study the effect of Normalized Differenced Vegetation Index (NDVI) on movement speed of vervet monkeys. More mechanistic approaches will likely yield to more detailed information about the movement process (Avgar et al. 2013a), different behavioral states (e.g., McClintock et al. 2012; Fleming et al. 2014a), and if animals, for example, follow resources over time (Bohrer et al. 2014). Hence, we see our framework as a first step to explore if movements are influenced by covariates, as done elsewhere for the influence of conspecifics (Delgado et al. 2014).

The proposed framework is based on the recently introduced semi-variance approach for analysis of animal telemetry data (Fleming et al. 2014a). The semi-variance framework offers a novel way to analyze telemetry data, by considering movement rates up to a given time lag. A major advance of the semi-variance framework is its ability to overcome the problem of missing data (Fleming et al. 2014a), which frequently occurs in telemetry studies (Schick et al. 2008). The issue of missing data has been recognized for a long time (e.g., Jonsen et al. 2003). Previously, missing data was addressed with state spaces models (reviewed by Patterson et al. 2008). Fitting state space models, however, can pose challenges to the practitioner because specifying the movement and error models, and fitting such models can be challenging.

Applying our framework requires two decisions to be made. First, a suitable maximum time lag must be decided on. The simulation study suggested that if the maximum time lag is too large the method fails to detect the signal of environmental covariates on movements (Fig. 6.3). If the maximum time lag is exceedingly large, the majority of data points in the analyses (such as illustrated in Fig. 6.2A) will be for large time lags and will not be informative. For example, the effect of habitat for larger time lags results in the animal exploring many habitats and hence the signal averages out. The optimal time lag will be a trade off between sufficient data to detect the signal and not too many data points to average out the signal.

Secondly, the researcher must decide which (environmental) covariates to include in the study, and how to summarize covariates within time lags. Decisions about the choice of covariates requires a good understanding of the life history and ecology of the study species and will depend hypotheses to be tested and data availability. When calculating

summary statistics within lags, any function can be applied either to the start and end point, or to any data point that lies in between. For example, in the simulation study we calculated the fraction of steps that fell into the matrix between the start and end point (using all points), and for temperature we calculated the mean between the start and end point. It would be easy to use the framework to test for the influence of non-environmental covariates. For example, models that test for the effect of internal processes (e.g., Ropert-Coudert and Wilson 2005) on animal movements using other biotelemetry sensors could be implemented using the framework.

Using simulations, we were able to illustrate the ability of the proposed statistical framework to identify the influence of environmental covariates on movements, if they were included in the simulation. Further, using a case study the framework led to the same conclusion after removing 15% of the location data, illustrating the ability of the proposed method to deal with incomplete movement tracks. However, the framework could be further improved by introducing additional statistical rigor and explicitly modelling the co-variance structure of semi-variance data. This would allow the use of information theoretic model selection, and 'standard' p-values, as well as, testing for many covariates simultaneously. Permutations, as currently in use, pose logistical and computational challenges for large numbers of covariates. Finally, our framework could be extended to include different behavioral modes in the model, to account for the reason the variogram was initially developed Fleming et al. (2014a).

We see two possible applications for our framework. First, summary statistics obtained from the semivariogram could be used to contrast real data with simulated data as an indication of model fit when fitting complex mechanistic movement models to data with Approximate Bayesian Computing approaches. Second, the method can be used to quickly (without the need of specifying a movement model) identify influences of environmental covariates on movement in the stage of exploratory data analysis. This is of potential value for scientists and practitioners that wish to understand the animal movement process better, as it is required for successful management planning and conservation Thirgood et al. (2004) and Schick et al. (2008).

## **6.6 Acknowledgments**

This chapter originated out of a working group involving Luca Börger, Kamil Barton, Thomas Cornulier, Maria Delgado, Pim Edelaar, Bill Fagan, Andrea Giometto, Roi Harel, Jenny Hodgson, Steve Palmer and Andreas Schmitz during the Aberdeen meeting on Movement and Dispersal Ecology in November 2013. We are grateful for comments from Kamil Barton, Luca Börger, Bill Fagan, Justin Calabrese and Chris Fleming on earlier versions of this draft. We are grateful to Horst Leitner and Hohen Tauern Nationalpark for allowing us to use the red deer data. The study was financially supported by the Academy of Finland (Grant no. 250444 to OO) and the Research Council of Norway (Centres of Excellence funding scheme, project number 223257).

## 6.7 References

- Avgar, T., R. Deardon, and J. M. Fryxell (2013a). “An empirically parameterized individual based model of animal movement, perception, and memory”. In: *Ecological Modelling* 251, pp. 158–172 (cit. on p. 91).
- Avgar, T., A. Mosser, G. S. Brown, and J. M. Fryxell (2013b). “Environmental and individual drivers of animal movement patterns across a wide geographical gradient”. In: *Journal of Animal Ecology* 82.1, pp. 96–106 (cit. on p. 80).
- Bohrer, G., P. Beck, S. M. Ngene, A. K. Skidmore, and I. Douglas-Hamilton (2014). “Elephant movement closely tracks precipitation-driven vegetation dynamics in a Kenyan forest-savanna landscape”. In: *Mov Ecology* 2.2 (cit. on pp. 80, 91).
- Buchin, K., S. Sijben, E. E. van Loon, N. Sapir, S. Mercier, T. J. M. Arseneau, and E. P. Willems (2015). “Deriving movement properties and the effect of the environment from the Brownian bridge movement model in monkeys and birds”. In: *Movement ecology* 3.1 (cit. on p. 91).
- Delgado, M. d. M., V. Penteriani, J. M. Morales, E. Gurarie, and O. Ovaskainen (2014). “A statistical framework for inferring the influence of conspecifics on movement behaviour”. In: *Methods in Ecology and Evolution* 5.2, pp. 183–189 (cit. on p. 91).
- Dodge, S., G. Bohrer, R. Weinzierl, S. C. Davidson, R. Kays, D. Douglas, S. Cruz, J. Han, D. Brandes, and M. Wikelski (2013). “The environmental-data automated track annotation (Env-DATA) system: linking animal tracks with environmental data”. In: *Movement Ecology* 1.3, pp. 1–14 (cit. on p. 80).
- Dunn, J. E. and P. S. Gipson (1977). “Analysis of radio telemetry data in studies of home range”. In: *Biometrics* 33.1, pp. 85–101 (cit. on p. 82).
- Fleming, C. H., J. M. Calabrese, T. Mueller, K. A. Olson, P. Leimgruber, and W. F. Fagan (2014a). “From fine-scale foraging to home ranges: A semi-variance approach to identifying movement modes across spatiotemporal scales”. In: *The American Naturalist* 183.5, E154–E167 (cit. on pp. 81, 82, 91, 92).
- Fleming, C. H., J. M. Calabrese, T. Mueller, K. A. Olson, P. Leimgruber, and W. F. Fagan (2014b). “Non-Markovian maximum likelihood estimation of autocorrelated movement processes”. In: *Methods in Ecology and Evolution* 5.5, pp. 462–472 (cit. on p. 81).
- Forester, J. D., A. R. Ives, M. G. Turner, D. P. Anderson, D. Fortin, H. L. Beyer, D. W. Smith, and M. S. Boyce (2007). “State-space models link elk movement patterns to landscape characteristics in Yellowstone National Park”. In: *Ecological Monographs* 77.2, pp. 285–299 (cit. on p. 80).
- Fortin, D., H. L. Beyer, M. S. Boyce, D. W. Smith, T. Duchesne, and J. S. Mao (2005). “Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park”. In: *Ecology* 86.5, pp. 1320–1330 (cit. on p. 81).
- Gurarie, E. and O. Ovaskainen (2011). “Characteristic spatial and temporal scales unify models of animal movement”. In: *The American Naturalist* 178.1, pp. 113–123 (cit. on p. 81).
- Horne, J. S., E. O. Garton, S. M. Krone, and J. S. Lewis (2007). “Analyzing animal movements using Brownian bridges”. In: *Ecology* 88.9, pp. 2354–2363 (cit. on p. 91).
- Jonsen, I. D., R. A. Myers, and J. M. Flemming (2003). “Meta-analysis of animal movement using state-space models”. In: *Ecology* 84.11, pp. 3055–3063 (cit. on p. 91).
- Kays, R., M. C. Crofoot, W. Jetz, and M. Wikelski (2015). “Terrestrial animal tracking as an eye on life and planet”. In: *Science* 348.6240, aaa2478 (cit. on p. 80).
- Kuenzer, C., M. Ottinger, M. Wegmann, H. Guo, C. Wang, J. Zhang, S. Dech, and M. Wikelski (2014). “Earth observation satellite sensors for biodiversity monitoring: potentials and bottlenecks”. In: *International Journal of Remote Sensing* 35.18, pp. 6599–6647 (cit. on p. 80).
- McClintock, B. T., R. King, L. Thomas, J. Matthiopoulos, B. J. McConnell, and J. M. Morales (2012). “A general discrete-time modeling framework for animal movement using multi-state random walks”. In: *Ecological Monographs* 82.3, pp. 335–349 (cit. on p. 91).
- Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell (2004). “Extracting more out of relocation data: building movement models as mixtures of random walks”. In: *Ecology* 85.9, pp. 2436–2445 (cit. on p. 80).
- Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse (2008). “A movement ecology paradigm for unifying organismal movement research”. In: *Proceedings of the National Academy of Sciences* 105.49, pp. 19052–19059 (cit. on p. 80).
- Ovaskainen, O. (2004). “Habitat-specific movement parameters estimated using mark-recapture data

- and a diffusion model”. In: *Ecology* 85.1, pp. 242–257 (cit. on p. 80).
- Ovaskainen, O., M. Luoto, I. Ikonen, H. Rekola, E. Meyke, and M. Kuussaari (2008). “An empirical test of a diffusion model: predicting clouded apollo movements in a novel environment”. In: *The American Naturalist* 171.5, pp. 610–619 (cit. on p. 80).
- Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos (2008). “State-space models of individual animal movement”. In: *Trends in ecology & evolution* 23.2, pp. 87–94 (cit. on pp. 80, 91).
- Roport-Coudert, Y. and R. P. Wilson (2005). “Trends and perspectives in animal-attached remote sensing”. In: *Frontiers in Ecology and the Environment* 3.8, pp. 437–444 (cit. on p. 92).
- Schick, R. S., S. R. Loarie, F. Colchero, B. D. Best, A. Boustany, D. A. Conde, P. N. Halpin, L. N. Joppa, C. M. McClellan, and J. S. Clark (2008). “Understanding movement data and movement processes: current and emerging directions”. In: *Ecology Letters* 11.12, pp. 1338–1350 (cit. on pp. 81, 91, 92).
- Thirgood, S., A. Mosser, S. Tham, G. Hopcraft, E. Mwangomo, T. Mlengeya, M. Kilewo, J. Fryxell, A. Sinclair, and M. Borner (2004). “Can parks protect migratory ungulates? The case of the Serengeti wildebeest”. In: *Animal Conservation* 7.02, pp. 113–120 (cit. on p. 92).
- Thurfjell, H., S. Ciuti, and M. S. Boyce (2014). “Applications of step-selection functions in ecology and conservation”. In: *Movement Ecology* 2.4 (cit. on p. 81).
- Urbano, F., F. Cagnacci, C. Calenge, H. Dettki, A. Cameron, and M. Neteler (2010). “Wildlife tracking data management: a new vision”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 365.1550, pp. 2177–2185 (cit. on p. 80).

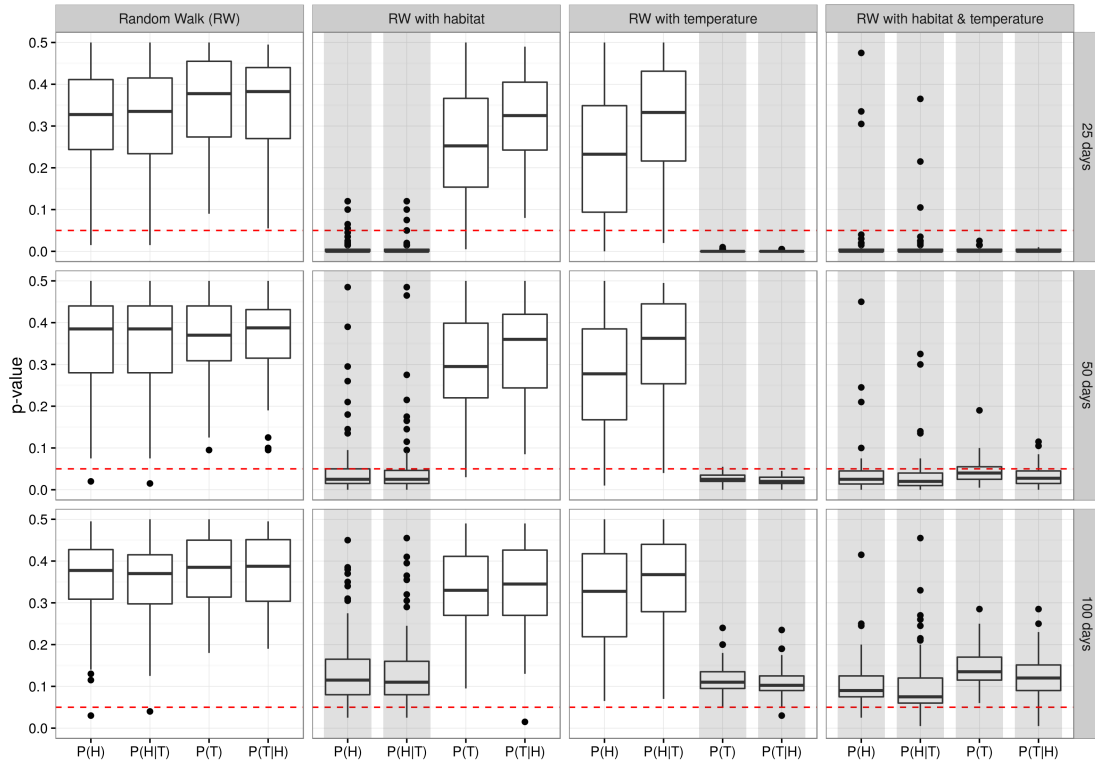
## 6.8 Appendix 1: Additional Figures and Tables

**Table 6.A1:** Error rates of different simulation scenarios and maximum time lags for the semivariogram. Type 1 error rate is defined as the portion of false positives, (test falsely detected a significant effect of habitat and temperature; individually or jointly), whereas Type 2 error rate is the proportion of false negatives, (test failed to detected a significant effect of habitat and temperature; individually or jointly). For the same data as a Figure, see Fig. 6.3 in the main text.

Simulation Scenario	Max. time lag (days)	Type 1 error rate	Type 2 error rate
Random Walk (RW)	50	0.0500	0.0000
Random Walk (RW)	100	0.0200	0.0000
Random Walk (RW)	150	0.0150	0.0000
Random Walk (RW)	200	0.0175	0.0000
RW with habitat	50	0.0100	0.2150
RW with habitat	100	0.0000	0.4300
RW with habitat	150	0.0000	0.6500
RW with habitat	200	0.0000	0.7800
RW with temperature	50	0.0950	0.0000
RW with temperature	100	0.0800	0.2550
RW with temperature	150	0.0650	0.8550
RW with temperature	200	0.0400	0.9650
RW with habitat & temperature	50	0.0000	0.0725

*Continued on next page*

6.8 Appendix 1: Additional Figures and Tables



**Figure 6.A1:** Power to detect an effect of environmental covariates on animal movement. Box plots of p-values from different hypothesis tests (Table 6.1) for different simulation scenarios (columns) and different values of the maximum lag in the semivariogram (rows). Results below the red dashed line are considered significant ( $\alpha = 0.05$ ). Tests for which we expected significant results are highlighted with gray vertical bars. Black dots indicate outliers. For the same data in numerical format, see Table 6.A2.

Table 6.A1 continued from previous page

Simulation Scenario	Max. time lag (days)	Type 1 error rate	Type 2 error rate
RW with habitat & temperature	100	0.0000	0.3775
RW with habitat & temperature	150	0.0000	0.6800
RW with habitat & temperature	200	0.0000	0.7750

**Table 6.A2:** Error rates of different simulation scenarios and maximum time lags for the semivariogram. Type 1 error rate is defined as the portion of false positives, (test falsely detected a significant effect of habitat and temperature; individually or jointly), whereas Type II error rate is the proportion of false negatives, (test failed to detected a significant effect of habitat and temperature; individually or jointly). For the same data as a Figure, see Fig. 6.A1 in the main text.

Simulation Scenario	Max. time lag (days)	Type 1 error rate	Type 2 error rate
Random Walk (RW)	50	0.00	0.0000
Random Walk (RW)	100	0.00	0.0000
Random Walk (RW)	150	0.00	0.0000
Random Walk (RW)	200	0.00	0.0000
RW with habitat	50	0.00	0.3400
RW with habitat	100	0.00	0.9500
RW with habitat	150	0.00	0.9950
RW with habitat	200	0.00	0.9950
RW with temperature	50	0.02	0.0150
RW with temperature	100	0.00	0.9950
RW with temperature	150	0.00	1.0000
RW with temperature	200	0.00	1.0000
RW with habitat & temperature	50	0.00	0.2950
RW with habitat & temperature	100	0.00	0.9300
RW with habitat & temperature	150	0.00	0.9875
RW with habitat & temperature	200	0.00	0.9975

**Table 6.A3:** Results based on applying the statistical framework on movement data of a red deer with different amounts of missing data. For each test the fraction of significant results is given with 10, 20, 50, and 80% of missing data (columns 2 through 5). Missing data were generated by randomly removing data points. The last column 'True p-value' provides the p-value from the full data.

Test	10%	20%	50%	80%	True p-value
P(A)	0.84	0.56	0.16	0.20	0.015
P(A+A <sup>2</sup> )	1.00	0.96	0.24	0.24	0.000
P(A+A <sup>2</sup>  R)	1.00	0.96	0.28	0.24	0.000
P(A+A <sup>2</sup>  R+T+T <sup>2</sup> )	1.00	1.00	0.44	0.28	0.000
P(A+A <sup>2</sup>  T+T <sup>2</sup> )	1.00	1.00	0.44	0.28	0.000
P(A R)	0.84	0.60	0.12	0.20	0.005

*Continued on next page*

## 6.8 Appendix 1: Additional Figures and Tables

Table 6.A3 continued from previous page

Test	10%	20%	50%	80%	True p-value
$P(A R+T+T^2)$	1.00	0.92	0.24	0.28	0.000
$P(A T+T^2)$	1.00	0.96	0.24	0.20	0.000
$P(R)$	0.00	0.00	0.04	0.04	0.315
$P(R A+A^2)$	0.00	0.00	0.00	0.04	0.240
$P(R A+A^2+T+T^2)$	0.00	0.00	0.00	0.04	0.345
$P(R T+T^2)$	0.00	0.00	0.08	0.12	0.490
$P(T)$	0.36	0.28	0.12	0.04	0.035
$P(T A+A^2)$	1.00	1.00	0.92	0.60	0.000
$P(T A+A^2+R)$	1.00	1.00	0.92	0.56	0.000
$P(T R)$	0.32	0.16	0.24	0.12	0.065
$P(T+T^2)$	1.00	1.00	1.00	0.80	0.000
$P(T+T^2 A+A^2)$	1.00	1.00	1.00	1.00	0.000
$P(T+T^2 A+A^2+R)$	0.96	0.80	0.32	0.04	0.000
$P(T+T^2 R)$	1.00	1.00	1.00	0.80	0.000





# 7

## Conclusions

In the final part of this thesis, I attempt to bring the individual chapters back to the bigger picture by evaluating technical and analytical aspects of the analysis of animal movement data. Finally, I close with an outlook on what may lie ahead for us in the field of animal biotelemetry.

### 7.1 Findings of this thesis

#### Technical considerations

A flexible data model is indispensable for managing and handling a large amount of movement data. In chapter 2, I introduced a data model for animal telemetry data that is implemented and used in the `rhr` package (chapter 3) for Program R (R Core Team 2015). The package aims to offer users an easy way to calculate home ranges as well as archiving data and parameter settings for later work and reproducibility. Further development of the package is currently concentrated on implementing a new graphical user interface that automatically generates `rmarkdown` (Allaire et al. 2015) code.

It is unlikely that standards for the analysis of home ranges (or movement data in general) will be feasible (as argued by Gula and Theuerkauf 2013), but I hope to provide a small step in this direction.

The wide availability of cloud computing and recent developments in web application frameworks within R (Cusumano 2010; Chang et al. 2015) open new opportunities for offering home range analysis as a service. The `rhr` package (discussed in chapter 3) is designed so that it can easily be deployed within a software as a service architecture.

#### Analytical considerations

Researchers should start analyses with clear and biologically meaningful research questions (Fieberg and Börger 2012) and use methods, such as home range estimation, as means for answers. Often, there is no single best method, especially if the theory is incomplete. For example, in chapter 4, I attempted comparing different methods of calculating the core area of an animal's home range. This turned out to be difficult because

a precise definition of what is considered to be a core area is not available. Similarly, the question of what is the best spatial representation of a home range is difficult to answer. Burt's (1943) commonly used definition of home ranges, as well as its related discussions (e.g., Powell 2000; Powell and Mitchell 2012), provide good theoretical reasoning of what an animal's home range might be, but they do not provide help in selecting a method that accurately delineates a home range in space. This leads to the question of what the spatial representation of an animal's home range is, and how it might be estimated. The simulation study in chapter 5 supports the suggestion that it may be advantageous to think of home ranges as a relative quantity of space use, rather than an absolute delimitable quantity in space. If home range is considered an index of space use, the value of interest becomes the relative change of this index, not the absolute value. The simulation study in chapter 5 illustrated that while absolute differences are substantial between estimators, relative differences are negligible. This suggests that more research should be done on how home ranges can be used as a relative index of space use to answer meaning full biological questions, rather than on finding the estimator that best delineates a home ranges in space.

The translation of methods from other disciplines to the context of the analysis of movement data allows researchers to analyze new questions with new tools, and to ask new questions. For example, Worton (1989) introduced kernel density estimation as a new method to estimate home ranges. More recently, Fleming et al. (2014) translated the concept of semivariograms from the geostatistical literature to movement ecology. By using stochastic processes, Fleming et al. (2014) were able to estimate biologically meaningful parameters (e.g., scales at which home ranging occurs). In chapter 6, we propose an extension to the semi-variance approach that allows researchers to test for the influence of environmental covariates on the movement rate of animals. With this extension, I add a new method to the tool box for testing the influence of environmental covariates on animal movement. The framework proposed in chapter 6 is able to handle irregularly sampled data and does not require the formulation of an explicit movement model.

### **Conceptual considerations**

This thesis touched only briefly on conceptual considerations of movement ecology. However, the movement ecology paradigm (Nathan et al. 2008) served as fundamental theoretical justification for much of the work presented here. This highlights the importance of the presence of conceptual frameworks, such as a the movement ecology paradigm, to advance analytical methods.

## **7.2 Other aspects of animal biotelemetry**

The methods used for analyzing home ranges are only a very small subset of the analytical tools available for the analyses of movement data. Home ranges are often used as a basic building block for studies that focus on space use of animal, habitat selection (Johnson

1980), or interactions between animals. However, many other analytical methods exist to address different biological questions, using data from animal movement data. For example, tools exist to infer different behavioral modes of animals (e.g., Gurarie et al. 2015), quantifying dynamic interactions between animals (Delgado et al. 2014; Langrock et al. 2014), or the effect on the whole ecosystem (Jeltsch et al. 2013) from movement data.

This thesis focuses on methods that are most commonly applied to large terrestrial mammals. However, comprehensive bodies of knowledge also exist for other systems (e.g., marine biotelemetry) or taxonomic groups (e.g., insects). In other systems, different methods or study designs are used. For example, aquatic studies often make use of sonar telemetry and the movement of other taxonomic groups like, for example, insects have been tracked using cameras (Dell et al. 2014).

## 7.3 Outlook

Animal biotelemetry has proven to be a useful method for many applied and theoretical questions. The emergence of two subject specific journals (Movement Ecology and Animal Biotelemetry) within the last 5 years highlight this (Nathan and Giuggioli 2013; Klimley 2013).

Technology will develop further and provide new tracking devices. The new devices will be smaller and lighter and have larger memory capacities (Wikelski et al. 2007; Bridge et al. 2011). This opens exciting new avenues of research that allow researchers to better integrate the temporal component of the movement into many analyses (e.g., Lyons et al. 2013; Long and Nelson 2012). The development of animal biotelemetry devices that can track other parameters than the position including, for example, the heart rate (Wilson et al. 2015) can help to understand other components of the movement ecology paradigm (e.g., the animal's internal state). A better understanding of such underlying (physiological) processes can help to parameterize mechanistic models (Jachowski and Singh 2015) in which the observed movement path is an emerging property.

It seems likely that the current rapid development of new, and often complex, methods will continue. However, a strong focus on methods can jeopardize the view on the wider biological picture because researchers may lose themselves in methodological comparisons. I hope that the results of this thesis will help researchers save time in handling and managing movement data (chapter 2 and 3), improve ease in selecting among methods (chapter 4 and 5), and making full use of available movement and environmental data (chapter 6). Asking good and interesting questions will ultimately help to improve our understanding of animal movement better, allow researchers to give scientifically based management recommendation to policy makers, and promote integration of movement ecology into management plans for the successful conservation of animals (Barton et al. 2015).

## 7.4 References

- Allaire, J., J. Cheng, Y. Xie, J. McPherson, W. Chang, J. Allen, H. Wickham, A. Atkins, and R. Hyndman (2015). *rmarkdown: Dynamic Documents for R*. R package version 0.8.1 (cit. on p. 99).
- Barton, P. S., P. E. Lentini, E. Alacs, S. Bau, Y. M. Buckley, E. L. Burns, D. A. Driscoll, L. K. Guja, H. Kujala, J. J. Lahoz-Monfort, et al. (2015). “Guidelines for Using Movement Science to Inform Biodiversity Policy”. In: *Environmental management* 56.4, pp. 791–801 (cit. on p. 101).
- Bridge, E. S., K. Thorup, M. S. Bowlin, P. B. Chilson, R. H. Diehl, R. W. Fléron, P. Hartl, R. Kays, J. F. Kelly, W. D. Robinson, et al. (2011). “Technology on the move: recent and forthcoming innovations for tracking migratory birds”. In: *BioScience* 61.9, pp. 689–698 (cit. on p. 101).
- Burt, W. H. (1943). “Territoriality and home range concepts as applied to mammals”. In: *Journal of Mammalogy* 24.3, pp. 346–352 (cit. on p. 100).
- Chang, W., J. Cheng, J. Allaire, Y. Xie, and J. McPherson (2015). *shiny: Web Application Framework for R*. R package version 0.12.2 (cit. on p. 99).
- Cusumano, M. (2010). “Cloud computing and SaaS as new computing platforms”. In: *Communications of the ACM* 53.4, pp. 27–29 (cit. on p. 99).
- Delgado, M. d. M., V. Penteriani, J. M. Morales, E. Gurarie, and O. Ovaskainen (2014). “A statistical framework for inferring the influence of conspecifics on movement behaviour”. In: *Methods in Ecology and Evolution* 5.2, pp. 183–189 (cit. on p. 101).
- Dell, A. I., J. A. Bender, K. Branson, I. D. Couzin, G. G. de Polavieja, L. P. Noldus, A. Pérez-Escudero, P. Perona, A. D. Straw, M. Wikelski, et al. (2014). “Automated image-based tracking and its application in ecology”. In: *Trends in ecology & evolution* 29.7, pp. 417–428 (cit. on p. 101).
- Fieberg, J. and L. Börger (2012). “Could you please phrase home range as a question?” In: *Journal of Mammalogy* 93.4, pp. 890–902 (cit. on p. 99).
- Fleming, C. H., J. M. Calabrese, T. Mueller, K. A. Olson, P. Leimgruber, and W. F. Fagan (2014). “From fine-scale foraging to home ranges: A semi-variance approach to identifying movement modes across spatiotemporal scales”. In: *The American Naturalist* 183.5, E154–E167 (cit. on p. 100).
- Gula, R. and J. Theuerkauf (2013). “The need for standardization in wildlife science: home range estimators as an example”. In: *European Journal of Wildlife Research* 59.5, pp. 713–718 (cit. on p. 99).
- Gurarie, E., C. Bracis, M. Delgado, T. D. Meckley, I. Kojola, and C. M. Wagner (2015). “What is the animal doing? Tools for exploring behavioral structure in animal movements”. In: *Journal of Animal Ecology*, in press (cit. on p. 101).
- Jachowski, D. S. and N. J. Singh (2015). “Toward a mechanistic understanding of animal migration: incorporating physiological measurements in the study of animal movement”. In: *Conservation Physiology* 3.1, cov035 (cit. on p. 101).
- Jeltsch, F., D. Bonte, G. Pe’er, B. Reineking, P. Leimgruber, N. Balkenhol, B. Schröder, C. M. Buchmann, T. Mueller, N. Blaum, et al. (2013). “Integrating movement ecology with biodiversity research—exploring new avenues to address spatiotemporal biodiversity dynamics”. In: *Movement Ecology* 1.6 (cit. on p. 101).
- Johnson, D. H. (1980). “The comparison of usage and availability measurements for evaluating resource preference”. In: *Ecology* 61.1, pp. 65–71 (cit. on p. 100).
- Klimley, A. P. (2013). “Why publish Animal Biotelemetry?” In: *Animal Biotelemetry* 1.1, p. 1 (cit. on p. 101).
- Langrock, R., J. G. C. Hopcraft, P. G. Blackwell, V. Goodall, R. King, M. Niu, T. A. Patterson, M. W. Pedersen, A. Skarin, and R. S. Schick (2014). “Modelling group dynamic animal movement”. In: *Methods in Ecology and Evolution* 5.2, pp. 190–199 (cit. on p. 101).
- Long, J. A. and T. A. Nelson (2012). “Time geography and wildlife home range delineation”. In: *The Journal of Wildlife Management* 76.2, pp. 407–413 (cit. on p. 101).
- Lyons, A. J., W. C. Turner, and W. M. Getz (2013). “Home range plus: a space-time characterization of movement over real landscapes”. In: *Movement Ecology* 1.1 (cit. on p. 101).
- Nathan, R., W. M. Getz, E. Revilla, M. Holyoak, R. Kadmon, D. Saltz, and P. E. Smouse (2008). “A movement ecology paradigm for unifying organismal movement research”. In: *Proceedings of the National Academy of Sciences* 105.49, pp. 19052–19059 (cit. on p. 100).
- Nathan, R. and L. Giuggioli (2013). “A milestone for movement ecology research”. In: *Move Ecol* 1.1 (cit. on p. 101).
- Powell, R. A. (2000). “Animal home ranges and territories and home range estimators”. In: *Re-*

- search techniques in animal ecology: controversies and consequences* 442 (cit. on p. 100).
- Powell, R. A. and M. S. Mitchell (2012). “What is a home range?” In: *Journal of Mammalogy* 93.4, pp. 948–958 (cit. on p. 100).
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria (cit. on p. 99).
- Wikelski, M., R. W. Kays, N. J. Kasdin, K. Thorup, J. A. Smith, and G. W. Swenson (2007). “Going wild: what a global small-animal tracking system could do for experimental biologists”. In: *Journal of Experimental Biology* 210.2, pp. 181–186 (cit. on p. 101).
- Wilson, A., M. Wikelski, R. Wilson, and S. Cooke (2015). “Utility of biological sensor tags in animal conservation”. In: *Conservation Biology* 29.4, pp. 1065–1075 (cit. on p. 101).
- Worton, B. J. (1989). “Kernel methods for estimating the utilization distribution in home-range studies”. In: *Ecology* 70.1, pp. 164–168 (cit. on p. 100).



# 8

## Curriculum vitae

## Education

<b>Georg-August-Universität Göttingen</b> <i>PhD Forest Science and Forest Ecology</i>	Göttingen, Germany 2012-2016
<b>Fachhochschule Wiener Neustadt</b> <i>M.Sc. Geoinformatics</i>	Wiener Neustadt, Austria 2008 - 2010
<b>University of East Anglia</b> <i>B.Sc. Ecology</i>	Norwich, England 2004 - 2007

## Work Experience

<b>Georg-August-Universität Göttingen</b> <i>Research Assistant</i>	Göttingen, Germany since 2012
<b>International Center for Tropical Agriculture (CIAT)</b> <i>Visiting Researcher</i>	Cali, Colombia 2010 - 2011
<b>ÖJAB Haus Greifenstein</b> <i>Civilian Service</i>	Höflein, Austria 2008

## Publications

### Submitted

- Edelhoff, H., **Signer, J.** & Balkenhol, N. Path segmentation for beginners: An overview of current methods for detecting changes in animal movement patterns *Movement Ecology*
- Koch, F., **Signer, J.**, Kapeller, P. & Fichtel, C. My home is my castle? The role of the residence effect on the outcome of intergroup encounters in Verreaux's sifakas *Scientific Reports*
- **Signer, J.**, & Ovaskainen, O. Detecting the influence of environmental covariates on animal movement: a semivariance approach *Methods in Ecology and Evolution*

### Accepted

- Koch, F., **Signer, J.**, Kapeller, P. & Fichtel, C. Intergroup encounters in Verreaux's sifakas (*Propithecus verreauxi*): who fights and why? *Behavioral Ecology and Sociobiology*

### Published

- Hofman, M., **Signer, J.**, Hayward, M. & Balkenhol, N. (2016) Spatial ecology of a herd of white-lipped peccaries (*Tayassu pecari*) in Belize using GPS telemetry: challenges and preliminary results *Therya*
- **Signer, J.**, Balkenhol, N., Ditmer, M. & Fieberg, J. (2015) Does estimator choice influence our ability to detect changes in home-range size? *Animal Biotelemetry*. DOI: 10.1186/s40317-015-0051-x.



- **Signer, J.** & Balkenhol, N. (2015). Reproducible home ranges (rhr): A new, user-friendly R package for analyses of wildlife telemetry data. *Wildlife Society Bulletin*.. DOI: 10.1002/wsb.539.
- Leibing, C., **Signer, J.**, Van Zonneveld, M., Jarvis, A. & Dvorak, W. (2013). Selection of provenances to adapt tropical pine forestry to climate change on the basis of climate analogs. *Forests*. DOI: 10.3390/f4010155.

#### Accepted Conference Proceedings

- **Signer, J.** & Balkenhol, N. Home range core areas: use and estimation *Tagungsband der Vereinigung der Wildbiologen und Jagdwissenschaftler*.

#### Conference Proceedings

- **Signer, J.** & Balkenhol, N. (2015). Error detection and handling in GPS data *Beiträge zur Jagd- & Wildforschung* 40.

*Chapter 8 Curriculum vitae*