Combining landscape genetics and movement ecology to assess functional connectivity for red deer (*Cervus elaphus*) in Schleswig-Holstein, Germany

Dissertation

to attain the doctoral degree (Dr. forest.)
of the Faculty of Forest Sciences and Forest Ecology
Georg-August-Universität Göttingen

Submitted by Hendrik Alfred Edelhoff
born on the 12th of August 1986 in Arnsberg

Göttingen, December 2019

1. Referee: Prof. Dr. Niko Balkenhol

2. Referee: Dr. Christian Roos

Co-referee: Prof. Dr. Kerstin Wiegand

Date of oral examination: 8th of July, 2020

Sponsored by the Scholarship Programme of the German Federal Environmental Foundation

DBU

Deutsche Bundesstiftung Umwelt

www.dbu.de

Table of Contents

Summary

Human-caused restrictions like the fragmentation of the landscape poses a major challenge to wildlife conservation. Large and mobile species such as red deer (*Cervus elaphus*) are subject to increasing effects of isolation and a decrease of primary habitats. This can result in a reduction of the exchange of individuals or even a long-term loss of gene flow. In order to counteract these negative effects and to promote genetic exchange, suitable approaches for estimating functional connectivity of the landscape are necessary.

In most cases, landscape models of functional connectivity for a given study species are based on expert knowledge, habitat suitability, or movement data. However, there is an ongoing debate whether these methods are representative of actual dispersal or effective gene flow. Landscape genetic analyses correlate estimates of genetic differentiation between populations or individuals with landscape composition. The advantage of genetic data is that it reflects both successful dispersal between populations, as well as subsequent reproduction with other individuals. Therefore, landscape genetics represent an innovative approach for assessing functional connectivity of the landscape matrix.

The aim of this dissertation is to compare different species-specific models of functional connectivity utilizing genetic and movement data. Using red deer in Northern Germany as an example, the methodological and conceptual differences of multiple approaches are demonstrated. Overall, the presented thesis provides important insights for applied conservation of wildlife and planning of corridors.

The **first chapter** provides a general introduction to the issue of landscape fragmentation and illustrates the effects on red deer in the study area of Schleswig-Holstein. Furthermore, the potential applications of landscape genetics and movement ecology to assess landscape connectivity are presented. For example, movement ecology provides an integral framework to explore the potential factors shaping the movements of organisms and the ecological consequences of these movements such as gene flow.

The **second chapter** comprises a study on the genetic diversity and structure of red deer populations in Northern Germany. The results indicate that local populations are best described as an hierarchical network of subpopulations with different levels of gene flow. Overall, genetic diversity of red deer from the study area is quite low compared to other populations from Central Europe. This underlines that a better understanding of the isolation effects caused by landscape fragmentation and species-specific assessment of landscape connectivity for red deer are needed to address the observed loss of genetic diversity.

One possible approach for estimating functional connectivity is by linking telemetry data with landscape variables in order to gain insights into the habitat requirements of a target species. However, habitat preferences are very likely to change with different movement behaviors. This represents an important point to consider when studying the effects of landscape composition on actual dispersal movements. The **third chapter** of this thesis presents an extensive overview on different methods for identifying behavioral patterns from movement data. Furthermore, it provides guidelines for deciding among the available methods of path-segmentation and shows how they can be applied to answer research questions within the movement ecology paradigm.

The study described in the **fourth chapter** utilizes such a path-segmentation method to detect potential dispersal movements from telemetry data of multiple red deer individuals. The observed movements are then linked to landscape variables in order to model functional connectivity based on landscape resistance towards dispersal of red deer throughout the study area. In addition, the study applies and compares different methodological approaches for modeling functional connectivity based on expert knowledge, habitat models and other analyses of movement data. A landscape genetic approach is used as a means to compare the resulting resistance models. Effective distances derived from the models are compared with estimates on genetic distance. The highest ranked models are further used to illustrate methodological differences in the designation of conservation corridors. The results show that for large scale dispersal red deer rely on primary habitat conditions within the landscape matrix. However, connectivity based on the identified dispersal movements showed that areas of poor habitat quality can be traversed by red deer at shorter distances.

Finally, in the **fifth chapter**, the results of the presented studies are summarized and discussed. In particular, the contribution of landscape genetics and movement ecology to applied conservation and landscape planning are elaborated. The results of this thesis could ultimately increase the effectiveness of conservation measures such as the placement of corridors.

Zusammenfassung

Die anthropogen bedingte Zerschneidung der Landschaft stellt eine wichtige Herausforderung für den Natur- und Artenschutz dar. Große Säugetiere, wie zum Beispiel der Rothirsch (*Cervus elaphus*) sind durch die Fragmentierung einer Verkleinerung und zunehmenden Isolierung der Lebensräume ausgesetzt. Dies kann weitreichende Folgen wie einen verringerten Austausch an Individuen und damit langfristig an Genen mit sich ziehen. Um diesen Folgen entgegenzuwirken und den genetischen Austausch zu verbessern sind objektive Beurteilungsverfahren über die Konnektivität der Landschaft notwendig.

Die Erfassung und Modellierung der funktionellen Landschaftskonnektivität für eine Zielart basiert häufig auf Grundlagen wie Expertenwissen, Habitatmodellen oder Bewegungsdaten. Allerdings werden diese Methoden hinsichtlich ihrer Repräsentativität für tatsächliche Abwanderungen oder effektivem Genfluss diskutiert. Im Rahmen von landschaftsgenetischen Analysen werden Informationen über den genetischen Austausch zwischen Populationen oder einzelnen Individuen mit entsprechenden Ausprägungen der Landschaft korreliert. Genetische Daten haben dabei den Vorteil, dass sie sowohl eine erfolgreiche Wanderung zwischen Verbreitungsgebieten als auch die anschließende Reproduktion mit anderen Individuen, widerspiegeln können. Daher stellt die Landschaftsgenetik eine innovative Ansatzmöglichkeit zur Beurteilung der funktionellen Landschaftskonnektivität dar.

Ziel der Dissertation ist die Konzipierung und Evaluierung von artspezifischen Modellen der Landschaftskonnektivität mit Hilfe von Gendaten und Telemetrie-Ergebnissen. Der Rothirsch in Schleswig-Holstein dient dabei als Beispielart, mit der die Unterschiede bezüglich der methodischen und konzeptionellen Herangehensweisen demonstriert werden sollen. Insbesondere für die naturschutzfachliche Praxis und Korridorplanung ist dies von grundlegender Bedeutung.

Im **ersten Kapitel** wird zunächst eine generelle Einleitung in die Problematik der Landschaftszerschneidung gegeben und anhand des Rothirschs in Schleswig-Holstein verdeutlicht. Anschließend werden die verschiedenen Ansatzmöglichkeiten der Landschaftsgenetik als auch der Bewegungsökologie zur Beurteilung der Landschaftskonnektivität dargestellt. Die Bewegungsökologie setzt sich unter anderem damit auseinander, welche Faktoren die Bewegungen von Organismen in ihrem Lebensraum beeinflussen. Durch die Verknüpfung von Bewegungsdaten mit Landschaftsvariablen lassen sich so wichtige Erkenntnisse über die Lebensraumansprüche einer Zielart gewinnen. Dabei können unter anderem die Habitatpräferenzen während unterschiedlicher Bewegungsmuster, wie zum Beispiel der Abwanderung in neue Gebiete, differenziert betrachtet werden.

Das **zweite Kapitel** befasst sich mit der genetischen Diversität und Differenzierung der lokalen Rothirschvorkommen in Schleswig-Holstein. Anhand der genetischen Daten wird dabei verdeutlicht, dass die regionalen Managementeinheiten (Hegeringe) nicht immer in sich geschlossene Populationen darstellen. Die Rothirschpopulationen weisen vielmehr eine hierarchische Struktur auf. Zum Beispiel ist der Genfluss, je nach Dichte der benachbarten Populationen, unterschiedlich stark ausgeprägt. Insgesamt konnte für mehrere Populationen eine im europäischen Vergleich geringe genetische Diversität festgestellt werden. Dies unterstreicht, dass ein besseres Verständnis über die Auswirkungen der Landschaftszerschneidung sowie eine Bewertung der Landschaftskonnektivität aus Sicht des Rothirschs notwendig ist, um dem Verlust an genetischer Vielfalt entgegenzuwirken.

Eine Möglichkeit die Landschaftskonnektivität zu bewerten stellt die Analyse von Telemetrie-Daten dar. Für die Auswertung von solchen Bewegungsdaten stehen eine Vielzahl an Methoden zur Verfügung. Im **dritten Kapitel** werden die verschiedenen Ansätze zur Differenzierung unterschiedlicher Bewegungsmuster aus Telemetrie-Daten zusammengestellt. Durch eine umfangreiche Methodenübersicht werden Entscheidungshilfen für die Anwendung solcher Pfad-Segmentierungen zur Beantwortung bestimmter Fragestellungen in der Bewegungsökologie gegeben.

Das **vierte Kapitel** greift unter anderem auf eine solche Methode der Pfad-Segmentierung zurück, um potentielle Ausbreitungsbewegungen innerhalb der Telemetrie-Daten von besenderten Rothirschen zu ermitteln. Diese Bewegungsdaten werden anschließend mit Landschaftsvariablen verknüpft und ein Modell abgeleitet, welches den Widerstand für Wanderbewegungen darstellt (Widerstandsmodell). Darüber hinaus werden in dieser Studie weitere methodische Ansätze zur Modellierung der funktionellen Landschaftskonnektivität verglichen. Diese basieren unter anderem auf Expertenwissen und Habitatmodellen sowie weiteren Auswertungsansätzen der Bewegungsdaten. Für den Vergleich der resultierenden Widerstandsmodelle wird die Landschaftsgenetik hinzugezogen. Dabei werden effektive Distanzen basierend auf den jeweiligen Modellen den genetischen Distanzmaßen gegenübergestellt. Die Modelle mit der höchsten Übereinstimmung werden ferner genutzt, um methodische Unterschiede in der Ausweisung von Korridoren darzustellen. Es zeigte sich, dass für weitreichende Abwanderungen die Rothirsche auf geeignete Habitatverhältnisse innerhalb der Landschaftsmatrix angewiesen sind. Die Auswertung der Bewegungsdaten ergab hingegen, dass für kürzere Distanzen auch suboptimale Gebiete durchquert werden können.

Abschließend werden im **fünften Kapitel** die Ergebnisse zusammengefasst und diskutiert. Besonderer Schwerpunkt liegt dabei auf dem Beitrag der Anwendung von Landschaftsgenetik und Bewegungsökologie im angewandten Naturschutz und welche Erkenntnisse für die Ausweisung und Effektivität von Korridoren gewonnen werden können.

CHAPTER 1

General Introduction

Fragmentation and Connectivity of the Landscape

Habitat loss along with changing climatic conditions are indisputably the leading drivers for biodiversity loss worldwide (Bellard et al. 2012; Devictor et al. 2012). Next to the net loss of suitable habitat, human development leads to an increased fragmentation of remaining habitat with various consequences for remaining populations residing in these more or less isolated habitat patches (Templeton et al. 1990; Keyghobadi 2007). Studying the genetic and demographic effects of fragmentation has become a central focus for nature conservation (Moilanen et al. 2005; Cushman et al. 2006; Epps et al. 2007).

While landscape fragmentation per se describes also natural processes of dividing, isolating and reducing of once continuous habitats (Fahrig 2003) the amount and speed at which these processes act under human action is not to be underestimated. The isolation of primary habitat and the restriction of dispersal and gene flow among those remaining habitats (Jaeger and Holderegger 2005; Balkenhol and Waits 2009) can cause the emergence of metapopulations (Opdam 1991; Hanski 1998; Hanski and Ovaskainen 2003) as well as long-term loss of genetic diversity and inbreeding (Andersen et al. 2004; Keyghobadi et al. 2005) and even the local extinction of affected populations (Merriam and Wegner 1992).

To mitigate these negative effects, knowledge on landscape connectivity for species of special concern is of major importance for an effective conservation management. Connectivity comprises to what degree the landscape matrix allows for the exchange of individuals among remaining habitat patches (Taylor et al. 1993; Crooks and Sanjayan 2006). This can either be due to spatial alignment of certain landscape features (i.e., structural connectivity; Kindlmann and Burel 2008) or because the landscape matrix (i.e., non-habitat) still facilitates movements of a given species (i.e., functional connectivity; Baguette and Van Dyck 2007). Understanding connectivity is hence vital for species in fragmented landscapes as maintaining dispersal movements and gene flow between habitats can counteract the negative consequences caused by fragmentation (Fahrig and Merriam 1994; Goodwin and Fahrig 2002; Kindlmann and Burel 2008). While one option for ensuring or reestablishing connectivity of the landscape is the delineation of conservation corridors (Chetkiewicz et al. 2006; Hilty et al. 2012) design and identification of the ideal locations for those corridors remains challenging.

Landscape models emerged as a central application in conservation and landscape planning to delineate areas of high connectivity between remaining habitats and to conserve their current composition (Tischendorf and Fahrig 2000; Rudnick et al. 2012). In addition, artificial linkages can be placed accordingly to compensate barrier effects originating from e.g. linear infrastructures (Epps et al. 2005; Balkenhol and Waits 2009). However, in order to increase the effectiveness and functionality of such rather structural mitigation measures, objective approaches for assessing functional connectivity for a target species need to be applied (Böttcher et al. 2004; Beier et al. 2008).

Assessing functional connectivity

A major challenge for researchers and practitioners remains the objective assessment of functional connectivity. Both, effects of fragmentation as well as the degree of connectivity, are highly species specific (Chetkiewicz et al. 2006; Beier et al. 2008) as landscape permeability is defined by an organisms movement capacity and perceptual range (Diniz et al. 2020) as well as habitat requirements during dispersal (Revilla and Wiegand 2008; Fattebert et al. 2015).

For all these reasons, there has been a plethora of methodological approaches proposed and applied in connectivity research and for corridor planning (Beier et al. 2008; Zeller et al. 2012). The majority of studies have relied either on expert opinion (Clevenger et al. 2002; Milanesi et al. 2016), on empirical data on species' space use such as habitat models based on occurrence data (Wang et al. 2008), or different types of resource-selection functions derived from movement data (reviewed in Zeller et al. 2012), as well as simulation models based on experimental data or a combination of different information sources (Vuilleumier and Metzger 2006; Aben et al. 2014). Depending on the focal species also experiments, such as translocations, can be applied (Volpe et al. 2014; Betts et al. 2015). Nevertheless, with all the different methods available there is an ongoing debate on which approaches are best suited, especially in terms of representing effective dispersal or actual gene flow (Spear et al. 2010; Richardson et al. 2016).

Models derived from expert opinion utilize previous studies and reviews of literature or from interviews on the expertise of local managers and experts (Jacobs et al. 2014; Reed et al. 2016). The main advantage of these approaches is their easy development, making them less time consuming and cost efficient as they do not depend on long term research and data acquisition (Murray et al. 2009; Milanesi et al. 2016). However, expert-opinion based models have been criticized for being subjective and hardly reproducible since assumptions or opinions are difficult to quantify (Epps et al. 2007; Beier et al. 2008). Furthermore, results from other studies cannot always simply be transferred and extrapolated to completely different systems or landscapes (Oyler-McCance et al. 2013; Richardson et al. 2016).

Next to expert opinion, there are various approaches depending on empirical data to derive functional connectivity (Spear et al. 2010; Zeller et al. 2012). For example, habitat models (also referred to as species distribution models; Franklin 2009; Guisan et al. 2013) describing the habitat requirements for a given species became a popular tool for estimating functional connectivity (Engler et al. 2014; Milanesi et al. 2016). There is a multitude of statistical models available to researches and practitioners for correlating habitat covariates with (often opportunistic) presence data (e.g., MAXENT, BIOMOD; Franklin 2009). However, a key assumption of the application of these models is that the target species requires the same habitat features during dispersal movements as for selecting resources and establishing a home range or maintaining a population in primary habitats (Spear et al. 2015, Zeller et al. 2012). Depending on functional grain and the species' niche, this can hold true for some species and certain study questions (Engler et al. 2014; Razgour 2015) but not for others (Wasserman et al. 2010; Mateo-Sanchez et al. 2015).

Information on actual movements of the study species (i.e., based on telemetry relocations) has been utilized extensively in the past decade for modeling landscape connectivity (Spear et al. 2010; Centeno-Cuadros et al. 2017). Most notably the establishment of the movement ecology paradigm by (Nathan et al. 2008) provided a unified framework for studying the causes and consequences of individual movements on different aspects of ecology. For example, one fundamental aspect of movement ecology is the influence of individual movements on effective dispersal and gene flow (Holyoak et al. 2008). Therefore, the movement ecology paradigm provides a means to model functional connectivity which can be accomplished at multiple scales (review in Zeller et al. 2012). First of all, resource selection functions can be estimated from relocations sampled via GPS telemetry (Manly et al. 1993; Boyce et al. 2002). This should provide comparable estimates on space use as habitat models which in return can be used to infer connectivity (Chetkiewicz et al. 2006; Shafer et al. 2012; Squires et al. 2013). Other models account for the actual composition of the observed movements and model space use at the step-level (Thurfjell et al. 2014) or path-level (Reding et al. 2013; Zeller et al. 2015). Finally, movement patterns and the underlying behaviors can be differentiated (Edelhoff et al. 2016, Chapter 3). As illustrated in Figure 1.1, habitat requirements of species during dispersal movements could differ from movements in primary habitat or within the established home range (Roever et al. 2013; Abrahms et al. 2017). However, movement data also holds some analytical restrictions and drawbacks that need to be accounted for to make full use of this data source. First, sample sizes are often restricted to a limited number of individuals. Second, such data is rarely derived from individuals doing actual dispersal since the main dispersal happens in offspring before their first (attempted) reproduction (i.e. natal dispersal, Colbert et al. 2001; Whitmee and Orme 2013). Hence most information from tracking data most likely quantifies habitat use but not dispersal movements (Spear et al. 2015; Centeno-Cuadros et al. 2017; Zeller et al. 2017).

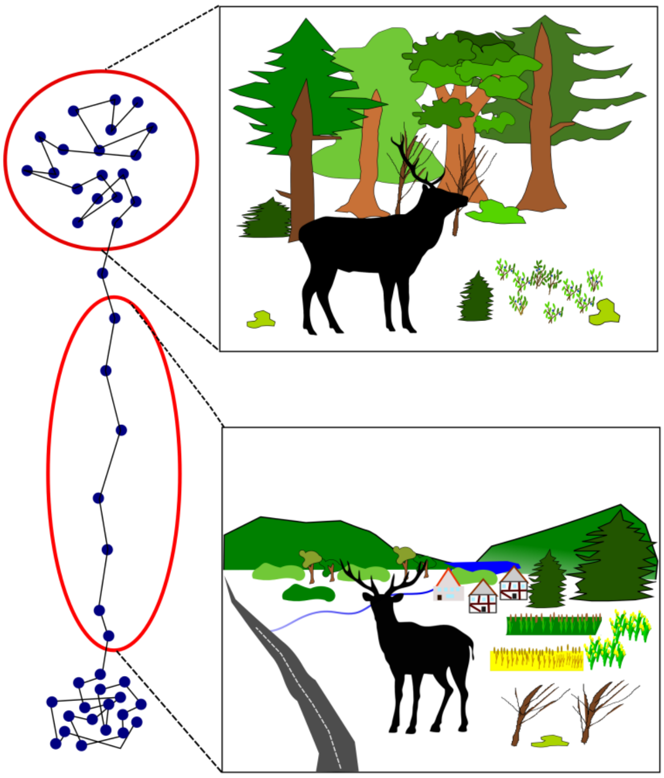


Figure 1.1 Illustration of different movement patterns derived from two separate behaviors: one being the regular movements in established home ranges or primary habitats characterized by many circular steps within a restricted area (top). The second one is a dispersal movement (middle) leaving the home range. The trajectory consists of long steps without any turns in direction. After the dispersal phase the individual returned to the normal movement behavior with area restricted movement patterns (bottom). Habitat requirements could substantially change during these two phases and deriving estimates on connectivity from all or only the regular movements could very likely be biased.

Information on gene flow among populations inhabiting different parts of a landscape can also be used to infer functional connectivity. Genetic data based on highly variable markers, such as microsatellites or single nuclear polymorphisms (SNPs; Sunnucks 2000), allow the estimation of effective dispersal through gene flow. Genetic data has the advantage over occurrence or movement information as it not only indicates the successful dispersal of individuals among populations but also their successful reproduction at these locations (Coulon et al. 2004). The data can be used to gain a better understanding of how well populations are connected effectively; therefore, one can derive estimates on functional connectivity. However using genetic data alone also has its restrictions as assessments of gene flow depend on theoretical assumptions on equilibrium (e.g., Wrights island model for all F statistics; Wright 1965) which are rarely met in empirical systems e.g. non-random mating, uneven effective population sizes, recent demographic changes and drift as well as population bottlenecks (Jost et al. 2018). Therefore, combining genetic data with multiple information from other data sources within a holistic analytical framework, could offer the most promising approach to assess functional landscape connectivity. The young field of landscape genetics aims for this multidisciplinary approach.

Landscape Genetics

The framework of combining genetic data with information on landscape composition was established by Manel et al. (2003) and has since then grown rapidly in application (Holderegger and Wagner 2006; Storfer et al. 2010; Manel and Holderegger 2013). Basically, landscape genetics comprise of two main components that are correlated with each other: a spatial and a genetic component. The latter quantifies genetic differentiation or gene flow between considered entities (populations or individuals in a given study area). The spatial component estimates the potential influence of landscape features as an effective distance often referred to as resistance or permeability (Storfer et al. 2007; Balkenhol et al. 2009). By this a major research focus of applying landscape genetic methods is to quantify the effects of fragmentation and to estimate functional connectivity and corridors for conservation management (Cushman et al. 2006; Epps et al. 2007; Kool et al. 2013). For example, this quantification includes the identification of specific barriers to dispersal (Frantz et al. 2010), the effects of both historical and anthropogenic landscape changes (Epps and Keyghobadi 2015), as well as the potential spread of diseases or invasive species (Storfer et al. 2010).

A vast number of evaluation methods have been developed and used for landscape genetic analyzes, which can be divided into three analytical steps (Balkenhol et al. 2009; Storfer et al. 2010, Spear et al. 2015) quantifying: 1) spatial (effective) distances, 2) genetic distances and 3) correlating both (Figure 1.3).

Spatial and effective distances

To assess spatial distances in a landscape genetic context three different theories are commonly used: isolation by distance (IBD), isolation by barrier (IBB), and isolation by resistance (IBR). The theory of isolation by distance (IBD; Wright 1943) hypothesizes that genetic distance is positively correlated with geographic distance among considered entities (Figure 1.2.a). For example, a positive correlation with IBD would indicate that effective dispersal (leading to gene flow) occurs only at shorter distances and therefore counteracts genetic drift at the regional scale (Hutchison and Templeton 1999). Here, a homogenous landscape without any restrictions is assumed and distances are measured as Euclidean distances among entities (Balkenhol et al. 2009; Spear et al. 2010; Ruiz-Gonzalez et al. 2015).

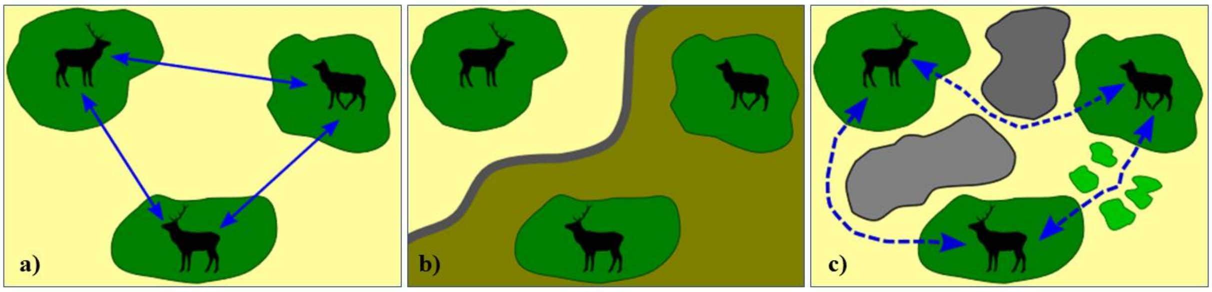


Figure 1.2 Schematic description of the three different models of spatial distances commonly applied in landscape genetics: a) isolation by distance (IBD) tests for correlation of gene flow with geographic / Euclidean distance between entities like red deer populations; b) isolation by barrier (IBB) accounts for the potential effects of barriers and distance is derived from the presence or lack thereof between two entities; c) isolation by resistance (IBR) models the effects of the landscape connectivity by calculating the effective distance between sites based on costs of movements through the landscape matrix.

The second theory, IBB, assumes effects of potential barriers or boundaries (Figure 1.2.b) which restrict gene flow and therefore increase genetic differentiation between entities located at separate sides of the putative barrier (Epps et al. 2005; Balkenhol et al. 2009). For example, two entities from the same side of the barrier are assigned a minimum effective distance value, whereas entities from two separate sides exhibit maximum effective distance.

The two former theories do not account for any restrictions to movement or gene flow derived from the composition of the landscape matrix (McRae 2006; van Strien et al. 2015). For this manner, the third theory explicitly refers to isolation by resistance (IBR; Figure 1.2.c) and correlates genetic distances with effective distances based on landscape heterogeneity (Ricketts 2001; Kindlmann and Burel 2008). The degree to which landscape features either impede or promote (effective) dispersal are summarized in so called resistance surfaces (Spear et al. 2010). Therefore, the approaches for estimating connectivity described above are transferable to the assessment of landscape resistance. Instead of permeability, resistance models assign values to different landscape features describing their "cost of movement" (Koen et al. 2012; Zeller et al. 2012; Spear et al. 2015).

For example, empirical data can be turned into resistance surfaces by taking the inverse of habitat suitability estimates or resource selection probabilities derived from movement data (Keeley et al. 2016). However, not always an inverse linear relationship holds true and different transformation functions between estimates of space use and resistance should be compared (see for example Zeller et al. 2018). Expert-opinion or hypotheses on landscape features potentially impacting resistance can be tested in a causal modeling framework (Cushman et al. 2006; Shirk et al. 2010). By this, single or multivariate landscape data are transformed into resistance using various functions and are iteratively compared to genetic distances (Wasserman et al. 2010; Cushman et al. 2013).

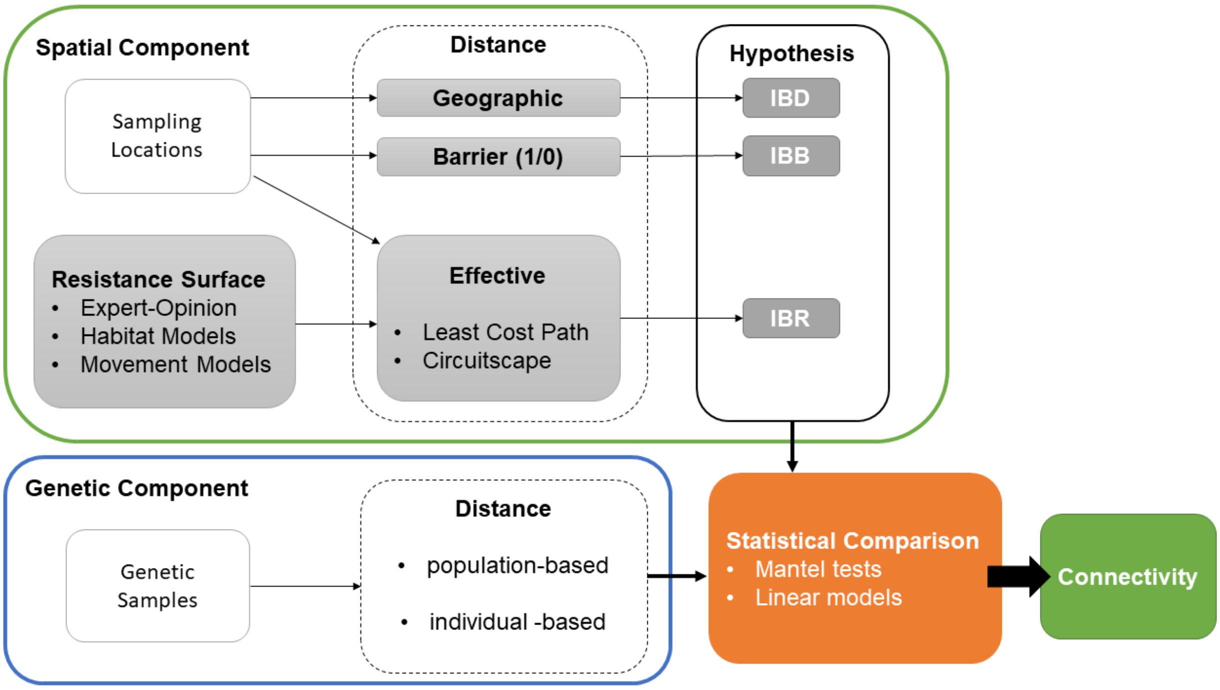


Figure 1.3 Flowchart of the major components of landscape genetic analyses: 1) spatial component: geographic distances between sampling locations are calculated to model isolation by distance (IBD), presence of barriers determines isolation by barrier (IBB) and resistance surfaces are used to derive effective distances to test for isolation by resistance (IBR). 2) Genetic component: either population- or individual-based distances are derived from genetic samples. 3) distances are correlated to each other to infer functional connectivity.

Effective distances based on the resistance models can be derived using least-cost (Adriaensen et al. 2003) or least-resistance (McRae 2006) algorithms (Figure 1.3). The first one calculates the distance based on a least-cost path (LCP) which is a single vector of minimal cumulative resistance (cost) between two locations (Graves et al. 2014). Although LCPs have been shown to provide effective estimates of dispersal (Driezen et al. 2007; Zeller et al. 2018) and gene flow (Stevens et al. 2006; Wang et al. 2008) their key assumption – that there is only a single best route known by the organism a priori – limits its biological significance and interpretation (McRae and Beier 2007). To overcome this caveat, McRae (2006) proposed to model effective distances based on circuit-theory which accounts for multiple randomized movement paths between two locations (McRae et al. 2008). This concept has been shown to outperform LCPs in some landscape-genetic studies, in particular for wide-ranging species (McRae and Beier 2007; Spear et al. 2015).

Genetic Distances

Most importantly, for landscape genetic analyses samples from multiple individuals and their spatial reference (location) are needed. In general, genetic data can be used to delineate genetic clusters or sub-populations (assignment tests like STRUCTURE; Wang 2017), for estimating genetic diversity and to infer gene flow among populations (Waits and Storfer 2015). For the majority of landscape genetic applications the determination of genetic distances between individuals or populations (Figure 1.3) is most relevant. Various parameters have been developed for this over time (Storfer et al. 2010; Shirk et al. 2017). Depending on the study or sampling design either population-based estimates of genetic differentiation e.g. F_{ST} (Wright 1949) and its several derivates such as Nei´s D (Nei 1972), or individual-based distances like Roussets´s a (Rousset 2000) or the proportion of shared alleles are being used (Bowcock et al. 1994).

Statistical Model Comparison

In the final step of a landscape genetic analysis the genetic distances are modeled or correlated with the spatial and effective distances (Figure 1.3). The three models of isolation (IBD, IBB, IBR) do not necessarily have to be mutually exclusive but can also be combined, e.g., by partialing out the effect of IBD when modeling IBR (Balkenhol et al. 2009). Again, there are multiple statistical approaches for accomplishing this (Storfer et al. 2007; Balkenhol et al. 2009; Shirk et al. 2018). Because of the pairwise comparison between the sampled entities the applied statistical tests are mostly based on correlations of distance matrices such as the Mantel tests (Mantel 1967; Perez et al. 2010). Linear models like multiple regressions on distance matrices (Legendre and Fortin 2010; Wang 2013) or mixed models accounting for the pairwise data structure (van Strien et al. 2012; Peterman et al. 2014; Shirk et al. 2018) are increasingly utilized, in particular, because they enable a multivariate model comparison or selection. However, there is still an ongoing debate on the appropriate statistical procedure to model pairwise distances (Guillot and Rousset 2013; Zeller et al. 2016).

In summary, it should have become clear that landscape genetics are a valuable tool for modeling species-specific connectivity, but there are a lot of forks in the road that need to be considered when applying them to new study systems. The myriad of techniques to parameterize resistance models calls for benchmark studies that compare their performance and the assumptions behind them (Reed et al. 2016; Zeller et al. 2018). In particular the pairing of movement and genetic data remains a profound challenge in order to define concrete conservation actions (Jeltsch et al. 2013).

The majority of studies examining functional connectivity for terrestrial animals using landscape genetics focused on large predators (Wasserman et al. 2013; Balkenhol et al. 2014; Mateo-Sanchez et al. 2015; Zeller et al. 2017). Therefore, more species need to be assessed in order to gain a better picture on the key factors that optimize landscape genetic studies for conservation purposes. In particular, large ungulates which are highly impacted by anthropogenic fragmentation (e.g., Frantz et al. 2012). For this reason, the presented thesis studies landscape genetics in combination with movement ecology of red deer in Northern Germany.

Red Deer in Northern Germany

Red deer (*Cervus elaphus*) are among the most widespread ungulates in Europe and one of the most iconic game species. They have been heavily impacted by anthropogenic influences such as habitat fragmentation, translocations and selective hunting for centuries (Hartl et al. 2003). As such, red deer have been the target of many population and conservation genetic studies analyzing the genetic diversity and population structure in human-dominated landscapes (Kuehn et al. 2003; Frantz et al. 2007; Zachos et al. 2007; Haanes et al. 2011; Fickel et al. 2012).

On the one hand, red deer are sensitive to human disturbance (Westekemper et al. 2018), but on the other hand they play a key role in shaping their habitats (Riesch et al. 2019) and as distribution vectors for plants (von Oheimb et al. 2005). Therefore, in regions where red deer are still capable to disperse this species can serve as an indicator for intact habitat networks with low levels of restrictions caused by humans (Tillmann and Reck 2003; Meißner et al. 2008).

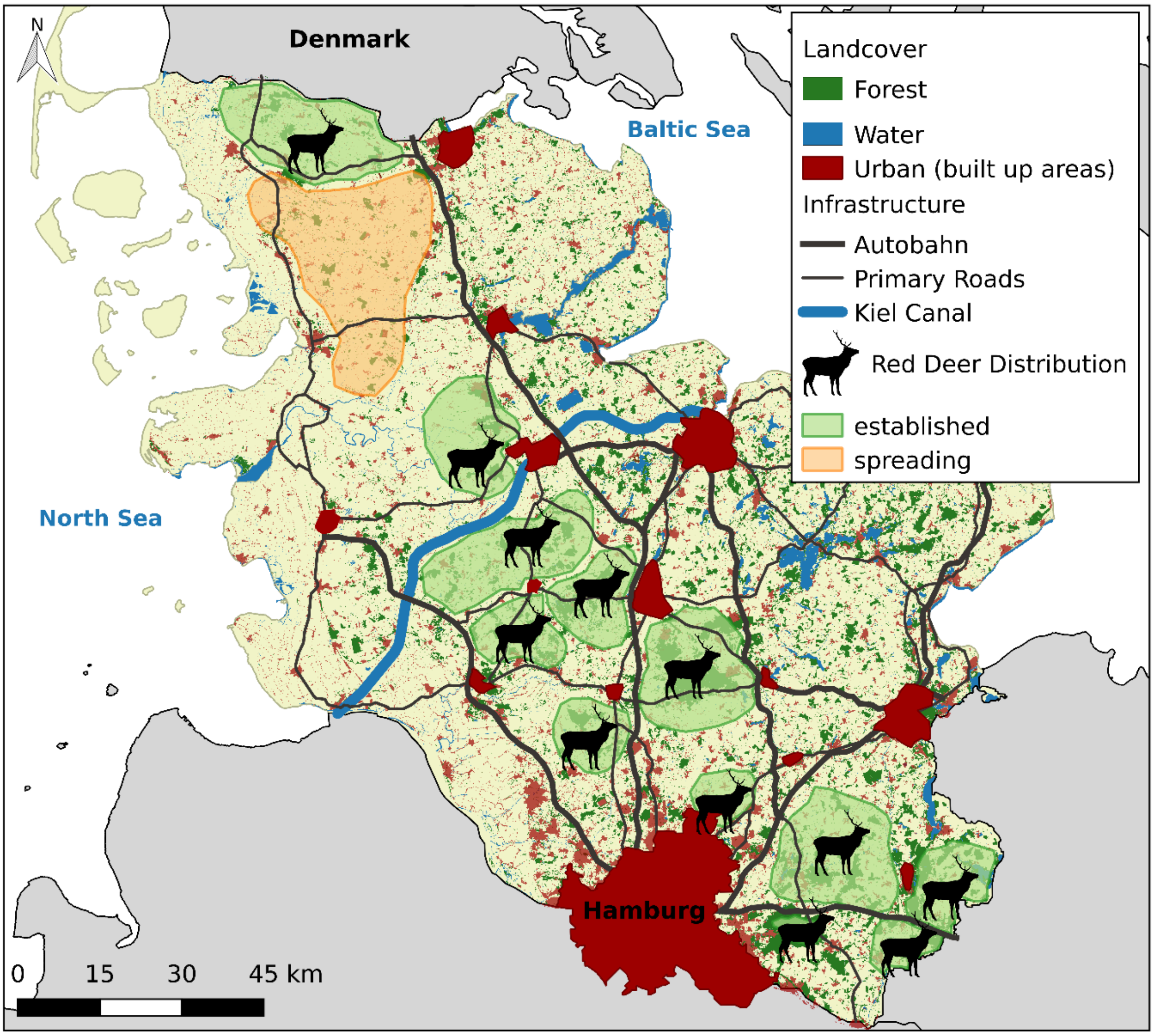


Figure 1.4: Distribution of red deer within the study area of Schleswig-Holstein. Red deer are divided into 12 management units (deer silhouettes). Since the last decade the species started to spread and establish in areas south of the border to Denmark. The map also shows important landcover features such as cultivated (urban) areas, forests and water bodies. Roads and canals form potential barriers to dispersal. Landcover data based on ATKIS (Official topographic and cartographic information system of Germany, http://www.atkis.de).

However, in Schleswig-Holstein, Germany’s northernmost federal state, red deer populations are unevenly distributed, mostly concentrating at the few larger patches of forest and remaining complexes of marshes, moor- and heathlands (Meißner et al. 2008, see Figure 1.4). Fragmentation caused by primary roads, canals and increasing cultivation along with other restrictions in the past (i.e., individuals were only allowed in designated deer areas; Wotschikowsky 2010) lead to a decrease in gene flow and, consequently, loss of genetic diversity. A previous study actually indicated first signs of inbreeding for one of the local populations (Zachos et al. 2007). In summary, these genetic conditions emphasize the need for an improved state-wide functional connectivity of this species. Given its role as indicator species, landscape genetic studies on red deer might ultimately help to improve the connectivity of many other species as well, hence serving an important role for conservation.

Aim of this thesis

Gaining a better understanding of the processes driving differentiation and loss of genetic diversity of populations inhabiting fragmented landscapes is integral for conservation and wildlife management (Cushman et al. 2010). The aim of the thesis is to utilize genetic information to study the consequences of fragmentation and other anthropogenic restrictions on red deer populations in Schleswig-Holstein. Mitigation measures for enhancing connectivity such as the delineation of conservation corridors need to be founded on objective information about the effects of landscape composition on dispersal and gene flow (Mateo-Sanchez et al. 2015).

Movement is a key factor when it comes to gaining a better understanding of how landscape features impede or facilitate dispersal of mobile species such as large ungulates (Diniz et al 2020). Therefore, a major focus of this thesis is to derive different models on red deer space use and test their performance in terms of describing functional connectivity. The key strength working on this system is the availability of high quality spatial and genetic information which provides ideal conditions to run extensive benchmark analyses for landscape genetic model construction and validation. In summary, the thesis provides important insights for applied conservation and management of wildlife in terms of counteracting the negative effects of anthropogenic fragmentation.

Structure of this thesis

Next to this general introduction the dissertation comprises three individual studies on separate topics of applied landscape genetics and movement ecology:

The first study depicted in **chapter two** focuses on the genetic diversity and structure of red deer populations in Schleswig-Holstein. An extensive overview on different methods for identifying behavioral patterns from movement data is presented in the **third chapter**. The final study shown in **chapter four** utilizes such a path-segmentation method along with other approaches to model landscape resistance and compares them in a landscape genetic framework. The **fifth chapter** summarizes and discusses the findings of the three studies with particular focus on the potential contribution to applied conservation.

References

Aben, J., Strubbe, D., Adriaensen, F., Palmer, S. C. F., Travis, J. M. J., Lens, L., and Matthysen, E. (2014). Simple individual-based models effectively represent Afrotropical forest bird movement in complex landscapes. *Journal of Applied Ecology*, 51, 693-702.

Abrahms, B., Sawyer, S. C., Jordan, N. R., McNutt, J. W., Wilson, A. M., and Brashares, J. S. (2017). Does wildlife resource selection accurately inform corridor conservation? *Journal of Applied Ecology*, *54*(2), 412–422.

Adriaensen, F., Chardon, J. P., De Blust, G., Swinnen, E., Villalba, S., Gulinck, H., and Matthysen, E. (2003). The application of ‘least-cost’ modelling as a functional landscape model. *Landscape and Urban Planning*, *64*(4), 233–247.

Andersen, L. W., Fog, K., and Damgaard, C. (2004). Habitat fragmentation causes bottlenecks and inbreeding in the European tree frog (Hyla arborea). *Proceedings of the Royal Society B: Biological Sciences*, *271*(1545), 1293–1302.

Baguette, M., and Van Dyck, H. (2007). Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. *Landscape Ecology*, *22*(8), 1117–1129.

Balkenhol, N., Holbrook, J. D., Onorato, D., Zager, P., White, C., and Waits, L. P. (2014). A multi-method approach for analyzing hierarchical genetic structures: A case study with cougars Puma concolor. *Ecography*, *37*(6), 552–563.

Balkenhol, N., and Waits, L. L. P. (2009). Molecular road ecology: exploring the potential of genetics for investigating transportation impacts on wildlife. *Molecular Ecology*, *18*(20), 4151–4164.

Balkenhol, N., Waits, L. P., and Dezzani, R. J. (2009). Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography*, *32*(5), 818–830.

Beier, P., Majka, D. R., and Spencer, W. D. (2008). Forks in the road: choices in procedures for designing wildland linkages. *Conservation Biology*, *22*(4), 836–851.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., and Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, *15*(4), 365–377.

Betts, M. G., Gutzwiller, K. J., Smith, M. J., Robinson, W. D., and Hadley, A. S. (2015). Improving inferences about functional connectivity from animal translocation experiments. *Landscape Ecology*, *30*(4), 585-593.

Böttcher, M., Reck, H., Hänel, K., and Winter, A. (2004). Lebensraumkorridore für Mensch und Natur. In *GAIA-Ecological Perspectives for Science* (Vol. 2).

Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R., and Cavalli-Sforza, L. L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, *368*(6470), 455-457.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., and Schmiegelow, F. K. . (2002). Evaluating resource selection functions. *Ecological Modelling*, *157*(2–3), 281–300.

Centeno-Cuadros, A., Hulva, P., Romportl, D., Santoro, S., Stříbná, T., Shohami, D., Evin, A., Tsoar, A., Benda, P., Horáček, I., and Nathan, R. (2017). Habitat use, but not gene flow, is influenced by human activities in two ecotypes of Egyptian fruit bat (Rousettus aegyptiacus). *Molecular Ecology*, *26*(22), 6224–6237.

Chetkiewicz, C.-L. B., St. Clair, C. C., and Boyce, M. S. (2006). Corridors for Conservation: Integrating Pattern and Process. *Annual Review of Ecology, Evolution, and Systematics*, *37*(1), 317–342.

Clevenger, A. P., Wierzchowski, J., Chruszcz, B., and Gunson, K. (2002). GIS-Generated, Expert-Based Models for Identifying Wildlife Habitat Linkages and Planning Mitigation Passages. *Conservation Biology*, *16*(2), 503–514.

Colbert, J., Danchin, E., Dondt, A., and Nichols, J. (2001). *Dispersal*. Oxford.

Coulon, A., Cosson, J., Angibault, J.-M., Cargnelutti, B., Galan, M., Morellet, N., Petit, E., Aulagnier S., and Hewison, A. (2004). Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual–based approach. *Molecular Ecology*, *13*, 2841–2850.

Crooks, K. R., and Sanjayan, M. (2006). *Connectivity conservation* (Vol. 14). Cambridge University Press.

Cushman, S. A., McKelvey, K., Hayden, J., and Schwartz, M. K. (2006). Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *The American Naturalist*, *168*(4), 486–499.

Cushman, S. A., Compton, B. W., and McGarigal, K. (2010). Habitat Fragmentation Effects Depend on Complex Interactions Between Population Size and Dispersal

Ability : Modeling Influences of Roads , Agriculture and Residential Development Across a Range of Life-History Characteristics. In *Spatial Complexity, Informatics, and Wildlife Conservation* (pp. 369–385). Springer, Tokyo.

Cushman, S. A., Wasserman, T. N., Landguth, E. L., and Shirk, A. J. (2013). Re-evaluating causal modeling with mantel tests in landscape genetics. *Diversity*, *5*(1), 51–72.

Devictor, V., Van Swaay, C., Brereton, T., Brotons, L., Chamberlain, D., Heliölö, J., Herrando, S., Julliard, R., Kuussaari, M., Lindström, Å., Reif, J., Roy, D. B., Schweiger, O., Settele, J., Stefanescu, C., Van Strien, A., Van Turnhout, C., Vermouzek, Z., WallisDeVries, M., Wynhoff, I., and Jiguet, F. (2012). Differences in the climatic debts of birds and butterflies at a continental scale. *Nature Climate Change*, *2*(2), 121–124.

Diniz, M. F., Cushman, S. A., Machado, R. B., and Júnior, P. D. M. (2020). Landscape connectivity modeling from the perspective of animal dispersal. *Landscape Ecology*, *35*(1), 41-58.

Driezen, K., Adriaensen, F., Rondinini, C., Doncaster, C. P., and Matthysen, E. (2007). Evaluating least-cost model predictions with empirical dispersal data: A case-study using radiotracking data of hedgehogs (Erinaceus europaeus). *Ecological Modelling*, *209*(2–4), 314–322.

Engler, J. O., Balkenhol, N., Filz, K. J., Habel, J. C., and Rödder, D. (2014). Comparative landscape genetics of three closely related sympatric Hesperid butterflies with diverging ecological traits. *PLoS ONE*, *9*(9).

Epps, C. W., Wehausen, J. J. D., Bleich, V. C., Torres, S. S. G., and Brashares, J. S. J. (2007). Optimizing dispersal and corridor models using landscape genetics. *Journal of Applied Ecology*, *44*(4), 714–724.

Epps, C. W., and Keyghobadi, N. (2015). Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Molecular Ecology*, *24*(24), 6021–6040.

Epps, C. W., Palsboll, P. J., Wehausen, J. D., Roderick, G. K., Ramey, R. R., and McCullough, D. R. (2005). Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters*, *8*(10), 1029–1038.

Fahrig, L. (2003). Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 487–515.

Fahrig, L., and Merriam, G. (1994). Conservation of Fragmented Populations. *Conservation Biology*, *8*(1), 50–59.

Fattebert, J., Robinson, H. S., Balme, G., Slotow, R., and Hunter, L. (2015). Structural habitat predicts functional dispersal habitat of a large carnivore: How leopards change spots. *Ecological Applications*, *25*(7), 1911–1921.

Fickel, J., Bubliy, O. A., Stache, A., Noventa, T., Jirsa, A., and Heurich, M. (2012). Crossing the border? Structure of the red deer (Cervus elaphus) population from the Bavarian-Bohemian forest ecosystem. *Mammalian Biology*, *77*(3), 211–220.

Franklin, J. (2009). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Frantz, A. C., Pope, L. C., Etherington, T. R., Wilson, G. J., and Burke, T. (2010). Using isolation-by-distance-based approaches to assess the barrier effect of linear landscape elements on badger (Meles meles) dispersal. *Molecular Ecology*, *19*(8), 1663–1674.

Frantz, A. C., Bertouille, S., Eloy, M. C., Licoppe, A., Chaumont, F., and Flamand, M. C. (2012). Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (Cervus elaphus), but not wild boars (Sus scrofa). *Molecular Ecology*, *21*(14), 3445–3457.

Frantz, A. C., Hamann, J. L., and Klein, F. (2007). Fine-scale genetic structure of red deer (Cervus elaphus) in a French temperate forest. *European Journal of Wildlife Research*, *54*(1), 44–52.

Goodwin, B., and Fahrig, L. (2002). How does landscape structure influence landscape connectivity? *Oikos*, *99*(3), 552–570.

Graves, T., Chandler, R. B., Royle, J. A., Beier, P., and Kendall, K. C. (2014). Estimating landscape resistance to dispersal. *Landscape Ecology*, *29*(7), 1201–1211.

Guillot, G., and Rousset, F. (2013). Dismantling the Mantel tests. *Methods in Ecology and Evolution*, *4*(4), 336–344.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M. R., Possingham, H. P., Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, *16*(12), 1424-1435.

Haanes, H., Røed, K. H., Perez-Espona, S., and Rosef, O. (2011). Low genetic variation support bottlenecks in Scandinavian red deer. *European Journal of Wildlife Research*, *57*(6), 1137–1150.

Hanski, I. (1998). Metapopulation dynamics. *Nature*, *396*, 41–49.

Hanski, I., and Ovaskainen, O. (2003). Metapopulation theory for fragmented landscapes. *Theoretical Population Biology*, *64*(1), 119–127.

Hartl, G. B., Zachos, F., and Nadlinger, K. (2003). Genetic diversity in European red deer (Cervus elaphus L.): anthropogenic influences on natural populations. *Comptes Rendus Biologies*, *326*, 37–42.

Hilty, J., Lidicker Jr, W. Z., and Merenlender, A. (2012). *Corridor ecology: the science and practice of linking landscapes for biodiversity conservation*. Island Press

Holderegger, R., and Wagner, H. H. (2006). A brief guide to Landscape Genetics. *Landscape Ecology*, *21*(6), 793–796.

Holyoak, M., Casagrandi, R., Nathan, R., Revilla, E., and Spiegel, O. (2008). Trends and missing parts in the study of movement ecology. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19060–19065.

Hutchison, D. W., and Templeton, A. R. (1999). Correlation of Pairwise Genetic and Geographic Distance Measures: Inferring the Relative Influences of Gene Flow and Drift on the Distribution of Genetic Variability. *Evolution*, *53*(6), 1898.

Jacobs, S., Burkhard, B., Van Daele, T., Staes, J., and Schneiders, A. (2014). 'The Matrix Reloaded': A review of expert knowledge use for mapping ecosystem services. *Ecological Modelling*, *295*, 21–30.

Jaeger, J., and Holderegger, R. (2005). Schwellenwerte der Landschaftszerschneidung. *GAIA*, *2*, 113–118.

Jeltsch, F., Bonte, D., Pe'er, G., Reineking, B., Leimgruber, P., Balkenhol, N., Schröder, B., Buchmann, C. M., Mueller, T., Blaum, N., Zurell, D., Böhning-Gaese, K., Wiegand, T., Eccard, J.A., Hofer, H., Reeg, J., Eggers, U., and Bauer, S. (2013). Integrating movement ecology with biodiversity research - exploring new avenues to address spatiotemporal biodiversity dynamics. *Movement Ecology*, *1*(1), 6.

Jost, L., Archer, F., Flanagan, S., Gaggiotti, O., Hoban, S., and Latch, E. (2018). Differentiation measures for conservation genetics. *Evolutionary Applications*, *11*(7), 1139–1148.

Keeley, A. T. H., Beier, P., and Gagnon, J. W. (2016). Estimating landscape resistance from habitat suitability: effects of data source and nonlinearities. *Landscape Ecology*, *31*(9), 2151–2162.

Keyghobadi, N. (2007). The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, *85*(10), 1049–1064.

Keyghobadi, N., Roland, J., and Strobeck, C. (2005). Genetic differentiation and gene flow among populations of the alpine butterfly, Parnassius smintheus, vary with landscape connectivity. *Molecular Ecology*, *14*(7), 1897–1909.

Kindlmann, P., and Burel, F. (2008). Connectivity measures: a review. *Landscape Ecology*, (23), 879–890.

Koen, E. L., Bowman, J., and Walpole, A. A. (2012). The effect of cost surface parameterization on landscape resistance estimates. *Molecular Ecology Resources*, *12*(4), 686–696.

Kool, J. T., Moilanen, A., and Treml, E. A. (2013). Population connectivity: Recent advances and new perspectives. *Landscape Ecology*, *28*(2), 165–185.

Kuehn, R., Schroeder, W., Pirchner, F., and Rottmann, O. (2003). Genetic diversity, gene flow and drift in Bavarian red deer populations (Cervus elaphus). *Conservation Genetics*, *4*(2), 157–166.

Legendre, P., and Fortin, M.-J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, *10*(5), 831–844.

Manel, Stéphanie, and Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology and Evolution*. *28*(10), 614-621.

Manel, Stephanie, Schwartz, M. K., Luikart, G., and Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, *18*(4), 189–197.

Manly, B.F.J., McDonald, L.L., and Thomas, D.L. (1993). Intro-duction to resource selection studies. In *Resource Selection by Animals: Statistical Design and Analysis for Field Studies* (pp. 1–13). Chapman & Hall, London, UK

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2), 209–220.

Mateo-Sanchez, M., Balkenhol, N., Cushman, S. A., T, P., Dominguez, A., and Saura, S. (2015). Estimating effective landscape distances and movement corridors : comparison of habitat and genetic data. *Ecosphere,* 6(4), 1-16..

McRae, B., Dickson, B., Keitt, T., and Shah, V. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology, 89*, 2712–2724.

McRae, B.H. (2006). Isolation by resistance. *Evolution*, *60*(8), 1551–1561.

McRae, Brad H., and Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(50), 19885–19890.

Meißner, M., Reinecke, H., Zachos, F., Corsmann, M., Ritter, T., and Wölfel, H. (2008). *Der Rothirsch in Schleswig-Holstein*. Göttingen.

Merriam, G., & Wegner, J. (1992). Local extinctions, habitat fragmentation, and ecotones. In *Landscape boundaries* (pp. 150-169). Springer, New York, NY.

Milanesi, P, Holderegger, R., Bollmann, K., Gugerli, F., and Zellweger, F. (2016). Three-dimensional habitat structure and landscape genetics: a step forward in estimating functional connectivity. *Ecology,* 18*3*(1), 67-79.

Milanesi, P., Holderegger, R., Caniglia, R., Fabbri, E., Galaverni, M., and Randi, E. (2017). Expert-based versus habitat-suitability models to develop resistance surfaces in landscape genetics. *Oecologia*, *183*(1), 67-79.

Moilanen, A., Franco, A. M. , Early, R. I., Fox, R., Wintle, B., and Thomas, C. D. (2005). Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings. Biological Sciences / The Royal Society*, *272*(1575), 1885–1891.

Murray, J. V., Goldizen, A. W., O'Leary, R. A., McAlpine, C. A., Possingham, H. P., and Choy, S. L. (2009). How useful is expert opinion for predicting the distribution of a species within and beyond the region of expertise? A case study using brush-tailed rock-wallabies Petrogale penicillata. *Journal of Applied Ecology*, *46*(4), 842–851.

Nathan, R., Getz, W. M., Revilla, E., Holoyak, M., Kadmon, R., Saltz, D., and Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, *105*(49), 19052–19059.

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*(949), 283–292.

Opdam, P. (1991). Metapopulation theory and habitat fragmentation: a review of holarctic breeding bird studies. *Landscape Ecology*, *5*(2), 93–106.

Oyler-McCance, S. J., Fedy, B. C., and Landguth, E. L. (2013). Sample design effects in landscape genetics. *Conservation Genetics*, *14*(2), 275–285.

Perez, S. I., Diniz-Filho, J. A. F., Bernal, V., and Gonzalez, P. N. (2010). Alternatives to the partial Mantel test in the study of environmental factors shaping human morphological variation. *Journal of Human Evolution*, *59*(6), 698–703.

Peterman, W. E., Connette, G. M., Semlitsch, R. D., and Eggert, L. S. (2014). Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Molecular Ecology*, *23*(10), 2402–2413.

Razgour, O. (2015). Beyond species distribution modelling: A landscape genetics approach to investigating range shifts under future climate change. *Ecological Informatics*, *30*, 250–256.

Reding, D. M., Cushman, S. A., Gosselink, T. E., and Clark, W. R. (2013). Linking movement behavior and fine-scale genetic structure to model landscape connectivity for bobcats (Lynx rufus). *Landscape Ecology*, *28*(3), 471–486.

Reed, G. C., Litvaitis, J. A., Callahan, C., Carroll, R. P., Litvaitis, M. K., and Broman, D. J. A. (2016). Modeling landscape connectivity for bobcats using expert-opinion and empirically derived models: how well do they work? *Animal Conservation*, *20*(4), 308-320.

Revilla, E., and Wiegand, T. (2008). Individual movement behavior, matrix heterogeneity, and the dynamics of spatially structured populations. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19120–19125.

Richardson, J. L., Brady, S. P., Wang, I. J., and Spear, S. F. (2016). Navigating the pitfalls and promise of landscape genetics. *Molecular Ecology*, *25*(4), 849–863.

Ricketts, T. H. (2001). The matrix matters: Effective isolation in fragmented landscapes. *American Naturalist*, *158*(1), 87–99.

Riesch, F., Tonn, B., Meißner, M., Balkenhol, N., and Isselstein, J. (2019). Grazing by wild red deer: Management options for the conservation of semi-natural open habitats. *Journal of Applied Ecology*, *56*(6), 1311–1321.

Roever, C. L., Beyer, H. L., Chase, M. J., and van Aarde, R. J. (2013). The pitfalls of ignoring behaviour when quantifying habitat selection. *Diversity and Distributions*, *20*(3), 322–333.

Rousset, F. (2000). Genetic differentiation between individuals. *Journal of Evolutionary Biology*, *13*(1), 58–62.

Rudnick, D. A., Ryan, S. J., Beier, P., Cushman, S. A., Dieffenbach, F., Epps, C. W., Gerber, L.R., Hartter, J. N., Jenness, J. S., Kintsch, J., Merenlender, A. M., Perkl, R. M., Periosi, D. V., and Trombulak, S. C. (2012). The role of landscape connectivity in planning and implementing conservation and restoration priorities. *Issues in Ecology*, (16), 1–23.

Ruiz-Gonzalez, A., Cushman, S. A., Madeira, M. J., Randi, E., and Gómez-Moliner, B. J. (2015). Isolation by distance, resistance and/or clusters? Lessons learned from a forest-dwelling carnivore inhabiting a heterogeneous landscape. *Molecular Ecology*, *24*(20), 5110–5129.

Shafer, A. B. A., Northrup, J. M., White, K. S., Boyce, M. S., Côté, S. D., and Coltman, D. W. (2012). Habitat selection predicts genetic relatedness in an alpine ungulate. *Ecology*, *93*(6), 1317-1329.

Shirk, A. J., Landguth, E. L., and Cushman, S. A. (2017). A comparison of individual-based genetic distance metrics for landscape genetics. *Molecular Ecology Resources*, *17*(6), 1308-1317.

Shirk, A. J., Wallin, D. O., Cushman, S. A., Rice, C. G., and Warheit, K. I. (2010). Inferring landscape effects on gene flow: a new model selection framework. *Molecular Ecology*, *19*(17), 3603–3619.

Shirk, Andrew J., Landguth, E. L., and Cushman, S. A. (2018). A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Molecular Ecology Resources*, *18*(1), 55–67.

Spear, S. F., Balkenhol, N., Fortin, M.-J., McRae, B. H., and Scribner, K. (2010). Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, *19*(17), 3576–3591.

Spear, S. F., Cushman, S. A., and McRae, B. H. (2015). Resistance Surface Modeling in Landscape Genetics. In *Landscape Genetics* (pp. 129–148).

Squires, J. R., DeCesare, N. J., Olson, L. E., Kolbe, J. A., Hebblewhite, M., and Parks, S. A. (2013). Combining resource selection and movement behavior to predict corridors for Canada lynx at their southern range periphery. *Biological Conservation*, *157*, 187–195.

Stevens, V. M., Verkenne, C., Vandewoestijne, S., Wesselingh, R. A., and Baguette, M. (2006). Gene flow and functional connectivity in the natterjack toad. *Molecular Ecology*, *15*(9), 2333–2344.

Storfer, A, Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., Dezzani, R., Delmelle, E., Vierling, L., and Waits, L. P. (2007). Putting the “landscape” in landscape genetics. *Heredity*, *98*(3), 128–142.

Storfer, Andrew, Murphy, M. A., Spear, S. F., Holderegger, R., and Waits, L. P. (2010). Landscape genetics: where are we now? *Molecular Ecology*, *19*(17), 3496–3514.

Sunnucks, P. (2000). Efficient genetic markers for population biology. *Trends in Ecology and Evolution*, *15*(5), 199–203.

Taylor, P. D., Fahrig, L., Henein, K., and Merriam, G. (1993). Connectivity is a vital element of landscape structure. *Oikos*, *68*(3), 571–573.

Templeton, A. R., Shaw, K., Routman, E., and Davis, S. K. (1990). The genetic consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden*, 13-27.

Thurfjell, H., Ciuti, S., and Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, *2*(1), 4.

Tillmann, J., and Reck, H. (2003). Zur Rolle des Rothirsches (Cervus elaphus L.) im Ökosystem und Empfehlungen für sein Management in Schleswig-Holstein.

Tischendorf, L., and Fahrig, L. (2000). How should we measure landscape connectivity. *Landscape Ecology*, *15*, 633–641.

van Strien, M. J., Holderegger, R., and Van Heck, H. J. (2015). Isolation-by-distance in landscapes: Considerations for landscape genetics. *Heredity*, *114*(1), 27–37.

van Strien, M. J., Keller, D., and Holderegger, R. (2012). A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Molecular Ecology*, *21*(16), 4010–4023.

Volpe, N. L., Hadley, A. S., Robinson, W. D., and Betts, M. G. (2014). Functional connectivity experiments reflect routine movement behavior of a tropical hummingbird species. *Ecological Applications*, *24*(8), 2122-2131.

von Oheimb, G., Schmidt, M., Kriebitzsch, W. U., and Ellenberg, H. (2005). Dispersal of vascular plants by game in northern Germany. Part II: Red deer (Cervus elaphus). *European Journal of Forest Research*, *124*(1), 55–65.

Vuilleumier, S., and Metzger, R. (2006). Animal dispersal modelling: Handling landscape features and related animal choices. *Ecological Modelling*, *190*, 159–170.

Waits, L. P., and Storfer, A. (2015). Basics of Population Genetics: Quantifying Neutral and Adaptive Genetic Variation for Landscape Genetic Studies. In *Landscape Genetics* (pp. 35–57).

Wang, I. J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: A multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution*, *67*(12), 3403–3411.

Wang, J. (2017). The computer program structure for assigning individuals to populations: easy to use but easier to misuse. *Molecular Ecology Resources*, *17*(5), 981–990.

Wang, Y. H., Yang, K. C., Bridgman, C. L., and Lin, L. K. (2008). Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*, *23*(8), 989–1000.

Wasserman, T. N., Cushman, S. A., Littell, J. S., Shirk, A. J., and Landguth, E. L. (2013). Population connectivity and genetic diversity of American marten (Martes americana) in the United States northern Rocky Mountains in a climate change context. *Conservation Genetics*, *14*(2), 529–541.

Wasserman, T. N., Cushman, S. A., Schwartz, M. K., and Wallin, D. O. (2010). Spatial scaling and multi-model inference in landscape genetics: Martes americana in northern Idaho. *Landscape Ecology*, *25*(10), 1601–1612.

Westekemper, K., Reinecke, H., Signer, J., Meißner, M., Herzog, S., and Balkenhol, N. (2018). Stay on trails – effects of human recreation on the spatiotemporal behavior of red deer Cervus elaphus in a German national park. *Wildlife Biology*, *2018*(1).

Whitmee, S., and Orme, C. D. L. (2013). Predicting dispersal distance in mammals: a trait-based approach. *The Journal of Animal Ecology*, *82*(1), 211–221.

Wotschikowsky, U. (2010). Ungulates and their management in Germany. In *European Ungulates and Their Management in the 21st Century* (pp. 201–222).

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114–138.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, *19*(3), 395–420.

Zachos, F. E., Althoff, C., Steynitz, Y. V., Eckert, I., and Hartl, G. B. (2007). Genetic analysis of an isolated red deer (Cervus elaphus) population showing signs of inbreeding depression. *European Journal of Wildlife Research*, *53*(1), 61–67.

Zeller, K. A., Creech, T. G., Millette, K. L., Crowhurst, R. S., Long, R. A., Wagner, H. H., Balkenhol, N., and Landguth, E. L. (2016). Using simulations to evaluate Mantel-based methods for assessing landscape resistance to gene flow. *Ecology and Evolution*, *6*(12), 4115-4128.

Zeller, K. A., Jennings, M. K., Vickers, T. W., Ernest, H. B., Cushman, S. A., and Boyce, W. M. (2018). Are all data types and connectivity models created equal?

Validating common connectivity approaches with dispersal data. *Diversity and Distributions*, *24*(7), 868–879.

Zeller, K. A., McGarigal, K., Cushman, S. A., Beier, P., Vickers, T. W., and Boyce, W. M. (2015). Using step and path selection functions for estimating resistance to movement: pumas as a case study. *Landscape Ecology*, *31*(6), 1319-1335.

Zeller, K. A., McGarigal, K., and Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. *Landscape Ecology*, *27*(6), 777–797.

Zeller, K. A., Vickers, T. W., Ernest, H. B., Boyce, W. M., Pollinger, J., and Ernest, H. (2017). Multi-level, multi-scale resource selection functions and resistance surfaces for conservation planning: Pumas as a case study. *PLOS ONE*, *12*(6), e0179570.

CHAPTER 2

Genetic analysis of red deer (*Cervus elaphus*) administrative management units in a human-dominated landscape – patterns of genetic diversity, population structure and gene flow

Published as: Edelhoff, H., Fickel, J., Epps, C. W., Zachos, F. E., and Balkenhol, N. (2020). Genetic analysis of red deer (*Cervus elaphus*) administrative management units in a human-dominated landscape. *Conservation Genetics*, *21*, 261-276

Abstract

Red deer (*Cervus elaphus*) throughout central Europe are impacted by different anthropogenic activities including habitat fragmentation, selective hunting and translocations. This has substantial influences on genetic diversity and the long-term conservation of local populations of this species. Here we use genetic samples from 480 red deer individuals to assess genetic diversity and differentiation of the 12 administrative management units located in Schleswig Holstein, the northernmost federal state in Germany.

We applied multiple analytical approaches and show that the history of local populations (i.e., translocations, culling of individuals outside of designated red deer zones, anthropogenic infrastructures) has led to comparably low levels of genetic diversity. Mean expected heterozygosity was below 0.6 and we observed on average 4.2 alleles across 12 microsatellite loci. Effective population sizes below the recommended level of 50 were estimated for multiple local populations.

Our estimates of genetic structure and gene flow show that red deer in northern Germany are best described as a complex network of asymmetrically connected subpopulations, with high genetic exchange among some local populations and reduced connectivity of others. Genetic diversity was also correlated with population densities of neighbouring management units.

Based on these findings, we suggest that connectivity among existing management units needs to be considered in the practical management of the species, which means that some administrative management units should be managed together, while the effective isolation of other units needs to be mitigated.

Keywords: meta-population, wildlife management, genetic connectivity, isolation, local populations

Introduction

Red deer (*Cervus elaphus*) are among the most widespread ungulates in Europe and one of the most iconic game species. They have been heavily impacted by anthropogenic influences such as habitat fragmentation, translocations and selective hunting for centuries (Hartl et al. 2003). As such, red deer have been the target of many population and conservation genetic studies analyzing the genetic diversity and population structure in human-dominated landscapes (e.g., Kuehn et al. 2003; Pérez-Espona et al. 2008, 2009; Fickel et al. 2012; Frantz et al. 2017). The aims of these studies varied, and included the quantification of genetic diversity in isolated and sometimes inbred populations (e.g., Zachos et al. 2007), estimating the amount and genetic consequences of translocations (e.g., Haanes et al. 2010), or characterizing the genetic impacts of postglacial recolonization (e.g., Krojerova-Prokesova et al. 2015).

In Schleswig-Holstein, Germany's northernmost federal state, red deer are distributed across the north, southeast, and center of the state (Figure 2.1). The local populations are managed in 12 administrative units. These units were not established on the basis of population structure, but rather were opportunistically located in areas with high red deer densities, mostly located around larger patches of forest (Meißner et al. 2008; Wotschikowsky 2010). Units located in close proximity to each other such as Barlohe (BAL), Iloo (ILO) and Schierenwald (SCW) are demarcated by spatial jurisdictions (e.g., municipalities or communities) or landmarks (e.g., rivers or roads), rather than by natural boundaries or population structure. Such administrative considerations are commonly included when delineating wildlife management units (Taylor and Dizon 1999).

From a genetic standpoint, populations should only be treated as separate management units when their genetic divergence is high enough to suggest demographic independence, meaning that the rate of dispersal among populations must be low (Palsbøll et al. 2007). To emphasize that red deer management units in our study area (Schleswig-Holstein) are not based on population structure, we refer to them as administrative management units (AMUs) and distinguish them from units defined by genetic divergence (i.e., genetic management units; GMUs).

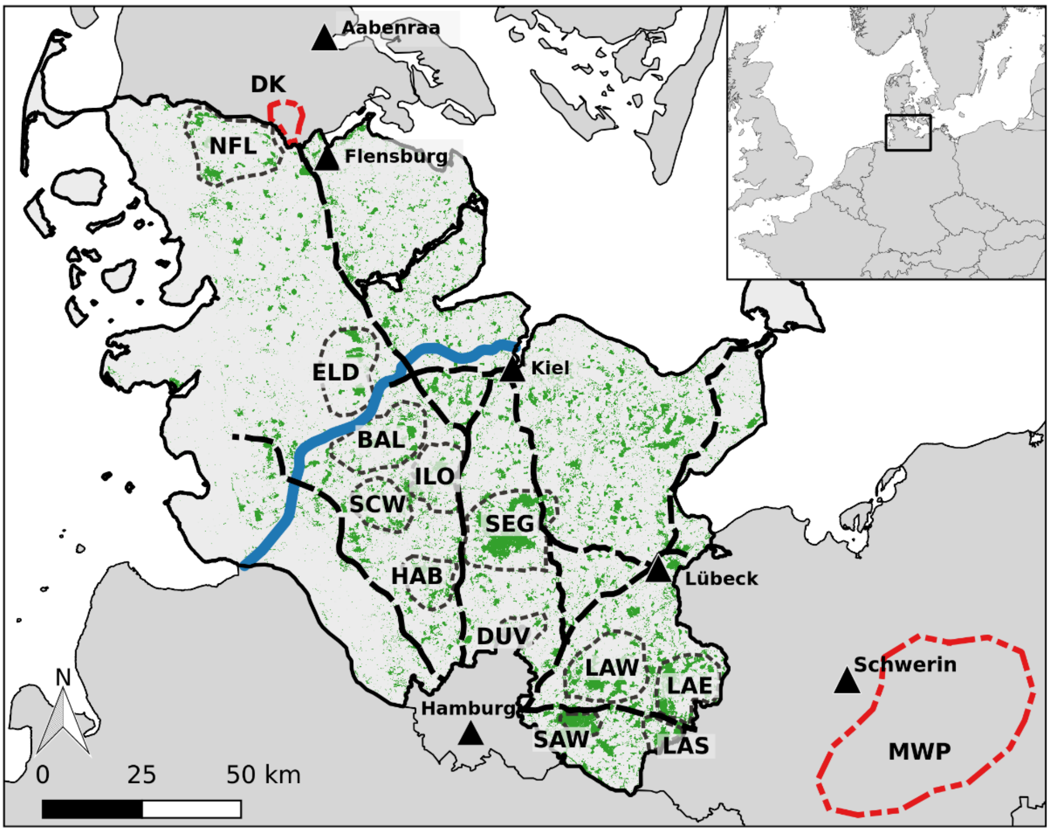


Figure 2.1: Map of Schleswig-Holstein (study area). Inset indicates location within Germany. The blue line in the center indicates the Kiel Canal. Broad dashed black lines represent major highways (Autobahn). Red deer management units are delineated with thin dashed black lines. Forested areas are indicated by dark green shading. Local deer management units of which samples were included are Northern Friesland (NFL), Elsdorf (ELD), Barlohe (BAL), Iloo (ILO), Schierenwald (SCW), Hasselbusch (HAB), Segeberger Heide (SEG), Duvenstedter Brook (DUV), Lauenburg West (LAW), East (LAE) and South (LAS) as well as Sachsenwald (SAW). The two reference areas Denmark (DK) and Mecklenburg-Western Pomerania (MWP) are delineated in red. Triangles represent larger cities throughout that area.

Historically, red deer within the AMUs in Schleswig-Holstein (SH) have been subject to various anthropogenic restrictions. For instance, until recently (i.e., 1980), red deer were only allowed to freely range in so called ‘designated red deer zones’ (Meißner et al. 2008; Wotschikowsky 2004, 2010). This policy was intended to prevent damages to crops and forests by red deer. Culling of all individuals outside these zones consequently limited gene flow between established populations (Ströhlein et al. 1993; Willems et al. 2016). Today, infrastructures such as fenced highways (Autobahn) or the Kiel Canal form potential barriers to gene flow across the entire state (Figure 2.1). Additionally, estimated population sizes vary greatly among the AMUs (range 35-530; see Table 2.1) and many of them contain fewer than 100 individuals. This population was founded by dispersed individuals from the Segeberger Heide (SEG) in the second half of the 19th century (ca. 1870), but has been isolated from its source for decades. More recently, a fenced highway has prevented any potential migration between SEG and HAB (Meißner et al. 2008). A previous study found low genetic diversity as well as the first signs of inbreeding for the Hasselbusch AMU (Zachos et al. 2007). For example, multiple animals with *brachygnathia inferior* (shortened lower jaw), a condition linked to inbreeding depression, have been found in the HAB population (Zachos et al. 2007). Furthermore, there are influences of translocations: the Duvenstedt (DUV) population is not native but goes back to an enclosure population founded with red deer from Austria, Hungary and Poland which was released in the 1950s (Jessen 1988; Meißner et al. 2008). Within the last decade, red deer have dispersed from Denmark, established themselves south of the German border and are increasing in numbers (Reinecke et al. 2013). As a consequence, the latest red deer AMU established in Schleswig-Holstein was the Nordfriesland unit (NFL). In the neighboring state of Mecklenburg-Western Pomerania (MWP) located south-east of Schleswig-Holstein, red deer are more abundant and have been roaming the state with less restrictions while occupying a large area (Kinser et al. 2015). Therefore, an exchange of individuals from these populations could result in higher levels of genetic diversity in the three AMUs located in the Lauenburg area (LAW, LAE, LAS).

Hunters and landowners participate in the management of red deer within the 12 AMUs in order to set different management goals such as hunting quotas (Wotschikowsky 2010). Therefore, managing the AMUs separately assumes that these units equate to GMUs, thus representing more or less disconnected (i.e., closed or genetically separated) populations that experience limited reproductive exchange of individuals with other populations (Moritz 1994). However, several recent studies have shown that if this implicit assumption is violated in wildlife management, actions in one management unit (MU) can substantially influence management effectiveness in neighboring units (Hemami et al. 2005; Robinson et al. 2008; Olea and Mateo-Tomás 2014; Stillfried et al. 2017). In such cases, management would need to be extended towards a larger spatial scale that includes multiple MUs and considers the degree of connectivity among them (e.g., Robinson et al. 2008; Wäber et al. 2013). Genetic approaches have been suggested for delineating more meaningful management units based on biological population entities (e.g., Moritz 1994, Palsbøll et al. 2007). Strong genetic sub-structuring or varying levels of genetic diversity among areas are still the metrics of choice commonly used to justify the separation of MUs (e.g., Wilting et al. 2015; Grosser et al. 2017; Gaillard et al. 2017). However, novel analytical tools now allow researchers to derive estimates of directed dispersal rates from genetic samples, which can provide important information on potential source-sink dynamics and gene flow (e.g., Draheim et al. 2016).

Overall, the history of red deer in SH and the different anthropogenic influences on the local populations raise the question of whether the current practice of managing each AMU as a separate, closed population is appropriate. In particular, it is questionable whether genetic diversity within AMUs is high enough in order to sustainably counteract genetic drift, thereby preventing a loss of genetic diversity and inbreeding. We expect some AMUs to be linked by dispersal and gene flow rates high enough to warrant management as a single unit. If this is the case, red deer AMUs in SH can be interpreted as a network of subpopulations where local populations are connected by gene flow of varying degrees (Pannell and Charlesworth 2000). If so, we should observe different levels of genetic exchange among AMUs and of genetic diversity within AMUs, with migration depending on connectivity among neighboring AMUs, and genetic diversity depending on a combination of connectivity and population size of neighboring AMUs.

To assess the genetic structure of red deer AMUs in Schleswig-Holstein, we make use of an extensive data set consisting of over 500 tissue samples collected over multiple years. Using those samples, we estimate different measures of genetic diversity and test the hypothesis that diversity will vary between the AMUs in Schleswig-Holstein but still be relatively low compared to other populations throughout Europe (Zachos et al. 2016). For this, we also added samples from two references areas located in the neighboring state of Denmark and the federal state Mecklenburg-Western Pomerania. By combining analyses of genetic differentiation and population structure with a novel approach of genetically-derived estimates of relative migration rates (Sundqvist et al. 2016), we also delineate clusters of AMUs that are connected by gene flow and thus should be managed as one GMU. In order to further confirm the genetic structure of the AMUs, we correlate observed patterns of genetic diversity, differentiation and gene flow to available information on current population size and density at the local and regional scale.

Table 2.1: Overview of sampled management units and reference areas including estimated population census size and number of samples (n) for each area. Population-based metrics of genetic diversity include expected (H_E) and observed (H_O) heterozygosities, allelic richness (AR; rarefaction approach based on 12 diploid individuals), as well as F_{IS} and N_E values with confidence intervals. NE values are presented for three different frequency thresholds (FT) accounting for the effects of rare alleles.

Administrative Management Unit	Abbreviation	Population (est.)	Size	Sample size	H_E	H_O	AR (CI)	F_{IS} (CI)	N_E (FT=0.01)	N_E (FT=0.02)	N_E (FT=0.05)
Barlohe	BAL	170		16	0.58	0.63	3.99 (3.5, 4.42)	-0.07 (-0.2, -0.02)	22.6 (14.8, 40.3)	22.6 (14.8, 40.3)	30.9 (17, 90)
Duvenstedter Brook	DUV	80		23	0.58	0.54	4.1 (3.58, 4.5)	0.05 (-0.06, 0.12)	20 (13.3, 33.5)	20 (13.3, 33.5)	19.8 (12.5, 36.4)
Elsdorf / Westermuehlen	ELD	50		46	0.56	0.57	3.89 (3.42, 4.42)	-0.01 (-0.08, 0.04)	21.2 (16.4, 28.1)	26.5 (19.8, 36.8)	20.8 (15.2, 29.2)
Hasselbusch	HAB	60		47	0.49	0.47	3.41 (2.92, 4)	0.05 (-0.02, 0.1)	27 (19.9, 37.8)	40.7 (27.7, 66.1)	34.3 (22.6, 57.9)
Iloo / Moerel	ILO	60		31	0.61	0.61	4.24 (3.83, 4.67)	0 (-0.08, 0.05)	210.9 (64.7, Inf)	76.9 (38.8, 444.4)	78.5 (36.9, 1153.3)
Lauenburg (east)	LAE	530		76	0.64	0.63	4.92 (4.42, 5.42)	0.01 (-0.03, 0.04)	138.9 (94.6, 239.2)	140.1 (92.9, 255.4)	187.5 (102.4, 662.5)
Lauenburg (south)	LAS	35		35	0.62	0.6	4.7 (4.25, 5.08)	0.05 (-0.04, 0.11)	53.8 (35.3, 97.1)	47.4 (32.1, 79.7)	44.8 (29.3, 80.4)
Lauenburg (west)	LAW	120		22	0.61	0.59	4.39 (3.92, 4.83)	0.03 (-0.08, 0.11)	133.8 (41.4, Inf)	133.8 (41.4, Inf)	91.8 (33.6, Inf)
Nordfriesland	NFL	30		12	0.57	0.54	3.99 (3.42, 4.33)	0.07 (-0.08, 0.11)	45.2 (14.5, Inf)	45.2 (14.5, Inf)	28.5 (10.7, Inf)
Sachsenwald	SAW	160		17	0.58	0.53	4.26 (3.67, 4.75)	0.07 (-0.06, 0.15)	45.6 (21, 908.1)	45.6 (21, 908.1)	47.5 (20.6, Inf)
Schierenwald / Steinburg	SCW	110		14	0.57	0.6	3.65 (3.33, 3.92)	-0.05 (-0.16, -0.01)	79.9 (18.1, Inf)	79.9 (18.1, Inf)	28.8 (11.4, Inf)
Segeberger Heide	SEG	300		73	0.62	0.63	4.34 (3.92, 4.75)	0 (-0.05, 0.04)	98.9 (68.2, 163.1)	105.4 (71.1, 181.8)	82.7 (56, 139)
Froslev (Denmark)	DK	~ 100		22	0.54	0.53	3.74 (3.25, 4.17)	0.04 (-0.09, 0.12)	98.9 (32.5, Inf)	98.9 (32.5, Inf)	41 (18.1, 730.6)
Mecklenburg-Western Pomerania	MWP	>700		46	0.66	0.64	5.12 (4.58, 5.58)	0.02 (-0.04, 0.07)	469.8 (143, Inf)	390.8 (130, Inf)	105.7 (60.1, 305.9)

Methods

Study Area

The study area extends over approximately 15,580 km² and covers the entire mainland of the federal state of Schleswig-Holstein in Northern Germany, south of the border with Denmark (Figure 2.1). The state comprises a mosaic of different types of land use, predominantly agriculture and pastures. Forested areas are scattered across the state but vary substantially in size and composition of tree species. Larger forest complexes such as the Segeberger Heide (SEG) form the core areas of the red deer distribution throughout the study area (Figure 2.1). Often, red deer habitats are further characterized by mixtures of marshes, heathlands and moors. Administrative management units vary in size from 13,000 up to 48,000 ha (Reinecke et al. 2013). Distances between AMUs range from a few kilometers (< 5km) up to 63km between the NFL and ELD units. Available information suggests that local populations range in size from 30 to nearly 600 individuals within the AMUs (Table 2.1). Schleswig-Holstein is not densely populated (182 people per km²; Statistisches Bundesamt 2018) compared to the German average (237 people per km²), with human settlements and villages scattered across the state. The landscape is fragmented by roads, major highways (Autobahn) and canals (e.g., the Kiel Canal), all of which form potential barriers to the movements of red deer (Pérez-Espona et al. 2008; Frantz et al. 2012).

Sampling

We obtained 279 genetic samples from red deer harvested during the hunting seasons of 2013 to 2015. In order to ensure a sufficient sample size across all 12 AMUs, we included 186 samples collected in previous studies (Zachos et al. 2007; Reinecke et al. 2013) during the years 2003 and 2004. Additionally, we used samples obtained from two reference areas for comparative purposes: 1) 34 samples from the Froslev forest located in Southern Denmark (DK) close to the German border , and 2) 46 samples from several forests within the federal state of Mecklenburg-Western Pomerania (MWP) neighboring Schleswig-Holstein in the Southeast (Figure 2.1). This lead to a total sample size of 545 (149 female, 104 male, 292 with no sex ID) red deer individuals (overview on sampling periods and sample sizes provided in supplement S2.3). Since free ranging red deer can live up to over 12 years (e.g., Guinness et al. 1978) the gap between the two sampling periods corresponds to a maximum of only one deer generation.

All samples were re-genotyped for our marker set in order to be fully comparable. We only considered samples for which the spatially referenced location of harvest (e.g., the forest complex) was reported. Individuals from MWP originated from areas not directly neighbouring our study area. Therefore, these samples were only included for comparative measures regarding genetic diversity whereas the DK samples were also used throughout the analyses on differentiation and gene flow.

DNA extraction and genotyping

DNA was extracted using the ‘all tissue DNA’ kit (Gen-Ial, Troisdorf, Germany) following the manufacturer’s instructions (final DNA-elution in 80 µl). DNA concentrations were measured spectrophotometrically using a NanoDrop1000 (PeqLab GmbH, Erlangen Germany). To genotype each individual, we used a panel of 14 microsatellite loci (see supplement S2.1). One primer of each of the 14 pairs was 5’-labelled with a fluorescent dye (6-FAM or HEX). To save time and costs, primers were combined (after optimization) in multiplex mixes (CerMix1 – CerMix4). CerMix1 contained primers for four loci (*INRA6*, *C143*, *T40*, and *T115*), CerMix2 combined three loci (*C105*, *C180*, and *C229*), CerMix3 combined four loci (*T107*, *Haut14*, *ILSTS06*, and *BM757*), and CerMix4 included three loci (*CSSM14*, *FSBH*, and *BM1818*). The genotyping reaction mixture (10 µl) consisted of 1× buffer (Promega, Germany), 2mM $MgCl_2$, 1 µl multiplex primer mix [final concentrations per forward and reverse primers varied and were either 0.25 µM (INRA6, T115, T40, C180, C105, C229), 0.3µM (T107, BM757), 0.5µM (C143, Haut14), 1µM (BM1818), 3.5µM (CSSM14), 4µM (FSHB), or 6µM (ILSTS06)], 150ng DNA, 0.25 U GoTaq polymerase (Promega, Germany) and 5.2 µl A.dest. (sterile). Cycling conditions were the same for all four multiplex mixes: 95°C 5 min, 5x (95°C 30s, touchdown beginning at 63°C, with a decrease of 2°C per cycle down to 55°C 90s, 72°C 30s), 40x (95°C 30s, 55°C 90s, 72°C 30s), final extension at 60°C for 30 min. Size of amplicons was determined by calibration using the GENESCAN™ 500 ROX™ size standard. Separation of fragments was carried out on an A3130*xl* automated capillary sequencer using the software GeneMapper v.3.7 for allele scoring (all Applied Biosystems).

Genotyping error estimation

Microsatellite amplicons were screened for genotyping errors (large allele dropouts, stutter bands) and probability of null alleles being present using *MICRO-CHECKER* (version 2.23, Van Oosterhout et al. 2004). We tested all loci across all populations for consistent patterns of deviation from Hardy-Weinberg expectations (HWE) using *GENEPOP* (version 4.5.1; Rousset 2008). All pairs of loci were further checked for linkage disequilibrium within all sampling units applying the algorithms implemented in *GENEPOP* and *ARLEQUIN* (version 3.5; Excoffier et al. 2005) including Bonferroni correction for multiple comparisons (Rice 1989). Additionally, we calculated the number of identified alleles and estimated expected and observed heterozygosities as well as the polymorphic information content (*PIC*) for each marker using the *adegenet* R package (Jombart 2008). Monomorphic markers were excluded from further analyses.

Estimating genetic diversity

All statistical analyses were performed using the R environment (R Core Team 2017). We assessed the amount of genetic variation within each AMU by estimating expected and observed heterozygosities (H_E, H_O), allelic richness (*AR*) and the degree of heterozygote deficiency (F_{IS}) in each management unit. Estimation of *AR* was based on rarefaction to correct for the smallest sample size (n=12). Confidence intervals for *AR* and F_{IS} metrics were obtained using bootstraps with 999 replications. All metrics were estimated applying the *diveRsity* package (Keenan et al. 2013). We estimated effective population sizes (N_E) for all administrative management units using the NeEstimator v2 software (Do et al. 2014). N_E values were based on the linkage disequilibrium method with bias correction developed by (Waples and Do 2008). The same critical thresholds (0.05, 0.02, 0.01) as in Zachos et al. (2016) were applied to correct for linkage of rare alleles with frequencies below these values. The NFL unit was excluded to avoid any potential bias in population size estimates due to low sample size below 15 individuals (Do et al. 2014).

Estimating genetic structure

We assessed genetic structure at the level of the AMUs based on pairwise F_{ST} values (Wright 1965) as well as the pairwise Jost's *D* metric (Jost 2008) using the *strataG* R package (Archer et al. 2017). While Jost's *D* is more appropriate for quantifying genetic (allelic) differentiation of populations showing varying levels of genetic diversity, F_{ST} better reflects past demographic processes and fixation (Whitlock 2011, Jost et al. 2018). Significance of differences in pairwise comparisons was estimated with 9,999 replications and subsequent Bonferroni correction.

To assess whether AMUs actually constituted genetically separate clusters, we applied a Bayesian clustering approach. Specifically, we used the program *STRUCTURE* (version 2.3.4, Pritchard et al. 2000) and tested for the presence of genotypic clusters (*K*), with the number of possible clusters ranging between *K*=1 and *K*=14, using an admixture model and correlated allele frequencies. After having checked for the likelihood to have converged, we estimated the probability for each *K*-value in five independent runs with 500,000 iterations as burn-in followed by 1,000,000 MCMC iterations. The optimal number of *K* was determined using log-likelihood plots and the Δ*K* method by Evanno et al. (2005) implemented in the *STRUCTURE Harvester* platform (Earl and vonHoldt 2012). Individual likelihoods of cluster memberships (*q*) were averaged over the five runs using the *CLUMPAK* online program (Kopelman et al. 2015).

We used *STRUCTURE* in a hierarchical framework by re-running the clustering algorithm for each of the detected genetic clusters in the previous analysis (Coulon et al. 2008; Balkenhol et al. 2014). The procedure was repeated until the optimal number of inferred genetic clusters was equal to one (*K*=1). By doing this, subtle structuring is more likely to be detected because the largest break in the dataset is reiteratively removed so that this strong signal does not blur a weaker signal at lower hierarchical levels (Janes et al. 2017). We performed the hierarchical *STRUCTURE* analysis with 'sampling location' (i.e., the AMU) as a prior (*locprior*; Hubisz et al. 2009). All AMUs from Schleswig-Holstein and the reference area from Denmark were included in this analysis, as these are the sampling areas among which gene flow can be substantial enough to form actual genetic clusters (i.e., MWP samples were excluded).

Estimating directional migration rates

Relative, directional migration was estimated using the *divMigrate* method (Sundqvist et al. 2016) which is implemented in the *diveRsity* R package (Keenan et al. 2013). While other, more complex algorithms are available for estimating asymmetric migration rates (e.g., *BayesAss*, Rannala 2007; *MIGRATE-N*, Beerli 2004), we chose *divMigrate* (Sundqvist et al. 2016) because it can be calculated from standard measures of genetic differentiation and does not require multiple additional parameters to be estimated (Sundqvist et al. 2016). The method tests for significant directionalities in gene flow between pairs of populations based on asymmetric distributions of allele frequencies and generates an output with relative migration rates scaled to values between 0 and 1 (Sundqvist et al. 2016).

We chose the G_{ST} measure of genetic differentiation (Nei 1972) from the options provided by *divMigrate* since it is similar to the F_{ST} values applied above (Whitlock 2011). Again, the analysis was performed for all AMUs within the study area as well as the Danish reference population, which we included because of suspected ongoing migration from Denmark into Germany. Based on the pairwise migration rates, we calculated the mean immigration (I) and emigration (E) rates as well as their ratio ($R_{I/E}$) for each AMU. $R_{I/E} > 1$ would indicate that the rate of immigration in a population is higher than the emigration rate and vice versa for $R_{I/E} < 1$. Finally, we note that the results of the *divMigrate* analysis do not necessarily represent actual migration but rather estimate the probability of the exchange of genes between two sampling locations (Marrotte et al. 2017, Bohling et al. 2019). Further, relative migration rates are estimated across all pairs of included populations and do not account for spatial context or distance between them.

Modeling of genetic patterns

In the next step, we used regression modelling to correlate genetic variation within and among the AMUs with available ecological and environmental information. Specifically, we tested whether genetic diversity, differentiation and migration rates can be explained by local population sizes (S_i) or densities (D_i) within each AMU i, or as a function of the cumulative sizes ($\sum_{j=1}^{n} S_j$) or cumulative densities ($\sum_{j=1}^{n} D_j$) of the three AMUs j (j = 1-3) closest to the focal AMU i. The first two indices, S_i (number of individuals in AMU i) and D_i (individuals per hectare in AMU i), assume that genetic patterns and migration are only influenced by local population characteristics (i.e., size or density). In contrast, the latter two indices essentially are metrics used to describe isolation of multiple, potentially connected populations, and assume that the existence of large or densely populated neighboring AMUs is important for explaining observed population structure (e.g., Balkenhol et al. 2013). The three closest AMUs were chosen to calculate the connectivity indices because this included, in all cases, all the directly neighboring management units that could potentially exchange dispersing individuals with the focal unit.

We used officially available population size estimates (Meißner et al. 2008; Ministerium für Energiewende, Landwirtschaft, Umwelt und ländliche Räume des Landes Schleswig-Holstein 2012; Reinecke et al. 2013) for each AMU to represent S, and estimated D by dividing population size by the area of potential red deer habitat in the AMU. Potential habitat for each AMU was based on official thematic landscape data (authoritative topographic cartographic information system, ATKIS) and included all patches of forest, heathland and moors within the range of each AMU (Reinecke et al. 2013).

We then modelled genetic diversity (AR), genetic differentiation (Jost's D) and mean immigration (I) as well as emigration (E) rates as a function of the four different indices, as well as a null model (intercept-only). We chose AR as a measure of genetic diversity as it was corrected for varying sampling sizes across AMUs. Similarly, we chose Jost's D as an estimate of genetic differentiation because it measures the fraction of allelic variation among populations and thus accounts for varying genetic diversities within AMUs (Jost et al. 2018). Finally, we chose immigration and emigration rates as measures of directional dispersal. To compare models, we used an information-theoretic approach based on Akaike's Information Criterion corrected for small sample size (AICc; Akaike 1973; Burnham and Anderson 2002). The model with the lowest AICc value was deemed best, but models with ΔAICc <= 2 were considered equally plausible (Burnham and Anderson 2002).

Genetic drift and isolation by distance

Following Jordan and Snell (2008) we tested for the potential effect of drift in isolation assuming that historic drift as represented in low genetic variation in smaller populations caused higher levels of differentiation. Therefore, we expect to see a negative relationship between the mean pairwise F_{ST} values of each AMU with all other AMUs and their expected heterozygosities H_E (i.e., AMUs with larger F_{ST} should show lower H_E values than AMUs with smaller F_{ST}-values). We further correlated mean pairwise F_{ST} values with allelic richness (*AR*) as the predictor variable (Whitley et al. 2010; Funk et al. 2016).

Finally, we tested for isolation-by-distance (IBD; Wright 1943) using a Mantel test between genetic distances (linearized F_{ST}, i.e. $F_{ST}/1\text{-}F_{ST}$; and Jost's *D* values) and the natural log of the geographic distance among AMUs (Slatkin 1993). A significant IBD pattern in both F_{ST} and Jost's *D* indicates that gene flow occurs among AMUs but is spatially limited, which hints at subpopulations connected via dispersal (Hutchison and Templeton 1999; Aguillon et al. 2017).

Results

We excluded 65 samples from further analyses because of insufficient numbers of successfully sequenced loci (≤ 11 markers). Therefore, the final dataset consisted of 480 samples including 68 individuals from the two reference regions located in Mecklenburg-Western Pomerania (46 samples) and Denmark (22 samples; Table 2.1). Two (*T40, C105*) of the original 14 microsatellite markers were dropped as they had only two alleles and were near monomorphic in the vast majority of samples with frequencies below 0.15 observed for one of the two alleles. The number of alleles of the remaining markers ranged between three and 14. Polymorphic information content ranged from 0.3 up to 0.86 with a mean *PIC* of 0.62 (SD=0.2) across all loci (more information on marker diversity is provided in the supplement; file S2.2). None of the retained markers showed issues with null alleles or consistent deviations from HWE. We did not find evidence for significant linkage for any of the compared pairs of loci across all sampling units. Private alleles were detected within samples from one reference area (MWP: three alleles) and from two management units (ILO and NFL one allele each).

Genetic diversity

We observed a mean expected heterozygosity of 0.59 (SD=0.04) and a mean allelic richness of 4.20 (SD=0.47) alleles with a minimum of 3.41 and a maximum of 5.12 alleles (based on 12 diploid individuals, see Table 2.1). The Hasselbusch administrative management unit (HAB) showed the lowest values regarding these two metrics. Samples from the two reference areas differed with regard to their genetic diversity with Denmark showing the lowest values of H_E and *AR* (Table 2.1). The samples from Mecklenburg-

Western Pomerania actually exhibited the highest estimates for all diversity metrics compared to DK and the AMUs from Schleswig-Holstein. We did not find any indications for significant heterozygote deficiency. With the exception of Barlohe (BAL) and Schierenwald (SCW), confidence intervals of all estimated F_{IS} values were low and overlapped with zero (Table 2.1), conforming with expectations for random mating within AMUs.

Genetic structure

We observed a global fixation (F_{ST}) value of 0.09 and a global Jost's D of 0.12 across all 12 AMUs of Northern Germany ($p<0.0001$ for both values). Pairwise estimates of F_{ST} and Jost's D ranged between 0.006 and 0.225 with an average of 0.1 for F_{ST} and 0.09 for Jost's D, respectively (Table 2.2). Overall, estimates of the two metrics agreed in most cases regarding the significant differentiation between the considered AMUs. However, not all AMUs were genetically differentiated. We were able to distinguish three groups of administrative management units which did not show significant structuring for both estimates. The first consists of BAL, ILO and SCW, the second one includes the three AMUs from the Lauenburg area (LAW, LAS and LAE), and the third group comprising NFL and DK where the lowest level of differentiation was observed (F_{ST}: 0.015; Jost's D: 0.007). In some pairwise comparisons, Jost's D estimates differed from F_{ST} values, e.g. for the SCW-HAB value (Table 2.2).

*Table 2.2: Pairwise estimates of F_{ST} (above diagonal) and Jost's D values (below diagonal); * marks significant values ($p < 0.05$) after Bonferroni correction for multiple comparisons. Discrepancies between the two metrics are indicated in italics for Jost's D values*

	BAL	**DK**	**DUV**	**ELD**	**HAB**	**ILO**	**LAE**	**LAS**	**LAW**	**NFL**	**SAW**	**SCW**	**SEG**
BAL	-	0.17*	0.14*	0.13*	0.2*	0.01	0.13*	0.16*	0.14*	0.14*	0.16*	0.04	0.11*
DK	0.04	-	0.15*	0.08*	0.15*	0.15*	0.08*	0.09*	0.07*	0.01	0.12*	0.12*	0.12*
DUV	0.10*	0.22*	-	0.13*	0.18*	0.12*	0.09	0.09*	0.08*	0.15*	0.12*	0.12*	0.10*
ELD	0.13*	0.05*	0.16*	-	0.14*	0.10*	0.09*	0.10*	0.10*	0.08*	0.14*	0.10*	0.12*
HAB	0.17*	0.03*	0.19*	0.05*	-	0.15*	0.11*	0.14*	0.13*	0.14*	0.15*	0.13*	0.09*
ILO	0.01	0.15*	0.02	0.06*	0.15*	-	0.1*	0.12*	0.11*	0.11*	0.12*	0.02	0.09*
LAE	0.13*	0.1*	0.13*	0.1*	0.13*	0.05*	-	0.01*	0.01	0.08*	0.05*	0.09*	0.06*
LAS	0.23*	0.1*	0.12*	0.11*	0.16*	0.16*	0.02*	-	0.02	0.09*	0.05*	0.1*	0.09*
LAW	0.13*	0.08*	0.11*	0.11*	0.02*	0.14*	0.01	0.02	-	0.07*	0.04*	0.1*	0.07*
NFL	0.18*	0.01	0.04	0.05*	0.03	0.07*	0.04*	0.06*	0.05	-	0.11*	0.1*	0.11*
SAW	0.21*	0.07*	0.15*	0.04	0.11*	0.15*	0.06*	0.06*	0.04*	0.13*	-	0.11*	0.09*
SCW	0.02	0.08*	0.13*	0.10*	0.01	0.03	0.06*	0.09*	0.1*	0.11*	0.12*	-	0.09*
SEG	0.12*	0.03*	0.13*	0.03*	0.03*	0.04*	0.09*	0.12*	0.1*	0.13*	0.1*	0.08*	-

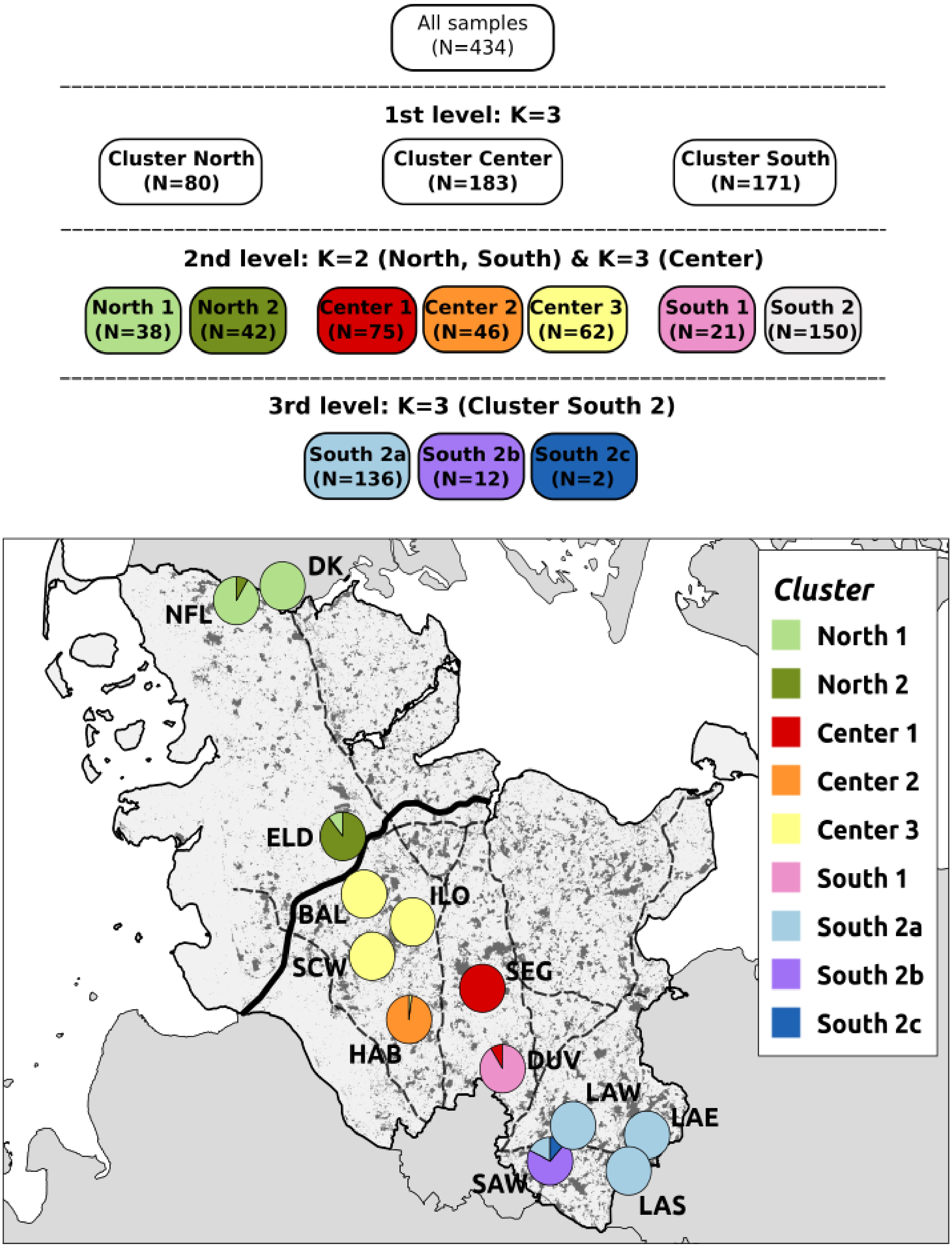


Figure 2.2: Results of hierarchical STRUCTURE analysis. Upper part shows partitioning among clusters. The map presents the final results for all MUs showing the proportions of the most likely origin of the sampled individuals. The overall sample size of this analysis is 434 out of 480 individuals. Samples from MWP (n = 46) were excluded since they originated from regions not directly neighboring the study area.

We observed a complex, hierarchical genetic structure of three different levels based on the *STRUCTURE* analysis. Using the ΔK method (Evanno et al. 2005) the optimal number of genetic clusters *K* at the first level was three, essentially dividing the individuals into a northern (Cluster North), central (Cluster Center) and southern (Cluster South) group of origin (Figure 3). The northern as well as the southern cluster was again split into another two subgroups whereas the central cluster comprised three different genetic groups at the second hierarchical level. Finally, we only found additional substructures of K=3 at the third level for one of the two southern clusters (South 2, Figure 3; see also supplemental file S2.4). The majority of individuals were clearly assigned to the different clusters with high ancestry values (q) above 0.7 (supplemental file S2.4).

Directional migration

Based on the *divMigrate* analysis, we observed variation in directionality and degree of gene flow among the AMUs and between some of them and the Danish reference area. Estimated rates of relative gene flow ranged from 0.04 up to 1 with an average of 0.15. A pairwise matrix with all directional estimates of gene flow is provided in the supplement (file S5). We observed the highest rates of directional gene flow (> 0.2) between AMUs in the southeastern region (LAW, LAE, and LAS) as well as the central region (BAL, ILO, SCW; Figure 3). The results further suggested that gene flow was more likely to occur from DK towards AMUs in the southern regions (e.g., the Lauenburg management units) than vice versa.

Differences with regard to directed migration rates were also detected by mean immigration and emigration rates (Table 2.3) with several AMUs either exhibiting similar rates of emigration (LAW, LAS, SAW) or very low values of overall gene flow (DUV, NFL). The HAB administrative deer management unit exhibited one of the lowest migration ratios ($R_{\mathrm{I/E}}$ =0.63), together with the reference area from Denmark ($R_{\mathrm{I/E}}$ =0.57). With an $R_{\mathrm{I/E}}$ value >1, five out of the 13 local deer populations potentially received more migrants than they produced (DUV, LAE, LAS, ILO, SEG)), while the eight remaining ones contributed more migrants than they received ($R_{\mathrm{I/E}}$ < 1:BAL, ELD, HAB, LAW, NFL, SAW, SCW and DK; Table 2.3).

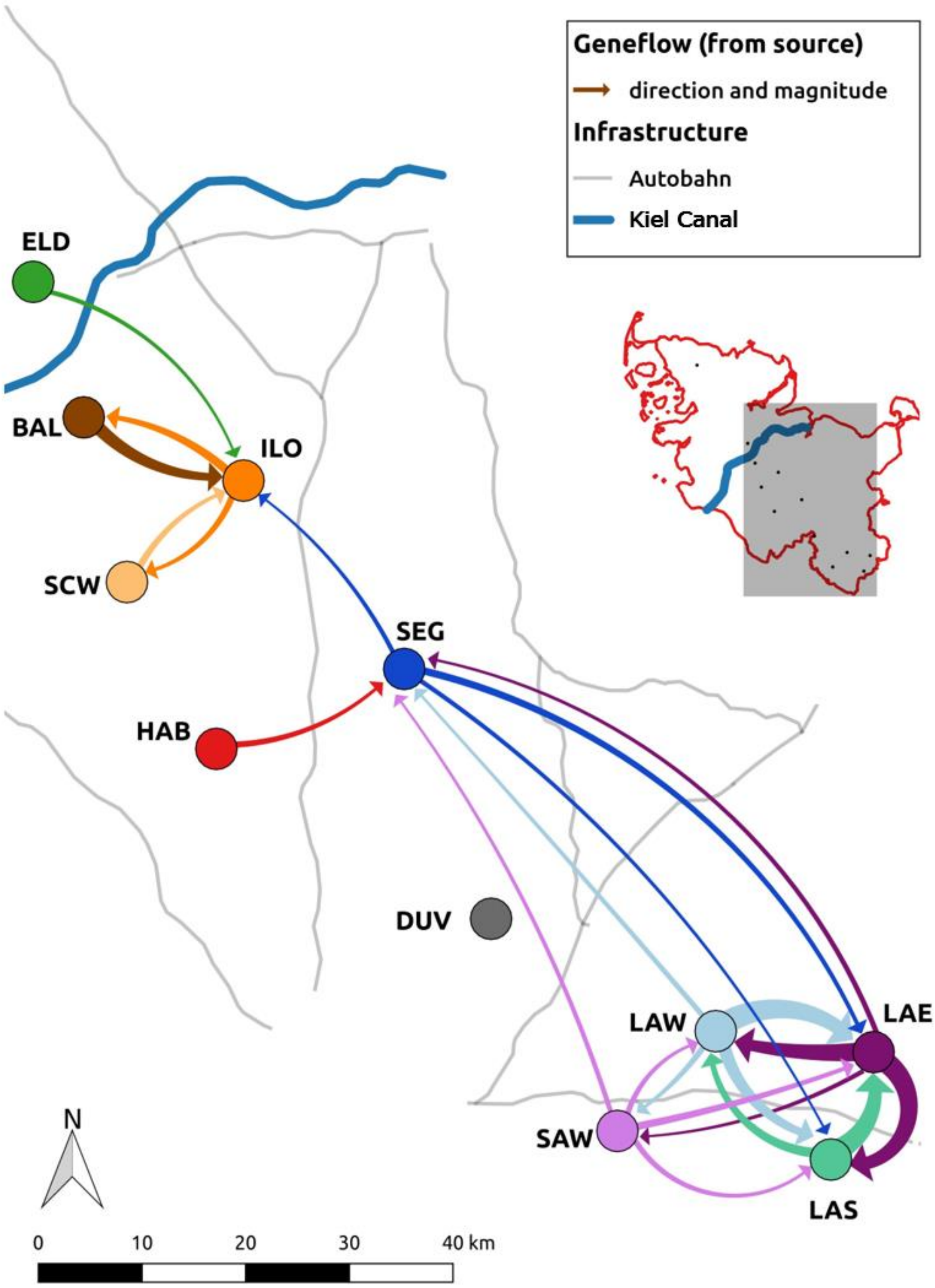


Figure 2.3: Direction and magnitude (indicated by arrow thickness) of estimated gene flow between management units (sources) based on the divMigrate analysis. Only results with migration rates above average (higher than 0.2) for the south-eastern region of Schleswig-Holstein without the DK reference area are shown here for illustrative purposes.

Table 2.3: Mean immigration (I) and emigration (E) rates as well as their ratio ($R_{I/E}$) estimated for all administrative deer management units in Schleswig-Holstein. The results summarize the pairwise estimates of directed gene flow between AMUs based on the divMigrate analysis. Values indicate whether a population is more likely to receive (migration ratio above one) or send out individuals (ratio below a value of one).

Management Unit	**Immigration rate (mean)**	**Emigration rate (mean)**	**migration ratio**
BAL	0.126	0.128	0.98
DUV	0.108	0.102	1.06
ELD	0.104	0.127	0.82
HAB	0.08	0.126	0.63
LAE	0.345	0.297	1.16
LAS	0.273	0.24	1.14
LAW	0.229	0.283	0.81
ILO	0.175	0.14	1.25
NFL	0.105	0.121	0.87
SAW	0.128	0.159	0.81
SCW	0.106	0.129	0.82
SEG	0.179	0.16	1.12
DK	0.09	0.157	0.57

Influence of population size and neighboring population densities

Our regression analysis (Table 2.4), allelic richness was best explained by the neighboring population densities (adj. $R^2 = 0.57$, $p = 0.003$), i.e., *AR* within AMUs increased with higher cumulative densities of red deer in the neighboring management units (Figure 5). Mean emigration rates (*E*) were best explained by both neighboring population size and by the density of the neighboring management units (Table 2.4). Neither mean Jost's *D* nor mean immigration rates (*I*) were explained by any of the indices.

Table 2.4: Meta-population study linking genetic metrics of diversity, differentiation and gene flow with estimates of meta-population structure. Only results for variables without null-model among candidates with ΔAICc smaller than two are shown.

Genetic Metric	Meta-Population Metric	Δ AICc	AICc Weight	adj. R^2	p-value
AR	neighbor pop.density	0.00	0.74	0.57	0.003
AR	neighbor pop.size	2.89	0.18	0.46	0.009
AR	focal pop.size	6.29	0.01	0.25	0.054
AR	nullmodel	7.73	0.02	0.00	-
AR	focal pop.density	7.97	0.03	0.17	0.099
JostD	nullmodel	0.00	0.58	0.00	-
JostD	focal pop.size	3.13	0.12	-0.05	0.527
mig.into	nullmodel	0.00	0.28	0.00	-
mig.into	neighbor pop.density	0.07	0.27	0.19	0.004
mig.from	neighbor pop.size	0.00	0.63	0.55	0.004
mig.from	Neighbor pop.density	1.40	0.31	0.49	0.007
mig.from	focal pop.size	6.11	0.03	0.25	0.058
mig.from	nullmodel	6.97	0.02	0.00	-
mig.from	focal pop.density	9.44	0.01	0.04	0.33

Genetic drift and isolation by distance

Genetic differentiation based on mean pairwise F_{ST} values was negatively correlated with higher estimates of genetic diversity. This indicates that drift is influencing genetic diversity and drives divergence between AMUs in our study area. For example, we observed the highest r^2 score of 0.73 ($p < 0.001$) between F_{ST} and expected heterozygosity (Figure 2.4). Allelic richness also significantly decreased with higher values of mean F_{ST} (r^2=0.58, p = 0.004). Further, we detected effects of spatially limited gene flow and significant isolation by distance. Results of the Mantel analyses indicated significant IBD among AMUs using both linearized F_{ST} (r = 0.42; p = 0.003) and Jost's *D* (r = 0.28; p = 0.015), respectively.

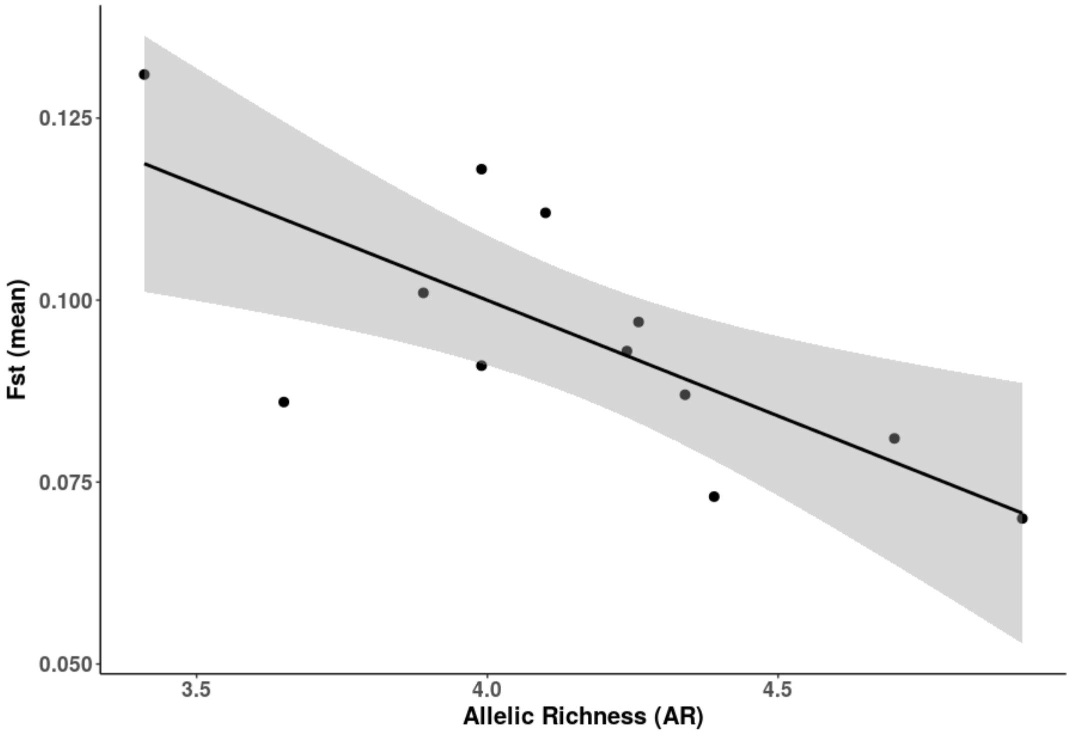


Figure 2.4: Scatterplot showing the significant decrease of mean pairwise F_{ST} values and genetic diversity of administrative deer management units in Schleswig-Holstein based on allelic richness (AR) Results are based on a linear regression model ($r^2=0.58$, $p = 0.004$).

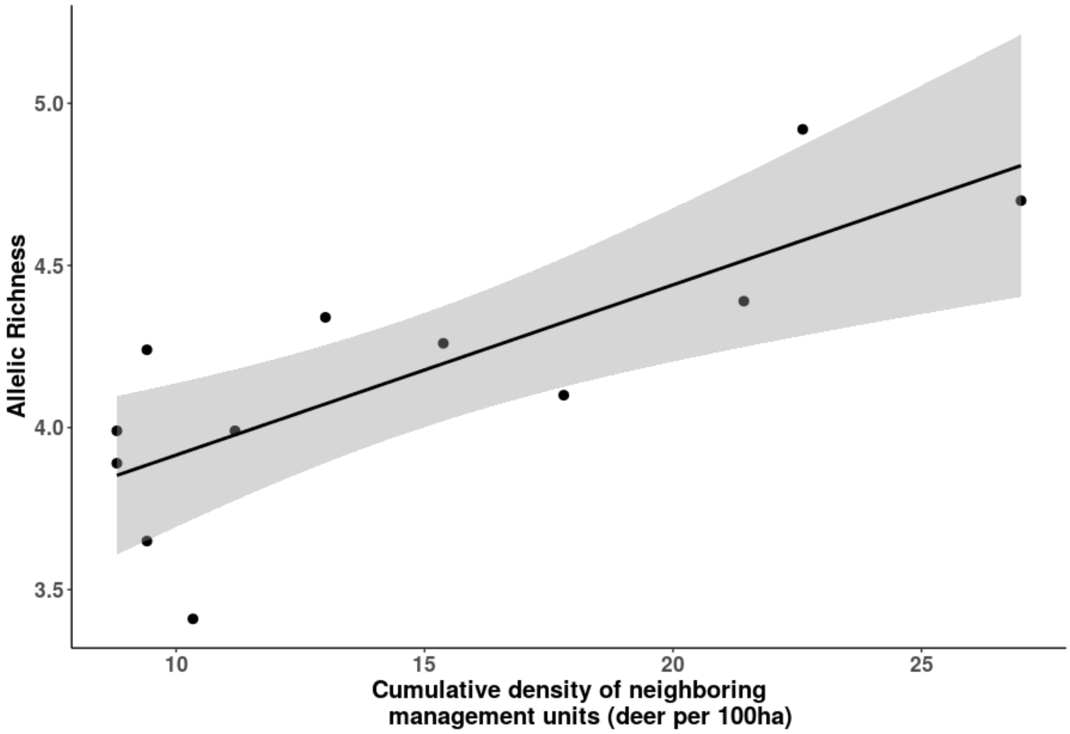


Figure 2.5: Linear regression model showing the significant increase (adj. $R^2 = 0.57$, $p = 0.003$) of allelic richness (AR) with higher cumulative densities of neighbouring deer management units (deer per 100 hectares).

Discussion

We investigated the genetic structure and differentiation of administrative management units to find out whether the practice of managing the local red deer populations as separate populations is effective or if future management should account for substructures and genetic exchange among them. This is particularly relevant when populations are low in abundance and experienced different types of restrictions in the past as in the presented study. Limited gene flow caused by anthropogenic fragmentation and management goals (culling of individuals outside of designated deer areas) lead to genetic drift and decreased genetic diversity.

We want to point out that comparisons of genetic diversity estimates across studies have to be performed with caution (Reiner et al. 2019). Although our data set shares only four loci with the most comprehensive microsatellite study of red deer in Europe to date (Zachos et al. 2016), our study nevertheless shows that there is a clear trend towards low genetic diversity in red deer from Northern Germany. The H_O values observed in our study are rare for red deer and usually only found in populations with long-term low effective population sizes such as the red deer from Sardinia or from Mesola in northern Italy (Hmwe and Zachos 2006). Both, $\mathrm{H_O}$ and $\mathrm{H_E}$ values, in the Hasselbusch AMU are among the lowest ever found in a population of this species (Zachos and Hartl 2011; Zachos et al. 2016). A very similar pattern can be seen in N_E values of all AMUs (Table 2.1). Although the N_E values for several of the northern German deer AMUs are within the range of reported values from other European populations, many of the Schleswig-Holstein populations, again including HAB, are clearly at the lower end and below the effective population size threshold of 50 individuals, a value below which inbreeding depression is likely to occur (Frankham et al. 2010).

However, observed F_{IS} values were quite low (Table 2.1) with no clear signs of heterozygote deficiency and fixation. We assume that existing gene flow at short ranges seems to compensate for drift effects on genetic diversity in some cases. This assumption is supported by significant isolation by distance which indicates that drift and gene flow are in equilibrium at regional scales (Hutchison and Templeton 1999, Jordan and Snell 2008).

Ultimately, the lack of such restrictions to dispersal as experienced by red deer in our study area should result in higher levels of genetic diversity. This was confirmed by relatively high values of genetic variability (H_E, N_E) in the reference population from MWP where red deer have not been restricted to declared red deer zones in the past and range throughout the state in higher abundance compared to Schleswig-Holstein (Kinser et al. 2010). Comparable values of genetic diversity were also confirmed for the three AMUs of the Lauenburg area (LAE, LAS, LAW) which could also be explained by the relatively large population size in Lauenburg in combination with gene flow from the east (Mecklenburg-Western Pomerania).

Furthermore, we observed a significant decrease in differentiation (F_{ST}) with higher levels of genetic diversity (H_E, *AR*). This could be due to potential effects of historic drift in isolation and small populations which is expected to be the predominant cause for genetic differentiation (Jordan and Snell 2008; Whiteley et al. 2010; Funk et al. 2016). Only AMUs in close proximity did not exhibit any significant values of differentiation based on both F_{ST} and Jost's D estimates (e.g., BAL, ILO and SCW or the Lauenburg populations).

Hierarchical structure and gene flow

We observed a hierarchical genetic structure comprising three main clusters: North, Center and South. The first cluster was located north of the Kiel Canal (an effective barrier to deer dispersal due to the steep embankments) and also included the Danish red deer. The small population of the NFL management unit was founded by red deer individuals dispersing from Denmark into northern Germany. The assignment of ELD to the northern cluster was surprising because its founders came from BAL crossing the Kiel Canal in the late 1960s (when the embankments were not yet in their present state; Meißner et al. 2008). However, low population size and genetic drift have apparently resulted in divergence from the central cluster, which is located just south of the canal. In addition, immigration from Denmark into the ELD population has been shown by means of genetic data (mtDNA sequences; Reinecke et al. 2013). Therefore, it seems in accordance with the population's history to consider ELD as a separate northern sub-cluster.

The sub-structure of the central cluster can also be explained based on geography and historic background. The three sample sites just south of the Kiel Canal (BAL, ILO, SCW) have historically been separated from the ones further south because of limited dispersal of red deer outside the ‘designated red deer zones’ (Wotschikowsky 2010, Reinecke et al. 2013). The HAB population was founded by dispersed individuals from SEG but a fenced highway has prevented any potential migration between these two AMUs. Due to its low census and effective sizes, drift has been high in HAB, which is mirrored by its substantial differentiation from SEG today. The separate status of Duvenstedt (DUV) is understandable since the population is not native but was founded with red deer from other parts of Europe (Jessen 1988; Meißner et al. 2008). In the southeast, LAS is separated from LAW and LAE by a highway, but this is a relatively recent barrier (completion during the 1990s), obviously not yet reflected in the gene pools on either side. Interestingly, the SAW population comprises three different subclusters, two of which were only found there (South 2b and 2c). A red deer hunting enclosure located in that area perhaps suggests a similar historical development as with the DUV population. Individuals from other parts of Europe introduced into the private enclosure could potentially have escaped the fenced area and established themselves within the local population (cf. Frantz et al. 2017).

As expected, we observed high levels of gene flow between AMUs with low differentiation (Figure 3), which were again the complex consisting of BAL, SCW, and ILO, as well as the local deer populations from the Lauenburg area in the south-eastern region of Schleswig-Holstein.

Overall, diversity within and gene flow among AMUs was best explained by size and density of the surrounding local populations as our modeling analyses (Table 2.4) showed. Deer populations that were adjacent to larger or higher-density populations had higher rates of gene flow and higher levels of diversity.

Most of the AMUs with lower mean rates of immigration as compared to emigration rates (Table 2.3) are characterized by either small population sizes, low densities or higher levels of isolation. Because of that they probably received fewer genes from other populations in the past. Relative to the other populations they are therefore more likely to send out individuals (Bohling et al. 2019). Genetic similarity, for example due to historical reasons, will also lead to positive values of inferred migration. Since the

HAB population was founded by migrants from SEG, migration values are not zero and reflect population history. This is in accordance with the Duvenstedt unit (DUV) showing no signs of migration to or from other populations because it was founded with non-native deer (Jessen 1988). Still, anecdotal reports of dispersing red deer further support the conclusion that there is some level of gene flow (Reinecke et al. 2013). Single individuals have been seen outside established population ranges, the Lauenburg red deer are known to be in contact with the neighboring populations in Mecklenburg-Western Pomerania to the east, and in 1986 and 1987 single stags migrated from Hasselbusch (HAB) to Barlohe (BAL) and from Duvenstedt (DUV) to Segeberg (Jessen 1988; Peters 2000; Zachos et al. 2007; Meißner et al. 2008). The latter is also supported by the results of the STRUCTURE analysis. Whether they successfully reproduced in SEG, however, is unknown.

Within the last decade red deer from Denmark have established themselves south of the German border and are increasing in numbers. We were able to detect first signs of genetic exchange between the NFL / DK population and the ELD management unit. This shows the high potential of the species to migrate throughout the state and establish new ranges.

Management Implications and Future Research

In summary, based on our analyses on genetic structure and gene flow we were able to distinguish two major groups of AMUs which essentially represent single GMUs: in the central part of Schleswig-Holstein the three AMUs of BAL, ILO and SCW form one genetically distinct cluster. The same holds for the AMUs in the Lauenburg area (LAE, LAS, LAW) in the south-east of the state. This indicates a discrepancy between the current administrative delineation of management units and actual levels of genetic exchange among these areas (see also Figure 2.3). Our results also show that observed genetic patterns (diversity and gene flow) in a local deer population are largely explained by the densities of populations in its close vicinity. Local management decisions that change local abundance could have genetic impacts not only on the local population but also on neighboring AMUs, especially if AMUs are interpreted as single GMUs when they are actually well connected to others. Therefore, future management of red deer populations in Schleswig-Holstein needs to incorporate parameters such as deer population sizes and habitat availability for neighboring administrative MUs. Data on dispersal or gene flow and population structure derived from genetic studies like ours should ideally be incorporated when new units for wildlife management are spatially delineated (Paetkau 1999; Lowe and Allendorf 2010).

Another important factor are temporal changes regarding age- and sex-structure of the local populations. Recording these parameters could help to gain a better understanding of potential source-sink dynamics (Draheim et al. 2016). In particular, younger males are more likely to disperse at higher local densities (Loe et al. 2009). Therefore, future research should also focus on the proportions of young males in local populations and how density-dependent dispersal could potentially influence gene flow and the genetic differentiation of the subpopulations. For example, estimating dispersal among localities using capture-mark-recapture or telemetry could be applied to assess the demographic effects of inter-population movements.

In particular, the exchange of individuals between isolated populations such as HAB needs to be enhanced in the near future to counteract the continuing loss of genetic diversity. HAB is not far away (approximately 10km) from the larger GMU formed by SCW, ILO and BAL. Still, we observe high levels of differentiation and hardly any gene flow. The STRUCTURE analysis assigned one single individual sampled in HAB to the cluster of SCW, ILO and BAL (Figure 2.2 .and Supplement S2.4). Similar patterns can be observed for DUV and SEG which are also not far apart (ca. 15km) but only two individuals sampled in DUV were assigned to the SEG cluster (Supplement S2.4). This leads to the conclusion that landscape characteristics between AMUs affect the genetic exchange among them and thus influence size and density of populations; we will need further analyses to identify landscape features that facilitate or impede natural dispersal among AMUs. Based on the results, migration corridors and locations for crossing-structures (e.g., green bridges) can then be identified to mitigate the effects of barriers and landscape resistance on the migratory movements of red deer.

Acknowledgements

We thank Tanja Noventa for her technical assistance with the genetic analyses and lab work. Further, we acknowledge the great support by the employees of the state forest service and hunters from Northern Germany and Denmark for providing the genetic samples of harvested animals. H. Edelhoff was funded by the scholarship program of the German Federal Environmental Foundation (Deutsche Bundesstiftung Umwelt, DBU).

Some of the genetic samples were originally obtained during the pilot study “Sicherung genetischer Diversität beim Rothirsch in der Kulturlandschaft” financed by the Federal Office for Agriculture and Food (Bundesanstalt für Landwirtschaft und Ernährung, BLE; project number: 07BM010).

Data Accessibility

Genetic data of individuals used throughout this study as well as information on the used samples are provided in DRYAD:

https://datadryad.org/stash/share/9HKXnc5CNkOFBfZt3kkq86KB0y1NSVmUYlctvsWIo1w

References

Aguillon, S. M., Fitzpatrick, J. W., Bowman, R., Schoech, S. J., Clark, A. G., Coop, G., and Chen, N. (2017). Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLOS Genetics*, *13*(8), e1006911.

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, *60*(2), 255–265.

Archer, F. I., Adams, P. E., and Schneiders, B. B. (2017). strataG: An R package for manipulating, summarizing, and analyzing population genetic data. *Molecular Ecology Resources*, *17*(1), 5–11.

Balkenhol, N., Holbrook, J. D., Onorato, D., Zager, P., White, C., and Waits, L. P. (2014). A multi-method approach for analyzing hierarchical genetic structures: A case study with cougars Puma concolor. *Ecography, 37*(6), 552–563.

Balkenhol, N., Pardini, R., Cornelius, C., Fernandes, F., and Sommer, S. (2013). Landscape-level comparison of genetic diversity and differentiation in a small mammal inhabiting different fragmented landscapes of the Brazilian Atlantic Forest. *Conservation Genetics*, *14*(2), 355–367.

Beerli, P. (2004). *MIGRATE: documentation and program, part of LAMARC. Version 2.0.*

Burnham, K., and Anderson, D. (2002). *Model selection and multimodel inference: a practical information-theoretic approach* (2nd ed.).

Coulon, A., Fitzpatrick, J. W., Bowman, R., Stith, B. M., Makarewich, C. A., Stenzler, L. M., and Lovette, I. J. (2008). Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (Aphelocoma coerulescens). *Molecular Ecology*, *17*(7), 1685–1701.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., and Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214.

Draheim, H. M., Moore, J. A., Etter, D., Winterstein, S. R., and Scribner, K. T. (2016). Detecting black bear source – sink dynamics using individual-based genetic graphs. *Proceedings of the Royal Society B, 283,* 2016002.

Earl, D. A., and vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*(2), 359–361.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology, 14*(8), 2611–2620.

Excoffier, L., Laval, G., and Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics, 1,* 47-50.

Fickel, J., Bubliy, O. A., Stache, A., Noventa, T., Jirsa, A., and Heurich, M. (2012). Crossing the border? Structure of the red deer (Cervus elaphus) population from the Bavarian-Bohemian forest ecosystem. *Mammalian Biology, 77*(3), 211–220.

Frankham, R., Briscoe, D., and Ballou, J. (2002). *Introduction to conservation genetics*. Cambridge university press.

Frantz, A. C., Bertouille, S., Eloy, M. C., Licoppe, A., Chaumont, F., and Flamand, M. C. (2012). Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (Cervus elaphus), but not wild boars (Sus scrofa). *Molecular Ecology, 21*(14), 3445–3457.

Frantz, A. C., Zachos, F. E., Bertouille, S., Eloy, M. C., Colyn, M., and Flamand, M. C. (2017). Using genetic tools to estimate the prevalence of non-native red deer (Cervus elaphus) in a Western European population. *Ecology and Evolution*, *7*(19), 7650–7660.

Funk, W. C., Lovich, R. E., Hohenlohe, P. A., Hofman, C. A., Morrison, S. A., Sillett, T. S., Ghalambor C. K., Maldonado J. E., Rick T. C., Day M. D., Polato, N. R., Fitzpatrick, S. W., Coonan, T. J., Crooks, K. R., Dillon, A., Garcelon, D.K., King, J. L., Boser, C. L., Gould, N., and Andelt, W. F. (2016). Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology, 25*(10), 2176–2194.

Gaillard, D., Ennen, J. R., Kreiser, B. R., Qualls, C. P., Sweat, S. C., Birkhead, R., Tuberville, T. D., Aresco, M., McCoy, E. D., Mushinsky, H. R., and Hentges, T. W. (2017). Range-Wide and Regional Patterns of Population Structure and Genetic Diversity in the Gopher Tortoise. *Journal of Fish and Wildlife Management, 8*(2), 497–512.

Grosser, S., Abdelkrim, J., Wing, J., Robertson, B. C., and Gemmell, N. J. (2017). Strong isolation by distance argues for separate population management of endangered blue duck (Hymenolaimus malacorhynchos). *Conservation Genetics, 18*(2), 327–341.

Haanes, H., Røed, K. H., Mysterud, A., Langvatn, R., and Rosef, O. (2010). Consequences for genetic diversity and population performance of introducing continental red deer into the northern distribution range. *Conservation Genetics, 11*(5), 1653–1665.

Hartl, G. B., Zachos, F., and Nadlinger, K. (2003). Genetic diversity in European red deer (Cervus elaphus L.): anthropogenic influences on natural populations. *Comptes Rendus Biologies*, *326*, 37–42.

Hemami, M.-R., Watkinson, A. R., and Dolman, P. M. (2005). Population densities and habitat associations of introduced muntjac Muntiacus reevesi and native roe deer Capreolus capreolus in a lowland pine forest. *Forest Ecology and Management*, *215*, 224–238.

Hmwe, S., and Zachos, F. (2006). Conservation genetics of the endangered red deer from Sardinia and Mesola with further remarks on the phylogeography of Cervus elaphus corsicanus. *Biological Journal of the Linnean Society*, *88*, 691–701

Hubisz, M. J., Falush, D., Stephens, M., and Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, *9*(5), 1322–1332.

Hutchison, D. W., and Templeton, A. R. (1999). Correlation of Pairwise Genetic and Geographic Distance Measures: Inferring the Relative Influences of Gene Flow and Drift on the Distribution of Genetic Variability. *Evolution*, *53*(6), 1898.

Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., and Andrew, R. L. (2017). The K = 2 conundrum. *Molecular Ecology*, *26*(14), 3594–3602.

Jessen, H. (1988). *Wild und Jagd in Schleswig-Holstein*.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405.

Jordan, M. A., and Snell, H. L. (2008). Historical fragmentation of islands and genetic drift in populations of Galápagos lava lizards (Microlophus albemarlensis complex). *Molecular Ecology*, *17*(5), 1224–1237.

Jost, L. (2008). GST and its relatives do not measure differentiation. *Molecular Ecology*, *17*(18), 4015–4026.

Jost, L., Archer, F., Flanagan, S., Gaggiotti, O., Hoban, S., and Latch, E. (2018). Differentiation measures for conservation genetics. *Evolutionary Applications*, *11*(7), 1139–1148.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., and Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*, 782–788.

Kinser, A., Koop, K., and von Münchhausen, H. (2010). Die Rotwildverbreitung in Deutschland. *Allgemeine Forstzeitung*, *5*, 32–34.

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., and Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 1179–1191. doi: 10.1111/1755-0998.12387

Krojerova-Prokesova, J., Barančeková, M., and Koubek, P. (2015). Admixture of eastern and western European red deer lineages as a result of postglacial recolonization of the Czech Republic (Central Europe). *Journal of Heredity*, *106*(4), 375–385.

Kuehn, R., Schroeder, W., Pirchner, F., and Rottmann, O. (2003). Genetic diversity, gene flow and drift in Bavarian red deer populations (Cervus elaphus). *Conservation Genetics*, *4*(2), 157–166.

Loe, L. E., Mysterud, A., Veiberg, V., and Langvatn, R. (2009). Negative density-dependent emigration of males in an increasing red deer population. *Proceedings. Biological Sciences / The Royal Society*, *276*(1667), 2581–2587.

Lowe, W. H., and Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, *19*(15), 3038–3051.

Marrotte, R. R., Gonzalez, A., and Millien, V. (2017). Functional connectivity of the white-footed mouse in Southern Quebec, Canada. *Landscape Ecology*, *32*(10), 1987–1998.

Meißner, M., Reinecke, H., Zachos, F., Corsmann, M., Ritter, T., and Wölfel, H. (2008). *Der Rothirsch in Schleswig-Holstein*. Göttingen.

Ministerium für Energiewende Landwirtschaft Umwelt und ländliche Räume des Landes. (2012). *Jagd und Artenschutz - Jahresbericht 2012*.

Moritz, C. (1994). Defining ‘Evolutionarily Significant Units’ for conservation. *Trends in Ecology and Evolution*, *9*(10), 373–375.

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*(949), 283–292.

Niedziałkowska, M., Jędrzejewska, B., Wójcik, J. M., and Goodman, S. J. (2012). Genetic structure of red deer population in northeastern poland in relation to the history of human interventions. *Journal of Wildlife Management*, *76*(6), 1264–1276.

Olea, P. P., and Mateo-Tomás, P. (2014). Living in risky landscapes: Delineating management units in multithreat environments for effective species conservation. *Journal of Applied Ecology*, *51*(1), 42–52.

Paetkau, D. (1999). Using genetics to identify intraspecific conservation units: A critique of current methods. *Conservation Biology*, 13, 1507–1509.

Palsbøll, P. J., Bérubé, M., and Allendorf, F. W. (2007). Identification of management units using population genetic data. *Trends in Ecology and Evolution*, *22*(1), 11–16.

Pannell, J. R., and Charlesworth, B. (2000). Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *355*(1404), 1851–1864.

Pérez-Espona, S, Pérez-Barbería, F. J., McLeod, J. E., Jiggins, C. D., Gordon, I. J., and Pemberton, J. M. (2008). Landscape features affect gene flow of Scottish Highland red deer (Cervus elaphus). *Molecular Ecology*, *17*(4), 981–996.

Pérez-Espona, Sílvia, Pemberton, J. M., and Putman, R. (2009). Red and sika deer in the British Isles, current management issues and management policy. *Mammalian Biology - Zeitschrift Für Säugetierkunde*, *74*(4), 247–262.

Peters, S. (2000). *Verbreitung der Schalenwildarten in Schleswig-Holstein*.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

R Core Team. (2017). *R: A language and environment for statistical computing*.

Rannala, B. (2007). *BayesAss edition 3.0 user’s manual*.

Reinecke, H., Meißner, M., Zachos, F., and Herzog, S. (2013). *Sicherung genetischer Diversität beim Rothirsch in der Kulturlandschaft.*

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution*, *43*(1), 223–225.

Robinson, H. S., Wielgus, R. B., Cooley, H. S., and Cooley, S. W. (2008). Sink populations in carnivore management: cougar demography and immigration in a hunted population. *Ecological Applications*, *18*(4), 1028–1037.

Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*(1), 264–279.

Statistisches Bundesamt. (2018). Fläche und Bevölkerung nach Ländern.

Stillfried, M., Fickel, J., Börner, K., Wittstatt, U., Heddergott, M., Ortmann, S., Kramer-Schadt, S., and Frantz, A. C. (2017). Do cities represent sources, sinks or isolated islands for urban wild boar population structure? *Journal of Applied Ecology*, *54*(1), 272–281.

Ströhlein, H., Herzog, S., Hecht, W., and Herzog, A. (1993). Biochemical genetic description of German and Swiss populations of red deer Cervus elaphus. *Acta Theriologica*, *38*, 153–161.

Sundqvist, L., Keenan, K., Zackrisson, M., Prodöhl, P., and Kleinhans, D. (2016). Directional genetic differentiation and asymmetric migration. *Ecology and Evolution*, *6*(11), 3461– 3475.

Taylor, B. L., and Dizon, A. E. (1999). First policy then science: why a management unit based solely on genetic criteria cannot work. *Molecular Ecology*, *8*(1), 11–16.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535–538.

Wäber, K., Spencer, J., and Dolman, P. M. (2013). Achieving landscape-scale deer management for biodiversity conservation: The need to consider sources and sinks. *Journal of Wildlife Management*, *77*(4), 726–736.

Waples, R. S., and Do, C. (2008). LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, *8*(4), 753–756.

Whiteley, A. R., Hastings, K., Wenburg, J. K., Frissell, C. A., Martin, J. C., and Allendorf, F. W. (2010). Genetic variation and effective population size in isolated populations of coastal cutthroat trout. *Conservation Genetics*, *11*(5), 1929–1943.

Whitlock, M. C. (2011). G'ST and D do not replace FST. *Molecular Ecology*, *20*(6), 1083–1091.

Willems, H., Welte, J., Hecht, W., and Reiner, G. (2016). Temporal variation of the genetic diversity of a German red deer population between 1960 and 2012. *European Journal of Wildlife Research*, *62*(3), 277-284.

Wilting, A., Courtiol, A., Christiansen, P., Niedballa, J., Scharf, A. K., Orlando, L., Balkenhol, N., Hofer, H., Kramer-Schadt, S., Fickel, J., and Kitchener, A. C. (2015). Planning tiger recovery: Understanding intraspecific variation for effective conservation. *Science Advances*, *1*, e1400175.

Wotschikowsky, U. (2004). *Ein Leitbild für das Rotwild- Management in Deutschland*.

Wotschikowsky, U. (2010). Ungulates and their management in Germany. In *European Ungulates and Their Management in the 21st Century* (pp. 201–222).

Wright, S. (1943). Isolation by distance. *Genetics*, *28*, 114-138.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*, *19*(3), 395–420.

Zachos, F. E., Althoff, C., Steynitz, Y. V., Eckert, I., and Hartl, G. B. (2007). Genetic analysis of an isolated red deer (Cervus elaphus) population showing signs of inbreeding depression. *European Journal of Wildlife Research*, *53*(1), 61–67.

Zachos, Frank E., and Hartl, G. B. (2011). Phylogeography, population genetics and conservation of the European red deer Cervus elaphus. *Mammal Review*, *41*(2), 138–150

Zachos, Frank E, Frantz, A. C., Kuehn, R., Bertouille, S., Colyn, M., Niedzialkowska, M., Pérez-González, J., Skog, A., Sprĕm, N., and Flamand, M.-C. (2016). Genetic structure and effective population sizes in European red deer (Cervus elaphus) at a continental scale: insights from microsatellite DNA. *Journal of Heredity*, *107*(4), 318-326.

Supplement

S2.1: information on microsatellite markers

- *INRA6* (Vaiman et al. 1994; Slate et al. 1998),
- *C105*, *C143*, *C180*, *C229, T40, T107, T115* (all tetranucleotide loci; Meredith et al. 2005),
- *HAUT14* (Thieven et al. 1995),
- *BM757, BM1818, ILSTS06* (Bishop et al. 1994),
- *CSSM14* (Moore et al. 1994),
- and *FSHB* (Moore et al. 1992).

Four of these loci were also included in the 13 loci of (Zachos et al. 2016): *Haut 14, ILSTS06, CSSM14 and BM1818.*

References

Bishop, M. D., Kappes, S. M., Keele, J. W., Stone, R. T., Sunden, S. L., Hawkins, G. A., Toldo, S. S., Fries, R., Grosz, M. D., and Yoo, J. (1994). A genetic linkage map for cattle. Genetics, 136(2), 619-639.

Meredith, E. P., Rodzen, J. A., Levine, K. F., and Banks, J. D. (2008). Characterization of an additional 14 microsatellite loci in California Elk (Cervus elaphus) for use in forensic and population applications. *Conservation Genetics*, *9*(6), 1715-1715.

Moore, S. S., Barendse, W., Berger, K. T., Armitage, S. M., and Hetzel, D. J. S. (1992). Bovine and ovine DNA microsatellites from the EMBL and GENBANK databases. *Animal Genetics*, *23*(5), 463-467.

Moore, S. S., Byrne, K., Berger, K. T., Barendse, W., McCarthy, F., Womack, J. E., & Hetzel, D. J. S. (1994). Characterization of 65 bovine microsatellites. *Mammalian Genome*, *5*(2), 84-90.

Slate, J., Coltman, D. W., Goodman, S. J., MacLean, I., Pemberton, J. M., and Williams, J. L. (1998). Bovine microsatellite loci are highly conserved in red deer (Cervus elaphus), sika deer (Cervus nippon) and Soay sheep (Ovis aries). *Animal Genetics*, *29*(4), 307-315.

Thieven, U., Harlizius, B., and Simon, D. (1995). Dinucleotide repeat polymorphism at the bovine HAUT1 and HAUT14 loci. *Animal Genetics*, *26*(2), 123–123.

Vaiman, D., Mercier, D., Moazami-Goudarzi, K., Eggen, A., Ciampolini, R., Lépingle, A., Martin, P., Levéziel, H., and Guérin, G. (1994). A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. *Mammalian Genome*, *5*(5), 288-297.

Zachos, Frank E, Frantz, A. C., Kuehn, R., Bertouille, S., Colyn, M., Niedzialkowska, M., Pérez-González, J., Skog, A., Sprĕm, N., and Flamand, M.-C. (2016). Genetic structure and effective population sizes in European red deer (Cervus elaphus) at a continental scale: insights from microsatellite DNA. *Journal of Heredity*, *107*(4), 318-326.

S2.2: marker diversity

Locus	Nr.Alleles	H.obs	H.exp	Ae	Fis	PIC
BM1818	8	0.758	0.825	5.718	0.082	0.803
BM757	14	0.687	0.752	4.037	0.087	0.725
C105	2	0.41	0.423	1.734	0.032	0.334
C143	6	0.396	0.464	1.867	0.147	0.437
C180	5	0.526	0.556	2.254	0.055	0.493
C229	6	0.275	0.333	1.5	0.175	0.302
CSSM14	3	0.409	0.446	1.804	0.083	0.365
FSHB	10	0.742	0.847	6.549	0.125	0.83
Haut14	11	0.672	0.855	6.894	0.214	0.839
ILSTS06	12	0.712	0.808	5.196	0.118	0.785
INRA6	5	0.536	0.621	2.636	0.136	0.574
T107	4	0.553	0.551	2.229	-0.002	0.493
T115	12	0.785	0.869	7.606	0.096	0.855
T40	2	0.033	0.041	1.042	0.183	0.04

S2.3: distribution of samples across red deer management units and sampling periods

Management Unit	Period 2003-2004	Period 2013-2015
Barlohe (BAL)	2	17
Denmark (DK)	22	0
Duvenstedter Brook (DUV)	13	10
Elsdorf (ELD)	21	25
Hasselbusch (HAB)	15	32
Lauenburg East (LAE)	46	30
Lauenburg South (LAS)	13	24
Lauenburg West (LAW)	0	22
Mecklenburg-Western Pomerania (MWP)	46	0
Moerel/Iloo (ILO)	10	21
Nordfriesland (NFL)	4	8
Sachsenwald (SAW)	15	2
Schierenwald (SCW)	0	14
Segeberger Heide (SEG)	9	64

S2.4: Results of hierarchical STRUCTURE analysis with locprior

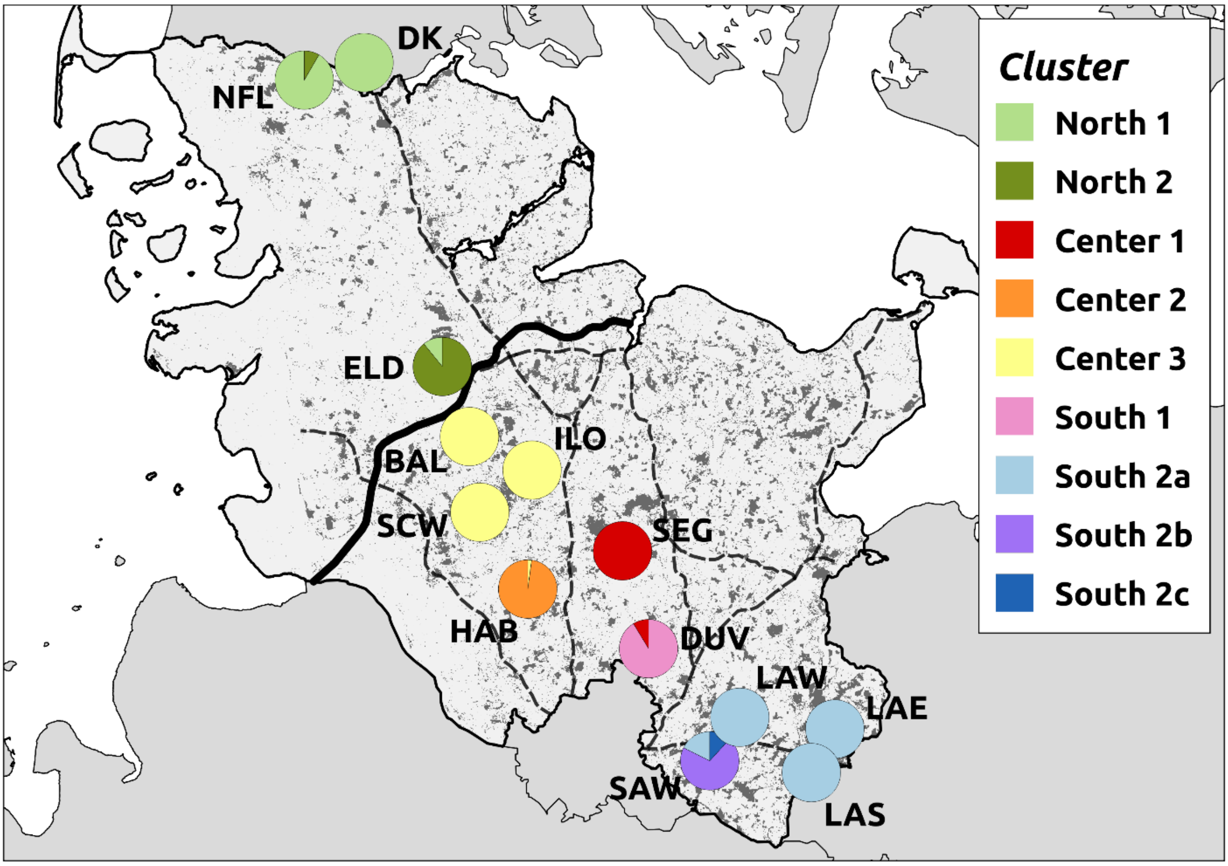


Figure S.2.4.1: results of hierarchical STRUCTURE analysis using sampling location (administrative management unit) as locprior.

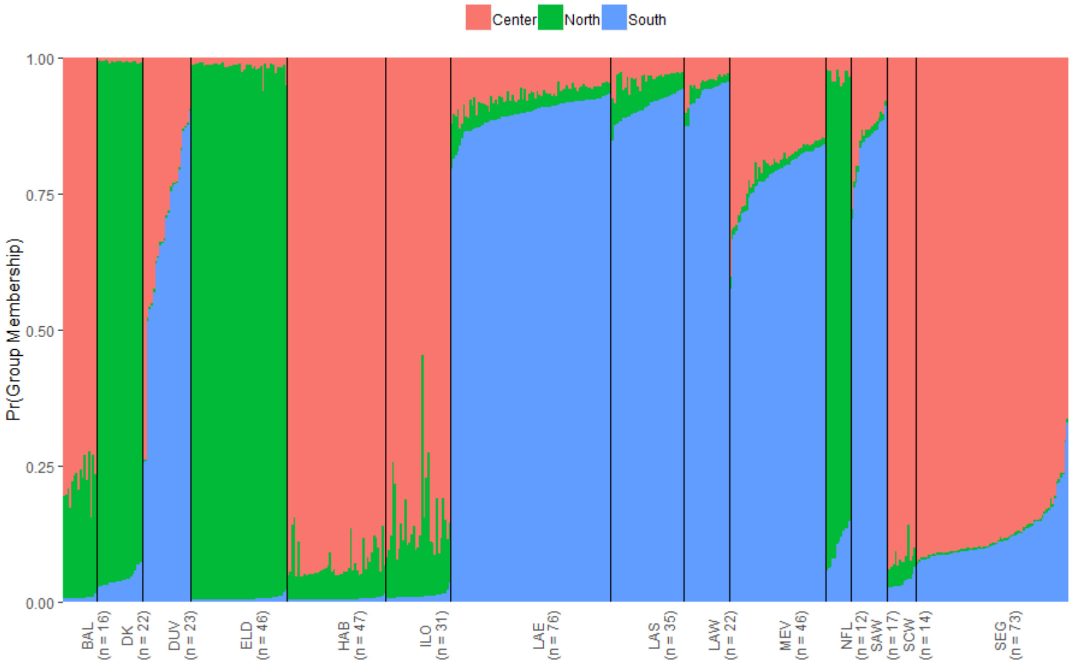


Figure S.2.4.2: First level of the hierarchical STRUCTURE analysis: the number of genetic clusters (K=3) *was determined with the Evanno method. Probabilities of group membership (Q-values) are presented for all individuals from the AMUs in Schleswig-Holstein and the reference area from Denmark (DK).*

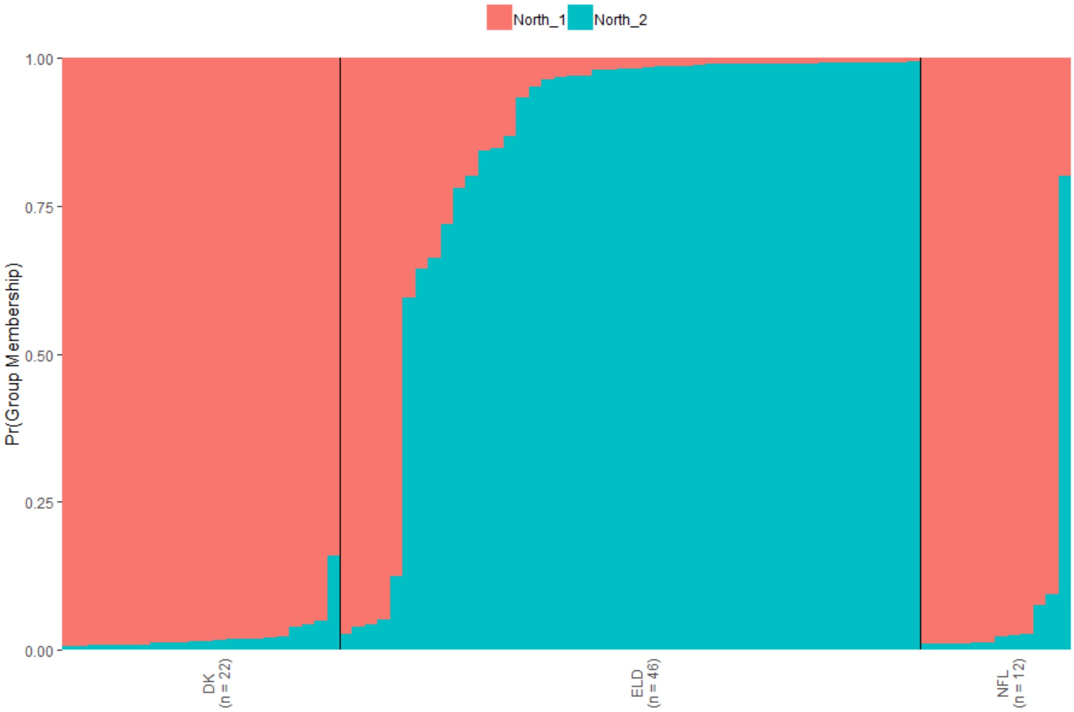


Figure S2.4.3: Results for the northern cluster at the second level of the hierarchical STRUCTURE analysis. Probabilities of cluster memberships are shown for the two AMUs from Schleswig-Holstein located north of the Kiel Canal and the reference area from Froslev (Denmark, DK).

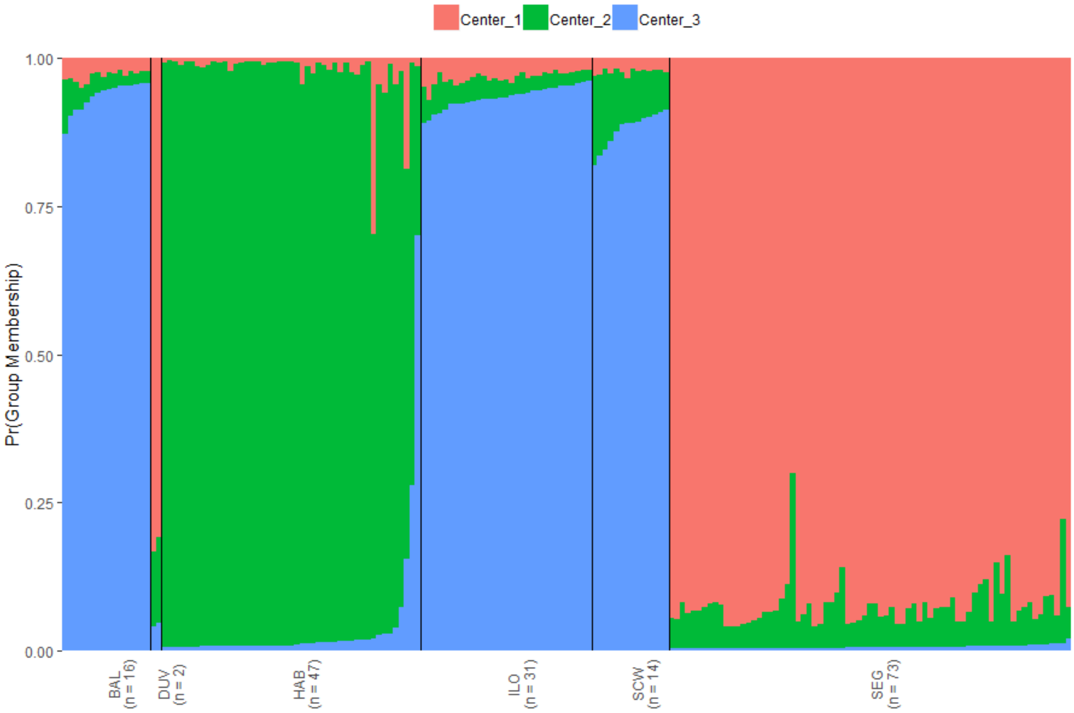


Figure S2.4.4: Results for the center cluster at the second level of the hierarchical STRUCTURE analysis. The Evanno method indicated the most likely value of K=3

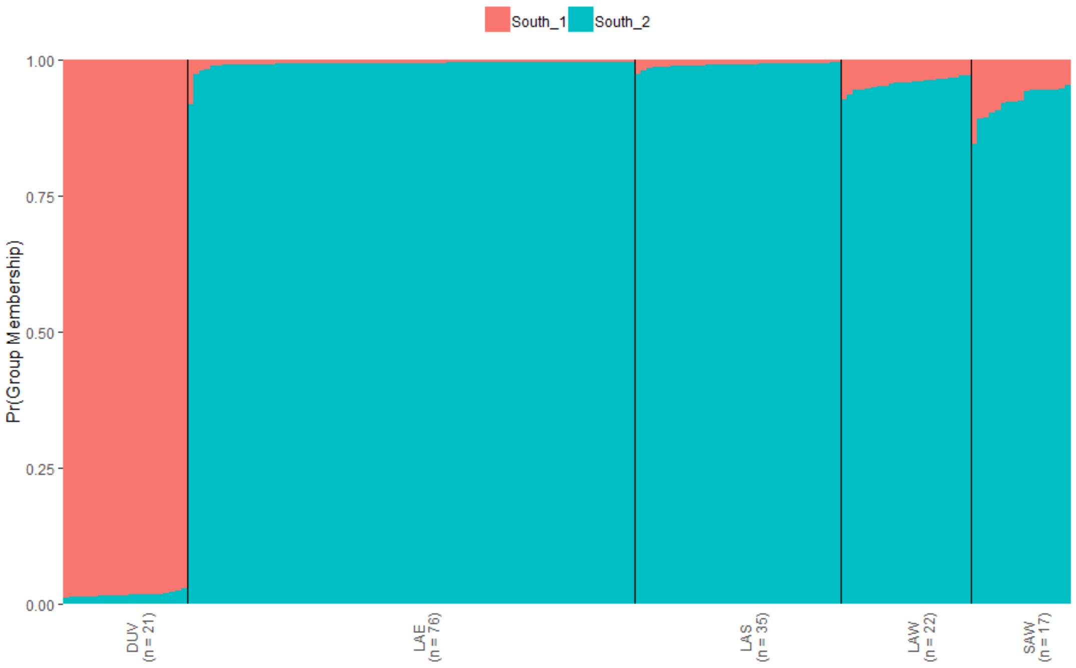


Figure S2.4.5: Results for the southern cluster at the second level of the hierarchical STRUCTURE analysis.

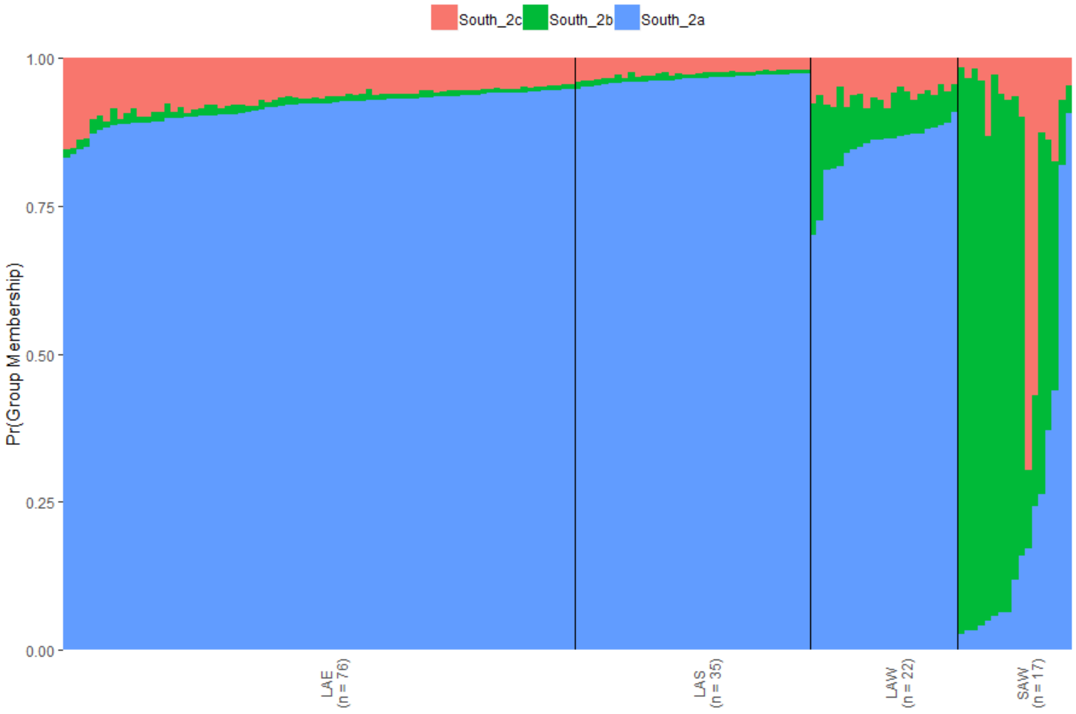


Figure S2.4.6: Results for the third level of our hierarchical STRUCTURE analysis. The second southern cluster was divided into another three different clusters (K=3, Evanno method).

S2.5: pairwise relative migration rates between administrative deer management units

Results of the *divmigrate* analysis: pairwise estimates of relative migration rates based on G_{ST} values between all AMUs in Schleswig-Holstein. Upper half represents migration rates into the AMU (immigration) whereas the lower half shows emigration values for each AMU.

	BAL	DUV	ELD	HAB	ILO	LAE	LAS	LAW	NFL	SAW	SCW	SEG
BAL		0.08	0.09	0.05	0.57	0.09	0.07	0.07	0.07	0.07	0.20	0.12
DUV	0.08		0.06	0.05	0.10	0.16	0.15	0.17	0.04	0.09	0.06	0.11
ELD	0.11	0.08		0.08	0.21	0.17	0.19	0.12	0.14	0.08	0.09	0.12
HAB	0.13	0.09	0.08		0.14	0.16	0.16	0.10	0.12	0.05	0.11	0.27
ILO	0.38	0.11	0.09	0.08		0.15	0.10	0.08	0.08	0.11	0.28	0.15
LAE	0.10	0.13	0.12	0.11	0.13		0.91	0.81	0.07	0.20	0.08	0.23
LAS	0.07	0.13	0.09	0.08	0.09	0.94		0.42	0.05	0.19	0.07	0.18
LAW	0.08	0.12	0.10	0.10	0.11	1.00	0.58		0.08	0.27	0.05	0.23
NFL	0.07	0.05	0.11	0.06	0.08	0.18	0.19	0.13		0.10	0.07	0.08
SAW	0.07	0.10	0.11	0.08	0.10	0.33	0.25	0.25	0.04		0.07	0.22
SCW	0.20	0.09	0.08	0.08	0.30	0.17	0.12	0.08	0.07	0.09		0.16
SEG	0.14	0.20	0.09	0.13	0.20	0.32	0.20	0.16	0.09	0.08	0.13	

CHAPTER 3

Path segmentation for beginners: An overview of current methods for detecting changes in animal movement patterns

Published as: Edelhoff, H., Signer, J., and Balkenhol, N. (2016). Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns. *Movement Ecology*, *4*(1), 21

Abstract

Increased availability of high-resolution movement data has led to the development of numerous methods for studying changes in animal movement behavior. Path segmentation methods provide basics for detecting movement changes and the behavioral mechanisms driving them. However, available path segmentation methods differ vastly with respect to underlying statistical assumptions and output produced. Consequently, it is currently difficult for researchers new to path segmentation to gain an overview of the different methods, and choose one that is appropriate for their data and research questions.

Here, we provide an overview of different methods for segmenting movement paths according to potential changes in underlying behavior. To structure our overview, we outline three broad types of research questions that are commonly addressed through path segmentation: 1) the quantitative description of movement patterns, 2) the detection of significant change-points, and 3) the identification of underlying processes or 'hidden states'. We discuss advantages and limitations of different approaches for addressing these research questions using path-level movement data, and present general guidelines for choosing methods based on data characteristics and questions. Our overview illustrates the large diversity of available path segmentation approaches, highlights the need for studies that compare the utility of different methods, and identifies opportunities for future developments in path-level data analysis.

Keywords: path topology, telemetry, GPS, animal behavior, state-space models, bio-logging, path segmentation, path-level analyses

Introduction

Movement is an important life history trait in organismal ecology. Individual movement decisions and capacities affect habitat-dependent space-use and foraging strategies, as well as dispersal and migration (Bowler and Benton 2005; Wilson et al. 2012). Changes in movement behavior impact individual fitness, reproductive success and survival (Kays et al. 2015; Owen-Smith et al. 2010), ultimately driving population dynamics and evolution of species. The importance of movement has led to the emergence of the movement ecology paradigm, which provides a fundamental conceptual framework for studying movement in a holistic and mechanistic manner (Nathan et al. 2008).

For animals, modern tracking devices (e.g., GPS or ARGOS) make it possible to gather relocation data at increasingly fine spatial and temporal resolutions, thereby providing the data necessary to address comprehensive questions about how individuals perceive, react to, utilize or even change their environment (Cagnacci et al. 2010; Schick et al. 2008).Traditionally, animal relocation data were used in different variants of point pattern analyses in order to describe space use and resource selection as well as home ranges and territorial behavior (Moorcroft and Barnett 2008; Powell 2000; Worton 1987). These methods are especially useful when relocations are sampled at low frequencies (e.g., several hours or days) or with large temporal gaps. However, researchers can now collect relocation data for mobile animals at intervals of minutes (e.g., Weber and Norman 2015) or even seconds (e.g., Thiebault and Tremblay 2013). Rather than analyzing such high-frequency data as mere point patterns, they are often treated as movement paths, which provide a temporal sequence of the steps an animal took through space (Cushman 2010). An important advantage of analyzing animal movements at the path-level is the enhanced opportunity to learn about the behavior driving the observed movement patterns.

Path segmentation methods are perhaps most widely-used for identifying behavioral states from path-level movement data. These methods essentially dissect movement paths into segments that are assumed to reflect different underlying behaviors. By defining behavioral states from the paths and then linking state-dependent movements to the environment, scientists can gain an enhanced understanding of the biological processes influencing the movement behavior of animals (Killeen et al. 2014; Roever et al. 2013).

Given the tremendous capabilities of path segmentation for movement ecology, it is not surprising that the number of approaches suggested for segmenting a path and detecting behavioral states is growing rapidly. However, many of these methods have their roots in non-ecological scientific disciplines and gaining a comprehensive understanding of the plethora of available methods can be time-consuming and even frustrating, which likely results in path-level analyses not being used as often and as efficiently as possible.

Here, we offer an overview of available methods for segmenting animal movement paths to detect underlying behavioral states. For this, we first introduce the basics of path-level analyses and relevant terms for distinguishing different movement types. Next, we outline some of the major differences between analytical approaches and suggest general considerations for matching available methods to three broad types of research questions: 1) the quantitative description of movement patterns, 2) the detection of significant change-points, or 3) the identification of underlying processes ("hidden states"). To illustrate our suggestions, we also apply multiple methods to a simulated dataset. We include examples of different ecologically relevant movement processes at varying temporal scales (e.g., diel and annual time scales), as well as behavioral responses to habitat configuration to provide more insight on the application of the presented segmentation approaches. Finally, we discuss remaining challenges and suggest future research avenues for path segmentation. Our overview is specifically intended as a starting point for beginners with little or no experience in path-level analysis of telemetry data, and we therefore avoid statistical details as much as possible. These details can be found in the supplement and also the references given for the individual methods.

Basics of path-level analyses

Movement paths and trajectories

Usually, we cannot observe the complete, continuous movement path of an animal. Instead, we sample a set of discrete relocations to approximate the animals' actual movement path (Calenge et al. 2009; Step 1 in Figure 3.1). The resulting sequence of consecutive records of the location of the animal (e.g., spatial coordinates, ordered by time) is termed a movement track or trajectory (Getz and Saltz 2008). How well a trajectory reflects the actual movement path of an animal depends on the sampling regime as well as the recording systems (GPS, Argos, VHF, light-level geolocation), which influences the spatial accuracy and frequency of relocations.

In path-level movement data, consecutive relocations are either sorted by an ordering factor, for example as the result of direct tracking or following of an animal (Fryxell et al. 2008; McKellar et al. 2014) or by the time at which the relocations were recorded (Calenge et al. 2009; Morales and Ellner 2002). Sampling frequency influences the resolution of the data and the level of inferential detail that can be obtained (Johnson and Ganskopp 2008; Nathan et al. 2008; Van Moorter et al. 2010). For example, shorter temporal intervals allow detailed insight into fine-scale behaviors, but are more sensitive to sampling errors (e.g., spatial inaccuracies of relocations). In contrast, movements sampled at longer temporal intervals can only be interpreted on a broader scale (e.g., encamped vs. dispersal movements). Additionally, recorded relocations can be spurious or lack spatial accuracy due to habitat induced sampling errors (Bradshaw et al. 2007; Hurford 2009; Jerde and Visscher 2005; Williams et al. 2012). Importantly, trajectories also differ with regard to their regularity of the time intervals between successive steps. Irregular data commonly results from missing relocation fixes or varying sampling frequencies throughout a study period (e.g., Graves and Waller 2006). Further, irregular intervals between relocation samples can stem from different behaviors of the study species. For example, relocation devices applied with marine animals can usually provide the measured position data only when the species is close to the surface (Gurarie et al. 2009; Jonsen et al. 2007; Laidre et al. 2004).

Figure 3.1: Overview of important steps throughout a segmentation analysis. In general, the actual continuous movement path of an organism is sampled as a set of consecutive relocations (Step 1; e.g., field work). Step 2: exploratory and descriptive analyses of path-characteristics; exploring and visualizing of the data structure. Step 3: applying one or several path segmentation method(s) to objectively distinguish different movement states. Step 4: Some methods require the use of clustering and summary statistics to quantify differences in distinguished movement states, and to facilitate biological interpretation in terms of behavioral modes.

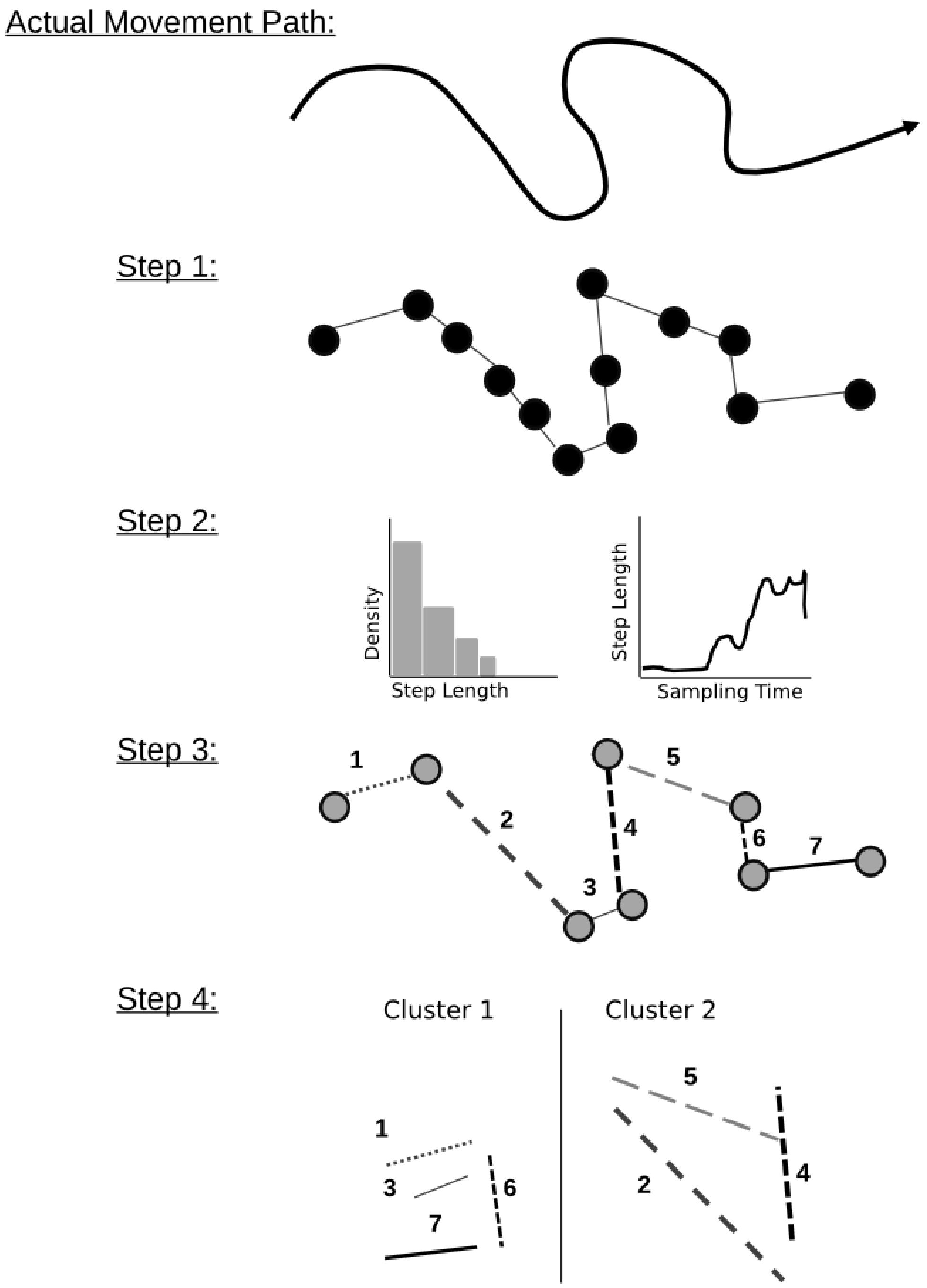


Table 3.1: Currently applied path-characteristics. Different signals or parameters can be calculated either based on consecutive relocations within a trajectory (“stepwise”) or for multiple relocations such as identified path-segments (“across multiple steps”).

Characteristic	Description	Type	Calculation	Reference
Displacement	Increment of the X and Y values between two consecutive relocations, change in absolute spatial position	primary	stepwise	(Calenge et al. 2009, Dodge et al. 2008, Brillinger et al.2004)
Time Lag	Duration / increment in time between consecutive relocations (usually determined by sampling regime)	primary	stepwise	(Calenge et al. 2009, Dodge et al. 2008)
Turning Angles / Heading	Relative and absolute turning angles between consecutive relocations, change in direction	primary	stepwise	(Calenge et al. 2009, Morales and Ellner 2002, Batchelet 1981, Turchin 1998)
Step Length	Euclidean distance between two consecutive relocations	primary	stepwise	(Calenge et al. 2009, Dodge et al. 2008)
Velocity / Speed	Distance traveled in a given time interval between two relocations; less sensitive to missing data than step length	primary	stepwise	(Calenge et al. 2009, Gurarie et al. 2009, Dodge et al. 2008)
Persistence / Turning Velocity	Transformations of speed and turning angle: persistence velocity represents the tendency and degree of a movement to persist in a certain direction. Turning velocity shows the tendency of a movement to turn in a perpendicular/opposite direction	secondary	stepwise	(Gurarie et al. 2009, Gurarie et al. 2016)
Net / Mean Squared Displacement	Squared displacement between the first and current relocation of the trajectory; applied to characterize diffusion behavior or migration patterns	secondary	stepwise	(Calenge et al. 2009, Morales and Ellner 2002, Gutenkunst et al. 2007)
First Passage Time	Time required for crossing a predefined endpoint based on a circle (radius) around a starting relocation. Sums the times of all forward and backwards relocations within the radius; index of area-restricted search behavior	secondary	stepwise	(Barraquand and Benhamou 2008, Fauchald and Tveraa 2003, McKenzie et al. 2009)

Residence Time	Extension of the first passage time accounting for returns of the animal in a given area. Sums the times of all relocations (backwards and forwards) of a trajectory within a given vicinity around a relocation.	secondary	stepwise	(Barraquand and Benhamou 2008)
Pseudo-Azimuth	Recalculates the basic azimuth value at the midpoint between two consecutive steps to range within 0 and 360. Can be used as indicators for movements with same or parallel directions.	primary	stepwise	(LaPoint et al. 2013)
Straightness Index	Ratio of Euclidean distance between the beginning and end of a trajectory and the total path length (sum of all step lengths)	secondary	across multiple steps	(Gurarie et al. 2016, McKenzie et al. 2009)
Sinuosity / Tortuosity	Adaptions of the straightness index analyzing the probabilistic distributions of the changes in the turning angles and the beeline distance between the start and end points of the trajectory; index of path orientation	secondary	across multiple steps	(Benhamou 2004, Bovet and Benhamou 1988)
Fractal Dimension	Measure of path tortuosity; non-Euclidean dimension of the trajectory varying between 1 (completely straight) and 2 (tortuous, completely spanning two-dimensional space); different implementations exist	secondary	across multiple steps	(Nams 1996, Dicke and Burrough 1988, Tremblay et al. 2007, Turchin 1996)
Multi-Scale Straightness Index	Repeated calculation of the straightness index of a trajectory over a range of different temporal scales	secondary	across multiple steps	(Postlethwaite et al. 2013)
Area Interest Index	Repeated calculation of the straightness index for a limited size of a sliding window along the trajectory. With each repetition, the number of relocations within the trajectory is reduced	secondary	across multiple steps	(Postlethwaite et al. 2013, Wilson et al. 2007)

Basics of path segmentation

We use the term segmentation as a general paraphrase for determining changes in an animal's movement behavior based on the observed trajectory. The process of segmentation involves the partitioning of a trajectory, τ, into a number of K subtrajectories (τ_1, τ_2, ..., τ_K) called segments (Steps 1-3 in Figure 3.1; see also Barraquand and Benhamou 2008; Buchin et al. 2011). Path segmentation can be accomplished directly, by designating each observation to different states or clusters (e.g., Franke et al. 2004; Van Moorter et al. 2010). However, path segmentation commonly relies on detecting significant changes (so called change- or breaking-points) in the trajectory as cut-offs for separating the trajectory into distinct segments (e.g., Gurarie et al. 2009). For this, a variety of path characteristics can be derived from the trajectory, for example the step length or velocity. These path characteristics should accurately capture movement patterns and allow the detection of changes in these patterns. Given the importance of these path characteristics for successfully segmenting movement paths, we discuss them in more detail in the next section.

Path characteristics

The various path characteristics used by current segmentation methods are summarized in Table 3.1. These characteristics have also been called movement metrics, movement parameters, path-signals or indices in the literature, and should convey relevant information about individual movement behavior (Barraquand and Benhamou 2008; Dodge et al. 2008; Gurarie et al. 2016). The majority of path characteristics are derived from consecutive relocations (stepwise), for example the speed of travel. However, some signals are calculated across multiple relocations, for example the straightness of a trajectory (Table 3.1).

Dodge et al. (2008) distinguished primitive path parameters from primary and secondary derived parameters. The information on the absolute spatial position (e.g., xy-coordinates) and the temporal dimension (time stamp) provide the primitive signals from which other parameters can be derived. For example, displacement and step length (see Table 3.1) are primary derivatives of the position parameter, whereas time lag (duration) is derived from the temporal primitive. Path-signals exclusively based on spatial criteria are particularly sensitive to sampling intervals and errors (Calenge et al. 2009; Van

Moorter et al. 2010). However, other signals such as the persistence or turning velocity avoid possible biases caused by varying sampling intervals by relating speed to the observed turning angles . Furthermore, signals such as the first passage (Fauchald and Tveraa 2003) and residence time (Barraquand and Benhamou 2008) constitute summary properties accounting for the temporal scales within the movement paths and can be seen as secondary derivatives of the distance and duration signals.

Table 3.1 also lists characteristics which are calculated over multiple relocations and can be applied to describe the signals of single segments, certain sub-samples of trajectories, or entire trajectories. Such summary signals like the straightness index (Batschelet 1981), sinuosity (Benhamou 2004) and the fractal dimension (Nams 1996) provide information on the spatial complexity of a given path segment and can be used to cluster segments into groups that are similar with respect to movement complexity (Step 4 in Figure 3.1). Sinuosity constitutes another example of a secondary derivative of the step length signal (Dodge et al. 2008).

Overall, a large number of different measures can be used to describe path characteristics and a chosen parameter should ideally convey relevant information about the underlying movement behavior (Barraquand and Benhamou 2008). This requires a good understanding of the species and a precise definition of research questions, and should also involve extensive exploratory analyses to understand the structure of obtained relocation data and to test the feasibility of different segmentation approaches (Step 2 in Figure 3.1; see also below and Gurarie et al. 2016)

Finding and interpreting segments

Regardless of how and which path characteristics are quantified, significant changes within these signals are then used to determine the K-1 break-points ($\tau^*_1, ..., \tau^*_{K-1}$) which can be used to divide the trajectory into K segments (Step 3 in Figure 3.1). Although preliminary visual analyses can provide useful indications about a meaningful value for K, an objective, data driven way is desirable. Therefore, path segmentation often involves quantitative approaches for detecting an unknown number of segments within a given trajectory, and many of these approaches have originated in non-ecological disciplines (e.g., Lavielle 1999). This is an important point, as many segmentation methods only provide information on significant change-points along the trajectory, without any further ecological context. Thus, it is often not trivial or even possible to directly associate the individual segments to specific activities and behaviors (Zhang et al. 2015). To facilitate the ecological and ethological interpretation of the defined segments, some methods require subsequent analyses to classify the determined segments based on different descriptive parameters or summary statistics (Step 4 in Figure 3.1). For example, either the mean values of stepwise characteristics or multi-step summary parameters, such as the straightness index (see Table 3.1), of the segments can be further analyzed in an additional classification analysis (e.g., Zhang et al. 2015). This generates clusters of segments that are similar with respect to relevant path parameters (e.g. calculated across multiple steps, Table 3.1), which can help to identify underlying movement patterns and associated behaviors. For example, short, meandering movement segments during within-patch foraging vs. long, straight segments during inter-patch movements (Madon and Hingrat 2014; Nams 2014). Other methods determine the state (also called class or cluster) of each individual relocation directly and no further classification is necessary (Franke et al. 2004; Van Moorter et al. 2010).

In sum, path segmentation involves at least three and sometimes four major steps (Figure 3.1). In the following, we focus on the third step, in which signals derived from trajectories are used to objectively define movement segments.

Overview of path segmentation methods

Types of methodological approaches

Methods for path segmentation can be distinguished or classified using many different criteria, for example based on their underlying statistical framework (e.g., maximum-likelihood versus Bayesian; parametric or non-parametric, inference-based etc.). Alternatively, Gurarie et al. (2016) recently classified broad types of movement analysis tools based on the analytical traditions they stem from. Since our overview is specifically intended for beginners wanting to apply path segmentation, we do not categorize methods based on their statistical properties or analytical traditions, but instead focus on the practical utility of the analyses, e.g., the research questions that can most readily be answered with a certain approach. Hence, we structure our overview based on three broad types of questions that are commonly addressed using path segmentation.

First, movement patterns within the trajectory can be quantified to test whether different movement components are identifiable within the data. For example, such ‘movement pattern description’ is used to distinguish active from resting phases (e.g., van Beest and Milner 2013), or encamped foraging from traveling movements (e.g. Dzialak et al. 2015). Second, path segmentation can also be used to locate significant changes in movement behavior and determine the timing of these changes. For example, such ‘change-point detection’ has been used to quantify behavioral responses to seasonal environmental changes (e.g., Garstang et al. 2014), or to identify the timing of migration events (e.g., Le Corre et al. 2014). Finally, path segmentation can be used to take a detailed look at the processes underlying observed movement patterns. Such ‘process identification’ can be used to examine the factors influencing diel variation in movement rates among individuals (e.g., Jonsen et al. 2006), or to quantify how sex and reproductive status influence the duration of, and transition among, different behavioral modes (van de Kerk et al. 2014). These three broad types of research questions can be matched to three basic categories of analytical approaches for path segmentation (Figure 3.2).

Topology-based approaches to describe movement patterns

If the study aim is to quantitatively describe movement patterns, one can use methods that focus on the description of geometric properties of the trajectory itself, or on one or several signals calculated from the trajectory. Based on this path topology, movement steps are then assorted into groups that are relatively similar with respect to these signals (Figure 2a). The exact way this is accomplished depends on the method, but can be achieved either by a) simply grouping individual movement steps based on similarity in topology-based signals, regardless of whether these steps are consecutive (e.g. thresholding or clustering; (Dzialak et al. 2015; Van Moorter et al. 2010); or b) identifying changes observed among the signals between successive relocations to detect so-called change-points (e.g., spatio-temporal criteria segmentation; Buchin et al. 2011). These change-points are assumed to correspond to changes in underlying movement behavior, therefore separating the trajectory into segments consisting of multiple consecutive steps based on pronounced changes in observed movement characteristics. These topology-based methods are mostly non-parametric and rather descriptive. Their application is usually based on predefined hypotheses on how movement behaviors might differ among habitats, seasons, times of day, sexes, social status, etc. .

Time-series analyses to detect significant change-points

If the goal of a study is to detect points in time when a significant change in the movement behavior occurs, path segmentation methods based on time-series analyses can be used. Such time-series analyses are widely used in ecology and related disciplines (see Lange 2006). In the context of path segmentation, these analyses treat signals calculated from consecutive movement steps as time-ordered observations. Essentially, the majority of these approaches try to find significant change-points along the time axis of the signal-sequence derived from the movement trajectory (Figure 3.2b). In contrast to the topology-based approaches that analyze the changes between temporally ordered relocations, most of the time-series methods treat movement patterns as a function of time and can directly account for the temporal correlations of the sequential signal data. The time-series approaches sometimes depend on certain information like the maximum number of change-points or the minimum length of the detected segments. However, they could also potentially be used to "blindly" search for all possible change-points of a given path-signal sequence.

State-space models to identify underlying processes

Finally, to increase our understanding of the behavioral processes underlying complex movement patterns, methods derived from the state-space modeling framework are most suitable. These state-space models represent a special type of time-series analysis (Patterson et al. 2008) and intend to identify latent or hidden behavioral states based on the observed movement data. The aim is to derive deeper insight into the underlying processes by formulating a movement model that explains observed movement patterns. Within these frameworks, the future state of a system is modeled to depend on its current state through a probabilistic model (see Figure 3.2c). Therefore, the models typically assume a so-called Markov process structure, meaning that a hidden future state depends on the state of the current step (Jonsen et al. 2013). Essentially, state-space models couple two stochastic time-series models, one based on an unobservable state process, and another based on a known observation process (Jonsen et al. 2013; Patterson et al. 2016). When applied to movement data, state-space models assume that animals have several 'hidden behavioral states' with certain characteristics (e.g., path-signals) that can be modeled using stochastic processes (e.g., correlated random walks; Morales et al. 2004). A basic result of a state-space model are the estimated transition probabilities between the considered states. Another outcome is the probability of a given relocation belonging to one of the hidden behavioral states. These probabilities are then used to assign steps to their most probable behavioral state (Figure 3.2c) and to segment the trajectory according to state memberships. Additionally, the transition probabilities can also be linked to different environmental factors to test various hypotheses on behavioral and ecological dependencies of the observed movement patterns (Beyer et al. 2013; DeRuiter et al. 2016; Morales et al. 2004). For example, the transition probabilities can be used to test whether switching between states depends on certain habitat characteristics (see simulation study below).

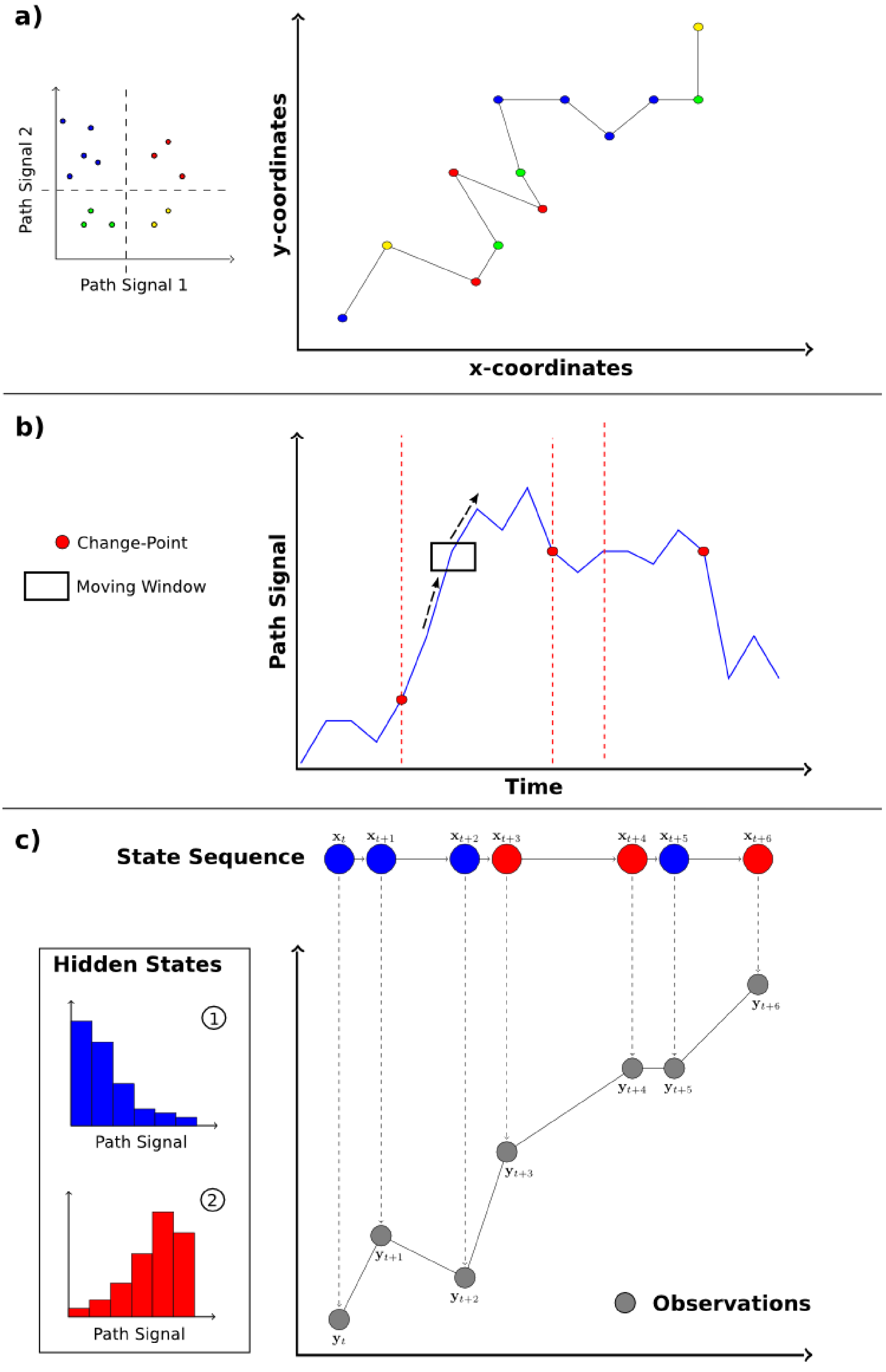


Figure 3.2: The main study aims of path segmentation and types of methods to answer them. a) Pattern description: Topology-based analyses rely directly on signals calculated from the movement trajectory (e.g. step length and bearing). They combine movement steps into groups based on similarity in the considered

path-signals, for example by applying clustering algorithms. b) Change-point detection: Time-series analyses assess a path-signal (y-axis) along its time-axis. For example, a moving window (rectangle) can be used to search for points along the time-series where local parameters (e.g. the mean) of the path-signal are significantly different from the global averages of these parameters. Significant change-points are assumed to indicate switches in underlying movement modes or behavioral states, and are used to separate the trajectory into segments (dashed lines). c) Process identification: The majority of the presented state-space models link two stochastic models describing the state process and its observation. For example, the state process could consist of two discrete behavioral states (red and blue). The process model describes how the hidden state (x) emerges based on a Markov process. Therefore, it accounts for the conditional probability of a future state depending on the one of the current relocation. The observation model links the actual observed data (y) at given points in time to the hidden state. As a result, the most probable state of each observation, the switching probabilities between the states, as well as the distributions of the measured path-signals within each state are provided.

Choosing among methods for path segmentation

Multiple methods for path segmentation exist within each of the three types of analytical approaches described above. Thus, multiple methods exist to answer each of the broad categories of research questions (study aims). Table 3.2 provides an overview of the available path segmentation methods and lists basic properties, and important background papers for each method. More detailed descriptions and further information on each path segmentation method, including implementations in the program R (R Core Team 2015), can be found in supplement S1.

Available path segmentation methods vary substantially with regard to their demands on data structure and underlying theory. This raises the question of how scientists can identify the most appropriate segmentation method(s) for their specific research goals. In the following, we provide some general guidelines for method selection. Additionally, the guidelines are visually summarized in Figure 3.3.

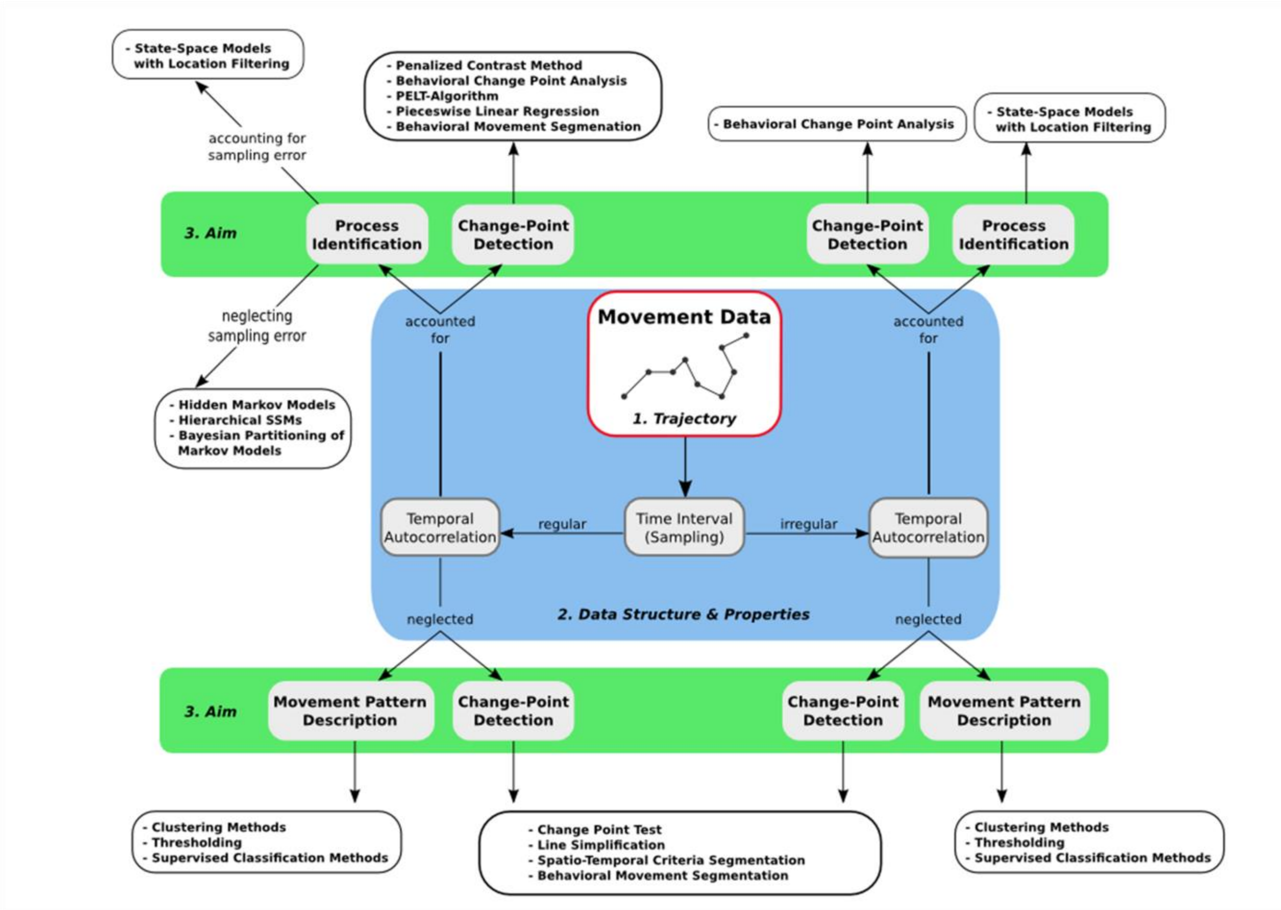


Figure 3.3: Decision guidelines for choosing appropriate segmentation methods. The process should begin with preliminary analyses of the trajectory data and derived path-parameters (1). Choosing among methods is then first directed by the data structure and sampling regime (2). Capability of the methods to account for temporal autocorrelation

Table 3.2: Characteristics of the methodological approaches for the three different categories of research questions. Different methods for answering the three type of broad research questions (study aims) are listed together with the analytical category they stem from, a short description of each method as well as the considered categories of input path-signals and important references.

Study Aim	Method	Analytical Category	Description	Input Signal	References
Movement Pattern Description	Thresholding	Topology-based	Applies thresholding schemes (cut-off values) to separate relocations into different groups based on single or multiple path parameters (e.g., short- vs. Long-range movements)	Primary and secondary signals	Gutenkunst et al. 2007, Tremblay et al. 2007, Dzialak et al. 2015, Sur et al. 2014
	Supervised Classification	Topology-based	Relocations (steps) of a trajectory are assigned to certain classes of movement behavior based on a classification scheme fitted with a training dataset	primary and secondary signals, additional information like activity data	Dodge et al. 2009, Shamoun-Baranes et al. 2012, Soleymani et al. 2014
	Clustering	Topology-based	Unsupervised classification for identifying distinctive groups within a multivariate set of path-signals	primary and secondary signals, additional information like activity data	Van Moorter et al. 2010, Garriga et al. 2016
	Bayesian Partitioning of Markov Models (BPMM)	Topology- and time-series based	Classification algorithm for determining the number and sequence of homogenous classes within a sequential path-signal (time series)	Primary and secondary signals	Gurarie et al. 2016, Gueguen 2000, Calenge 2011
Change-Point Detection	Line Simplification	Topology- or time-series based	Tests whether reducing the number of vertices in a trajecotry significantly impacts path topology to determine change points (can also be applied with graphs of sequential path-signals)	primitive signals (spatial position)	Thiebault and Tremblay 2013, Douglas and Peucker 1973
	Change Point Test	Topology-based	Detects significant changes in the observed movement direction (orientation) between the starting point and an attraction point of a trajectory	primitive signals (spatial position)	Noser and Byrne 2013, Byrne et al. 2009
	Spatio-Temporal Criteria Segmentation	Topology-based	Special type of thresholding seeking optimal segmentation of a trajectory based on monotone criteria: relocations are included in a segment as long as they fullfill certain predefined requirements	Primitive, primary and secondary signals	Buchin et al. 2011, Buchin et al. 2013
	Piecewise Regression	Time-series analysis	Splits time-series model into representative segments based on a signficant change-point (fits a polynomial model for each segment)	primary and secondary signals	Byrne et al. 2009, Buchin et al. 2013

	Penalized Contrast Method (PCM)	Time-series analysis	Non-parametric segmentation of a path-signal: the unknown number of segments is estimated by minimizing a penalized contrast function	mostly secondary signals	Barraquand and Benhamou 2008, Lavielle 1999, Johnson et al. 2002
	Behavioral Change Point Analysis (BCPA)	Time-series analysis	Likelihood-based method for detecting significant change points; applies moving window over continuous autocorrelated time series of a path-signal	mostly secondary signals	Gurarie et al. 2009, Gurarie et al. 2016
	Pruned Exact Linear Time (PELT) Algorithm	Time-series analysis	Search method for detecting optimal number and locations of change points minimizing different cost and penalty functions	primary and secondary signals	Madon and Hingrat 2014, Liminana et al. 2007
	Behavioral Movement Segmentation (BMS)	Time-series analysis	Combined search algorithm which optimizes segmentation based on parsimony and subsequent clustering for assigning segments to similar behaviors	primary and secondary signals, additional information like activity data	Nams 2014
Process Identification	Hidden-Markov Models (HMM)	State-space models	Estimate the sequence and composition of a predifined number of discrete states (e.g., movement behaviors) as well as the switching-probabilities between these states	Primary signals, additional information like activity data	Franke et al. 2004, van de Kerk et al. 2014, Patterson et al. 2016, Morales et al. 2004, Beyer et al. 2013
	State-Space Models with Location Filtering	State-space models	More complex models which can model hidden movement states and also correct for errors in the observation process (e.g., GPS errors)	Primitive (spatial position) and primary signals, additional information like activity data	Patterson et al. 2008, Jonsen et al. 2013, Forester et al. 2007, Dowd and Joy 2011, Sibert et al. 2003, Jonsen et al. 2005
	Hierarchical State-Space Models	State-space models	Hierarchical models accounting for variability of number and composition of movement states between individuals (further making inferences at population level)	Primary signals	Jonsen et al. 2006 , Jonsen et al. 2013, Mills et al. 2010
	Bayesian Partitioning of Markov Models (BPMM)	Topology- and time-series based	Can also be used as partinioning algorithm determining the number and sequence of homogenous models („states“) within a sequential path-signal	primary and secondary signals	Gurarie 2016, Gueguen 2000, Calenge 2011

Preliminary Data Analyses

Because the structure and composition of movement data dictate the applicability of certain methods (Figure 3.3; blue panel), the first step in any segmentation study should be a preliminary analysis of the available location data. Various analyses can be carried out to gain a better understanding of data properties, but a preliminary analysis for path segmentation should contain at least the following four steps.

1) Sampling Regime

Movement data usually varies substantially with regard to the sampling regime, spatial accuracy and temporal resolution. Therefore, preliminary analyses should include checking for regularity of time-intervals between relocations as well as testing for temporal autocorrelation of the path-parameter data (Calenge et al. 2009; Dray et al. 2010). Depending on the results of these analyses, several segmentation methods may no longer be suitable (Figure 3.3).

2) Data Regularity

Irregular data can be the product of missing relocation fixes and varying sampling regimes which can be a challenge, as some of the statistics used to analyze movement paths assume regular intervals within the trajectory and are valid only under those circumstances (Gurarie et al. 2009). Different processing tools can be applied to relocation data in order to fulfill the assumptions of regularity. For example, trajectories can be re-discretized (Benhamou 2004; Calenge et al. 2009), which means that relocations can be removed until the remaining data fulfills the requirement of temporal regularity (“thinning”). Alternatively, missing relocations can be replaced by applying techniques such as spatial interpolation (Lonergan et al. 2009; Thiebault and Tremblay 2013) or dead-reckoning (Bidder et al. 2015; Liu et al. 2015; Wensveen et al. 2015). Furthermore, only subsets limited to continuous and regularly sampled relocations of the original trajectory can be selected for further analyses (Benhamou 2004; Calenge et al. 2009). Approaches modeling movement in continuous time are also capable of dealing with irregular data structures (e.g., Johnson et al. 2008; Patterson et al. 2016).

Additionally, habitat induced sampling errors and spatial inaccuracies can occur and need to be addressed throughout the preliminary analyses (Hurford 2009; Jerde and

Visscher, 2005; Williams et al. 2012). This includes checking the data for extreme outliers or estimating the error of the applied tracking technology (e.g., provided by ARGOS systems; (Lowther et al. 2015). Some types of state-space models include location filtering where such information can be implemented as a prior in order to estimate the true positions of erroneous relocation data (e.g., Kalman Filtering; Austin et al. 2003; Sibert et al. 2003; Silva et al. 2014).

3) Data Visualization and Signal Distributions

Visual inspection of the movement trajectory can already indicate the existence of different behavioral modes (Brillinger et al. 2004; Demšar et al. 2015; Shamoun-Baranes et al. 2012). Also, in order to choose appropriate path-signals conveying information on potential changes within the movement behavior, investigations of their variability and distributions (e.g., histograms) should be considered. For example, multi-modality within the path-signal distributions can also indicate the potential existence of different behavioral modes (see applied examples). Further, depending on the intended segmentation method, knowledge on the parameter distributions is also needed for fitting of movement models within the various types of state-space models (Codling et al. 2008; Morales et al. 2004). As a substantial part of the methods stem from the time-series framework, time-ordered plotting of the path-signals can indicate the existence of changes in the sequence over time (see applied examples below). Visual inspection of the variation of the signals over time can provide insight on the ranging and movement behavior. For example, Bunnefeld et al. (2011) and Killeen et al. (2014) inspected time-ordered values of net-squared displacement (Table 3.1) for single or multiple modality in order to detect potential migratory individuals. Further, the visual inspection of movement trajectories can help to identify unusual relocations and movements (Demšar et al. 2015; Shamoun-Baranes et al. 2012). Thus, visual inspection of the trajectory is important for error checking and can help to refine biological hypotheses to be tested with a given data set.

4) Scales of Movement and Data Transformation

Detectability and observability of changes in movement behavior can also change with temporal and spatial scale (Fryxell et al. 2008; Gurarie and Ovaskainen 2011). There are multiple options of indexes and transformations providing information on the varying spatial and temporal scales of the path-characteristics (e.g., trigonometric circle space;

Thiebault and Tremblay 2013). Further, sub-sampling, re-discretization or moving-windows can be applied to alter the temporal grain (e.g. daily, nocturnal, weekly or monthly relocations) in order to summarize the means or variances of path-parameters (Ganskopp and Johnson 2007; Laidre et al. 2004; Long et al. 2013; Postlethwaite and Dennis 2013). Also, multi-step signals (see Table 3.1) such as the simple straightness index (Batschelet 1981) and its different extensions (Postlethwaite et al. 2013; Wilson et al. 2007) can be applied to investigate the variation of path straightness within a trajectory over time and multiple temporal resolutions. Path-parameters such as the first passage or residence time (Table 3.1) can be calculated at varying spatial and temporal scales and allow further insight in underlying spatial and temporal scales of individual movement behavior (Barraquand and Benhamou 2008; Byrne and Chamberlain 2012; Frair et al. 2005). Finally, different transformations of the path parameters can be applied to determine dominant and constant periodic frequency patterns in the movement data. For example, Fourier and wavelet transformations provide valuable insight in periodic structures of movement, such as circadian, seasonal or diurnal rhythms (Gaucherel 2011; Polansky et al. 2013; Sur et al. 2014; Wittemyer et al. 2008).

Study aims

After the preliminary analysis of the data structure and relevant path characteristics, choosing appropriate segmentation methods is mostly influenced by the aims of the study (Figure 3.3; green panels). Thus, depending on the study aims and data structure, different methods can be applied.

1) Movement pattern description

The majority of appropriate methods for quantitatively describing movement patterns are based on the path-topology approaches such as simple threshold or multivariate classification algorithms (detailed information for each method in supplement S3.1). These approaches are least demanding with regard to data properties like regularity and do not require any data transformations as they make minimal assumptions about underlying data structures, movement models, or behavioral states. However, they can be valuable exploratory tools for determining the potential number of different behavioral states within the observed movement data (e.g., Dzialak et al. 2015; Gutenkunst et al. 2007; Van Moorter et al. 2010). Furthermore, the methods can be applied for testing certain hypotheses on how particular path-signals change with different behaviors or at certain time-periods. Therefore, for some study aims it might be sufficient to split movements into two or more different behavioral states (e.g., long- vs. short-range movements) based on a threshold within a selected path-signal (e.g., step length; Zeller et al. 2014). Similarly, the time when the relocations were recorded could be used to distinguish different types of behavior (e.g., daytime vs. nocturnal movements).

In sum, methods for pattern description can be applied to gain insight on potential behavioral states and even for detecting potential drivers of the observed patterns (e.g., nocturnal movement behaviors with longer step length). However, the considered path-signals have to be chosen carefully and according to expected changes in movement behaviors and underlying behaviors (Gurarie et al. 2016; Van Moorter et al. 2010). Furthermore, due to their relative simplicity, topology-based methods offer little explanatory power and are usually not suitable for analyzing complex movement patterns (Gurarie et al. 2016).

2) Change-point detection

The second example of general study aims is the determination of important (significant) change-points in the movement behavior or trajectory of an animal. The presented approaches either focus on the path-topology or on a time-series of a path-signal. In both cases, the sequential relationship between consecutive relocations is accounted for.

The relevant topology-based methods either focus on the changes within the absolute spatial position (e.g., the change point test; Byrne et al. 2009; Table 3.2) or different path-signals and their shape along the trajectory (e.g., using Spatio-Temporal Criteria Segmentation; Figure 3.3). However, the change-points resulting from the topology-based methods usually do not provide any information on the significance of the observed changes within the data composition. If identifying significant change-points is the aim, for example, to detect the onset of migratory events, then methods from the time-series category are the better choice, as they specifically estimate the significance of changes within a time-ordered data sequence (Figure 3.2b). The majority of time-series approaches are capable of accounting for temporal autocorrelation within the data sequence which can be an important advantage, because non-independence of relocations is a challenge for many standard statistics (Gurarie et al. 2009). As can be seen in our example, the autocorrelation structure of the data can also contain valuable information about the underlying behavioral states (Cushman 2010). As a caveat, most time-series methods show higher demands on data properties, especially regularity of the time intervals between relocations (an exception is the behavioral change-point analysis; BCPA). Furthermore, many of the appropriate time-series methods listed in Table 3.2 depend on one or multiple parameters which need to be defined prior to the analyses such as the size of a moving window (e.g., for the behavioral change point analysis; Gurarie et al. 2009) or the minimum number of relocations within a determined segment (e.g., for the penalized contrast method; Lavielle 1999). Therefore, several assumptions, about the number of potential changes or the length of a behavioral state, need to be made before setting these parameters, which increases the susceptibility to errors and bias and limits reproducibility.

In contrast to that, topology-based methods for change-point detection are less dependent on such parameter settings and mostly focus on changes within the spatial composition of the trajectory. However, the scale at which these methods can detect changes in movement behavior is highly dependent on the temporal resolution of the data. Relocations recorded at higher frequencies can provide more detailed information on fine-scale behaviors. Low frequencies usually limit the scale at which the topology-based algorithms can determine changes in the underlying behavior (Byrne et al. 2009; Getz and Saltz 2008).

Time-series approaches are usually less sensitive to the temporal sampling frequency of the data for detecting change-points when appropriate input signals conveying meaningful information are used (e.g., persistence velocity; Gurarie et al. 2009). However, time-series based methods need to be chosen carefully as their assumptions on data distributions (e.g., Gaussian vs. non-Gaussian time-series) and applied statistics can differ (see Supplement S3.1 for more details).

3) Underlying process identification

To identify processes underlying complex movement behaviors, various types of state-space models (SSM) are suitable choices. SSMs intend to identify latent states or hidden models based on the observed movement data. In this context, hidden states represent different behavioral modes, assuming that they can be described with different parametric distributions of the path characteristics. The majority of SSMs can be interpreted as a multi-state random walk and are usually based on assumptions about the density functions of the step length and turning angle distributions (Gurarie et al. 2016; Morales et al. 2004). Hierarchical approaches can be used to estimate different numbers and compositions of behavioral states for each of the studied individuals and further draw model inferences at the population level (Jonsen et al. 2013; Jonsen et al. 2005; Flemming et al. 2010; Morales et al. 2004). Another advantage of these models is that some can account explicitly for issues of animal movement data, such as irregularities caused by missing relocations and measurement errors (e.g., location filtering; Jonsen et al. 2013; Patterson et al. 2008). In particular, SSMs fitted with Bayesian estimation techniques allow the integration of prior knowledge on sampling errors (Jerde and Visscher, 2005; Jonsen et al. 2013; Jonsen et al. 2005). For example, information on the accuracy and quality of the acquired relocation data as provided by the ARGOS system can be implemented in the observational model of such a SSM framework (Jonsen et al. 2005; Flemming et al. 2010; Silva et al. 2014). Importantly, state-space models can integrate the influence of habitat features and other environmental information, such as sea depth or temperature obtained from electronic tagging data, on behavioral changes (Beyer et al. 2013; Dowd and Joy 2011; Patterson et al. 2008). Therefore, they provide a valuable framework for estimating and comparing the responses of state compositions and their transition probabilities to different covariates (DeRuiter et al. 2016; Morales et al. 2004; van de Kerk et al. 2014). Furthermore, due to their mechanistic basis, many of the SSM methods provide information on the differences in the estimated parameter distributions of the considered movement models. Thus, state-space models can also be used to simulate or predict movement patterns under varying environmental settings (Patterson et al. 2008). The biggest challenge of using state-space models is the necessity to estimate the various model parameters, which can require mathematically and computationally complex procedures (Jonsen et al. 2005; Patterson et al. 2016). In summary, state-space models offer much flexibility towards a mechanistic understanding of animal movements, because the process models make it possible to fit specific underlying movement patterns (e.g., different correlated-random walks) to the observed movements (Jonsen et al. 2005; Patterson et al. 2008).

However, the number of potential states considered within the models usually needs to be determined prior the application (Patterson et al. 2016). Also, the general composition of the considered movement models within the states has to be predefined. This limits SSM mostly to variations of discrete correlated random walks (Morales et al. 2004).

Another option for identifying “hidden states” with different compositions of movement parameters is the Bayesian partitioning of Markov models (Gueguen 2000; Gurarie et al. 2016). Technically, this approach is not a state-space model but it represents a simple solution for detecting different models within the observed movement data. The method estimates the distributions of a path-signal for a given number of potential states and assigns each relocation to one of them (Calenge 2011; Gueguen 2000). However, BPMM does not provide any information on the potential processes, the transition probabilities between the detected states, or the potential influence of covariates.

Illustration using simulated data

To illustrate the three types of research questions and related analytical approaches, we next apply one method of each type of analytical approaches to a single data set. For this, we used a simple individual-based simulation model to generate the annual movement track of an animal in R (R Core Team 2015). Details on the simulations and all relevant parameters can be found in supplement S3.2. In essence, we simulated an animal that is more active during the day than during the night, moved faster in its habitat than in the matrix (unfavorable habitat) and migrated between two centers of activity (e.g., summering and wintering range). We simulated a movement track for 12 months with relocations taken every hour in a landscape consisting of 400 * 400 cells (Figure 3.4a).

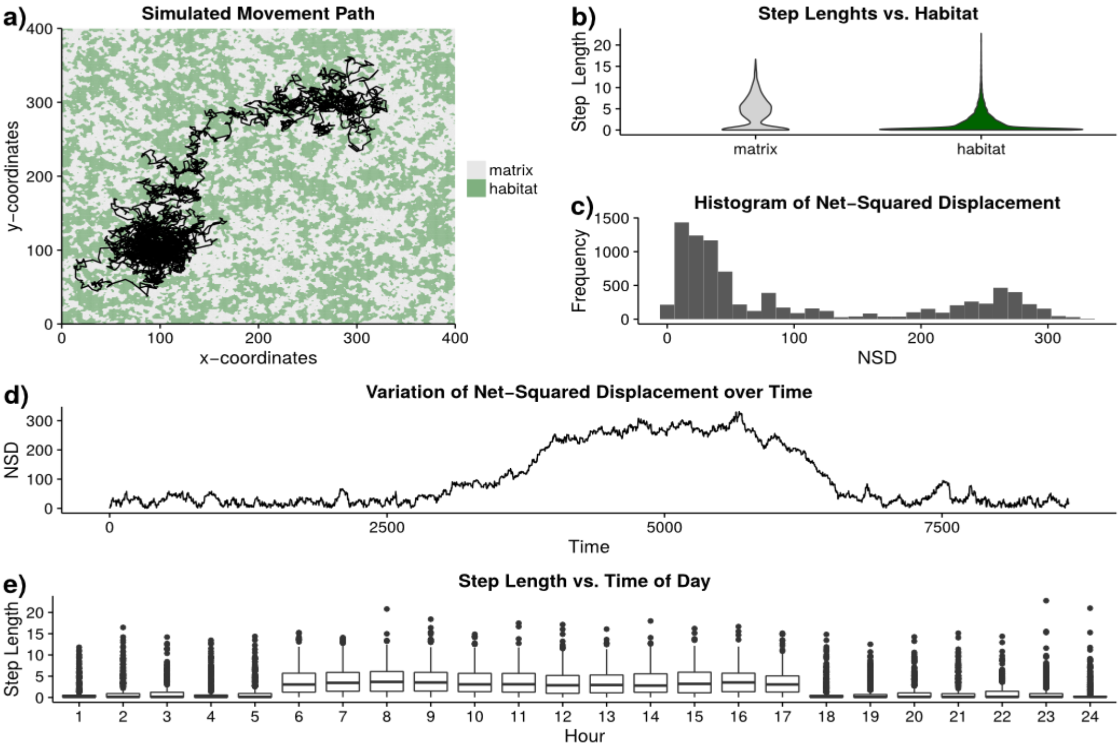


Figure 3.4: Simulated trajectory and results of preliminary analyses. a) overview of the simulated movement path and habitat configuration. b) distributions of observed step lengths within and outside the habitat (matrix) of the tracked animal. Results of preliminary analyses for the net-squared displacement signal including the distribution (c) and the time-series across the entire tracking period (d). Distributions of observed step lengths at different hours of the day (e).

For this data set, we were interested in three different research questions. First, we evaluated the hypothesis that the movement intensity of the animal somehow differed between its habitat and the (potentially hostile) matrix, sensu stricto non-habitat. To address this question, we chose a topology-based method using a threshold to distinguish short- from long-range movements and compared the proportions of these two stages within the habitat and matrix. Second, we wanted to assess whether the animal showed a seasonal migration pattern and, if so, to detect the times when migration movements occurred throughout the year. For this, we applied a time-series analysis to segment the movement data based on changes in an observed path-signal. Finally, we assessed whether two different behavioral states could be distinguished and whether the switching probability between those two states could be linked to time of day and habitat. To answer this research question, we used a state-space model approach with two discrete states differing with regard to their distributions of certain path parameters. Before addressing these research questions, we performed different preliminary analyses to gain insight about data properties and guide the decision process on meaningful path-signals and an appropriate segmentation method for each question (Figure 3.3).

Results

Preliminary analyses

As pointed out above, preliminary analyses are a fundamental part of path-level analyses and should be performed thoroughly prior to the application of any segmentation approach. As our simulated data set consisted of relocation data sampled at an hourly interval, we did not test for regularity of the sampling regime. However, such tests can be performed by inspecting the distribution of the time-lags between the sampled relocations (e.g., using histograms). More analyses for checking the regularity of a trajectory or testing the independence of missing data points are implemented in the *adehabitatLT* package (Calenge 2011). In the next step, one should test for potential correlation structures within the observed movement data. We applied different tests based on Dray et al. (2010) and detected significant correlations between consecutive measures of the step length and also turning angles up to a time lag of five relocations. Therefore, following our guidelines (Figure 3.3), we chose among methods accounting for such temporal autocorrelations.

Meaningful path parameters conveying relevant information about potential changes in movement behavior are essential for a sound path-segmentation analysis. Thus, comparisons of different signals (e.g., primary and secondary derivatives, Table 3.1) with regard to their distributions and variation over time should be performed in the preliminary analysis. We applied several exploratory analyses for the step length (due to the hourly sampling regime this is also the speed signal), turning angles and net-squared displacement (NSD) signals (more details in supplement S3.2). For example, Figure 3.4 shows the distributions of NSD and step length as well as their variation over time. The NSD signal provides meaningful information on the ranging behavior of an animal as it represents the distance to the point where the tracking period started. Inspection of this signal over the entire sampling period revealed that there was a steep increase in the values of this parameter followed by a plateau and decrease until the values were in the same range as at the beginning (Figure 3.4d). Further, we observed a trend for a bimodal distribution of NSD (Figure 3.4d). As described above, behavioral changes might be detectable at different temporal scales. Plotting the distribution of step lengths against the time of the day they were recorded revealed that the animal was potentially more active during the day as during the night (Figure 3.4e). Finally, we used all three path signals, step length (in our case equivalent with speed), turning angles and NSD for the different segmentation approaches.

Habitat-specific movement patters

We applied a thresholding method to distinguish two different movement patterns within the simulated dataset. A simple cut-off value was used to split relocations into short-range (e.g., encamped) and long-range (e.g., roaming or dispersing) movements. Relocations with an observed step length shorter than 2 units were considered short-range movements whereas those with a longer step length were classified as long-range movements. As can be seen in Figure 3.5a, the proportion of the two movement behaviors varied between habitat and non-habitat. For example, the majority of short-range movements (about 73.3%) occurred within the habitat of the animal. More than half of the movements (about 58.5%) outside the habitat stemmed from the long-range behavioral state. Further, a chi-square test indicated a significant (non-random) distribution of the two stages between habitat and non-habitat ($p < 0.001$). Clearly, results highly depend on the chosen threshold value. Therefore, cut-off values need to be chosen carefully and based on well-reasoned inferences, especially when they are applied with real movement data (see examples in Dzialak et al. 2015; Zeller et al. 2014).

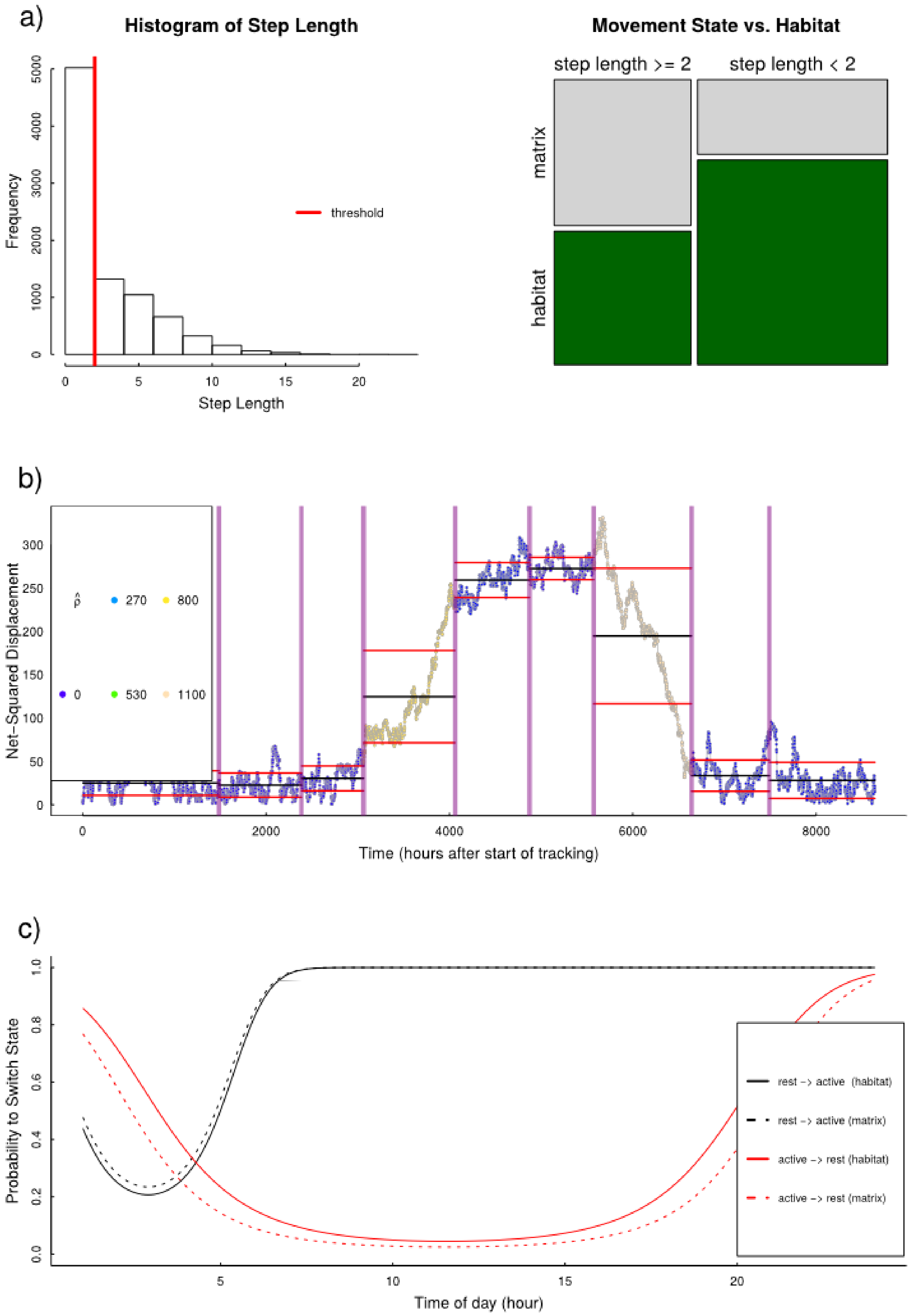


Figure 3.5 Results of three different segmentation methods using the simulated movement data. a) the left panel shows the distribution of the observed step lengths as well as the applied cut-off value (threshold = 2 units). The proportions of the resulting behavioral states (short- and long-range movements) within and outside of the habitat are shown in the right panel. b) Results from the behavioral change point analyses

applied with the net-squared displacement signal. The observed time-series was segmented at significant change-points (vertical lines) to distinguish movements within the main ranges of the animal and two migratory periods. The color of the estimated parameter ρ^ indicates the level of temporal autocorrelation. c) Change in switching probabilities between the two states (resting vs. active) dependent on the different hours of the day. Switching probabilities also differed with regard to whether the animal was in its habitat or not. Black lines indicate the switches from the resting state to the active state. Red lines are showing the switching probabilities from active to resting state.

Timing of migration

In our applied example, we chose the behavioral change-point analysis (BCPA Gurarie et al. 2009; see Table 3.2) to demonstrate how significant changes can be detected within a time-series of a path-signal in order to find segments of potential migratory behavior. We chose the sequence of the net-squared displacement parameter (NSD, Table 3.1) as the model input. As can be seen in Figure 3.5b) the BCPA determined multiple segments with comparably low net-squared displacement prior to the simulated migration event (from 0 to 3000 hours after the start of tracking). That period is followed by a segment with increasing displacement and also higher autocorrelation which can be interpreted as potentially migratory behavior. The plateau within the NSD time-series (around 4000 to 5500 hours after start of tracking) marks the arrival of the simulated animal track in its second range (e.g., summering grounds). The second migratory event is once again detected by a segment with decreasing NSD but also high autocorrelation values. Finally, the last two segments have low values of NSD comparable to the beginning indicating that the animal has returned to the first range where the tracking was started (e.g., wintering grounds). In summary, the time-series based analysis was successful at determining multiple segments, including a distinction of within-range movements from migratory movements, as well as an identification of the starting time of migration.

Underlying processes

In the third example, we addressed the question whether the switches between different movement states could be linked to two covariates, the time of the day and whether the animal was within or outside its habitat. We applied a hidden Markov model (HMM; Table 3.2) with two discrete behavioral states which differed with regard to their means of the step length and turning angle parameter distributions (more details are presented in Supplement S2). The model was fitted using the moveHMM package (Michelot et al. 2016). The first state consisted of relocations with very low step length values (mean of 0.11 units) and mostly negative turning angles. Therefore, this state was considered to represent resting or sedentary movement behavior. In contrast, the second state comprised of relocations with longer step lengths (mean of 3.4 units) and positive turning angles potentially representing active movement phases. The probability for the animal to switch from the resting to the active state was lower during the beginning of the day and increased with daytime (Figure 3.5c). The switching-probability from active to resting decreased during daytime and was higher during the night. Further, the probability to switch from resting to active was slightly higher when the animal was outside its habitat. Complementary to that, the animal was less probable to switch from active to resting when it was in non-habitat (Figure 3.5c). Overall, the model output represents the simulated movement behavior which consisted of higher movement activity during the daytime and faster movements outside the habitat. This underlines the high potential of different state-space model approaches for gaining a better understanding of the processes and mechanisms potentially driving the observed movement patterns (Gurarie et al. 2016; Patterson et al. 2008).

Discussion

The aim of movement ecology is to gain a deeper understanding of the mechanisms and ecological processes shaping organismal movement patterns and their consequences for ecology and evolution (Kays et al. 2015; Nathan et al. 2008). The methods presented here can be applied to define behavioral states from the observed movement paths and link these behavioral states to different environmental covariates to gain an enhanced understanding of the biological processes influencing the movement behavior of animals (Killeen et al. 2014; Roever et al. 2013). However, there is no single method that can be universally applied to any kind of study scenario. As illustrated above, path segmentation methods vary substantially with regard to their demands on data structure and underlying theory. Given this analytical variability, there are certainly several possibilities on how to group and categorize the different methods for path segmentation (Gurarie et al. 2016). Here, we chose to contrast different analytical approaches with regard to their applicability for answering certain research questions, rather than their underlying statistical frameworks. Nevertheless, we encourage researchers interested in applying path segmentation methods to read about the statistical details of the different methods (supplement S3.1) and consult the original method papers to fully understand the statistical properties of the method(s) they intend to apply.

We used a simulated dataset to demonstrate how our proposed decision process (Figure 3.3) can be performed to answer different research questions using methods from the three analytical categories of topology-based, time-series and state-space analyses. Certainly, each of these categories have advantages and disadvantages one has to account for when choosing among them.

The majority of methods focusing on path-topology (Table 3.2) are purely descriptive and usually just draw new observations based on the tracked movement pattern (Franke et al. 2004; Gurarie et al. 2016). However, for certain analyses this might already be sufficient to answer the defined research questions. For example, we showed how a relatively simple thresholding approach can be used to distinguish between two extrema of a potential movement behavior (short- vs long-range movements) based on a path characteristic and linked them to different habitat configurations. Thus, topology-based approaches are useful when specific hypotheses regarding movement patterns can be formulated a priori (Van Moorter et al. 2010). Also, topology-based methods are least demanding in terms of data composition and regularity, as they make no specific assumptions about data properties or the distribution of the considered path characteristics. Furthermore, they are analytically the most straightforward and can serve as exploratory tools e.g., for determining the number of potential movement states that could be further analyzed in a more inference- or process-based approach such as a SSM (Franke et al. 2004). However, these methods should not generally be applied as end-point analyses since they are mostly ignoring other valuable information like the serial autocorrelation of path parameters.

Time-series based approaches are usually more demanding with regard to data composition but provide deeper insight to significant changes in movement behaviors and account for important correlation structures present in movement data (Gurarie et al. 2009). Such methods can easily be used for finding single or multiple change-points in a trajectory to determine the moment of important changes in movement behavior.

State-space models are arguably the most powerful way for analyzing animal movement data, providing a “bottom-up” (holistic) approach where behavioral states and switching probabilities between them are modeled within the same process (Beyer et al. 2013; Jonsen et al. 2013; Patterson et al. 2008). However, the estimated state configurations are also based on certain model assumptions about the movement properties (e.g., variants of correlated random walks) and the observed pattern in the considered data (Patterson et al. 2016). Therefore, SSMs do not ultimately convey a biological meaningful differentiation between different (“true”) movement behaviors (Beyer et al. 2013; Patterson et al. 2016). Furthermore, many of the presented SSMs are quite complex and hence perhaps the most challenging to apply to empirical data. In order to foster the application of state-space models in movement ecology, we encourage biologists to cooperate with statisticians and modelers when designing studies and analyzing data. Such interdisciplinary research teams should refer to the growing number of R packages for fitting state-space models (e.g., Albertsen et al. 2015; Michelot et al. 2016; see supplement S3.1), and to the increasing number of papers providing practical advice for using these models (e.g., Jonsen et al. 2013; Patterson et al. 2008; Pedersen et al. 2011).

Finally, the majority of the presented methods of the time-series and state-space analyses are based on discrete-time models and therefore require regular sampling regimes (Figure 3.3; McClintock et al. 2014). Such data regularity is not always possible to obtain, even though various procedures reaching regular sampling are available (see above). However, there are multiple approaches using diffusion processes which model movements in continuous time and are capable of dealing with irregular data compositions (Fleming et al. 2014; Patterson et al. 2016). For example, highly infrequently sampled movement data can be analyzed using a spatial HMM with a discrete space structure (Jonsen et al. 2013; Pedersen et al. 2011). Furthermore, methods implementing continuous time processes and estimating switches between different behavioral states were presented by (Blackwell 2003; Blackwell et al. 2015; Hanks et al. 2012; Harris and Blackwell 2013; Johnson et al. 2008).

As highlighted by Gurarie et al. (2015), preliminary data analysis is a very important part of working with movement data, and we emphasize that it will often result in a much deeper understanding of observed patterns, can help to identify optimal analytical approaches for a given data set, and can eventually lead to more meaningful conclusions. A main focus should be to determine what characteristic of the movement is changing in order to choose optimal path-signals representing these changes. Further, the functional relevant time frames at which the observed species moves and potentially changes its behavior needs to be assessed carefully (Benhamou 2014; Postlethwaite and

Dennis 2013). In general, there are multiple path-signals that are commonly used for certain segmentation methods only. For example, in the literature the penalized contrast method (Lavielle 2005) is almost exclusively applied with either the first passage or residence time parameters (e.g. Barraquand and Benhamou 2008; Henry et al .2016; Le Corre et al. 2014). However, as outlined above (Table 3.1) there are multiple options for drawing information from the observed trajectory using different path parameters. We suggest that new combinations of path-signals or hybrids of different techniques might lead to valuable insights on movement behavior. For instance, instead of the typically used velocity measures for the BCPA (e.g., persistence velocity; Gurarie et al. 2009) we chose the net-squared displacement parameter as the in input signal to determine the timing of migratory behaviors in our simulated dataset. Different analytical methods can also be combined in a multi-stage approach where, in a first step, a movement path is segmented using one of the methods for detecting change-points within the movement data (e.g., a time-series approach like BCPA). In a second step, a clustering algorithm could be applied for determining groups of segments with the potentially same behavior (e.g., Step 4 in Figure 3.1). In a final step, the segments of the different clusters of movement behavior could be linked to various types of environmental data (e.g., using a step-selection analysis (Thurfjell et al. 2014; Zeller et al. 2015). For example, Zhang et al. (2015) applied such a multi-stage approach to determine a number of distinct behaviors within the movement data of little penguins (*Eudyptula minor*) and compared the location and timing of the behavioral switches between the sampled individuals. However, throughout this “top-down” process uncertainties of the chosen segmentation method are potentially projected on to the results of the subsequent analyses which could lead to biased results and interpretations. Currently, it is not clear how severe such uncertainties are for subsequent analyses and ecological inferences.

Future Research Needs

The continuing improvement of tracking devices will provide researchers with long-term movement data at high spatial and temporal resolutions (Cagnacci et al. 2010). Additionally, the establishment of collaborative projects and data collections will continue to facilitate analyses across many individuals, species, and study areas (Kays et al. 2015; Urbano et al. 2010). To fully realize the potential of this abundant high quality data, powerful analytical techniques are needed. While a substantial variety of methods for path segmentation already exists, we have only just begun to explore the analytical options for path-level movement data, and many more methods will likely be developed in the future. Ideally, these future methods will allow us to quantitatively compare multiple trajectories within and among individuals, so that we can gain a better understanding of the drivers of individual movement paths and underlying behaviors across time and space. For example, this could be accomplished by new topology-based methods using similarity comparisons (Long and Nelson 2013) and pattern recognition (Gudmundsson et al. 2004), as well as data mining of either time-series or the original trajectory data (Fu 2011; Wang et al. 2013).

Future methods should also combine path characteristics with other relevant information such as activity, metabolic and acceleration data (Brown et al. 2013) or information on body temperature derived from bio-logging devices (Bestley et al. 2010). Furthermore, the effects of habitat and weather on individual movement behavior could be incorporated into path-level analyses using high resolution environmental and climate data (Dodge et al. 2013; Sapir et al. 2014).

Regardless of how path segmentation will be improved in the future, a crucial aspect is the evaluation and comparison of available approaches, and the development of guidelines for matching methods to specific research questions. We have provided general suggestions for choosing among methods for three broad types of research questions. However, we feel that it is currently not yet possible to provide a detailed assessment of each of the listed methods we identified for path segmentation (Table 3.2). For this, it would be necessary to analyze multiple data sets with different characteristics and with different research questions in mind. While suitable data sets for this can probably be identified, we also encourage researchers to make stronger use of individual-based simulations to compare and evaluate segmentation approaches (e.g., Getz and Saltz 2008; Hooten and Wikle 2010). Such validation and accuracy assessment of different methods could also be improved by direct observations (McKellar et al. 2014), via unmanned aerial vehicles (UAVs; Ditmer et al. 2015), or other animal-born logging devices such as video cameras (Gómez-Laich et al. 2015; Moll et al. 2007).

Conclusions

Overall, future studies will likely provide a more detailed understanding of the advantages and limitations of different methods for path segmentation. However, given the complexity of segmentation analyses, and considering the variety of research questions that can be addressed with them, it is unlikely that a single method will universally be ‘best’ for all questions and data sets. Hence, while method development and evaluation are clearly crucial, the most important aspect of working with movement data is to define precise research questions (Fieberg and Börger 2012). We hope that our overview of currently available segmentation methods provides a first starting point for researchers interested in applying these approaches, so that they can dedicate even more time and energy to defining meaningful questions related to individual movement behavior.

Authors' contributions

HE prepared the overview and conceptualized the manuscript. JS and NB have substantially been involved in drafting and revising the manuscript. JS further provided the applied examples, simulated data and R code. All authors read and approved the final manuscript.

Acknowledgements

We thank Jan Engler, Femke Pflüger, Maarten Hofman, and Diana Dishman for their valuable input and discussions improving our manuscript. Furthermore, we like to thank Eliezer Gurarie and two anonymous reviewers for their constructive comments and guidance which significantly improved the direction and quality of the manuscript. We acknowledge support by the Open Access Publication Funds of Göttingen University.

References

Albertsen, C. M., Whoriskey, K., Yurkowski, D., Nielsen, A., and Flemming, J. M. (2015). Fast fitting of non-Gaussian state-space models to animal movement data via Template Model Builder. *Ecology*, *96*(10), 2598–2604.

Austin, D., McMillan, J. I., and Bowen, W. D. (2003). A Three-Stage Algorithm for Filtering Erroneous Argos Satellite Locations. *Marine Mammal Science*, *19*(2), 371–383.

Barraquand, F., and Benhamou, S. (2008). Animal movements in heterogeneous landscapes: identifying profitable places and homogeneous movement bouts. *Ecology*, *89*(12), 3336–3348.

Batschelet, E. (1981). Circular statistics in biology.

Benhamou, S. (2004). How to reliably estimate the tortuosity of an animal's path: straightness, sinuosity, or fractal dimension? *Journal of Theoretical Biology*, *229*(2), 209–220.

Benhamou, S. (2014). Of scales and stationarity in animal movements. *Ecology Letters*, *17*(3), 261–272.

Bestley, S., Patterson, T. A., Hindell, M. A., and Gunn, J. S. (2010). Predicting feeding success in a migratory predator: Integrating telemetry, environment, and modeling techniques. *Ecology*, *91*(8), 2373–2384.

Beyer, H. L., Morales, J. M., Murray, D., and Fortin, M.-J. (2013). The effectiveness of Bayesian state-space models for estimating behavioural states from movement paths. *Methods in Ecology and Evolution*, *4*(5), 433–441.

Bidder, O. R., Walker, J. S., Jones, M. W., Holton, M. D., Urge, P., Scantlebury, D. M., Marks, N. J., Magowan, E. A., Maguire, I. E., and Wilson, R. P. (2015). Step by step: reconstruction of terrestrial animal movement paths by dead-reckoning. *Movement Ecology*, *3*(1), 23.

Blackwell, P. G. (2003). Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika*, *90*(3), 613–627.

Blackwell, P. G., Niu, M., Lambert, M. S., and Lapoint, S. D. (2015). Exact Bayesian Inference for Animal Movement in Continuous Time. *Methods in Ecology and Evolution*, *12*.

Bowler, D. E., and Benton, T. G. (2005). Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, *80*(2), 205–225.

Bradshaw, C., Sims, D., and Hays, G. (2007). Measurement error causes scale-dependent threshold erosion of biological signals in animal movement data. *Ecological Applications*, *17*(2), 628–638.

Brillinger, D. R., Preisler, H. K., Ager, A. A., and Kie, J. G. (2004). An exploratory data analysis (EDA) of the paths of moving animals. *Journal of Statistical Planning and Inference*, *122*(1–2), 43–63.

Brown, D. D., Kays, R., Wikelski, M., Wilson, R., and Klimley, P. (2013). Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*, *1*(1), 20.

Buchin, M., Driemel, A., Van Kreveld, M., and Sacristan, V. (2011). Segmenting trajectories: A framework and algorithms using spatiotemporal criteria. *Journal of Spatial Information Science*, *3*(3), 33–63.

Bunnefeld, N., Börger, L., van Moorter, B., Rolandsen, C. M., Dettki, H., Solberg, E. J., and Ericsson, G. (2011). A model-driven approach to quantify migration patterns: individual, regional and yearly differences. *The Journal of Animal Ecology*, *80*(2), 466–476.

Byrne, M. E., and Chamberlain, M. J. (2012). Using first-passage time to link behaviour and habitat in foraging paths of a terrestrial predator, the racoon. *Animal Behaviour*, *84*(3), 593–601.

Byrne, R. W., Noser, R., Bates, L. A., and Jupp, P. E. (2009). How did they get here from there? Detecting changes of direction in terrestrial ranging. *Animal Behaviour*, *77*(3), 619–631.

Cagnacci, F., Boitani, L., Powell, R. A., and Boyce, M. S. (2010). Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society of London Series B - Biological Sciences*, *365*(1550), 2157–2162.

Calenge, C. (2011). Analysis of Animal Movements in R: the adehabitatLT Package

Calenge, C., Dray, S., and Royer-Carenzi, M. (2009). The concept of animals’ trajectories from a data analysis perspective. *Ecological Informatics*, *4*(1), 34–41.

Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal society interface*, *5*(25), 813-834.

Cushman, S. A. (2010). Animal movement data: GPS telemetry, autocorrelation and the need for path-level analysis. In *Spatial complexity, informatics, and wildlife conservation* (pp. 131-149). Springer, Tokyo.

Demšar, U., Buchin, K., Cagnacci, F., Safi, K., Speckmann, B., Van de Weghe, N.,Weiskopf, D., and Weibel, R. (2015). Analysis and visualisation of movement: an interdisciplinary review. *Movement Ecology*, *3*(1), 1–24.

DeRuiter, S. L., Langrock, R., Skirbutas, T., Goldbogen, J. A., Chalambokidis, J., Friedlaender, A. S., and Southall, B. L. (2016). A multivariate mixed hidden Markov model to analyze blue whale diving behaviour during controlled sound exposures. *The Annals of Applied Statistics*, *11*(1), 362–392.

Ditmer, M. A., Vincent, J. B., Werden, L. K., Tanner, J. C., Laske, T. G., Iaizzo, P. A., Garshelis, D. L., and Fieberg, J. R. (2015). Bears show a physiological but limited behavioral response to unmanned aerial vehicles. *Current Biology*, *25*(17), 2278-2283.

Dodge, S., Bohrer, G., Weinzierl, R., Davidson, S. C., Kays, R., Douglas, D., Cruz, S., Han, J., Brandes, D., and Wikelski, M. (2013). The environmental-data automated track annotation (Env-DATA) system: linking animal tracks with environmental data. *Movement Ecology*, *1*(1), 3.

Dodge, S., Weibel, R., and Lautenschütz, A.-K. (2008). Towards a taxonomy of movement patterns. *Information Visualization*, *7*(3–4), 240–252.

Dowd, M., and Joy, R. (2011). Estimating behavioral parameters in animal movement models using a state-augmented particle filter. *Ecology*, *92*(3), 568–575.

Dray, S., Royer-Carenzi, M., and Calenge, C. (2010). The exploratory analysis of autocorrelation in animal-movement studies. *Ecological Research*, *25*(3), 673–681.

Dzialak, M. R., Olson, C. V, Webb, S. L., Harju, S. M., and Winstead, J. B. (2015). Incorporating within- and between-patch resource selection in identification of critical habitat for brood-rearing greater sage-grouse. *Ecological Processes*, *4*(5).

Fauchald, P., and Tveraa, T. (2003). Using first-passage time in the analysis of area-restricted search and habitat selection. *Ecology*, *84*(2), 282–288.

Fieberg, J., and Börger, L. (2012). Could you please phrase “home range” as a question? *Journal of Mammalogy*, *93*(4), 890–902.

Fleming, C. H., Calabrese, J. M., Mueller, T., Olson, K. A., Leimgruber, P., and Fagan, W. F. (2014). Non-Markovian maximum likelihood estimation of autocorrelated movement processes. *Methods in Ecology and Evolution*, *5*(5), 462–472.

Flemming, J. M., Jonsen, I. D., Myers, R. A., and Field, C. A. (2010). Hierarchical state-space estimation of leatherback turtle navigation ability. *PLoS One*, *5*(12), e14245.

Frair, J. L., Merrill, E. H., Visscher, D. R., Fortin, D., Beyer, H. L., and Morales, J. M. (2005). Scales of movement by elk (Cervus elaphus) in response to heterogeneity in forage resources and predation risk. *Landscape Ecology*, *20*(3), 273–287.

Franke, A., Caelli, T., and Hudson, R. J. (2004). Analysis of movements and behavior of caribou (Rangifer tarandus) using hidden Markov models. *Ecological Modelling*, *173*(2–3), 259–270.

Fryxell, J. M., Hazell, M., Börger, L., Dalziel, B. D., Haydon, D. T., Morales, J. M., McIntosh, T., and Rosatte, R. C. (2008). Multiple movement modes by large herbivores at multiple spatiotemporal scales. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19114–19119.

Fu, T. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, *24*(1), 164–181.

Ganskopp, D. C., and Johnson, D. D. (2007). GPS Error in Studies Addressing Animal Movements and Activities. *Rangeland Ecology and Management*, *60*(4), 350–358.

Garstang, M., Davis, R. E., Leggett, K., Frauenfeld, O. W., Greco, S., Zipser, E., and Peterson, M. (2014). Response of African Elephants (Loxodonta africana) to Seasonal Changes in Rainfall. *PLoS ONE*, *9*(10), e108736.

Gaucherel, C. (2011). Wavelet analysis to detect regime shifts in animal movement. *Computational Ecology and Software*, *1*(2), 69–85.

Getz, W. M., and Saltz, D. (2008). A framework for generating and analyzing movement paths on ecological landscapes. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19066–19071.

Gómez-Laich, A., Yoda, K., Zavalaga, C., and Quintana, F. (2015). Selfies of Imperial Cormorants (Phalacrocorax atriceps): What Is Happening Underwater? *Plos One*, *10*(9), e0136980.

Graves, T. A., and Waller, J. S. (2006). Understanding the Causes of Missed Global Positioning System Telemetry Fixes. *Journal of Wildlife Management*, *70*(3), 844–851.

Gudmundsson, J., van Kreveld, M., and Speckmann, B. (2004). Efficient detection of motion patterns in spatio-temporal data sets. *Proceedings of the 12th Annual ACM International Workshop on Geographic Information Systems - GIS '04*, 250.

Gueguen, L. (2000). Segmentation by maximal predictive partitioning according to composition biases. In *International Conference on Biology, Informatics, and Mathematics* (pp. 32-44). Springer, Berlin, Heidelberg.

Gurarie, E., Andrews, R. D., and Laidre, K. L. (2009). A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, *12*(5), 395–408.

Gurarie, E., Bracis, C., Delgado, M., Meckley, T. D., Kojola, I., and Wagner, C. M. (2016). What is the animal doing? Tools for exploring behavioral structure in animal movements. *Journal of Animal Ecology*, *85*(1), 69-84.

Gurarie, E., and Ovaskainen, O. (2011). Characteristic spatial and temporal scales unify models of animal movement. *The American Naturalist*, *178*(1), 113–123.

Gutenkunst, R., Newlands, N., Lutcavage, M., and Edelstein-Keshet, L. (2007). Inferring resource distributions from Atlantic bluefin tuna movements: an analysis based on net displacement and length of track. *Journal of Theoretical Biology*, *245*(2), 243–257.

Hanks, E. M., Hooten, M. B., and Alldredge, M. W. (2012). Continuous-Time Discrete-Space Models for Animal Movement Data. *The Annals of Applied Statistics*, *9*(1), 145-165.

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

Henry, D. A. W., Ament, J. M., and Cumming, G. S. (2016). Exploring the environmental drivers of waterfowl movement in arid landscapes using first-passage time analysis. *Movement Ecology*, *4*(1), 8.

Hooten, M. B., and Wikle, C. K. (2010). Statistical Agent-Based Models for Discrete Spatio-Temporal Systems. *Journal of the American Statistical Association*, *105*(489), 236–248.

Hurford, A. (2009). GPS measurement error gives rise to spurious 180 degree turning angles and strong directional biases in animal movement data. *PloS One*, *4*(5), e5632.

Jerde, C., and Visscher, D. (2005). GPS measurement error influences on movement model parameterization. *Ecological Applications*, *15*(3), 806–810.

Johnson, D., and Ganskopp, D. (2008). GPS collar sampling frequency: effects on measures of resource use. *Rangeland Ecology and Management*, *61*(3), 226–231.

Johnson, D. S., London, J. M., Lea, M.-A., and Durban, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, *89*(5), 1208–1215.

Jonsen, I. D., Basson, M., Bestley, S., Bravington, M. V., Patterson, T. A., Pedersen, M. W., Thomson, R., Thygesen, U.H., and Wotherspoon, S. J. (2013). State-space models for bio-loggers: A methodological road map. *Deep Sea Research Part II: Topical Studies in Oceanography, 88–89,* 34–46.

Jonsen, I. D., Myers, R. A., and James, M. C. (2006). Robust hierarchical state-space models reveal diel variation in travel rates of migrating leatherback turtles. *Journal of Animal Ecology, 75*(5), 1046–1057.

Jonsen, I. D., Flemming, J. M., and Myers, R. (2005). Robust State-Space Modeling of Animal Movement Data. *Ecology*, *86*(11), 2874–2880.

Jonsen, I. D., Myers, R. A., and James, M. C. (2007). Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. *Marine Ecology Progress Series*, *337*, 255–264.

Kays, R., Crofoot, M. C., Jetz, W., and Wikelski, M. (2015). Terrestrial animal tracking as an eye on life and planet. *Science*, *348*(6240), aaa2478.

Killeen, J., Thurfjell, H., Ciuti, S., Paton, D., Musiani, M., and Boyce, M. S. (2014). Habitat selection during ungulate dispersal and exploratory movement at broad and fine scale with implications for conservation management. *Movement Ecology*, *2*, 13.

Laidre, K. L., Heide-Jørgensen, M. P., Logsdon, M. L., Hobbs, R. C., Dietz, R., and VanBlaricom, G. R. (2004). Fractal analysis of narwhal space use patterns. *Zoology*, *107*, 3–11.

Lange, H. (2006). Time-series Analysis in Ecology. *ELS*.

Lavielle, M. (1999). Detection of multiple changes in a sequence of dependent variables. *Stochastic Processes and Their Applications*, *83*(1), 79–102.

Lavielle, M. (2005). Using penalized contrasts for the change-point problem. *Signal Processing*, *85*(8), 1501–1510.

Le Corre, M., Dussault, C., and Côté, S. D. (2014). Detecting changes in the annual movements of terrestrial migratory species: using the first-passage time to document the spring migration of caribou. *Movement Ecology*, *2*(1), 19.

Liu, Y., Battaile, B. C., Trites, A. W., and Zidek, J. V. (2015). Bias correction and uncertainty characterization of Dead-Reckoned paths of marine mammals. *Animal Biotelemetry*, *3*(51), 1:11.

Lonergan, M., Fedak, M., and McConnell, B. (2009). The effects of interpolation error and location quality on animal track reconstruction. *Marine Mammal Science*, *25*(2), 275–282.

Long, E. S., Jacobsen, T. C., Nelson, B. J., and Steensma, K. M. (2013). Conditional daily and seasonal movement strategies of male black-tailed deer (Odocoileus hemionus columbianus). *Canadian Journal of Zoology*, 91, 679-688.

Long, J. A., and Nelson, T. A. (2013). A review of quantitative methods for movement data. *International Journal of Geographical Information Science*, *27*(2), 292–318

Lowther, A. D., Lydersen, C., Fedak, M. a., Lovell, P., and Kovacs, K. M. (2015). The Argos-CLS Kalman Filter: Error Structures and State-Space Modelling Relative to Fastloc GPS Data. *Plos One*, *10*(4), e0124754.

Madon, B., and Hingrat, Y. (2014). Deciphering behavioral changes in animal movement with a “multiple change point algorithm- classification tree” framework. *Frontiers in Ecology and Evolution*, *2*, 1–9.

McClintock, B. T., Johnson, D. S., Hooten, M. B., Ver Hoef, J. M., and Morales, J. M. (2014). When to be discrete: the importance of time formulation in understanding animal movement. *Movement Ecology*, *2*(1), 21.

McKellar, A. E., Langrock, R., Walters, J. R., and Kesler, D. C. (2014). Using mixed hidden Markov models to examine behavioral states in a cooperatively breeding bird. *Behavioral Ecology*, *26*(1), 148-157.

Michelot, T., Langrock, R., and Patterson, T. A. (2016). moveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. *Methods in Ecology and Evolution*, *7*(11), 1308-1315.

Moll, R. J., Millspaugh, J. J., Beringer, J., Sartwell, J., and He, Z. (2007). A new “view” of ecology and conservation through animal-borne video systems. *Trends in Ecology and Evolution*, *22*(12), 660–668.

Moorcroft, P. R., and Barnett, A. (2008). Mechanistic home range models and resource selection analysis: a reconciliation and unification. *Ecology*, *89*(4), 1112–1119.

Morales, J., and Ellner, S. (2002). Scaling up animal movements in heterogeneous landscapes: the importance of behavior. *Ecology*, *83*(8), 2240–2247.

Morales, J., Haydon, D., and Frair, J. (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, *85*(9), 2436–2445

Nams, V. (1996). The VFractal: a new estimator for fractal dimension of animal movement paths. *Landscape Ecology*, *11*(5), 289–297.

Nams, V. O. (2014). Combining animal movements and behavioural data to detect behavioural states. *Ecology Letters*, 1228–1237.

Nathan, R., Getz, W. M., Revilla, E., Holoyak, M., Kadmon, R., Saltz, D., and Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, *105*(49), 19052–19059.

Owen-Smith, N., Fryxell, J. M., and Merrill, E. H. (2010). Foraging theory upscaled: the behavioural ecology of herbivore movement. *Philosophical Transactions of the Royal Society of London Series B - Biological Sciences*, *365*(1550), 2267–2278.

Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L., King, R., and King, R. (2016). Statistical modelling of animal movement: a myopic review and a discussion of good practice. *ArXiv Preprint, arXiv:1603*.

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. (2008). State-space models of individual animal movement. *Trends in Ecology and Evolution*, *23*(2), 87–94.

Pedersen, M. W., Berg, C. W., Thygesen, U. H., Nielsen, A., and Madsen, H. (2011). Estimation methods for nonlinear state-space models in ecology. *Ecological Modelling*, *222*(8), 1394–1400.

Polansky, L., Douglas-Hamilton, I., and Wittemyer, G. (2013). Using diel movement behavior to infer foraging strategies related to ecological and social factors in elephants. *Movement Ecology*, *1*(1), 13.

Postlethwaite, C. M., Brown, P., and Dennis, T. E. (2013). A new multi-scale measure for analysing animal movement data. *Journal of Theoretical Biology*, *317*, 175–185.

Postlethwaite, C. M., and Dennis, T. E. (2013). Effects of temporal resolution on an inferential model of animal movement. *PloS One*, *8*(5), e57640.

Powell, R. (2000). Animal home ranges and territories and home range estimators. In L. Boitani and T. K. Fuller (Eds.), *Research Techniques in Animal Ecology : Controversies and Consequences* (1st ed., p. 476).

R Core Team. (2017). *R: A language and environment for statistical computing*. Retrieved from http://www.r-project.org/

Roever, C. L., Beyer, H. L., Chase, M. J., and van Aarde, R. J. (2013). The pitfalls of ignoring behaviour when quantifying habitat selection. *Diversity and Distributions*, *20*(3), 322–333.

Sapir, N., Horvitz, N., Dechmann, D. K. N., Fahr, J., and Wikelski, M. (2014). Commuting fruit bats beneficially modulate their flight in relation to wind. *Proceedings of the Royal Society B*, *281*(1782), 20140018.

Schick, R. S., Loarie, S. R., Colchero, F., Best, B. D., Boustany, A., Conde, D. A., Halpin, P.N., Joppa, L. N., McClellan, C. M., and Clark, J. S. (2008). Understanding movement data and movement processes: current and emerging directions. *Ecology Letters*, *11*(12), 1338–1350.

Shamoun-Baranes, J., van Loon, E. E., Purves, R. S., Speckmann, B., Weiskopf, D., and Camphuysen, C. J. (2012). Analysis and visualization of animal movement. *Biology Letters*, *8*(1), 6–9.

Sibert, J. R., Musyl, M. K., and Brill, R. W. (2003). Horizontal movements of bigeye tuna (Thunnus obesus) near Hawaii determined by Kalman filter analysis of archival tagging data. *Fisheries Oceanography*, *12*, 141–151.

Silva, M. A., Jonsen, I., Russell, D. J. F., Prieto, R., Thompson, D., and Baumgartner, M. F. (2014). Assessing performance of bayesian state-space models fit to argos satellite telemetry locations processed with kalman filtering. *PloS One*, *9*(3), e92277.

Sur, M., Skidmore, A. K., Exo, K.-M., Wang, T., J. Ens, B., and Toxopeus, A. G. (2014). Change detection in animal movement using discrete wavelet analysis. *Ecological Informatics*, *20*, 47–57.

Thiebault, A., and Tremblay, Y. (2013). Splitting animal trajectories into fine-scale behaviorally consistent movement units: breaking points relate to external stimuli in a foraging seabird. *Behavioral Ecology and Sociobiology*, *67*(6), 1013–1026.

Thurfjell, H., Ciuti, S., and Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, *2*(1), 4.

Urbano, F., Cagnacci, F., Calenge, C., Dettki, H., Cameron, A., and Neteler, M. (2010). Wildlife tracking data management: a new vision. *Philosophical Transactions of* *the Royal Society of London. Series B, Biological Sciences, 365*(1550), 2177–2185.

van Beest, F. M., and Milner, J. M. (2013). Behavioural responses to thermal conditions affect seasonal mass change in a heat-sensitive northern ungulate. *PloS One, 8*(6), e65972.

van de Kerk, M., Onorato, D. P., Criffield, M. A., Bolker, B. M., Augustine, B. C., McKinley, S. A., and Oli, M. K. (2014). Hidden semi-Markov models reveal multiphasic movement of the endangered Florida panther. *The Journal of Animal Ecology*.

Van Moorter, B., Visscher, D. R., Jerde, C. L., Frair, J. L., and Merrill, E. H. (2010). Identifying Movement States From Location Data Using Cluster Analysis. *Journal of Wildlife Management*, *74*(3), 588–594.

Wang, Y., Luo, Z., Qin, G., Zhou, Y., Guo, D., and Yan, B. (2013). Mining Common Spatial-Temporal Periodic Patterns of Animal Movement. *2013 IEEE 9th International Conference on E-Science*, 17–26.

Weber, T., and Norman, J. (2015). Functional connectivity modeling and optimal siting of conservation networks in the Midwest USA. *Ecological Informatics*. 30, 277-283.

Wensveen, P. J., Thomas, L., and Miller, P. J. O. (2015). A path reconstruction method integrating dead-reckoning and position fixes applied to humpback whales. *Movement Ecology*, *3*(1), 31.

Williams, D. M., Dechen Quinn, A., and Porter, W. F. (2012). Impact of habitat-specific GPS positional error on detection of movement scales by first-passage time analysis. *PloS One*, *7*(11), e48439.

Wilson, R. P., Liebsch, N., Davies, I. M., Quintana, F., Weimerskirch, H., Storch, S., Lucke, K., Siebert, U., Zankl, S., Müller, G., Zimmer, I., Scolaro, A., Campagna, C., Plötz, J., Bornemann, H., Teilmann, J., and McMahon, C. R. (2007). All at sea with animal tracks; methodological and analytical solutions for the resolution of movement. *Deep Sea Research Part II: Topical Studies in Oceanography*, *54*(3–4), 193–210.

Wilson, R. R., Gilbert-Norton, L., and Gese, E. M. (2012). Beyond use versus availability: behaviour-explicit resource selection. *Wildlife Biology*, *18*(4), 424–430.

Wittemyer, G., Polansky, L., Douglas-Hamilton, I., and Getz, W. M. (2008). Disentangling the effects of forage, social rank, and risk on movement autocorrelation of elephants using Fourier and wavelet analyses. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19108–19113.

Worton, B. J. (1987). A review of models of home range for animal movement. *Ecological Modelling*, *38*(3–4), 277–298.

Zeller, K. A., McGarigal, K., Beier, P., Cushman, S. A., Vickers, T. W., and Boyce, W. M. (2014). Sensitivity of landscape resistance estimates based on point selection functions to scale and behavioral state: pumas as a case study. *Landscape Ecology*, *29*(3), 541–557.

Zeller, K. A., McGarigal, K., Cushman, S. A., Beier, P., Vickers, T. W., and Boyce, W. M. (2015). Using step and path selection functions for estimating resistance to movement: pumas as a case study. *Landscape Ecology*. *31*(6), 1319-1335.

Zhang, J., O'Reilly, K. M., Perry, G. L. W., Taylor, G. A., and Dennis, T. E. (2015). Extending the Functionality of Behavioural Change-Point Analysis with k-Means Clustering: A Case Study with the Little Penguin (Eudyptula minor). *Plos One, 10*(4), e0122811.

Supplement

S3.1: Description of individual segmentation methods

Introduction

Here we provide more detailed information on available methods for path segmentation listed in the publication. Further, we cite literature with applied examples to illustrate the utility of the different methods. Table S3.1.1 summarizes basic statistical properties of the discussed methods and lists background papers and availability of code for implementing methods in the program *R* (R Core Team 2015).

As outlined in the main article, the presented methods could generally be distinguished based on their analytical background. For a better overview, we assigned the presented methods to three different categories based on whether they focus predominantly on path-topology or apply different time-series based analyses. Within the latter, one can further distinguish state-space modeling approaches from other general time-series analyses which focus on the detection of significant changes in within a time-ordered data sequence.

Table S3.1.1: Statistical characteristics of the different methodological approaches within the three categories of segmentation methods SI indicates the required sampling interval, which can be either irregular (-), strictly regular (+) or both (-/+). AC provides information on whether a method accounts for (+), neglects (-) or only partly implements (~) estimates of temporal autocorrelation. Further, for each method an outline of the analytical approach, necessary specifications (e.g., parameter settings) and the generated output (results) are listed.

Table shown on the following pages.

Category	Method	SI	AC	Statistical Approach	Specifications	Result	References	R Implementation
Topology-Based Analyses	Thresholding	-/+	-	Indexing/classifying each relocation based on thresholding scheme	Thresholding scheme (absolute or relative threshold values for single or multiple path parameters)	assigned index of each relocation	Tremblay et al. 2007, Dodge et al. 2008, Sur et al. 2014, Gutenkunst et al. 2007	see Supplement S2
	Change Point Test	-/+	-	Permutation test for detecting significant change points in movement direction between a starting point and an attraction point	subtrajectory defining the potential point of attraction and a starting point; parameters for permutation test	positions of significant change points	Byrne et al. 2009, Noser and Byrne 2013	R code provided in Byrne et al. 2007
	Spatio-Temporal Criteria Segmentation	-/+	-	Search algorithm for optimal segmentation based on monotone criteria; segment size is increased as long as predefined criteria are fulfilled	monotone criteria (based on path parameter)	positions of significant change points	Buchin et al. 2013, Buchin et al. 2011	NA, theoretical and MATLAB code provided in Buchin et al. 2013
	Supervised Classification	-/+	-	Regression trees, vector support machines, random forests, machine learning algorithms; tests for assessing classification accuracy	Multivariate set of path parameters (signals); training set for fitting the classification algorithm	assigned class of each relocation	Dodge et al. 2009, Shamoun-Baranes et al. 2012, Soleymani et al. 2014	e.g. *e1071*(Meyer et al. 2015), *tree* (Ripley, 2015) packages
	Clustering	-/+	-	K-means Clustering, Expectation-Maximization Binary Clustering; e.g. gap-statistics for determining optimal number of clusters	Multivariate set of path parameters (signals); number of clusters, tolerance value	assigned cluster of each relocation	Van Moorter et al. 2010, Garriga et al. 2016	*EMbC* (Garriga and Bartumeus), *cluster* (Maechler et al. 2015) packages
	Line Simplification	-/+	-	Reduction of the trajectory dimension by removing relocations which do not add to variation e.g. using the Douglas-Peucker Algorithm	tolerance parameter defining the sensitivity for removing a relocation	position of significant change points (i.e., remaining relocations)	Douglas and Peucker 1973, Thiebault and Tremblay 2013	*Rgeos* package (Bivand et al. 2015)
Time-Series Analyses	Bayesian Partitioning of Markov Models (BPMM)	+	~	Partitioning and classification algorithm for determining the number and sequence of homogenous processes (models) within a signal (time series) using randomized likelihood estimation	input path signal (time series); number and distributions of candidate models (processes)	estimated state of each relocation and distributions of state sequences	Gurarie et al. 2016. Gueguen 2000, Calenge 2011	adehabitatLT (Calenge, 2011) package
	Piecewise Regression	-/+	+	Split linear (time) series into most representative segments and fit a polynomial model for each segment	incorporated path signal and model settings	positions of significant change points	Johnson et al. 2002, Liminana et al. 2007	*segmented* (Muggeo, 2015)package

Time-Series Analyses	Penalized Contrast Method (PCM)	+	~	Non-parametric segmentation of signal; the unknown number of segments is estimated by minimizing a penalized contrast function	input either based on mean, var or mean and var of input signal; minimum segment length; maximum number of change points	positions of significant change points	Lavielle 2005, Lavielle 1999, Le Corre et al. 2014	adehabitatLT (Calenge, 2011) package
	Behavioral Change Point Analysis (BCPA)	-/+	+	Likelihood-based method for detecting significant change points; applies moving window over continuous autocorrelated time series of input signal	input path signal (time series); window size; sensitivity parameter (BIC restriction); minimum number of change point detections	positions of significant change points	Gurarie et al. 2016, Gurarie et al. 2009	bcpa (Gurarie, 2015) package
	Pruned Exact Linear Time (PELT) Algorithm	+	~	Pruning search method for detecting optimal number and locations of change points minimizing given cost and linear penalty functions	input path signal; type of expected data changes (mean, var, mean and var); penalty method (SIC, BIC)	positions of significant change points	Madon and Hingrat 2014, Killick et al. 2011	changepoint (Jonsen, Myers and Fleming, 2003)
	Behavioral Movement Segmentation (BMS)	+	~	Combined search algorithm using BIC values for optimizing segmentation based on parsimony and subsequent K-means clustering for assigning segments to similar behaviors	minimum segment length, sampling resolution, number of behavioral clusters	path-segments and their behavioral classification	Nams 2014	Mathematical code available from author (Nams, 2014)
State-Space Models	Hidden-Markov Models (HMM)	+	+	Expectation-Maximization, maximum likelihood or Bayesian likelihood estimation; Viterbi and Baum-Welch algorithms for estimating state sequences	number of hidden states, input path parameters, assumed parameter distributions	assigned hidden state of each relocation	Franke et al. 2004, Morales et al. 2004, Lanrock and Zucchini 2011	*moveHMM* (Michelot et al. 2015); R code provided in (Beyer et al. 2013; Patterson et al. 2009)
	State-Space Models with Location Filtering	-/+	+	Bayesian Likelihood Estimation, Particle Filtering, Kalman Filtering	distributions of path parameters for random walk models, number of states, distributions of switching probabilities (state-space model structure, nr. of particles, weighting scheme)	assigned hidden model of each relocation; estimated parameters of movement model and states	Breed et al. 2012, Jonsen et al. 2013, Patterson et al. 2008, Forester et al. 2007, Dowd and Joy 2011	bsam (Jonsen, 2014), argosTrack (Albertsen et al.2015) packages, R code provided in (Albertsen et al.2015; Jonsen, Flemming and Myers, 2005; Zeller et al. 2015)

Thresholding

In this most basic segmentation approach, movement modes can be defined directly from observed values of path-signals. Commonly, a set of thresholds is needed as a filtering scheme to separate the relocations into different groups of movement behavior. In many cases, thresholding is used to partition path-signals into either high or low values (Franke et al. 2004; Zeller et al. 2015), or to differentiate between localized and long-range movements (Tremblay et al. 2007; Dzialak et al. 2015). The applied thresholds can either be absolute or relative values based on certain observations or hypotheses. The selection of one or more path-signals is mostly based on the research question and data resolution and could be any kind of spatial or temporal property of the movement track (primary or secondary derivatives, see *Table 3.1* in publication). Further, no data regularity is required in case that signals of relative displacement (e.g. velocity or persistence velocity) are chosen. *Absolute* thresholds usually constitute a cut-off value where a signal is split into two different groups. For example, Zeller et al. (2015) defined relocations with a step length less than 200m as “resource use” whereas a step length larger than this threshold was interpreted as actual “movement” (e.g., dispersal). Similarly, Gutenkunst et al. (2007) applied a low-pass filter on the ratio between the net-squared displacement and the total length of a movement track of Atlantic bluefin tuna (*Thunnus thynnus*). A predefined threshold of this ratio was used to distinguish localized from long-ranged movements. In contrast to that, *relative* thresholds are often based on the distribution of the considered path-signal, for example by testing whether the observed values are higher or lower than those contained within the 95% confidence interval across all observations (Sur et al. 2014).

Thresholding can also be extended to multiple path-signals summarized around one or more relocations (e.g., using a moving-window or circular neighborhood) which in the next step are classified according to a thresholding scheme. For example, LaPoint et al. (2013) identified potential corridor use behavior of fishers (*Martes pennanti*) based on multiple relocations which were parallel and comparably linear in direction at a certain speed. To calculate this kind of parallelism of multiple movements, they introduced a path parameter called “pseudo-azimuth” (Table 3.1 in the main manuscript) which is based on a buffer around midpoints between consecutive relocations.

Supervised Classification

These algorithms have been applied to assign relocations (steps) to different classes of movement behavior based on multiple path-signals. For this, individual steps of a subset of available data (e.g., a training dataset) are assigned to certain classes of movement behavior either visually or by applying a threshold approach as described above. The remaining data sets are then fitted to this classification scheme using either decision trees (Soleymani et al. 2014), support vector machines (Dodge et al. 2009) or classification trees (Shamoun-Baranes et al. 2012).

Clustering

Clustering can be regarded as a type of unsupervised classification, where no training data is used to define the groups that the data should be assigned to. In the context of movement data, clustering methods aim to identify distinctive groups within a multivariate set of path-signals without any prior assumptions on the underlying processes (Van Moorter et al. 2010). For cluster analyses in general, test statistics have been developed to assess classification accuracy and to find the optimal value for the *k* number of clusters that should be distinguished (e.g. Steinley 2006; Legendre and Legendre, 2012). Further, algorithm performance will depend on the distribution of the parameter values used for clustering (usually one ore multiple path-signals). For example, Van Moorter et al. (2010) used a classic *k*-means clustering approach with several parameters including step-length, turning angles and activity data to group movements of elk (*Cervus elaphus*) into within and between feeding patch behaviors. The expectation-maximization binary clustering algorithm (EmBC; Garriga et al. 2016) was used by Louzao et al. (2015) to distinguish four different behavioral modes in the movements of wandering albatrosses (*Diomedea exulans*). This method essential splits the relocations into different groups based on a combination of either high or low values for two different path-signals (e.g., speed and turning angles).

Spatio-Temporal Criteria Segmentation

This special type of thresholding relies on a search algorithm that extends an initial segment as long as path-signals at each step fulfill a certain criterion (Buchin et al. 2011, 2013). Thus, the approach essentially attempts to obtain an optimal segmentation of a trajectory, in terms of a minimum number of homogeneous segments. For example, path-signals can be compared to predefined ranges of values such as upper and lower bounds for movement speeds or directional changes that are expected to occur during known movement modes (Buchin et al. 2013). Consecutive steps are then included in the same segment as long as observed values fall within these bounds, but separated if outside of expected values. In contrast to simple thresholding, spatio-temporal criteria segmentation is based on the concept of monotone criteria, which means that within each segment defined by certain criteria, any subsegment must also fulfill the criteria (e.g., speed and heading within predefined bounds). For example, Buchin et al. (2013) applied this algorithm to differ segments of migration flights from stopovers within the trajectories of white-fronted geese (*Anser albifrons*). The monotone criterion for a segment to be determined as migration flight behavior was that all consecutive relocations had to have bounded headings (angles) of around 120°. Therefore, these segments comprised of approximately linear movements. In contrast to that, segments were identified as stopovers when they fulfilled the criterion of containing relocations that where within a disk (radius) of 30km and remained within this disk for a duration of minimum of 48 hours (Buchin et al. 2013). Finally, the change-points, where the trajectories switched between one of these behavioral states were detected and linked to their recorded timing of the year.

Change Point Test

The method detects significant changes in the observed movement direction or orientation of a trajectory (Byrne et al. 2007). For this, a subset of the trajectory based on a potential attraction point (e.g., food source) and the previous relocations back to a starting point (e.g., den or roosting spot) is used as an input. Each of the previous relocations prior to the attraction point is tested “backwards in time” for a change in total direction (Byrne et al. 2007). The collinearity of the movement vectors before and after a potential change-point are calculated to assess whether movements after a given point are aligned with movements before that point. The significance of the change in directionality is tested using a permutation test, which avoids any assumptions about the distributions of turning angles (Byrne et al. 2007). The approach is most useful when attraction points can be defined *a priori*. For example, Noser and Byrne (2014) applied the change point test to daily travel routes of baboons (*Papio ursinus*) and were able to identify locations where the animals decided to return back towards their sleeping sites, and locations were they adjusted their movements due to important landmarks (e.g., change of topographic slope or tire tracks).

Line Simplification

Line simplification is an approach commonly used in cartography and geographic information science to reduce the number of vertices in geometric objects while maintaining their basic structure (Saalfeld 1999; Douglas and Peucker 1973). For movement data, this method can be applied to test whether simplifying a trajectory by deleting relocations has a significant impact on the topology of the trajectory. Consecutive relocations that do not change path-topology when being removed can be grouped into the same segment. In contrast, change-points are indicated if their exclusion strongly alters path-topology. As the most prevalent method, the Douglas-Peucker algorithm (Douglas and Peucker 1973) excludes points which do not add variation along a simplified line between two non-consecutive relocations. For example, Thiebault and Tremblay (2013) used this algorithm to segment movement paths of Cape gannets (*Morus capensis*) by calculating the distance between the original path and the simplified, straight line connection of relocations before and after a potential change-point has been removed. If the distance between true and simplified paths was larger than a specified threshold, a change-point was detected. Since small threshold values lead to small-scale segmentation, and high values to broader-scale segmentation, multiple threshold values should be assessed and compared (Theibault and Tremblay 2013). The cited example shows that line simplification can also be applied to segment time-ordered data, such that these methods are at the convergence between the two categories of topology-based and time-series analyses.

Bayesian Partitioning of Markov Models (BPMM)

This algorithm can also be interpreted as a hybrid between a method focusing on path-topology on the one hand and accounting for sequential time-series data on the other hand. It is originally derived from a DNA classification method developed by Guéguen (2001) and applies randomized likelihood estimation for determining the optimal number and sequence of a list of candidate Markov models (Calenge 2011). The input path-signal for the candidate models needs to be ordered in time and derived from a regular trajectory. The input data could be any primary or secondary signal conveying spatial or temporal information (e.g., step length). The candidate models, for example, could have Gaussian distributions with a range of different means while keeping a constant variance (Gurarie et al. 2016). As a result, the trajectory is split into homogeneous segments based on the optimal sequence of Markov models. The BPMM method assumes that the path-signals within these segments are independent, an assumption that is often violated for movement data (Gurarie et al. 2016). Additionally, for each relocation (step) the associated candidate model is specified (Calenge 2011). Therefore, BPMM has been viewed as a sophisticated classification algorithm (e.g., Gurarie et al. 2016). However, the list of candidate models could also be interpreted as "hidden states" and therefore we point out that this method could also be potentially applied for identifying hidden processes (see Table 3.2 of main article).

Methods based on time-series analyses

Piecewise Regression

This approach is also termed "broken-stick" or "segmented" regression and is essentially a type of curve fitting (Neter et al. 1985). Basically, the approach finds breakpoints where the relationships between the dependent variable and the independent variable change abruptly. The data are then split at these breakpoints and a separate regression line is fit in each interval. For movement data, the dependent data is a path-signal of interest (e.g., primary or secondary descriptors like net-squared displacement), which is analyzed as a function of time. Detected breakpoints can be interpreted as a change in movement behavior, so that the trajectory can be segmented at that given point in time. For example, Liminana et al. (2007) used piecewise linear regression to detect the start of the migratory phase in the movement paths of Montagu's harriers (*Circus pygargus*). Similar, non-linear types of curve-fitting approaches have been used to determine breaks regarding individual scales of movements (Johnson et al. 2002; Sibly et al. 1990; Saher and Schmiegelow 2005).

Penalized Contrast Method (PCM)

This method developed by Lavielle (1999, 2005) has been widely applied in animal movement analyses (e.g. Sur et al. 2014; Barraquand and Benhamou 2008). The optimal number of segments is determined by minimizing a contrast function which rates the differences between signals of the entire trajectory versus the signals of the segmented series. The contrast functions are either based on the mean, standard deviation or a combination of both. The method implies that the contrast function decreases with increasing numbers of segments in the series (Lavielle 1999). In order to avoid visual (and potentially subjective) splitting of the trajectory, Lavielle (2005) proposed to use the second derivative of the contrast function and the value at which it reaches a certain threshold. Le Corre et al. (2014) used this approach to objectively determine departure and arrival dates in migration patterns of caribou, (*Rangifer tarandus).* In addition, the method requires the definition of the minimum length of resulting segments to avoid over-splitting, and a maximum amount of possible segments in order to limit processing time (Calenge 2011). PCM is also less susceptible to biases from temporal autocorrelation (Lavielle 1999, Barraquand and Benhamou 2008). The majority of studies applying the PCM algorithm used either the first passage- or residence-time (see Table1 of publication) as the input signal (e.g., Sommerfeld et al. 2013). However, potentially any primary or secondary signal conveying spatial or temporal information on movement properties (e.g., step length) could be applied.

Behavioral Change Point Analysis (BCPA)

The behavioral change point analysis introduced by Gurarie et al. (2009) consists of several consecutive analytical steps. First, either the persistence or turning velocity is chosen as the input signal as these parameters are less sensitive to irregular sampling (see Table 3.1 of publication). The signal is modeled as a continuous autocorrelated time-series with three local components (e.g., mean, variance, and temporal autocorrelation). In a second step, the likelihood of a significant change-point within the three local parameters is estimated for a subsample (window) of the time-series (Gurarie et al. 2009). Subsequently, the window is moved forward along the entire time-series. Whether a relocation is a change-point is then evaluated based on a Bayesian Information Criterion (BIC; Burnham and Anderson 2012), which compares different model assumptions ranging from a null-model (no changes) to one, any two, or all three parameters changing at a potential change point. The BCPA does not depend on regular sampling and is able to cope with missing data because primary descriptive features of movements are captured in the velocity signal and the continuous-time modeling framework (Gurarie et al. 2009; Zhang et al. 2015). However, a set of multiple input parameters, such as the window size and the minimum number of detections of each change-point, have to be set prior to the analysis (see Table S1). For example, Zhang et al. (2015) applied the BCPA to detect different behaviors (e.g., foraging) in the movement tracks of little penguins (*Eudyptula minor*).

Pruned Exact Linear Time (PELT) Algorithm

This algorithm searches for an optimal combination of the number and locations of change-points along a time-series (Killick et al. 2011). Simply put, the algorithm treats the segmentations produced by different change-points as competing models, and assesses which model best fits the mean, variance, or a combination of both within the produced segments. Optimality of any set of change-points is defined by a cost function that needs to be minimized and with a penalty term to avoid over-splitting (for example via BIC). Madon and Hingrat (2014) used the PELT algorithm to segment movement paths of Macqueen's bustards (*Chlamydotis macqueenii*) and subsequently classified the identified segments into migratory, non-migratory, and staging movements. Similar to the BCPA, the PELT approach is able to detect a set of change-points in an individual movement signal without any *a priori* knowledge on the total number of behavioral modes and switches (see Table S1). Any primary or secondary derivative of path-signals could be used as an input. However, the data is assumed to follow a normal distribution with constant mean, and relocations are assumed to be independent (non-autocorrelated).

Behavioral Movement Segmentation (BMS)

This approach characterizes a behavioral state by a specific mean for one or several path-signals which can be estimated from the data (e.g., any primary or secondary derivative parameter or even acceleration data; Nams 2014). The positions of change-points are also treated as a parameter that can be estimated from the data. The BMS approach attempts to find the most parsimonious set of these two parameters and again uses the BIC (Burnham and Anderson 2012) to quantify parsimony. For this, a series of different combinations for the number of segments and number of behavioral states is compared and the combination with lowest BIC is chosen. A cluster analysis is then performed to group similar segments and infer behavioral states. One major advantage of the approach is that the estimation of the most likely number and location of behavioral switches can easily be extended to include data other than movement signals. Additionally, the grouping of the resulting segments into clusters of potentially similar movement behaviors is already implemented within the algorithm. For example, Nams (2014) combined GPS-relocations and accelerometer (activity) data within the BMS approach to analyze movement behavior of a fisher (*Martes pennant*). When only using movement speed and turning angles to distinguish behavioral states, four different movement stages could be identified. When additionally accounting for acceleration data, seven behavioral stages could be distinguished, revealing greater complexity in movement behavior than could be inferred from the trajectory alone.

Methods based on state-space models

Methods within this category stem from the broad state-space modeling (SSM) framework. From a statistical perspective, state-space models are special types of time-series analyses also accounting for the correlation structure of consecutive measurements (Patterson et al. 2008). In general, within this framework the future state of a system is estimated from its previous state(s) through a probabilistic model. For this, two stochastic time-series models, one based on an unobservable state process, and another based on a known observation process are coupled (Jonsen et al. 2013; Patterson et al. 2008). SSMs differ with regard to the number and composition of the state variables (e.g., discrete vs. continuous), the statistical estimation technique as well as the structure of the main components, the two stochastic process and observation models. Nomenclature for differentiating is unfortunately inconsistent in the literature (e.g., Jonsen et al. 2013; Patterson et al. 2008; Gurarie et al. 2016; Patterson et al. 2016).

In order to provide better guidance for deciding among different SSMs we distinguished three general classes of state-spaces modeling approaches. For example, Hidden Markov models are based on a predefined number of discrete states and typically neglect observation errors within the data. In contrast to that, state-space models can also be extended to include a location filtering component essentially estimating probabilities of different parameters of the movement process including the probable relocations of error-prone movement data. Further, they can work with an undefined number of either discrete or even continuous behavioral states and fit various movement models (e.g., different variants of a correlated random walk; Gurarie et al. 2016; Patterson et al. 2008, 2016). Therefore, certain state-space models can also be applied in hierarchical and meta-analyses accounting for individual variations in the number and composition of the hidden states (Jonsen et al. 2003; Jonsen et al. 2006; Eckert and Moore 2008; Flemming et al. 2010).

Hidden Markov Models (HMM)

These are special cases of SSMs that estimate properties of a fixed set of discrete hidden states (Franke et al. 2004; Zucchini et al. 2008). In an HMM, state transitions are usually driven by first order Markovian processes, which means that a state depends only on the previous state. However, State transitions in HMMs can be modified, so that the switching probability can also depend on several previous states or their durations (Patterson et al. 2009; Langrock et al. 2012), environmental and social factors (Bergman et al. 2008), as well as habitat data (Morales et al. 2004; Beyer et al. 2013). Some HMMs integrate extensions of the random walk framework (see Codling et al. 2008 for more details) as part of their process model (Gurarie et al. 2016; Patterson et al. 2016). For example, Morales et al. (2004) used a Bayesian approach to model movements of elk (*Cervus elaphus*) as a mixture of different random walks, and found that elk movements were either exploratory or encamped, with the latter occurring in open habitats during foraging. Recently, HMMs have also been applied to model behavior based on bio-logger data (e.g., Patterson et al. 2009; Dean et al. 2013). In summary, the focus of HMM methods is on the estimation of switching probabilities between states, the most likely sequence of the hidden states as well as their length (Franke et al. 2004; Zucchini et al. 2008). Parameters in HMMs can be estimated through various statistical techniques, including expectation-maximization (Franke et al. 2004; Rabiner 1989), likelihood-maximization (Patterson et al. 2009; Zucchini et al. 2008; Langrock, et al. 2015) or Bayesian likelihood estimation (Morales et al. 2004; Beyer et al. 2013; Roberts et al. 2004). Similarly, several test statistics for evaluating the association between the observed data sequences and the estimated HMM have been suggested, including correct percentage statistics or pseudo-residuals (Franke et al. 2004; Patterson et al. 2009, Jonsen et al. 2013).

SSMs with Location Filtering

The majority of SSMs applied in movement-based studies are multi-state random walks integrating different forms of movement models, such as extensions of the random walk framework, as part of their process component (Patterson et al. 2008; Gurarie et al. 2016). However, in contrast to HMMs some SSMs do not neglect potential sampling errors of the relocation data but account for them within their observation model structure (Jonsen et al. 2013, Patterson et al. 2016). Typically, these models include both continuous (e.g., estimated true locations) and one or several discrete behavioral states in the process component (Jonsen et al. 2005; Hopcraft et al. 2014).

Different Bayesian estimation techniques (e.g., Markov Chain Monte Carlo, MCMC; Jonsen et al. 2005) are usually required since multiple probability distributions need to be integrated along with non-linear structures within the models. Further, MCMC estimation also allows for non-Gaussian error structures in the observation model (Jonsen et al. 2013; Patterson et al. 2008). Besides MCMC, there are also different so called filtering methods for fitting such SSMs. Filtering methods can be applied to obtain parameters of linear state-space models (Patterson et al. 2016, Anderson-Sprecher and Ledolter 1991) as well as to estimate the most likely position of missing or biased relocations (Sibert et al. 2003; Austin et al. 2003). For example, Kalman Filtering (KF; Kalman and Bucy 1961) provides unbiased estimates of a first-order autoregressive model (e.g., the diffusion coefficient in a random walk model) given that the movement model is strictly linear with a Gaussian error distribution (Royer et al. 2005). Kalman filtering can further account for the influence of environmental covariates (Forester et al. 2007). However, the KF is not applicable for estimation of time-varying and discrete behavioral states. Particle Filters (PF) represent Bayesian approaches which overcome those limitations and can also fit non-linear SSMs. They are based on Sequential Monte Carlo sampling using ensembles of random sampling units (particles) which are moved forward by the process model of the state-space framework (Patterson et al. 2008, 2016). The importance of each particle is weighted in order to estimate the likelihood or posterior distributions of the model parameters. PFs are able to estimate non-linear and non-stationary movement models and can also implement non-Gaussian error structures in the process model (Dowd and Joy 2011; Royer et al. 2005; Andersen et al. 2007). Overall, these types of SSMs are highly useful for error-prone or incomplete relocation data with large gaps in sampling frequency. For example, many studies use different SSMs with location filtering in order to account for errors in Argos telemetry data (Patterson et al. 2010; Costa et al. 2010; Silva et al. 2014). However, all of these algorithms can be quite complex and computational intensive. More detailed explanations on different statistical options, including model fitting and diagnostics are provided in Jonsen et al. (2013). Different examples for implementations and applications in *R* are presented in (Dowd and Joy 2011; Albertsen et al. 2015; Jonsen et al. 2005; Pedersen et al. 2011).

References

Albertsen, C.M., Whoriskey, K., Yurkowski, D., Nielsen, A., and Flemming, J.M. (2015). Fast fitting of non-Gaussian state-space models to animal movement data via Template Model Builder. *Ecology, 96*(10), 2598–2604.

Andersen, K. H., Nielsen, A., Thygesen, U. H., Hinrichsen, H. H., and Neuenfeldt, S. (2007). Using the particle filter to geolocate Atlantic cod (Gadus morhua) in the Baltic Sea, with special emphasis on determining uncertainty. *Canadian Journal of Fisheries and Aquatic Sciences*, *64*(4), 618-627.

Anderson-Sprecher, R., and Ledolter, J. (1991). State-space analysis of wildlife telemetry data. *Journal of the American Statistical Association, 86*(415), 596-602.

Austin, D., McMillan, J. I., and Bowen, W. D. (2003). A three-stage algorithm for filtering erroneous Argos satellite locations. *Marine Mammal Science*, *19*(2), 371-383.

Barraquand, F., and Benhamou, S. (2008). Animal movements in heterogeneous landscapes: identifying profitable places and homogeneous movement bouts. *Ecology*, *89*(12), 3336-3348.

Bergman, C. M., Fryxell, J. M., Gates, C. C., and Fortin, D. (2001). Ungulate foraging strategies: energy maximizing or time minimizing? *Journal of Animal Ecology*, *70*(2), 289-300.

Beyer, H. L., Morales, J. M., Murray, D., and Fortin, M. J. (2013). The effectiveness of Bayesian state-space models for estimating behavioural states from movement paths. *Methods in Ecology and Evolution*, *4*(5), 433-441.

Bivand, R., and Rundel, C. (2015). rgeos: interface to geometry engine-open source (GEOS). *R package version 0.3-15*.

Breed, G. A., Costa, D. P., Jonsen, I. D., Robinson, P. W., and Flemming, J. M. (2012). State-space methods for more completely capturing behavioral dynamics from animal tracks. *Ecological Modelling*, *235*, 49-58.

Buchin, M., Driemel, A., Van Kreveld, M., and Sacristán, V. (2011). Segmenting trajectories: A framework and algorithms using spatiotemporal criteria. *Journal of Spatial Information Science*, *2011*(3), 33-63.

Buchin, M., Kruckenberg, H., and Kölzsch, A. (2013). Segmenting trajectories by movement states. In *Advances in spatial data handling* (pp. 15-25). Springer, Berlin, Heidelberg.

Burnham, K. P., and Anderson, D. R. (2002). A practical information-theoretic approach. *Model selection and multimodel inference*.

Byrne, R. W., Noser, R., Bates, L. A., and Jupp, P. E. (2009). How did they get here from there? Detecting changes of direction in terrestrial ranging. *Animal Behaviour*, *77*(3), 619-631.

Calenge, C. (2011). Analysis of animal movements in R: the adehabitatLT package. *R Foundation for Statistical Computing, Vienna*.

Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal society interface, 5*(25), 813-834.

Costa, D. P., Robinson, P. W., Arnould, J. P., Harrison, A. L., Simmons, S. E., Hassrick, J. L., ... and Crocker, D. E. (2010). Accuracy of ARGOS locations of pinnipeds at-sea estimated using Fastloc GPS. *PloS one*, *5*(1), e8677.

Dean, B., Freeman, R., Kirk, H., Leonard, K., Phillips, R. A., Perrins, C. M., and Guilford, T. (2013). Behavioural mapping of a pelagic seabird: combining multiple sensors and a hidden Markov model reveals the distribution of at-sea behaviour. *Journal of the Royal Society Interface*, *10*(78), 20120570.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C. C., and Lin, C. C. (2015). Misc functions of the department of statistics, probability theory group (formerly: E1071). *Package e1071. TU Wien*.

Dodge, S., Weibel, R., and Forootan, E. (2009). Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects. *Computers, Environment and Urban Systems, 33*(6), 419-434.

Dodge, S., Weibel, R., and Lautenschütz, A. K. (2008). Towards a taxonomy of movement patterns. *Information visualization*, *7*(3-4), 240-252.

Douglas, D. H., and Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: the international journal for geographic information and geovisualization*, *10*(2), 112-122.

Dowd, M., and Joy, R. (2011). Estimating behavioral parameters in animal movement models using a state-augmented particle filter. *Ecology*, *92*(3), 568-575.

Dzialak, M. R., Olson, C. V., Webb, S. L., Harju, S. M., and Winstead, J. B. (2015). Incorporating within-and between-patch resource selection in identification of critical habitat for brood-rearing greater sage-grouse. *Ecological Processes, 4*(1), 1-15.

Eckert, S. A., Moore, J. E., Dunn, D. C., van Buiten, R. S., Eckert, K. L., and Halpin, P. N. (2008). Modeling loggerhead turtle movement in the Mediterranean: importance of body size and oceanography. *Ecological Applications*, *18*(2), 290-308.

Forester, J. D., Ives, A. R., Turner, M. G., Anderson, D. P., Fortin, D., Beyer, H. L., Smith, D.W., and Boyce, M. S. (2007). State–space models link elk movement patterns to landscape characteristics in Yellowstone National Park. *Ecological Monographs*, *77*(2), 285-299.

Franke, A., Caelli, T., and Hudson, R. J. (2004). Analysis of movements and behavior of caribou (Rangifer tarandus) using hidden Markov models. *Ecological Modelling*, *173*(2-3), 259-270.

Garriga, J., and Bartumeus, F. The EMbC R-package: quick reference. *cran.mtu.edu* .

Garriga, J., Palmer, J. R., Oltra, A., and Bartumeus, F. (2016). Expectation-maximization binary clustering for behavioural annotation. *PloS one*, *11*(3), e0151984.

Gueguen, L. (2000). Segmentation by maximal predictive partitioning according to composition biases. In *International Conference on Biology, Informatics, and Mathematics* (pp. 32-44). Springer, Berlin, Heidelberg.

Gurarie, E., Andrews, R. D., and Laidre, K. L. (2009). A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, *12*(5), 395-408.

Gurarie, E., Bracis, C., Delgado, M., Meckley, T. D., Kojola, I., and Wagner, C. M. (2016). What is the animal doing? Tools for exploring behavioural structure in animal movements. *Journal of Animal Ecology*, *85*(1), 69-84.

Gurarie, E. (2015). bcpa: Behavioral change point analysis of animal movement. R package version 1.1.

Gutenkunst, R., Newlands, N., Lutcavage, M., and Edelstein-Keshet, L. (2007). Inferring resource distributions from Atlantic bluefin tuna movements: An analysis based on net displacement and length of track. *Journal of Theoretical Biology*, *245*(2), 243-257.

Hopcraft, J. G. C., Morales, J. M., Beyer, H. L., Borner, M., Mwangomo, E., Sinclair, A. R. E., Olff, H., and Haydon, D. T. (2014). Competition, predation, and migration: individual choice patterns of Serengeti migrants captured by hierarchical models. *Ecological Monographs*, *84*(3), 355-372.

Johnson, C. J., Parker, K. L., Heard, D. C., and Gillingham, M. P. (2002). Movement parameters of ungulates and scale-specific responses to the environment. *Journal of Animal Ecology*, *71*(2), 225-235.

Jonsen, I. D., Myers, R. A., and Flemming, J. M. (2003). Meta-analysis of animal movement using state-space models. *Ecology*, *84*(11), 3055-3063.

Jonsen, I. D. (2014). bsam: Bayesian State-space Models for Animal Movement. R Package Version 0.43-1.

Jonsen, I. D., Basson, M., Bestley, S., Bravington, M. V., Patterson, T. A., Pedersen, M. W., Thomson, R., Thygesen, U. H. and Wotherspoon, S. J. (2013). State-space models for bio-loggers: A methodological road map. *Deep Sea Research Part II: Topical Studies in Oceanography*, *88*, 34-46.

Jonsen, I. D., Flemming, J. M., and Myers, R. A. (2005). Robust state–space modeling of animal movement data. *Ecology*, *86*(11), 2874-2880.

Jonsen, I. D., Myers, R. A., and James, M. C. (2006). Robust hierarchical state–space models reveal diel variation in travel rates of migrating leatherback turtles. *Journal of Animal Ecology*, *75*(5), 1046-1057.

Kalman, R. E., and Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Journal of Basic Engineering*, *83*, 95-108.

Killick, R., Fearnhead, P., and Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590-1598.

Killick, R., Haynes, K. J., Eckley, I. A., Fearnhead, P., and Lee, J. (2015). changepoint: Methods for Changepoint Detection. R package version 2.2.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., and Morales, J. M. (2012). Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, *93*(11), 2336-2342.

Langrock, R., Kneib, T., Sohn, A., and DeRuiter, S. L. (2015). Nonparametric inference in hidden Markov models using P-splines. *Biometrics*, *71*(2), 520-528.

Langrock, R., and Zucchini, W. (2011). Hidden Markov models with arbitrary state dwell-time distributions. *Computational Statistics and Data Analysis*, *55*(1), 715-724.

LaPoint, S., Gallery, P., Wikelski, M., and Kays, R. (2013). Animal behavior, cost-based corridor models, and real corridors. *Landscape Ecology*, *28*(8), 1615-1630.

Lavielle, M. (1999). Detection of multiple changes in a sequence of dependent variables. *Stochastic Processes and their applications*, *83*(1), 79-102.

Lavielle, M. (2005). Using penalized contrasts for the change-point problem. *Signal Processing*, *85*(8), 1501-1510.

Le Corre, M., Dussault, C., and Côté, S. D. (2014). Detecting changes in the annual movements of terrestrial migratory species: using the first-passage time to document the spring migration of caribou. *Movement Ecology*, *2*(1), 1-11.

Legendre, P., and Legendre, L. (2012). *Numerical ecology*. Elsevier.

Liminana, R., Soutullo, A., and Urios, V. (2007). Autumn migration of Montagu's harriers Circus pygargus tracked by satellite telemetry. *Journal of Ornithology*, *148*(4), 517-523.

Louzao, M., Wiegand, T., Bartumeus, F., and Weimerskirch, H. (2014). Coupling instantaneous energy-budget models and behavioural mode analysis to estimate optimal foraging strategy: an example with wandering albatrosses. *Movement Ecology*, *2*(1), 1-15.

Madon, B., and Hingrat, Y. (2014). Deciphering behavioral changes in animal movement with a 'multiple change point algorithm-classification tree'framework. *Frontiers in Ecology and Evolution*, *2*, 30.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., Studer, M., and Roudier, P. (2015). cluster: "Finding Groups in Data": Cluster Analysis Extended. R package version 2.0.3.

Michelot, T., Langrock, R., Patterson, T., and Rexstad, E. (2015). moveHMM: animal movement modelling using Hidden Markov Models. *R package version, 1*.

Flemming, J. M., Jonsen, I. D., Myers, R. A., and Field, C. A. (2010). Hierarchical state-space estimation of leatherback turtle navigation ability. *PLoS One*, *5*(12), e14245.

Morales, J. M., Haydon, D. T., Frair, J., Holsinger, K. E., and Fryxell, J. M. (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, *85*(9), 2436-2445.

Muggeo, V. M. R. (2015). Regression models with breakpoints/changepoints estimation. *R package version 0.5-1.4*.

Nams, V. O. (2014). Combining animal movements and behavioural data to detect behavioural states. *Ecology Letters*, *17*(10), 1228-1237.

Neter, J., Wasserman, W., and Kutner, M. H. (1985). Applied linear statistical models: Regression. *Analysis of Variance, and Experimental Designs, 2nd Edition, Homewood: Richard D.*

Noser, R., and Byrne, R. W. (2014). Change point analysis of travel routes reveals novel insights into foraging strategies and cognitive maps of wild baboons. *American Journal of Primatology*, *76*(5), 399-409.

Patterson, T. A., Basson, M., Bravington, M. V., and Gunn, J. S. (2009). Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *Journal of Animal Ecology*, *78*(6), 1113-1123.

Patterson, T. A., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. (2008). State–space models of individual animal movement. *Trends in Ecology and Evolution*, *23*(2), 87-94.

Patterson, T. A., McConnell, B. J., Fedak, M. A., Bravington, M. V., and Hindell, M. A. (2010). Using GPS data to evaluate the accuracy of state–space methods for correction of Argos satellite telemetry error. *Ecology*, *91*(1), 273-285.

Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L., and King, R. (2016). Statistical modelling of animal movement: a myopic review and a discussion of good practice. *arXiv preprint arXiv:1603.07511*.

Pedersen, M. W., Berg, C. W., Thygesen, U. H., Nielsen, A., and Madsen, H. (2011). Estimation methods for nonlinear state-space models in ecology. *Ecological Modelling*, *222*(8), 1394-1400.

R Core Team (2015). R: A language and environment for statistical computing.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257-286.

Ripley, B. (2015). tree: classification and regression trees. R package version 1.0-36.

Roberts, S., Guilford, T., Rezek, I., and Biro, D. (2004). Positional entropy during pigeon homing I: application of Bayesian latent state modelling. *Journal of Theoretical Biology*, *227*(1), 39-50.

Royer, F., Fromentin, J. M., and Gaspar, P. (2005). A state–space model to derive bluefin tuna movement and habitat from archival tags. *Oikos*, *109*(3), 473-484.

Saalfeld, A. (1999). Topologically consistent line simplification with the Douglas-Peucker algorithm. *Cartography and Geographic Information Science*, *26*(1), 7-18.

Saher, D. J., and Schmiegelow, F. K. (2005). Movement pathways and habitat selection by woodland caribou during spring migration. *Rangifer*, 143-154.

Shamoun-Baranes, J., Bom, R., van Loon, E. E., Ens, B. J., Oosterbeek, K., and Bouten, W. (2012). From sensor data to animal behaviour: an oystercatcher example. *PloS one*, *7*(5), e37997.

Sibert, J. R., Musyl, M. K., and Brill, R. W. (2003). Horizontal movements of bigeye tuna (Thunnus obesus) near Hawaii determined by Kalman filter analysis of archival tagging data. *Fisheries Oceanography*, *12*(3), 141-151.

Sibly, R. M., Nott, H. M. R., and Fletcher, D. J. (1990). Splitting behaviour into bouts. *Animal Behaviour*, *39*(1), 63-69.

Silva, M. A., Jonsen, I., Russell, D. J., Prieto, R., Thompson, D., and Baumgartner, M. F. (2014). Assessing performance of Bayesian state-space models fit to Argos satellite telemetry locations processed with Kalman filtering. *PloS one, 9*(3), e92277.

Soleymani, A., Cachat, J., Robinson, K., Dodge, S., Kalueff, A., and Weibel, R. (2014). Integrating cross-scale analysis in the spatial and temporal domains for classification of behavioral movement. *Journal of Spatial Information Science, 2014*(8), 1-25.

Soleymani, A., van Loon, E. E., and Weibel, R. (2014). Capability of movement features extracted from GPS trajectories for the classification of fine-grained behaviors. *AGILE*. 3-16.

Sommerfeld, J., Kato, A., Ropert-Coudert, Y., Garthe, S., and Hindell, M. A. (2013). Foraging parameters influencing the detection and interpretation of area-restricted search behaviour in marine predators: a case study with the masked booby. *PloS one, 8*(5), e63742.

Steinley, D. (2006). K-means clustering: a half-century synthesis. *British Journal of Mathematical and Statistical Psychology, 59*(1), 1-34.

Sur, M., Skidmore, A. K., Exo, K. M., Wang, T., Ens, B. J., and Toxopeus, A. G. (2014). Change detection in animal movement using discrete wavelet analysis. *Ecological Informatics, 20,* 47-57.

Thiebault, A., and Tremblay, Y. (2013). Splitting animal trajectories into fine-scale behaviorally consistent movement units: breaking points relate to external stimuli in a foraging seabird. *Behavioral Ecology and Sociobiology, 67*(6), 1013-1026.

Tremblay, Y., Roberts, A. J., and Costa, D. P. (2007). Fractal landscape method: an alternative approach to measuring area-restricted searching behavior. *Journal of Experimental Biology, 210*(6), 935-945.

Van Moorter, B., Visscher, D. R., Jerde, C. L., Frair, J. L., and Merrill, E. H. (2010). Identifying movement states from location data using cluster analysis. *The Journal of Wildlife Management, 74*(3), 588-594.

Zeller, K. A., McGarigal, K., Cushman, S. A., Beier, P., Vickers, T. W., and Boyce, W. M. (2016). Using step and path selection functions for estimating resistance to movement: pumas as a case study. *Landscape Ecology, 31*(6), 1319-1335.

Zhang, J., O'Reilly, K. M., Perry, G. L., Taylor, G. A., and Dennis, T. E. (2015). Extending the functionality of behavioural change-point analysis with k-means clustering: a case study with the little penguin (eudyptula minor). *PloS one, 10*(4), e0122811.

Zucchini, W., Raubenheimer, D., and MacDonald, I. L. (2008). Modeling time series of animal behavior by means of a latent-state model with feedback. *Biometrics, 64*(3), 807-815.

S2: Applied Illustration of Path Segmentation Approaches

Aim of this supplement

In this supplement we illustrate the basic approach of how to segment a path. We simulate tracking data, go through explorative steps, and then apply three different segmentation methods.

Set up

```
       library(secr)
library(raster)
library(moveHMM)
library(adehabitatLT)
library(lubridate)
library(bcpa)


set.seed(12322211)

genHabitat <- function(nc, p = 0.1, A = 0.1) {
  r <- make.mask(nx = nc, ny = nc, spacing = 1)
  h <- randomHabitat(r, p = p, A = A)
  r <- raster(xmn=0, xmx=nc, ymn=0, ymx=nc, ncols=nc, nrows=nc)
  r <- rasterize(data.frame(h), r, field=1, background=0)
  r <- as.matrix(r)
  r <- list(hab = r, p = p, A = A, nc = nc)
  class(r) <- c("hab", "matrix")
  r
}
```

Data generation

We use a simple individual based simulation model for 12 months with 24 relocations per day. The movement of the animal was influenced by the time of the day, habitat and attraction to a temporally varying home range center. Turning angles where uniformily distributed between $-pi$ and pi, and step lengths where drawn from a Gamma distribution with scale and shape of 2.9 and 0.9 for habitat and 1.2 and 5 for the matrix respectively Animals moved with a probability of 0.3 during the night and a probability of 1 during the day. At each location t, the animal chose 20 candidate locations and chose one at random with probably weighted towards the home range center.

```
       n_months <- 12
n <- 24 * 30 * n_months  # one relocation every hour

pm_day <- 1
pm_night <- 0.3

b0 <- rep(c(rep(pm_night, 5), seq(pm_night, pm_day, length.out = 2), rep(pm_day, 10),
            seq(pm_day, pm_night, length.out = 2), rep(pm_night, 5)), 30 * n_months)
```

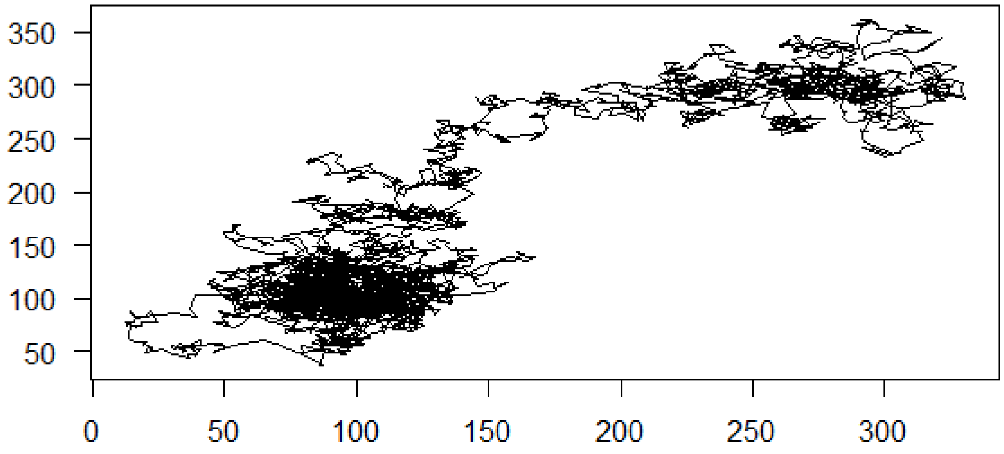
```r
b1 <- rep(c(1, 0, 1), each = n/3) * 1
b2 <- as.numeric(b1 == 0) * 1

b3 <- 10 # hab matrix

xy0 <- c(100, 100)

x_c1 <- 100
y_c1 <- 100

x_c2 <- 300
y_c2 <- 300

tpl <- raster(xmn = 0, xmx = 400, ymn = 0, ymx = 400, res = 1)
hab <- raster(genHabitat(400, A = 0.5, p = 0.5)$hab, template = tpl)
```

```
       ## Loading required namespace: igraph

       hab0 <- hab
```

```r
hab <- hab0
hab[] <- ifelse(hab[] == 0, 1, 2)

set.seed(2090160703)
xy <- matrix(NA, nrow = n, ncol = 2)
xy[1, ] <- xy0


for (i in 2:n) {
  # cand locations
  if (runif(1) < b0[i]) {
    ta <- runif(20, -pi, pi)

    if (raster::extract(hab, xy[i-1, ,drop = FALSE]) == 2) {
      scl <- 2.9
      shp <- 0.9
    } else {
      scl <- 1.2
      shp <- 5
    }
    slen <- rgamma(20, scale = scl, shape = shp)
    x1 <- xy[i - 1, 1] + (cos(ta) * slen)
    y1 <- xy[i - 1, 2] + (sin(ta) * slen)

    d1 <- sqrt((x1 - x_c1)^2 + (y1 - y_c1)^2)
    d2 <- sqrt((x1 - x_c2)^2 + (y1 - y_c2)^2)
    w <- (dexp(d1, rate = 1/25) * b1[i] +
            dexp(d2, rate = 1/25) * b2[i])
    w[is.na(w)] <- 0
    w <- sample(20, 1, prob = w)
    xy[i, ] <- c(x1[w], y1[w])
  } else {
    xy[i, ] <- xy[i-1, ]
  }
}

xy[, 1] <- xy[, 1] + runif(nrow(xy), -0.1, 0.1)
xy[, 2] <- xy[, 2] + runif(nrow(xy), -0.1, 0.1)
xy <- xy[, 1:2]
xy <- data.frame(xy)
names(xy) <- c("x", "y")
xy$time <- ymd_hm("2000-01-01 00:00") + hours(0:(n-1))
xy$hour <- rep(1:24, n/24) # hour of the day
xy$id <- 1:nrow(xy)
xy$hab <- raster::extract(hab, xy[, 1:2])
```

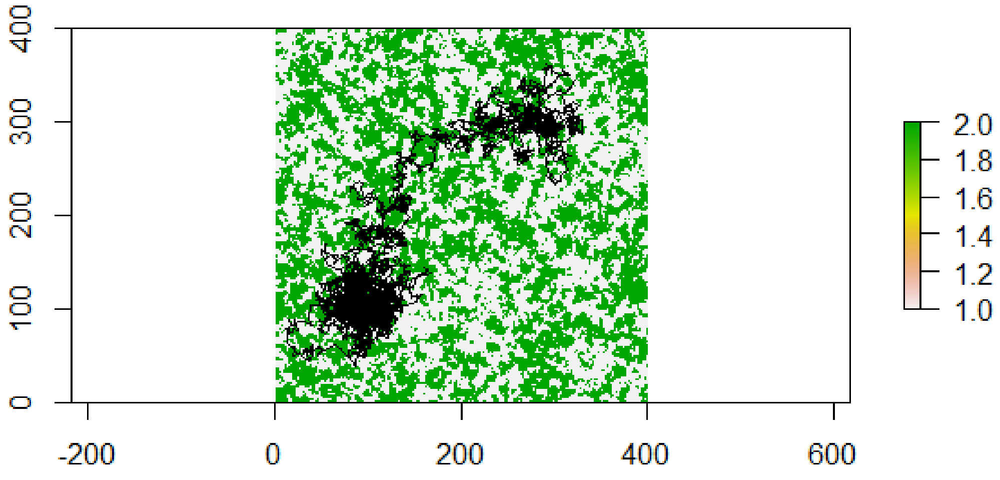
The simulated data:

```
        plot(xy[, 1:2], pch = 20, xlab = "", ylab = "", las = 1, type = "l", lwd = 0.5
)
```



Simulated movement path and habitat composition:

```
        plot(hab)
lines(xy[,1:2])
```



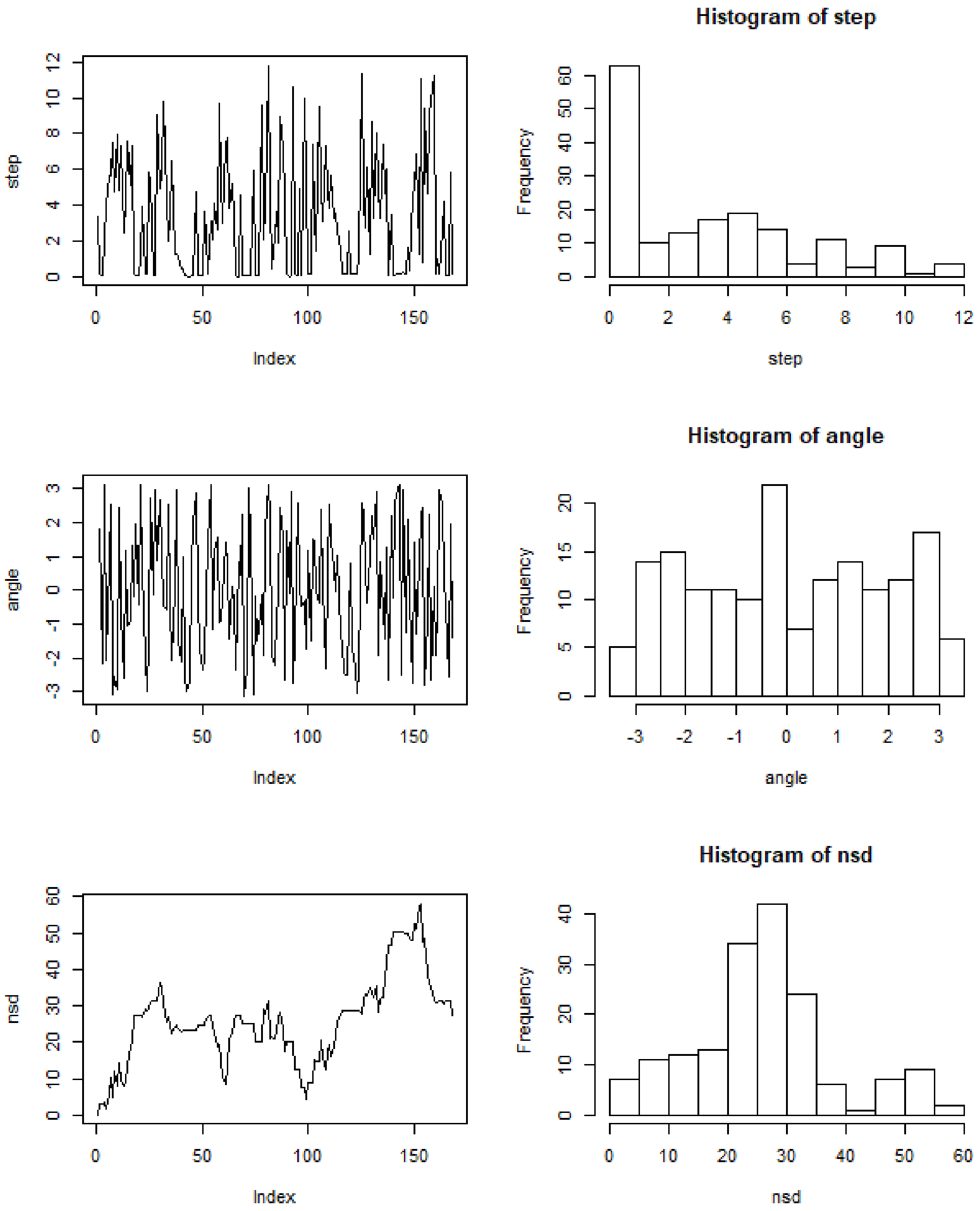
Path metrics

As a next step we calculate and inspect three path metrics. Namely, the step length, the net squared displacement and finally the turning anlges.

```
        xy$sl <- c(NA, with(xy, sqrt((head(x, -1) - tail(x, -1)^2 + (head(y, -1) - tail(y, -1)^2))
xy$nsd <- with(xy, sqrt((x - xy0[1])^2 + (y - xy0[2])^2)
dat <- prepData(xy[, c("x", "y", "hour", "nsd", "hab")], type = "UTM")
```

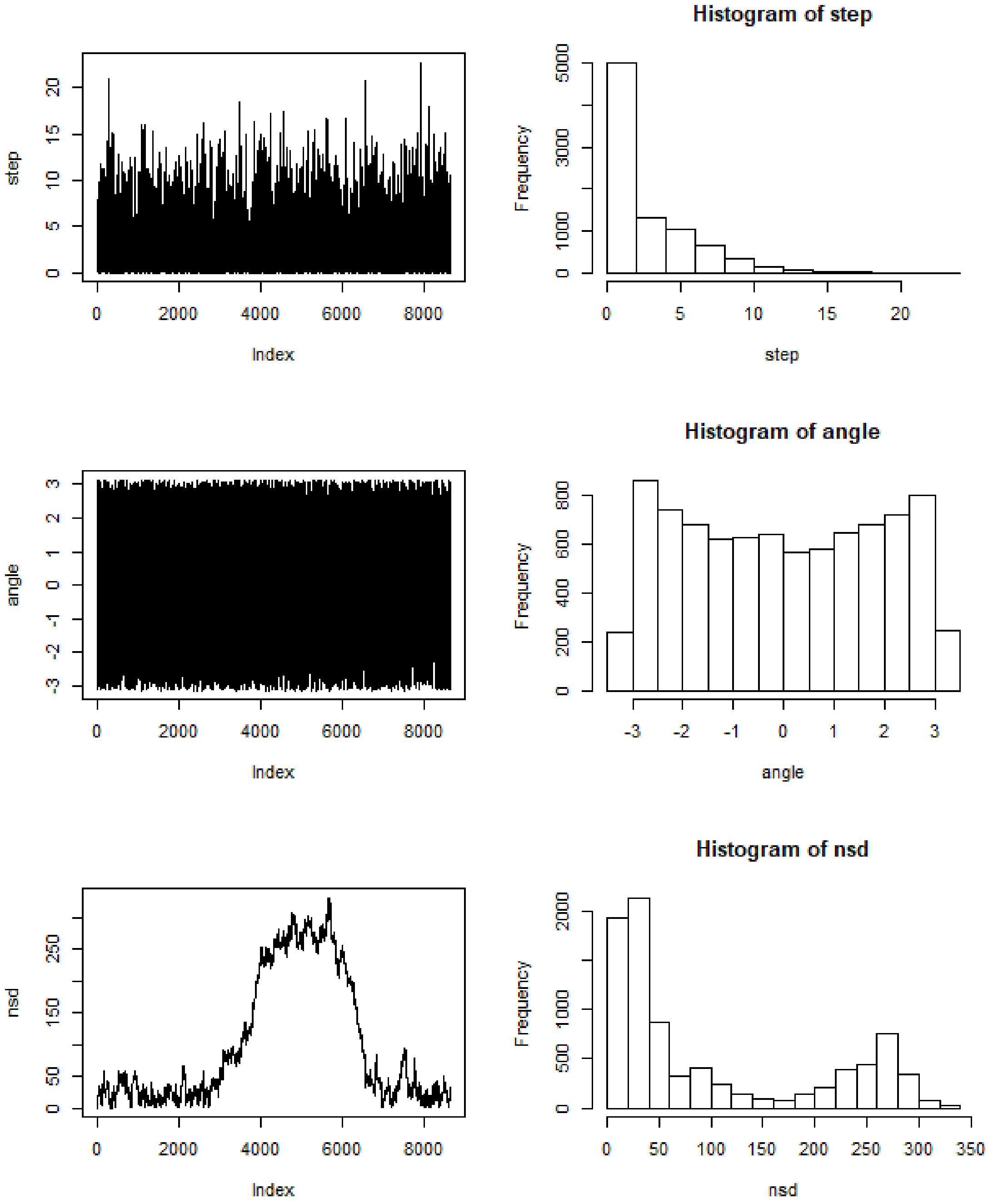
Plotting path metrics as time series and histograms. First we look at the signals for 1 week.

```
        par(mfrow = c(3, 2)
with(dat[1:(24 * 7), ], plot(step, type = "l")
with(dat[1:(24 * 7), ], hist(step)
with(dat[1:(24 * 7), ], plot(angle, type = "l")
with(dat[1:(24 * 7), ], hist(angle)
with(dat[1:(24 * 7), ], plot(nsd, type = "l")
with(dat[1:(24 * 7), ], hist(nsd)
```



Next we look at the path signals for a whole year.

```
        par(mfrow = c(3, 2))
with(dat, plot(step, type = "l")
with(dat, hist(step)
with(dat, plot(angle, type = "l")
with(dat, hist(angle)
with(dat, plot(nsd, type = "l")
with(dat, hist(nsd)
```

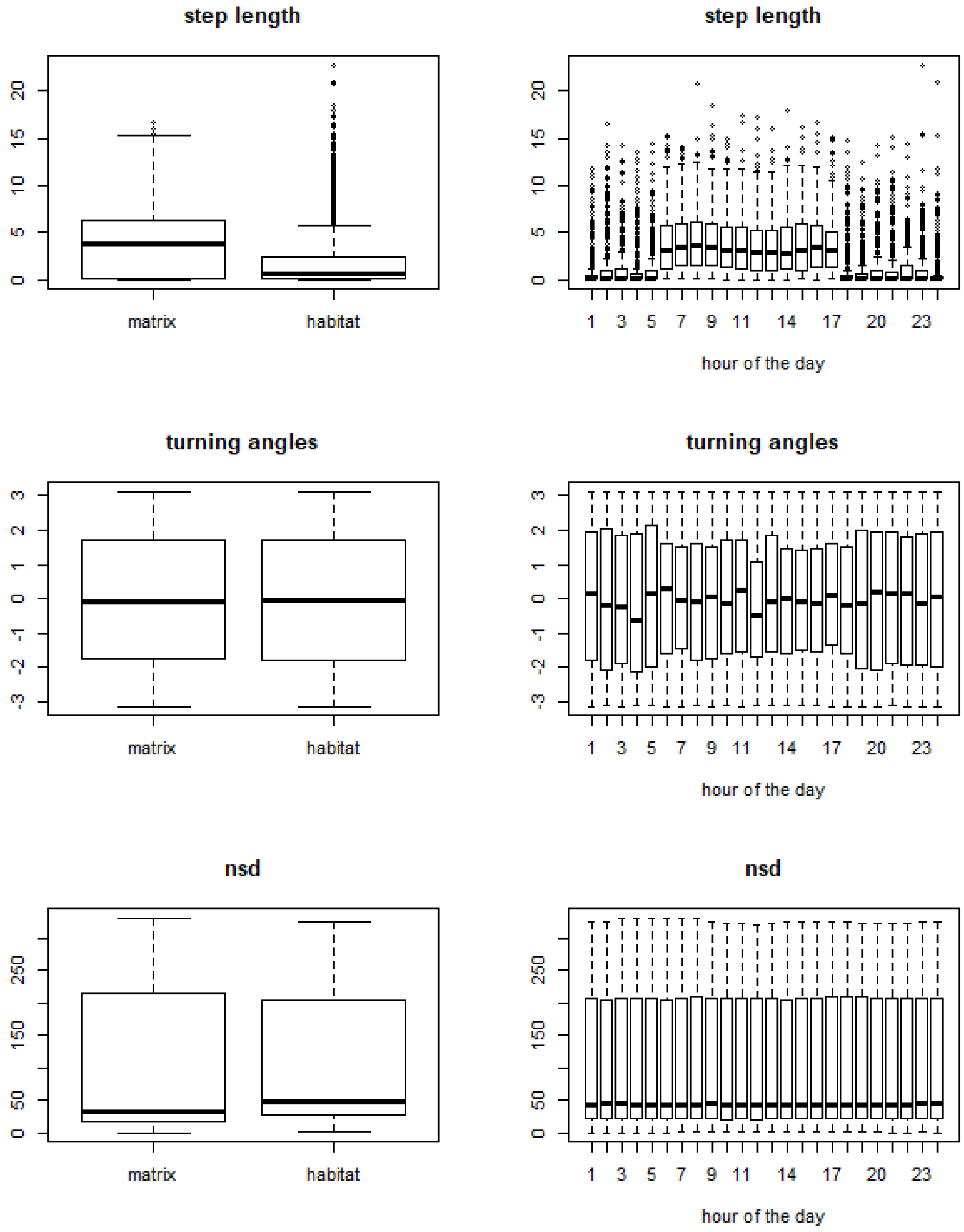


Note, that when plotting path signals for one week we can observe diurnal patterns, while when plotting the signals for one year we observe a migratory pattern, especially for the net-squared displacement.

Finally, we can look at the same path signal, but as a funciton of different covariates (here habitat and matrix, left column; hour of the day right column).

```
       par(mfrow = c(3, 2)
with(dat, boxplot(step ~ hab, main = "step length", xaxt = "n")
axis(1, at = 1:2, labels = c("matrix", "habitat")
```

```
with(dat, boxplot(step ~ hour, main = "step length", xlab = "hour of the day")
with(dat, boxplot(angle ~ hab, main = "turning angles", xaxt = "n")
axis(1, at = 1:2, labels = c("matrix", "habitat")
with(dat, boxplot(angle ~ hour, main = "turning angles", xlab = "hour of the day")
with(dat, boxplot(nsd ~ hab, main = "nsd", xaxt = "n")
axis(1, at = 1:2, labels = c("matrix", "habitat")
with(dat, boxplot(nsd ~ hour, main = "nsd", xlab = "hour of the day")
```



```
       par(mfrow = c(1, 1)
```

Explorative analyses

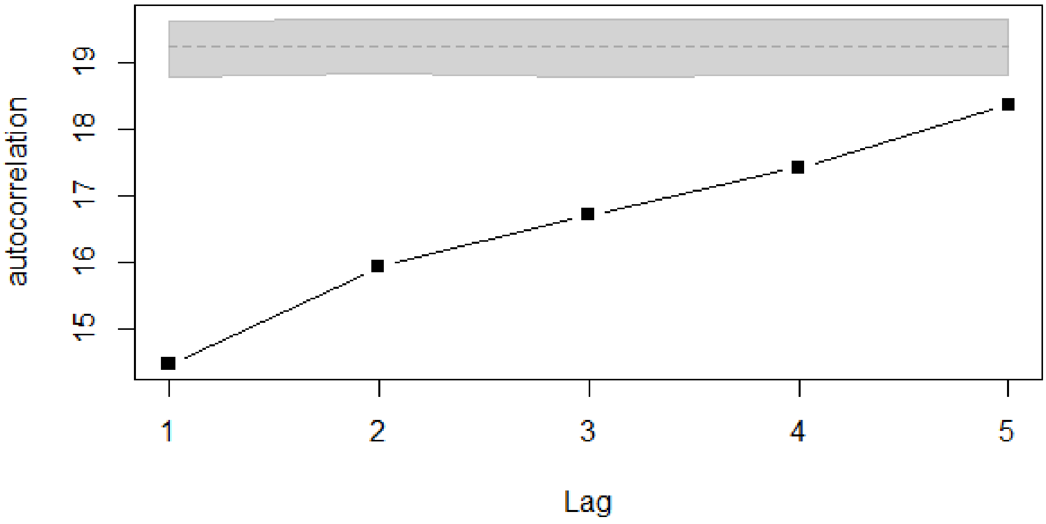
First we check if the data is regularly sampled:

```
        table(diff(xy$time)

        ##
##    1
## 8639
```

All of the 8639 relocations have a time lag of 1 hour. Therefore, the data has a regular sampling regime.

Next we test for independence in the consecutive path-signals:

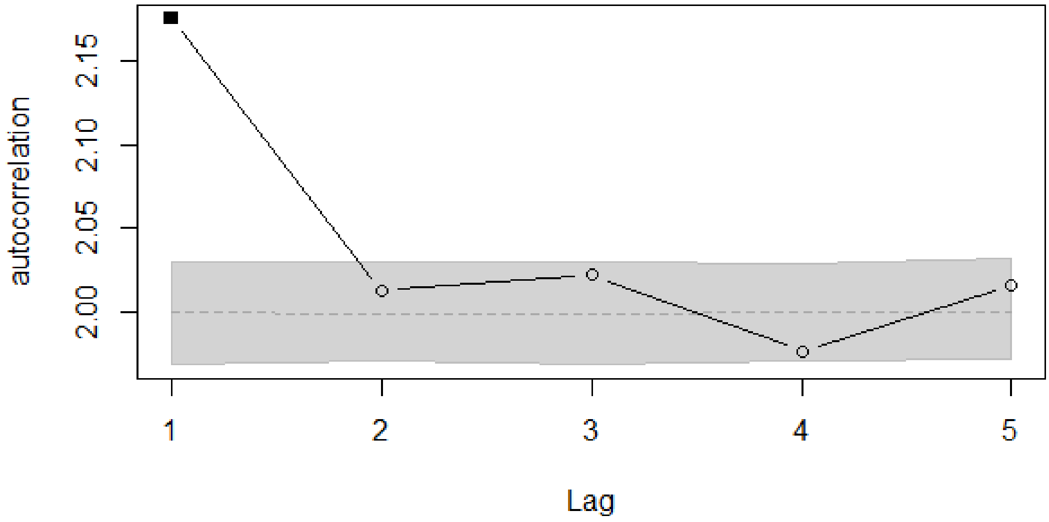
```
        # Wald-Wolfowitz Test of Randomness

wawotest(dat$step)

        ## 1 NA removed

        ##         a         ea         va         za          p
## 2143.4821    -1.0000 8632.8562   23.0805     0.0000

        wawotest(dat$nsd)

        ##          a          ea          va          za          p
## 8635.65169    -1.00000 8636.99627    92.93173    0.00000

        # correlogram for angular and linear descriptors of a movement path

ind <- acfdist.ltraj(as.ltraj(xy[, 1:2], date = xy$time, id = "a1"), lag = 5)
```



```
        ind
```

```
## [[1]]
##          lag.1    lag.2    lag.3    lag.4    lag.5
## obs   14.46598 15.92845 16.71358 17.43126 18.36498
## 2.5%  18.77615 18.83905 18.79915 18.82473 18.82640
## 50%   19.24214 19.24873 19.24953 19.25300 19.23892
## 97.5% 19.63601 19.65166 19.65655 19.64178 19.65752
```

```
ind2 <- acfang.ltraj(as.ltraj(xy[, 1:2], date = xy$time, id = "a1"), lag = 5)
```



```
ind2
```

```
## [[1]]
##          lag.1    lag.2    lag.3    lag.4    lag.5
## obs   2.175379 2.012483 2.022091 1.976375 2.016311
## 2.5%  1.969032 1.970675 1.968645 1.971118 1.971807
## 50%   2.000121 1.998833 1.998663 2.000100 2.000147
## 97.5% 2.029612 2.029967 2.030055 2.028320 2.032133
```

both tests suggest for correlated data structures.

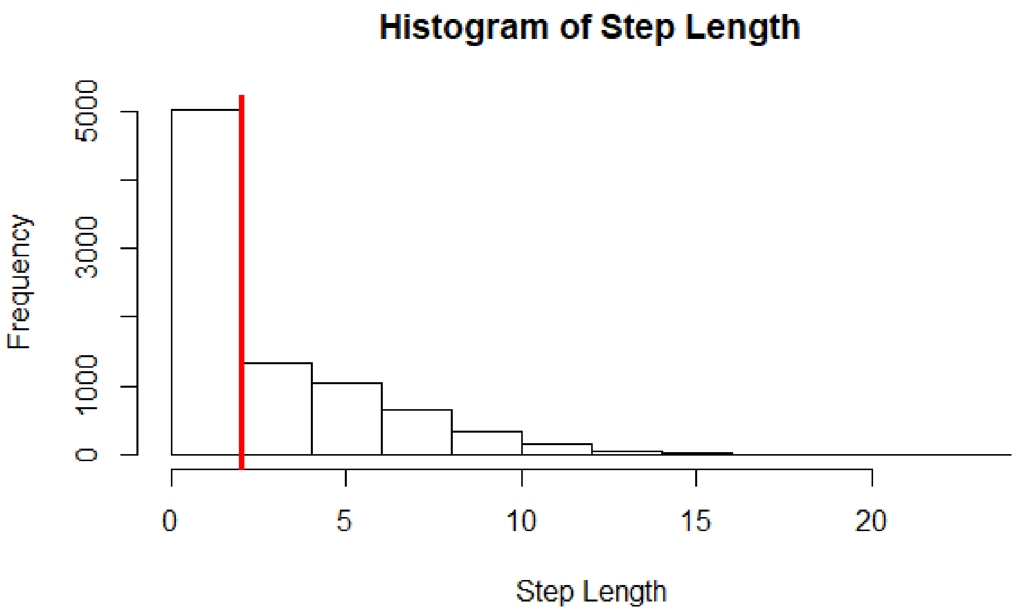
Path segmentation

We consider three methods: 1. Thresholding, 2. Behavioural Change Point Analysis, and 3. Hidden Markov Models:

Threshholding

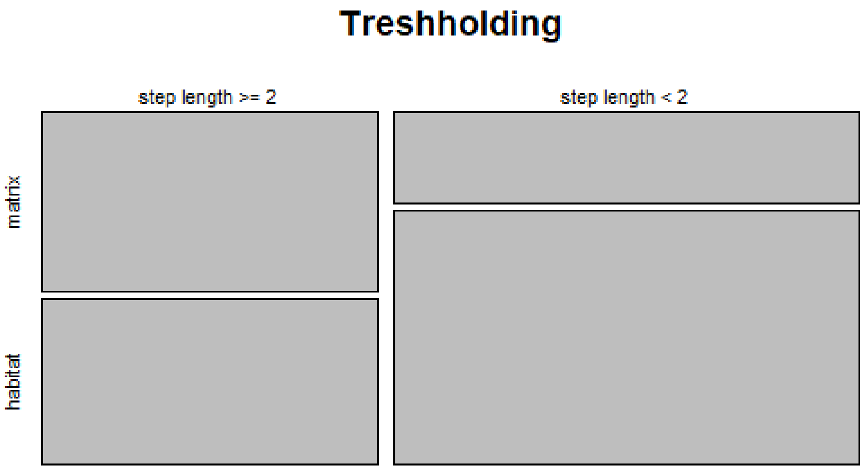
An ecologist may assume that due to biological reasoning a step length threshhold of 2 is indicating an important biological phenomena.

```
hist(xy$sl, main="Histogram of Step Length", xlab="Step Length")
abline(v=2, col="red", lwd=3)
```



This threshold can than be plotted against the habitat type.

```
        tb <- table(xy$sl < 2, xy$hab)
row.names(tb) <- c("step length >= 2", "step length < 2")
colnames(tb) <- c("matrix", "habitat")
mosaicplot(tb , main = "Treshholding")
```



Finally, we can apply a Pearson's Chi-squared Test to determine if the observed proportions of each state within the habitat components are signficantly different from a random distribution.

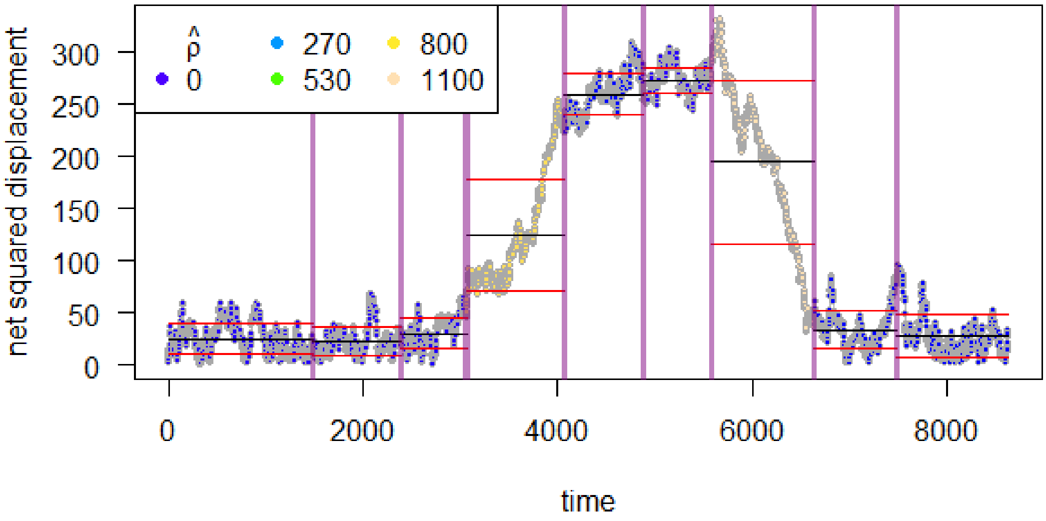
```
       prop.table(tb, 1)

       ##
##                       matrix   habitat
##   step length >= 2 0.5222560 0.4777440
##   step length < 2  0.2672242 0.7327758

       prop.table(tb, 2)

       ##
##                       matrix   habitat
##   step length >= 2 0.5846487 0.3195266
##   step length < 2  0.4153513 0.6804734

       chisq.test(tb)

       ##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  tb
## X-squared = 583.03, df = 1, p-value < 2.2e-16
```

The test points out that the observed pattern of the state proportions is significant.

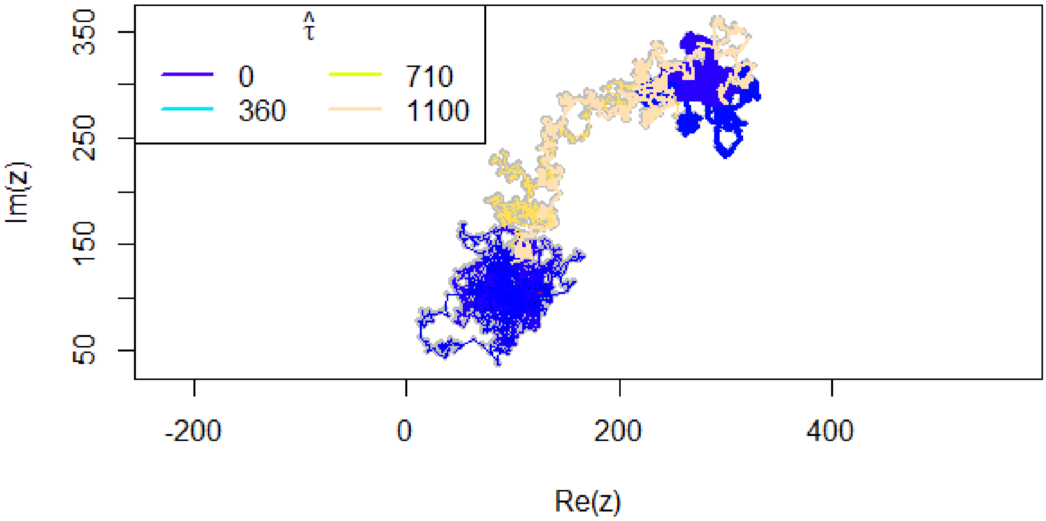
Behavioural change point analysis

Next, we are interested in finding the points where the animal starts to migrate.

```
       xy$Time <- 1:nrow(xy)
path_char <- bcpa::MakeTrack(xy$x, xy$y, xy$Time)
path_char <- bcpa::GetVT(path_char)
path_char$nsd <- xy$nsd[-(1:2)]

# run the bcpa
ws <- WindowSweep(path_char, "nsd", windowsize = 30, progress = FALSE)

plot(ws, type = "flat", clusterwidth = 24 * 7, xlab = "time", las = 1, ylab = "net sq
uared displacement")
```



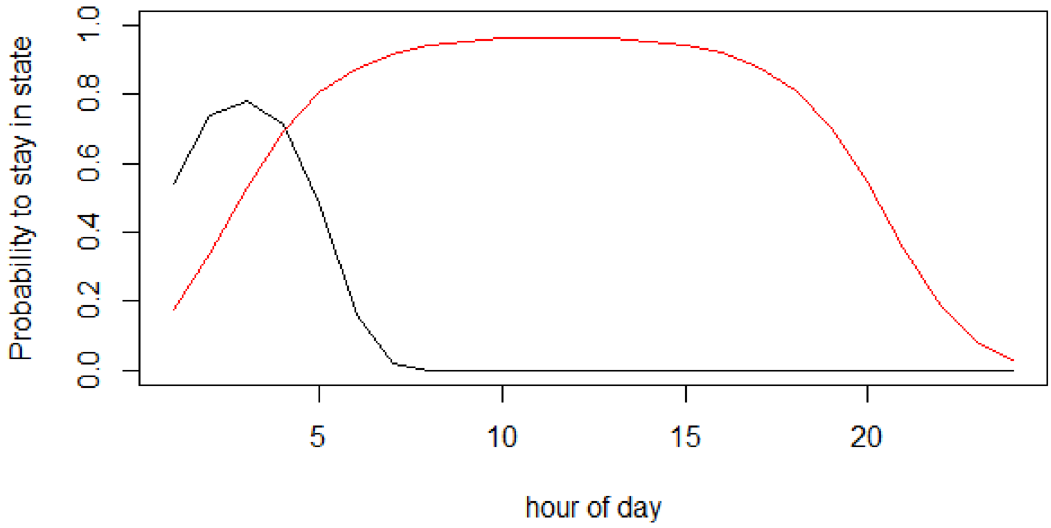
```
        # add trajectory plot
xy_bc <- data.frame(Time=xy$Time, X=xy$x, Y=xy$y)
PathPlot(xy_bc,ws, type = "flat",clusterwidth = 24 * 7, plotlegend = T, tauwhere = "t
opleft", n.legend = 4, ncol.legend = 2, bty.legend = T)
```



Hidden Markov Models

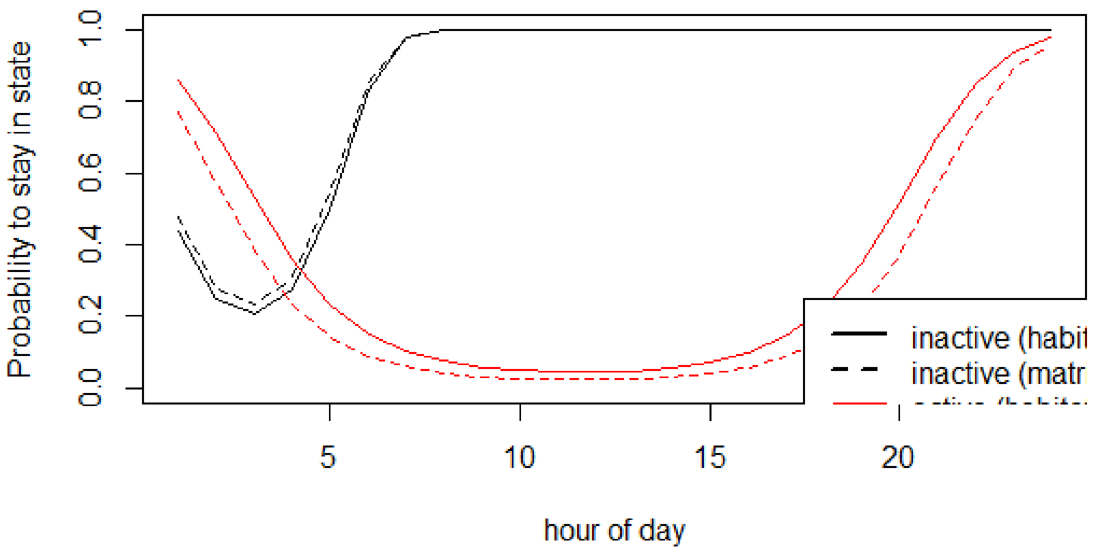
In the last step we want to find two different states of the animal (e.g., activ and resting) and model the transition probabilities as a function of habitat tand time of the day.

```
       ## moveHMM
mu0 <- c(0.1, 1)
sigma0 <- c(0.1, 1)
zeromass0 <- c(0.1, 0.05)
stepPar0 <- c(mu0, sigma0)
angleMean0 <- c(pi, 0)
kappa0 <- c(1, 1)
anglePar0 <- c(angleMean0, kappa0)
dat$hab <- factor(dat$hab)

m0 <- fitHMM(data = dat, nbStates = 2, stepPar0 = stepPar0, anglePar0 = anglePar0,
             formula = ~ 1)

m1 <- fitHMM(data = dat, nbStates = 2, stepPar0 = stepPar0, anglePar0 = anglePar0,
             formula = ~ hour + I(hour^2)

m2 <- fitHMM(data = dat, nbStates = 2, stepPar0 = stepPar0, anglePar0 = anglePar0,
             formula = ~ hour + I(hour^2) + hab)

AIC(m0, m1, m2)

       ##   Model      AIC
## 1     m0 56095.59
## 2     m2 56161.20
## 3     m1 56193.31

       # plots
z <- 1:24


plot(0, 0, type = "n", xlim = range(z), ylim = c(0, 1), xlab = "hour of day", ylab =
"Probability to stay in state")
lines(z, 1 - plogis(m1$mle$beta[1, 1] + m1$mle$beta[2, 1] * z + m1$mle$beta[3, 1] * z
^2)
lines(z, 1 - plogis(m1$mle$beta[1, 2] + m1$mle$beta[2, 2] * z + m1$mle$beta[3, 2] * z
^2), col = "red")
```



```
       plot(0, 0, type = "n", xlim = range(z), ylim = c(0, 1), xlab = "hour of day",
ylab = "Probability to stay in state")
lines(z, plogis(m2$mle$beta[1, 1] + m2$mle$beta[2, 1] * z + m2$mle$beta[3, 1] * z^2 +
```

```
m2$mle$beta[4,1] * 1)
lines(z, plogis(m2$mle$beta[1, 1] + m2$mle$beta[2, 1] * z + m2$mle$beta[3, 1] * z^2 +
m2$mle$beta[4,1] * 0), lty = 2)

lines(z, plogis(m2$mle$beta[1, 2] + m2$mle$beta[2, 2] * z + m2$mle$beta[3, 2] * z^2 +
m2$mle$beta[4, 2] * 1), col = "red")
lines(z, plogis(m2$mle$beta[1, 2] + m2$mle$beta[2, 2] * z + m2$mle$beta[3, 2] * z^2 +
m2$mle$beta[4, 2] * 0), lty = 2, col = "red")

legend(17.5,0.25, c("inactive (habitat)", "inactive (matrix)", "active (habitat)", "a
ctive (matrix)"),
       lty = c(1,2,1,2), lwd = c(2,2,2,2), col=c("black","black", "red","red")
```



CHAPTER 4

It`s all in the matrix: Comparing models of functional connectivity for red deer (*Cervus elaphus*) in Northern Germany using landscape genetics

Abstract

Estimating functional connectivity emerged as a central tool for conservation given its huge potential to quantify corridors and barriers to gene flow in increasingly fragmented landscapes. A common approach for deriving connectivity is to model resistance of the landscape matrix for a given study species. However, there is a plethora of approaches that have been proposed to assess landscape resistance. In particular, these methods vary with regard to the key assumptions they are established on, which can significantly affect how functional connectivity is quantified.

Here, we apply a landscape genetic approach to compare different models of landscape resistance using red deer (*Cervus elaphus*) in Northern Germany as a case study. In order to derive information on optimal placement of conservation corridors we utilize an extensive data set consisting of over 400 genetic samples as well as telemetry data from 20 red deer individuals. Using a multi-step model selection framework we account for a wide range of methodological decisions in quantifying i) resistance surfaces (expert-, habitat-, and movement-informed), ii) effective distances (circuit theory and least cost paths) as well as iii) correlations with genetic distances (Mantel test, linear mixed models, and multiple regression on distance matrices).

First, we selected a final model among each of the three approaches for quantifying resistance using Mantel tests to correlate the derived effective distances with genetic distances. Additionally, we combined the preselection of resistance models into ensemble models. The set of final models served as different hypotheses on potential causes of isolation by resistance (IBR). We hypothesized that 1) the ensemble approach outperforms other resistance models 2) movement-informed resistance models correlate better with observed genetic distance as compared to habitat-informed models and 3) a weak performance of models based on expert-knowledge. Additionally, we tested for potential effects of isolation by geographic distance (IBD) and modeled isolation by barrier (IBB) based on putative barriers such as primary roads in our study area. Finally, we fitted univariate and multivariate linear mixed models (MLPE) to correlate pairwise genetic distances with effective distances derived from the formulated hypotheses (variants of IBR, IBD and IBB). We used a model selection framework to evaluate model performance and also compared the results with multiple regression on distance matrices (MRM).

Resistance model performance heavily relied on how effective distance was quantified and how it was compared against genetic distances. Results regarding the highest-ranked (best-performing) resistance model were depending on the applied effective distance (circuitscape vs least cost path) as well as the statistical approach for comparison. The MLPE method indicated a high correlation of the observed genetic distances with circuit distances based on an habitat-informed model. For effective distances derived from least cost paths we identified a movement-informed approach to perform best. Model comparison based on MRM showed that for both distance algorithms an ensemble model works well to describe genetic patterns. Model performance improved in all cases when IBD and IBB were included.

For corridor design at a small scale, model-based corridor locations overlapped significantly. Our results of a movement-informed model approach indicated that red deer are capable of moving through less suitable habitat at short distances in comparison to habitat requirements at the home range scale determined with habitat suitability models. On the other hand, for dispersal over longer distances, suitable habitat conditions are required. For deriving large scale conservation corridors we recommend to apply ensembles of multiple resistance surfaces to overcome limitations of single methodological approaches. In our case, placement of short range corridors was less impacted by the choice of resistance model. Still we caution to apply models with valid assumptions and appropriate data such as (movement-informed) resource selection functions based on observed dispersal events.

Keywords: effective distance, fragmentation, landscape genetics, red deer, ensemble model

Introduction

Assessing landscape connectivity between patches of primary habitat is an important, yet challenging, task in wildlife conservation. Overall, landscape connectivity describes the degree to which a landscape matrix still enables the movement of individuals between remaining patches of suitable habitat (Crooks and Sanjayan 2006; Taylor et al. 1993). In particular, anthropogenic development and fragmentation have been shown to impede such movements which limits effective dispersal and gene flow for certain study species (Fahrig 2003; Templeton et al. 1990). Such restrictions can lead to the emergence of metapopulations (Hanski 1998; Opdam 1991), a decrease in effective population sizes (Keyghobadi 2007) and long-term loss of genetic diversity and inbreeding (Andersen et al. 2004; Haddad et al. 2016; Proctor et al. 2005). In contrast, landscape connectivity can facilitate the exchange of individuals (i.e., genes) and thus counteracts the negative effects of fragmentation which have been shown to be one of the greatest threats to mobile species in human dominated landscapes (Epps et al. 2005; Fahrig and Merriam 1994; Proctor et al. 2005). Therefore, understanding connectivity is important for maintaining the long-term viability of populations in fragmented landscapes (Cushman et al. 2011; Flather and Bevers 2002; Hanski and Ovaskainen 2003).

A distinction is made in the literature between two types of landscape connectivity: 1) structural connectivity which only refers to the physical alignment of habitable parts of the landscape matrix (Betts et al. 2015; Kindlmann and Burel 2008) and 2) functional connectivity, which accounts for a species' capability to move through a less favorable landscape matrix where habitat features are not always structurally connected (Baguette and Van Dyck 2007; Taylor et al. 1993; With and Crist 1995). In terms of conservation, functional connectivity is particularly relevant as it describes to what degree a landscape matrix still allows for individuals to disperse (Fahrig 2007; Ricketts 2001). Landscape models describing the functional connectivity of an area of interest have become a fundamental tool in applied conservation for delineating corridors in order to either maintain, facilitate or re-establish dispersal (Beier and Noss 2008; Hilty et al. 2012; Rudnick et al. 2012). However, estimating functional connectivity for a given target species remains to be one of the major challenges when identifying conservation corridors objectively (Abrahms et al. 2017; Beier et al. 2008; Naidoo et al. 2018). Most notably, this is due to the fact that functional connectivity is highly species specific: the influence of different landscape components on dispersal is depending, among other things, on the movement capacity, perceptual range and habitat requirements of an organism (Bélisle 2005; Diniz et al. 2020; Fletcher et al. 2012).

Functional connectivity is ultimately linked to gene flow as higher probability of movement through the landscape matrix should result in effective dispersal and eventually reproduction (Baguette et al. 2013). Therefore, information on the degree of genetic differentiation and current or historic gene flow derived from genetic data allows conclusions to be drawn about recent or contemporary landscape composition (Coulon et al. 2004; Keyghobadi et al. 2005; Stevens et al. 2006). Landscape genetic studies which link such patterns of genetic differentiation to patterns of the landscape matrix (Holderegger and Wagner 2006; Manel et al. 2003; Storfer et al. 2010) have been shown to be valid framework for assessing functional connectivity and deriving conservation corridors (Braunisch et al. 2010; Keller et al. 2012; Ruiz-González et al. 2014). Most frequently, landscape genetics correlate information on genetic differentiation (i.e., estimates of gene flow) with measures of landscape composition among populations or individuals to test hypotheses on 1) isolation by distance, 2) isolation by barriers, or 3) isolation by effective distance i.e. landscape resistance (Storfer et al. 2007, Balkenhol et al. 2009). Isolation by distance (IBD) assumes that genetic differentiation is correlated to geographic distance between compared entities (Hutchison and Templeton 1999; Wright 1943). IBD is commonly tested for as a null model (Balkenhol et al. 2009) since it ignores any potential effects of the landscape matrix. Isolation by barrier (IBB) accounts for the presence of putative barriers or boundaries being the main cause of differentiation between populations or individuals (Epps et al. 2005; Frantz et al. 2012; Ward et al. 2015). And finally, isolation by resistance (IBR) hypothesizes that effective distance derived from the resistance of the landscape matrix between considered entities and thus best describing functional connectivity (McRae 2006; Segelbacher et al. 2010). However, one of the biggest challenges in landscape genetic analyses remains to be the parameterization of resistance values for different landscape features, i.e. how to weight the cost of movement through the landscape matrix (Spear et al. 2010, 2015). As a result, there is a multitude of methods available for modeling landscape resistance and deriving effective distances from these models.

Modelling Landscape Resistance

First of all, there are important differences with regard to the methodological approaches that are available. Next to the data foundation these approaches vary also with hypotheses and assumptions behind them. A multitude of studies have relied on expert-opinion for parameterizing resistance surfaces for a given study species (Beier et al. 2008; Jacobs et al. 2014; Reed et al. 2016). This is often done by assigning resistance values to different classes of landcover based on previous studies, literature reviews, or experience of local managers (e.g., Clevenger et al. 2002; Broquet et al. 2006). However, defining these resistance effects can be complex and difficult to quantify accurately. Therefore, expert-informed analyses have been criticized for being subjective and non-transparent (Rayfield et al. 2009; Spear et al. 2010). On the other hand, such approaches do not rely on empirical data and could provide a less time-consuming and effective solution for inferring functional connectivity (Milanesi et al. 2016a; Garroway et al. 2011; Reed et al. 2016).

Other methodological approaches for estimating landscape resistance depend on empirical data such as information on habitat use or species distribution (Razgour, 2015; Stevenson-Holt et al. 2014). For example, habitat models (also referred to as species distribution models) are widely applied in ecology and conservation to infer primary habitat requirements of a target species (Guisan and Zimmermann 2000; Franklin 2009). These models rely on information on the species' presence (occurrence) which can be obtained for example from direct sightings, or indirect detections and signs of species occurrence via feces or feathers (Braunisch et al. 2010; Mateo Sánchez et al. 2013), and camera traps (O'Connell et al. 2010). Subsequently, resistance can be derived from habitat suitability or species distribution models (SDMs) through taking the inverse or other transformations of the suitability estimate (Keeley et al. 2016). Essentially, this approach assumes an inverse relationship between habitat suitability and landscape resistance meaning that the species needs good habitat conditions during movement through the matrix (Spear et al. 2010; Zeller et al. 2012). However, previous studies have challenged this assumption by showing that some study species are capable of moving through poor habitat making habitat models weak proxies of functional connectivity (Wasserman et al. 2010; Shirk et al. 2010; Mateo-Sánchez et al. 2015). On the other hand, this can hold true depending on the modeled scale and ecological niche as well as the dispersal abilities of the study species (Engler et al. 2014; Razgour 2015; Wang et al. 2008)

Since the last decade movement data derived from telemetry relocations has increasingly been used for inferring landscape resistance (Cushman and Lewis 2010; Reding et al. 2013; Zeller et al. 2017).The information about the realized movements of tracked individuals can be linked to landscape features using different forms of resource selection functions (Boyce et al. 2002; Manly et al. 1993; Zeller et al. 2012). Similar to habitat suitability models, transformations such as the inverse of the fitted resource selection functions can then be used to infer landscape resistance for a given study species (Squires et al. 2013; Zeller et al. 2017). In particular, analyzing relocations at the step- and path-level has been shown to produce more realistic estimates for modeling resistance of landscape features towards movements of a given study species (Coulon et al. 2004; Reding et al. 2013; Zeller et al. 2015). Still, the majority of movement-informed resistance models provide comparable results to habitat-informed approaches as observed relocations are the result of primary habitat selection or within home range movements (Chetkiewicz et al. 2006; Shafer et al. 2012; Spear et al. 2015). However, habitat requirements during actual dispersal events could potentially differ substantially from these primary habitat preferences (Diniz et al. 2020). Therefore, identifying different movement behaviors and in particular actual dispersal movements prior to fitting step- or path-level resource selection functions has been shown to significantly improve derived estimates on functional landscape connectivity (Abrahms et al. 2017; Roever et al. 2013; Zeller et al. 2018).

Besides the differences in methodologies there are also two main concepts for deriving effective distances between locations based on the applied resistance surface: least-cost analysis (Adriaensen et al. 2003) and circuit-theory analysis (McRae 2006). The former estimates a single path of minimal cumulated landscape resistance connecting two entities in a landscape assuming that there is only a single best route (Singleton et al. 2002). In contrast, the circuit-theory based analysis estimates effective distances incorporating less informed random walks as alternative pathways between two entities (McRae and Beier 2007). Finally, in landscape genetic analyses these effective distances are linked to estimates of genetic differentiation using for example distance matrix correlations (Legendre and Fortin 2010; Mantel 1967; Storfer et al. 2010) or adaptions of linear mixed models (Row et al. 2017; Shirk et al. 2018).

Benchmarking Resistance Models

In order to apply landscape genetics for gaining a thorough understanding of functional connectivity and delineating conservation corridors for a given study species one is faced with deciding among a plethora of available methods to model landscape resistance, derive effective distances and comparing competing hypotheses on IBR, IBD and IBB with genetic data. The majority of landscape genetic studies on terrestrial mammals so far focused on large carnivores (Balkenhol et al. 2014; Mateo-Sánchez et al. 2015; Wasserman et al. 2013). Ungulates, especially from Central Europe, are to our knowledge highly underrepresented in studies using landscape genetics for inferring functional connectivity. Since estimates on connectivity are highly specific to the movement capacities, perceptual range, and ecological niche of a given organism, insight from the extensive literature on fitting resistance models or previous comparative studies on multiple approaches (e.g., Cushman et al. 2006; Squires et al. 2013; Zeller et al. 2018) are not directly transferable to a new target species.

For all these reasons, we provide a comparative analysis on estimating functional connectivity for red deer (*Cervus elaphus*) in Northern Germany. Red deer represent an interesting model species as they are one of the larges ungulates in Europe inhabiting various habitats across the continent (Borkowski and Ukalska 2008; Clutton-Brock et al. 1982; Kamler et al. 2008; Lande et al. 2013) and showing a high movement capacity with potential dispersal for long distances (Catchpole et al. 2004; Jarnemo 2007; Skog et al. 2009). However, anthropogenic fragmentation, game management and hunting as well as other restrictions caused by humans heavily impact this iconic game species (Hartl et al. 2003; Milner et al. 2006; Zachos et al. 2016). For example, previous studies have found indications for limited gene flow between local populations of red deer in our study area leading to a substantial loss of genetic diversity and even inbreeding (Zachos et al. 2007); Edelhoff et al. 2020; Chapter 2). Therefore, it is necessary to gain a better understanding of landscape effects on functional connectivity in order to enhance dispersal between local populations.

Previous studies have compared either a couple or up to multiple methodological approaches for estimating functional connectivity (Milanesi et al. 2016b; Reed et al. 2016; Squires et al. 2013; Zeller et al. 2018). However, these studies mostly utilize landscape genetics as one potential approach using e.g., causal-modelling for deriving a single resistance surface (Cushman, et al. 2006) and rarely as a means to validate the different models. Here, we utilize an extensive data set consisting of genetic samples and telemetry data to fit multiple resistance surfaces using different methodological approaches and apply a landscape genetic framework to compare them. Specifically, we distinguish three major categories of methodological approaches which are commonly used for modeling landscape resistance: expert-informed, habitat-informed and movement-informed models. From each category we applied multiple methods to estimate functional connectivity for red deer in Schleswig-Holstein. Ultimately, we aim to benchmark the various methodological approaches for inferring functional connectivity for red deer in our study area. This allows us to directly test the assumptions behind these models in a hypothesis framework using genetic data as a means to identifying the ecological processes (e.g., dispersal or habitat selection) that most likely influence functional connectivity for our target species (Spear et al. 2010).

For this, we developed a multi-step model selection framework: First, we apply both algorithms (circuit-theory and least-cost paths) to derive effective distances and correlate them to genetic distances. We rank the individual models within each category using Mantel tests and select only the highest-ranked models for the subsequent analyses. Further, the selected models are also combined into ensemble models which have been shown to perform well as a combination of multiple hypotheses and to overcome limitations of single methodological approaches (Araújo and New 2007). Second, we use model selection to compare the selected models as well as the ensemble models as alternative hypotheses on isolation by resistance while also accounting for isolation by distance and barrier as well as combinations of all competing hypotheses (Balkenhol et al. 2009).

Specifically we hypothesize that: 1) the ensemble resistance models essentially combining information from multiple approaches outperforms single-method resistance models in terms of explaining genetic distances. 2) That red deer are not needing primary habitat for dispersal and are capable of moving through unsuitable landscapes for certain distances. Therefore, we expect movement-based models to perform better in correlating effective distances to genetic distances as compared to habitat-informed models. 3) That corridor placement is highly depending on applied resistance surface and that the ensemble approach should ideally represent the information of multiple resistance surfaces to provide an effective tool in applied conservation and landscape planning.

Methods

Species and Study System

Our study focused on red deer from the federal state of Schleswig-Holstein (SH) located in Northern Germany (Figure 4.1). Despite their general wide distribution the local red deer populations are mostly restricted to smaller areas within this range where they are managed in administrative units (Edelhoff et al. 2020; see Chapter 2 of the dissertation). Primary deer habitat within these units comprises a high amount of forests as well as heath, moor- and wetlands (Meißner et al. 2008). Forested areas are mostly concentrated in patches which differ in size, level of fragmentation and tree composition. Outside the forests, the landscape consists of a mosaic of intensively used land forms, predominantly agriculture including crop-land, plantations, and pastures. A regional peculiarity are mound hedges, so called “Knicks”, which are a relic of extensive cultural land use in the past (Reif and Achtziger 2004). They are still found throughout the entire state and constitute linear features connecting forested areas in a landscape otherwise dominated by agriculture. Therefore, they are assumed to play an important role as structural element providing cover to red deer (Meißner et al. 2008; Davies and Pullin 2007). In terms of permeability for red deer the landscape is also impacted by settlements, urban areas and other forms of cultivation. Potential barriers such as primary roads and fenced highways (Autobahn) as well as larger waterbodies and canals, most prominently the Kiel Canal, are spread across the entire state (Figure 4.1).

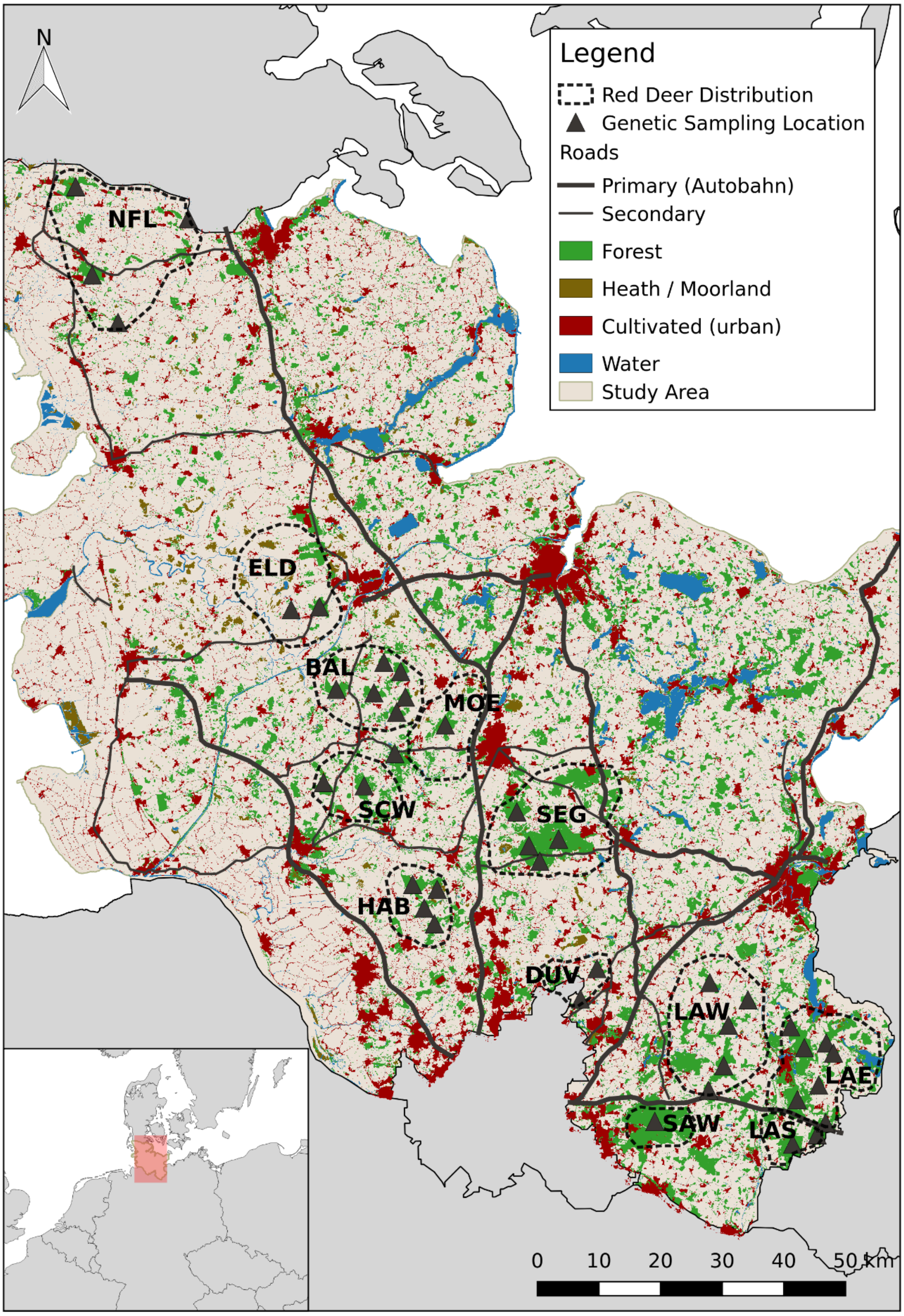


Figure 4.1 Map depicting the study area, main red deer distributions and their administrative management units (dashed outlines and abbreviations). The shown landscape features include forested areas, other habitats like heath (moorland), developed (cultivated or urban) areas and roads. The rest of the landscape is primarily dominated by agriculture (crop-land, pastures). Triangles indicate locations of genetic sampling. The inlet in the lower left corner provides information on the location of the study area in Central Europe.

Telemetry Data

The movement data utilized throughout the presented study included 61,532 telemetry relocations of 20 red deer individuals (14 male, 6 female). The original telemetry study was performed from 2009 to 2012 but overall runtime differed among individuals (mean runtime of 606 days). All animals were darted and equipped with GPS-collars (*Vectronic Aerospace GmbH, Berlin, Germany*). The considered individuals were located in different parts of the study area and covered regions where the majority of genetic samples were also obtained (Figure 4.2). Positions were recorded every 4 to 6 hours (on average every 5.2 hours). Detailed summary provided in the supplement (S4.1).

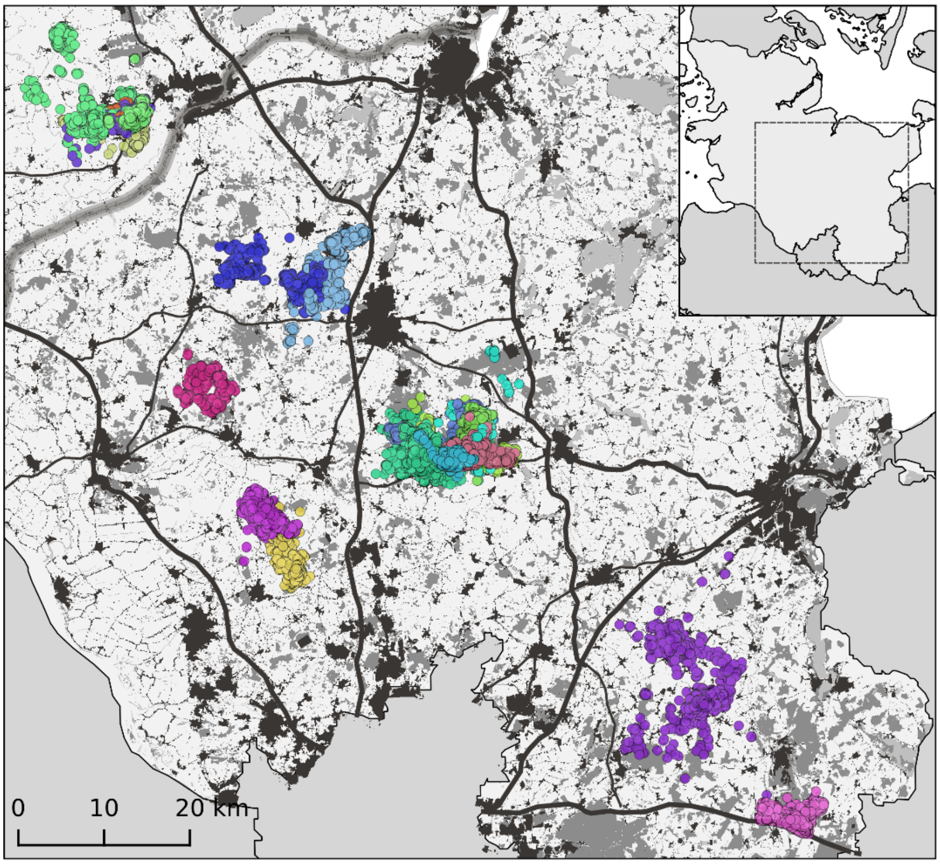


Figure 4.2 GPS-relocations of 20 red deer individuals across different regions of the study area. Different individuals are depicted in different colors. Inlet in upper right corner indicates the position of the mapped section of Schleswig-Holstein.

Genetic Data

A total of 434 genetic samples from red deer individuals across the entire distribution within Schleswig-Holstein were used for our analyses (Figure 4.1). Tissue samples were obtained from harvested animals during regular hunting seasons during the two periods of 2003/2004 and 2013/2014. For each sample we recorded the specific forest patch or areal unit where the individual was shot and used the centroid of the patch as the input location for our spatial analyses. We used 12 variable microsatellite marker to genotype the extracted DNA samples (supplement S4.2). All samples were screened for inconsistencies such as scoring errors, null alleles, significant deviation from Hardy-Weinberg equilibrium and linkage disequilibrium across each (Van Oosterhout et al. 2004; Excoffier et al. 2005; Rousset 2008). Detailed information on lab procedures, sequencing and applied microsatellite loci can be found in Edelhoff et al. (2020; Chapter 2).

Table 4.1: Summary of genetic samples used throughout the analyses. Information includes the main red deer distribution area (administrative management unit), number of red deer sampled in that area, and number of forest patches as origin of samples.

Area	**Abbreviation**	**Nr. Samples**	**Sampled Patches**
Barlohe	BAL	16	6
Duvenstedter Brook	DUV	23	2
Elsdorf / Westermuehlen	ELD	46	2
Hasselbusch	HAB	47	4
Moerel / Iloo	MOE	31	2
Lauenburg (east)	LAE	76	7
Lauenburg (south)	LAS	35	2
Lauenburg (west)	LAW	22	5
Nordfriesland	NFL	12	4
Sachsenwald	SAW	17	1
Schierenwald / Steinburg	SCW	14	2
Segeberger Heide	SEG	73	4

Landscape Variables

We chose a set of different input variables to model resistance of the landscape matrix in Schleswig-Holstein. The variables were based on multiple types of landcover representing either natural or anthropogenic features of the landscape (Table 4.2). Information on the extend of each landcover type was retrieved from ATKIS (authoritative topographic cartographic information system of Germany, www.adv-online.de) and turned into raster grids of 30m resolution. We transformed the categorical landcover data into continuous variables by calculating the distance to the closest feature. Additionally, we assessed the proportion (percentage) of each landcover type at radiuses of differing size: 100m, 200m, 500m, and 1000m. This allowed us to account for varying scales upon which the landcover types could influence landscape resistance (Baguette and Van Dyck 2007; Boyce et al. 2003; McGarigal et al. 2016). All grid calculations and preparations of spatial data were accomplished using *GRASS GIS* (Neteler et al. 2012).

Natural landcover types were hypothesized to decrease resistance and enhance permeability of the landscape for red deer whereas anthropogenic features should increase resistance and impact permeability negatively. In addition, we expected that linear (anthropogenic) structures, such as roads, have a negative impact due to their fragmentation effect. Mound hedges on the other hand, could provide cover and serve as a connecting element through their linear structure (Meißner et al. 2008). To account for the potential impacts of linear units consisting of either settlements (urban areas) or forest, we performed a morphological spatial pattern analysis (MSPA; Soille and Vogt 2009) for the forest and urban landcover type. The MSPA is implemented in the *GUIDOS* software (Vogt and Riitters 2017) and requires a binary map of a given landcover type (e.g., forest vs. non-forest) to partition the landscape into exclusive categories of different patterns (also see Soille and Vogt 2009). We chose the two linear categories (bridges and branches) and pooled them to form an additional landcover type for urban and forested areas (Table 4.2).

Table 4.2: List of variables used for modeling landscape resistance. Landcover classes were divided into a natural (environmental) and an anthropogenic category. Input variables were either implemented as percentage at different scales or as the distance to the nearest feature of a given landcover class.

Category	Landcover	Input Variable (units)
natural	forest	scale: percentage in radius (%)
		distance to nearest feature (meters)
	forest (linear feature)*	scale: percentage in radius (%)
	Knicks / hedgerows	scale: percentage in radius (%)
		distance to nearest feature (meters)
	heath/moorland	scale: percentage in radius (%)
		distance to nearest feature (meters)
	water (lakes, rivers, canals)	scale: percentage in radius (%)
		distance to nearest feature (meters)
anthropogenic	agriculture (crop-land, pastures)	scale: percentage in radius (%)
		distance to nearest feature (meters)
	urban (settlements, cultivated areas)	scale: percentage in radius (%)
		distance to nearest feature (meters)
	urban (linear features)*	scale: percentage in radius (%)
	roads (primary roads: autobahn; secondary roads)	scale: percentage in radius (%)
		distance to nearest feature (meters)

***(based on MSPA analysis implemented in the GUIDOS software)**

Modeling Landscape Resistance

All statistical analyses and data processing were performed using the R software package (R Core Team 2017) and various extensions for the specific modeling tasks.

Expert-Informed Resistance Models

We used two different resistance models based on expert-opinion. The first model (BUFFER) was originally developed by Meißner et al. (2008) and inferred connectivity based on a categorical map with varying buffer zones for different landcover types. The expert model distinguished between neutral (e.g., agriculture), positive (e.g., forests) and negative (e.g., urban) categories of landcover in terms of their permeability for red deer.

Additionally, the functional zone of influence of these categories was extended into the landscape with varying buffer sizes. The zone of influence (i.e., buffer) was based on the size of the respective areal unit of a given landcover category: for example small settlements (built-up areas) were assigned a small negative buffer zone whereas large urban patches where buffered with a negative zone of a larger extent (up to 1000m). In turn, positive buffer zones were placed around forested areas (more detail provided in appendix S4.3). In order to turn the resulting buffermap into a resistance surface we ranked the different categories and their buffer zones to values ranging from 0 (low resistance) to 1 (high resistance).

The second model using expert knowledge was based on a landscape mosaic analysis (LS_MOSAIC). For this we divided the landscape of our study area into three categories: natural, neutral, and developed areas. The division into these three categories was based on our assumptions on habitat requirements by red deer. We used the landscape mosaic analysis (Riitters et al. 2009; Wickham et al. 1994) implemented in the *GUIDOS* software (Vogt and Riitters 2017) to assign each grid cell of our landscape raster a new value accounting also for the categories of the neighboring grid cells. The analysis weighs the amount of each category in the surrounding cells and defines new clusters based on a trigonometric gradient of influence of the natural, neutral, or developed category (see supplement for further detail). In the next step, we used these mosaic clusters as a foundation for a linear weighting scheme (Clevenger et al. 2002). Resistance values of each mosaic cluster were linearly weighted based on the amount of natural, neutral and developed area within the cluster. Finally, we assigned resistance values to each mosaic cluster based on the combination of the three weights (table provided in supplement S4.3). This resulted in a final landscape model with resistance values ranging from zero to one.

Habitat-Informed Resistance Models

There were no records of presence or occurrence for red deer based on a systematic study available for our analyses. Therefore, we simulated presence data by drawing random subsamples of the GPS relocations using a grid-sampling approach (details provided in supplement S4.4). We also produced a similar number of random absence points. Among the multitude of available methods for modeling habitat suitability we chose the MAXENT algorithm (Phillips et al. 2004) as a presence-only model. Additionally, we fitted both, a generalized linear model (GLM) and a random forest model (RF; Breiman 2001) to our sets of pseudo presence and absence data. In addition, an SDM ensemble model was produced by calculating the mean of the three fitted habitat models (Stohlgren et al. 2010; Grenouillet et al. 2011). We first performed a preselection procedure to determine the final set of landscape variables in each of the three models: using univariate model comparison we selected the most relevant radius (scale) of the percentage-based variables and accounted for covariation among all considered variables (details provided in supplement S4.4). Finally, the remaining variables were applied in a multivariate, multi-scale model. The inverse of each model prediction was used to produce four different models of landscape resistance based on habitat suitability (Wang et al. 2008; Keeley et al. 2016).

Movement-informed Resistance Models

We used the GPS relocations of the 20 red deer individuals to estimate resource selection at different scales and during different types of movements in order to infer landscape resistance from these models. First, we estimated resource selection (RSF) at the home range level (third order RSF; Johnson 1980) using the telemetry relocations in a point-based model (Zeller et al. 2012). Therefore, we compared observed relocations within the home ranges of each individual with a set of weighted random points distributed within the same home range boundary. Coefficients for resource selection were estimated using logistic regression (Manly et al. 1993; Boyce et al. 2002).

As a second method based on movement data we applied a step-selection function (SSF) to account for the actual structure of observed movements (distribution of step length and change in direction of movements) when applying a used-availability study design (Fortin et al. 2005; Thurfjell et al. 2014). Observed movement steps were compared to random movement steps starting from the same origin. Placement of the random steps was sampled from the observed distributions of step-lengths and turning angles of each individual. Effects of landscape variables on the step-selection process were estimated with a conditional logistic regression model (Avgar et al. 2017; Duchesne et al. 2010). We further accounted for mixed effects of model coefficients caused by comparing multiple individuals (Craiu et al. 2011).

For the third movement-informed model we first performed a behavioral change point analysis (BCPA; Gurarie et al. 2009) to partition the individual movement paths into segments of potentially different behaviors (Gurarie et al. 2016). We clustered the resulting segments based on multiple path-characteristics (see Edelhoff et al. 2016; Chapter 3) and identified dispersal-like movements based on net-squared displacement (NSD; Morelle et al. 2017). Subsequently, we applied a step-selection function but this time only considering the relocations from cluster of path-segments with high levels of NSD. This way, we aimed to model habitat selection during movement phases which can be interpreted as dispersal which should provide a better estimate for landscape resistance compared to resource selection at the home-range level or based on all observed steps pooled together (Zeller et al. 2012, 2018). The model is referred to as BCPA_SSF.

Similar to the habitat suitability models we preselected the landscape variables for each of the two SSFs and the RSF based on a univariate model selection and tested for significant covariation. Subsequently, a final multivariate, multiscale model was fit for each of the three selection functions (Zeller et al. 2017). Based on the fitted coefficients the probability of selection (usage within home range or selection during movements) was predicted for the entire study area. Finally, the inverse of the selection probability was used as a resistance surface (Spear et al. 2015; Keeley et al. 2016). More detail on the three analytical procedures is provided in the supplement (S4.5).

Resistance Transformation

All of the compiled resistance surfaces resulted in values ranging between 0 and 1 as they were primarily derived from probability estimates of either habitat suitability or resource selection during movements. Other studies have shown that the relationship between these probability estimates and landscape resistance are not always strictly linear and intermediate conditions regarding, for example suitability, could either already indicate high levels of resistance and vice versa (see also Keeley et al. 2016; Zeller et al. 2018). Therefore, we applied three different transformation functions to each model in order to obtain the final resistance values ranging between 0 and 100. Specifically, we used a linear (lin), an inverse-reverse monomolecular (invrev-mono) and a monomolecular (mono) transformation function (Figure 4.3). All transformations were performed with the *ResistanceGA* package (Peterman 2014).

We further evaluated the distribution of the resistance values of each of the models as well as their transformations by placing a grid of 1000 sample points across the resulting resistance surfaces and extracting the values at each spatial location. Subsequently, we modeled the observed resistance values as a function of either the methodological group (EXPERT, HABITAT, MOVEMENT) or type of transformation (lin, mono, invrev-mono). We used generalized linear models and ranked them based on AIC values and explained deviance (Arnold 2010). In addition, we used a Spearman rank correlation to test the level of correlation among all models.

Isolation by resistance

Effective distances for each of the resistance surfaces were calculated using 1) circuit-theory based distance (CS) applying the *CIRCUITSCAPE* software implementation (McRae et al. 2016) and 2) least-cost path distance (LCP; Adriaensen et al. 2003) using the *gdistance* package (van Etten 2017). In order to decrease computational time all resistance surfaces were rescaled to 100m resolution using a bilinear interpolation prior the calculation of effective distances (Cushman and Landguth 2010b). We applied a patch-based sampling design and calculated pairwise distances between all centroids of the sampled forest patches (total of 41 locations; see Figure 4.1). Individuals sampled within the same focal patch were assigned a distance value of zero (Garroway et al. 2011). The resulting effective distances were used to test for isolation by resistance in our multi-step model selection.

Isolation by distance and barrier

We calculated pairwise geographic distances (Euclidean distances) between all sampled forest patches in order to account for isolation by distance (Balkenhol et al. 2009). The potential limitation of gene flow due to barriers was implemented in an additional pairwise distance matrix. We assumed that the Kiel Canal and all primary roads with fences (predominantly the Autobahn) constitute barriers to red deer (Figure 4.1 and supplement S4.7). Individuals sampled within the same area enclosed by either of this linear features where assigned a distance value of zero. Pairwise distances between individuals outside the same area were assigned distance values based on the number of barriers between their locations (e.g., one primary road between sampling locations = distance value of one). The resulting distance matrix was used as our isolation by barrier (IBB) hypothesis in our model selection framework (Figure 4.4).

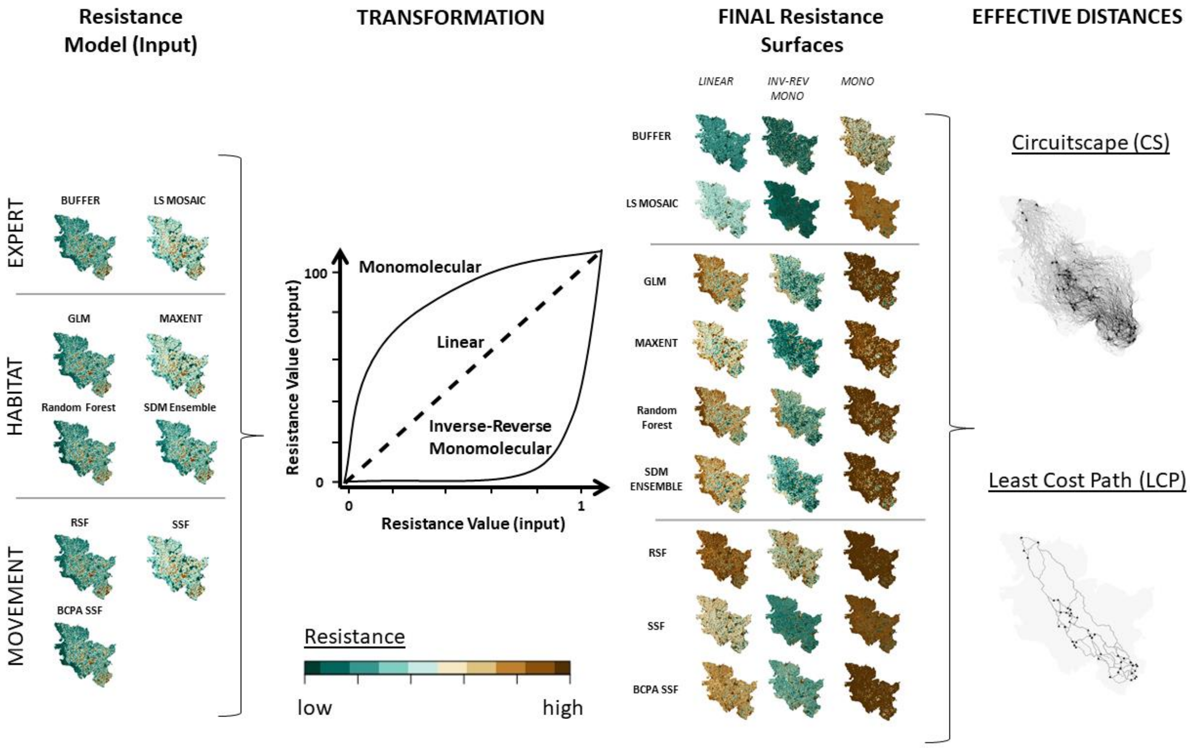


Figure 4.3: Flowchart depicting the applied steps for modeling landscape resistance based on different methodological approaches (expert-, habitat-, and movement-informed) followed by three different transformation functions. Finally, effective distances between sampling locations were calculated using circuit distance and least cost path distance (adopted from Zeller et al. 2018).

Genetic Distances

We utilized two different measures of pairwise genetic distances between all 434 sampled individuals. One being the inverse of the proportion of shared alleles (PSA; Bowcock et al. 1994) calculated as *PSA = 1 - (proportion of shared alleles)*. The second one was Rousset's a (Rousset 2000) calculated with the *SPAGEDI* software (Hardy and Vekemans 2002). Both represent individual-based genetic distances commonly used in landscape genetics (Shirk et al. 2017).

Landscape Genetic Model Comparison

We used a multi-step model selection framework for identifying the best performing resistance models (Figure 4.4). First, effective distances derived from all final resistance surfaces of each methodological group (Figure 4.3) were correlated to pairwise genetic distance using the Mantel test implemented in the *ecodist* package (Goslee and Urban 2007) with 1000 bootstraps for estimation of confidence intervals. The tests were performed separately for the two effective distance algorithms (CS, LCP) and the two measures of genetic distance: PSA and Rousset's a. The aim was to find the model with the highest correlation within each methodological group (EXPERT, HABITAT, MOVEMENT) and for each effective distance algorithm (step one of Figure 4.4). Subsequently, the three resulting models were combined into one ensemble model as the sum of the three resistance grids and effective distances were again calculated using the two ensemble models (ENSEMBLE_CS, ENSEMBLE_LCP) and the corresponding algorithm.

Second, we applied linear mixed effects models to link pairwise genetic distance with pairwise values based on effective distance (IBR), geographic distance (IBD) and barrier distance (IBB). As effective distances we used the three final models determined in the first step (EXPERT, HABITAT, MOVEMENT) as well as their two ensemble models. Again, we applied this procedure separately for CS and LCP based measures (step 2 in Figure 4). We fitted univariate as well as multivariate combinations of the three hypotheses IBR, IBD and IBB. Linear mixed models were fit with maximum likelihood population effects (MLPE; Clarke et al. 2002) implemented in the *resistanceGA* package (Peterman 2014). MLPE accounts for the pairwise structure of the input data and has been shown to outperform other statistical tests for correlating multiple distance matrices (Row et al. 2017; Shirk et al. 2018). Restricted maximum likelihood was set to false in order to make a valid comparison based on AICc values possible (Shirk et al. 2018). Models were compared using AICc values and the marginal R^2 coefficient for fixed effects (van Strien et al. 2012; Nakagawa and Schielzeth 2013). AIC weights were calculated separately for effective distances based on CS and LCP. Additionally, we used multiple regression on distance matrices (MRM; Legendre and Fortin 2010) to compare the same univariate and multivariate models as with the MLPE approach. MRM models were fitted with the *ecodist* package (Goslee and Urban 2007). Model performance of MRM was assessed based on R^2 (Shirk et al. 2018).

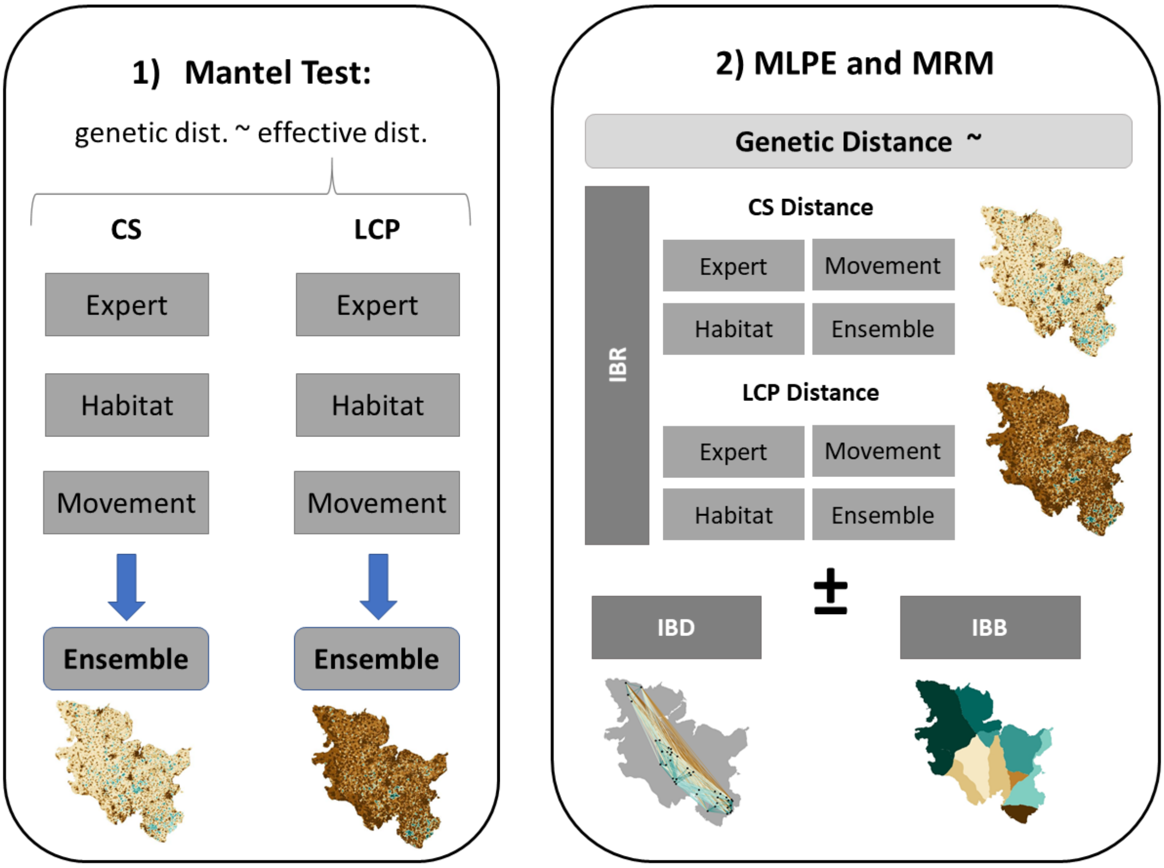


Figure 4.4: Illustration of the applied multi-step model selection framework. 1) selection of the highest performing resistance models of the expert-, habitat- and movement-informed approaches using Mantel correlation tests. The selection was done separately for circuitscape (CS) and least-cost path (LCP) measures of effective distance. 2) Mixed models with maximum likelihood population effects (MLPE) and multiple regression on distance matrices (MRM) were applied to compare models of isolation by resistance based on the highest ranked resistance surfaces along with the added ensemble model. IBD and IBB were also included and tested for individually and also in combination with each other and in addition to IBR models.

Corridor Placement

In order to compare the impact of applied resistance models on the placement of potential conservation corridors we assessed the overlap between corridors derived from the three highest ranked resistance models determined in the first step of our selection framework (Figure 4.4) as well as their ensemble combinations. Results for both effective distances (CS, LCP) were compared separately since the derived corridors are directly depending on the two applied algorithms. First, we calculated the correlation of the four cumulative conductance surfaces (EXPERT, HABITAT, MOVEMENT, ENSEMBLE) resulting from the CS analysis. However, conductance values of different resistance models are not directly comparable (Rudnick et al. 2012; Poor et al. 2012). Therefore, we additionally partitioned each conductance grid into quantiles and calculated the amount of overlap (in percent) for the same quantiles of separate models (Maiorano et al. 2017). Concurrently, we assessed the level of overlap between the corridors based on the resistance models considered for the least-cost analysis. We calculated LCPs between all focal patches using the *gdistance* package and placed a buffer of 100m around them. Subsequently, we calculated the overlap between LCPs of the three input models. In another step, we merged all buffered LCPs of the three resistance models into one single buffer-network and calculated the amount of overlap with buffered LCPs derived from the retrospective ensemble model. Mantel tests were applied to correlate the amount of overlap (calculated in percent) of buffered LCPs with geographic distance between patches.

Finally, we turned all buffered LCPs of the expert, habitat, and movement model into a binary raster with values of zero (outside of buffered LCP) and one (part of buffered LCP). The sum of all three raster grids was calculated to determine the level of overlap of the three LCP sets which resulted in values between zero (no LCP) to three (LCPs of all three models present). In turn, we correlated this raster with the four conductance surfaces and also calculated the percentage of overlap with their quantiles as described above. The aim of these analyses was to determine the level of agreement of the corridors derived from the different resistance models and the two algorithms CS and LCP.

Results

Resistance Surfaces

In total we compared 27 resistance surfaces derived from 9 different methodological approaches (Figure 4.4). The selected landscape variables as well as their inferred scales differed between all input models (see also results in supplements S4.3 – S4.5) but their overall influence on connectivity (model coefficients) were comparable among all fitted models (e.g., Table 4.7). Resistance values of the two expert-informed models were lower as compared to models based on the other two approaches (Figure 4.5). However, the highest amount of variation among resistance surfaces was caused by the transformation functions (see supplement S4.6 for more detail). Based on the univariate model comparison among a sample of 1000 random points variation in derived resistance values was best explained by the influence of the three transformation functions (D^2=0.457). Overall, the three transformation functions lead to similar shifts of resistance values across all models. Most noticeable, the monomolecular transformation resulted in an major increase of resistance whereas the inverse-reverse monomolecular transformation resulted in an decrease of the overall resistance (Figure 4.5).

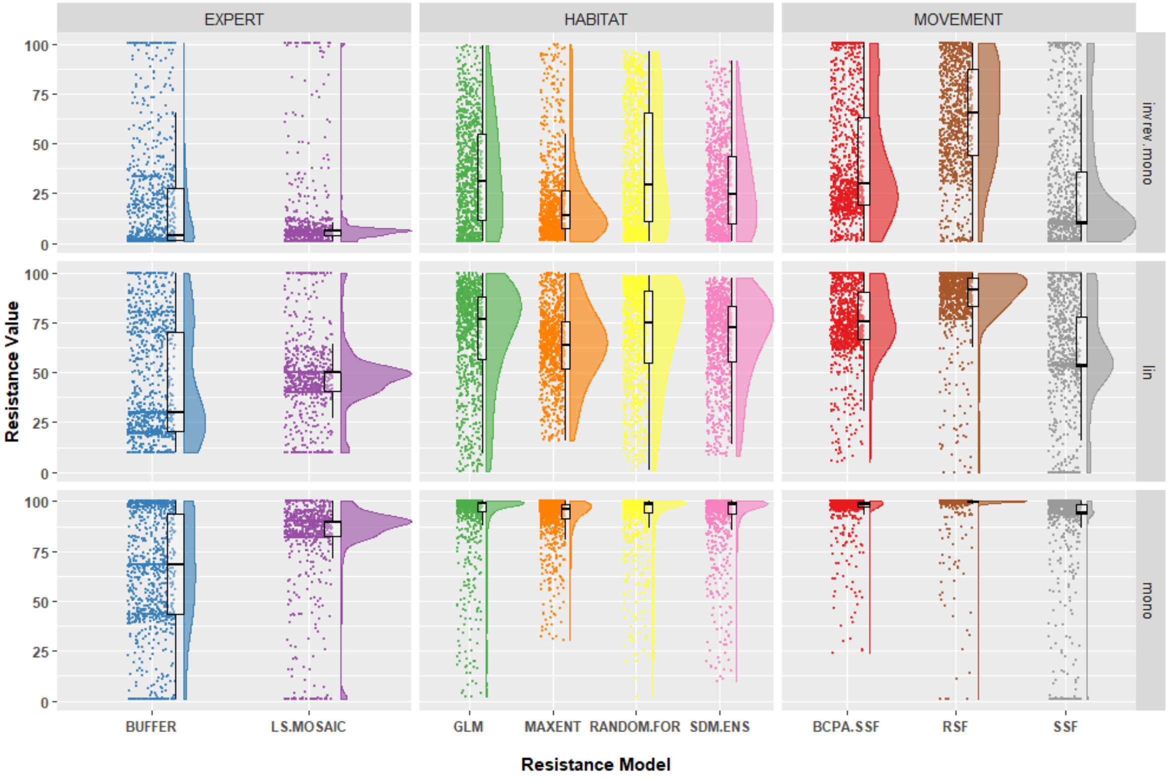


Figure 4.5: Distribution of resistance values for all compared models. Distributions are divided according to the three methodological categories (columns) and the three transformation functions (rows). Values were extracted from each resistance surface using 1000 sample points.

Multi-step model selection

The highest Mantel correlation between effective distances based on CS and both genetic distances was observed for the BUFFER model with *inverse-reverse monomolecular* [invrev_mono] transformation (Table 4.3). Comparable results were found for the movement-based model using the BCPA_SSF approach and a *monomolecular* transformation [mono]. For the habitat based approach the MAXENT model with *monomolecular* transformation showed the highest correlation using the CS distance whereas the *invrev_mono* transformation was higher correlated with LCP distance. Based on these results we calculated one ensemble model for CS distance using the sum of the BUFFER[invrev_mono], MAXENT[mono] and BCPA_SSF[mono] resistance surfaces

(ENSEMBLE_CS). We repeated this procedure for the LCP based ensemble (ENSEMBLE_LCP) but using the *invrev_mono* transformation of the MAXENT model. However, The two ENSEMBLE models did not perform better for any of the effective and genetic distances based on the Mantel test results (Table 4.3). Overall, many of the correlation coefficients differed only at the third digit (results for all models are presented in the supplement S4.7).

Table 4.3: Results of Mantel tests correlating pairwise genetic distances (PSA, Rousset's a) with effective distances based on Circuitscape (CS) and least-cost path distances (LCP). Only Mantel correlation coefficients (r2 values) for the highest ranked models of each methodological category (expert-informed, habitat-informed, movement-informed) as well as the combination of them (ENSEMBLE) are shown. All values were significantly different from zero based on a bootstrap test.

Resistance Distance	**PSA**	**Rousset's a**
Circuitscape (CS):		
Expert: BUFFER [invrev-mono]	0.246	0.219
Habitat: MAXENT [mono]	0.236	0.218
Movement: BCPA_SSF [mono]	0.241	0.215
ENSEMBLE_CS	0.241	0.218
Least Cost Path (LCP):		
Expert: BUFFER [invrev-mono]	0.216	0.191
Habitat: MAXENT [invrev-mono]	0.212	0.187
Movement : BCPA_SSF [mono]	0.208	0.185
ENSEMBLE_LCP	0.208	0.183

Next, we applied a model selection approach for comparing the top ranked resistance surfaces (EXPERT, HABITAT, MOVEMENT, ENSEMBLE) as independent hypotheses testing for isolation by resistance (IBR). Overall, the two genetic distances produced similar results. Therefore, only data for models using PSA as the dependent variable (genetic distance) are shown here.

The model selection approach based on linear mixed models (MLPE) and AICc ranked the habitat-informed model (MAXENT [mono]) combined with IBD and IBB highest for the CS distance (Table 4.4). The ENSEMBLE_CS resistance model in combination with IBD and IBB was ranked third but already with an ΔAICc of 160. Marginal R^2 values of all MLPE models were very similar (Table 4.4). The highest correlations ($R^2_{glmm'm}$ = 0.122) were observed for the habitat-informed model both with and without accounting for IBD (*habitat; habitat+ibd*). For LCP distance the highest agreement based on AICc with genetic distance was observed for the movement-informed model (BCPA_SSF [mono]) in combination with IBD and IBB (Table 4.4). Again, the *ensemble+ibd+ibb* model combination was ranked third but with an ΔAICc of 741. Marginal R^2 coefficients did not indicate a lot of variation. The highest correlation ($R^2_{glmm'm}$ = 0.143) also supported the movement-informed model but only in combination with IBD (Table 4.4).

Results based on the correlation coefficient of the matrix regression models (MRM; Table 4.4) differed from the MLPE results. For both, CS and LC distances, the *ensemble+ibd+ibb* as well as the *expert+ibd+ibb* model combinations showed the highest level of correlation. However, correlation coefficients across all models differed mostly at the third digit and did not indicate a lot of variation in terms of the level of agreement between genetic and effective distances.

Table 4.4: Model selection results for effective distances derived with Circuitscape (CS) and least-cost paths (LCP). Coefficients, AICc, ΔAICc, weights and marginal correlation coefficient (R2glmm’m) values are shown for the linear mixed models with maximum likelihood population effects (MLPE). Additionally, correlation coefficients R2 based on multiple regression on distance matrices (MRM) are presented.

Model	**Coefficients**			**MLPE**				**MRM**
	geo.dist	ibb.dist	eff.dist	AICc	ΔAICc	weight	$R^2_{glmm'm}$	R^2
<u>Circuitscape</u>								
habitat+ibd+ibb	-0.01	0.007	0.034	-214299	0	1	0.115	0.063
habitat+ibb		0.005	0.027	-214214	85	0	0.117	0.061
ensemble+ibd+ibb	-0.007	0.007	0.030	-214139	160	0	0.112	0.066
habitat+ibd	-0.002		0.034	-214124	175	0	0.122	0.056
habitat			0.032	-214121	178	0	0.122	0.056
ensemble+ibb		0.005	0.026	-214098	201	0	0.114	0.062
movement+ibd+ibb	-0.002	0.007	0.025	-213979	320	0	0.110	0.065
movement+ibb		0.007	0.024	-213978	321	0	0.111	0.063
ensemble			0.032	-213969	330	0	0.118	0.058
ensemble+ibd	0.001		0.031	-213969	330	0	0.118	0.059
movement+ibd	0.006		0.025	-213804	495	0	0.117	0.058
movement			0.031	-213738	561	0	0.116	0.058
expert+ibd+ibb	-0.008	0.005	0.033	-213711	588	0	0.112	0.066
expert+ibb		0.004	0.027	-213675	624	0	0.111	0.064
expert+ibd	-0.004		0.034	-213638	661	0	0.117	0.061
expert			0.031	-213630	669	0	0.115	0.061
ibd+ibb	0.022	0.008		-212871	1428	0	0.100	0.053
ibd (null)	0.030			-212693	1606	0	0.107	0.046
ibb		0.023		-211918	2381	0	0.081	0.049
<u>Least-Cost Path</u>								
movement+ibd+ibb	-0.138	0.015	0.158	-214024	0	1	0.135	0.053
movement+ibd	-0.085		0.120	-213435	589	0	0.143	0.049
ensemble+ibd+ibb	-0.089	0.010	0.111	-213283	741	0	0.120	0.054
movement+ibb		0.007	0.024	-213180	845	0	0.109	0.053
expert+ibd+ibb	-0.032	0.005	0.057	-213167	858	0	0.106	0.054
expert+ibd	-0.034		0.064	-213091	934	0	0.111	0.047
expert+ibb		0.005	0.024	-213067	957	0	0.104	0.053
movement			0.032	-213032	993	0	0.118	0.043
ensemble+ibb		0.007	0.023	-213022	1003	0	0.106	0.053
habitat+ibd+ibb	-0.004	0.007	0.026	-212993	1031	0	0.102	0.053
habitat+ibb		0.007	0.022	-212993	1032	0	0.102	0.053
expert			0.030	-212981	1043	0	0.110	0.047
ensemble+ibd	-0.056		0.089	-212971	1054	0	0.128	0.050
ibd+ibb	0.022	0.008		-212871	1153	0	0.100	0.053
ensemble			0.031	-212854	1170	0	0.115	0.043
habitat+ibd	0.004		0.027	-212821	1203	0	0.110	0.046
habitat			0.030	-212821	1204	0	0.110	0.045
ibd (null)	0.030			-212693	1332	0	0.107	0.046
ibb		0.023		-211918	2106	0	0.081	0.049

Corridor Comparison

Placement of corridors was depending on the applied methodological approach (Figure 4.6). Detailed results of conductance surfaces and LCPs for the three highest ranked models are presented in supplements S4.10 and S4.11. In case of the conductance surfaces the areas of high conductance were more restricted for the BUFFER [invrev-mono] model as compared to the MAXENT [mono] and BCPA_SSF [mono] models (see maps in supplement S4.10).

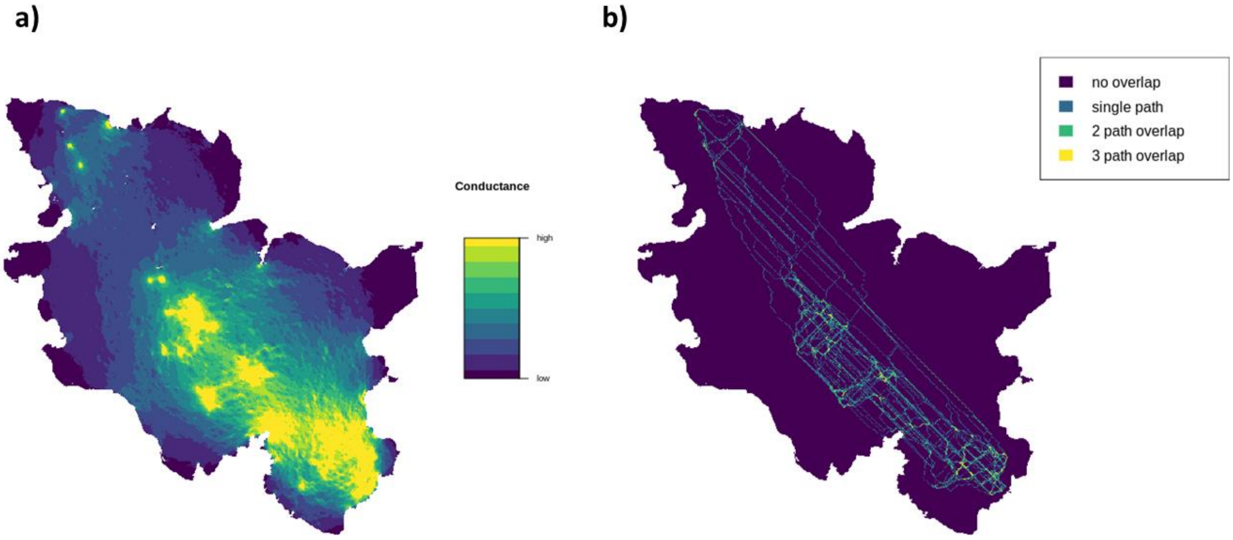


Figure 4.6: a) comparison of cumulative conductance surface (result from Circuitscape) based on the ENSEMBLE_CS resistance model and b) the overlap of the least-cost paths for all three models selected based on LCP distance: BUFFER[invrev-mono], MAXENT[invrev-mono], and BCPA_SSF[mono].

Quantiles of the ENSEMBLE_CS conductance surface largely agreed with those of the individual models. In particular, for areas of the lowest and highest quantile we observed an overlap of over 90% between either the habitat- or movement-informed models and the ensemble (Table 4.5). The expert-informed model (BUFFER[invrev-mono]) showed lower levels of agreement with the output of ENSEMBLE_CS model. Here, only up to 52% of the areas within the lowest overlapped with the ensemble results and the degree of overlap was even lower within the remaining quantiles (Table 4.5).

Comparable results were derived from the Spearman rank correlation (supplement S4.12): ENSEMBLE_CS conductance was highly correlated with the habitat- and movement-informed models (R_S = 0.97) whereas correlation with the conductance surface resulting from the expert-informed model was lower (R_S = 0.71).

Table 4.5 Results of the overlap analysis between the quantiles of the conductance surfaces of the expert-, habitat- and movement- informed resistance models and their ensemble model. Additionally, the sum of overlapping LCPs based on the three highest ranked resistance models and the ensemble conductance are shown at the bottom of the table.

Quantile:	**1**	**2**	**3**	**4**
BUFFER[invrev-mono] vs. ENSEMBLE _CS	52%	19%	20%	48%
MAXENT[mono] vs. ENSEMBLE _CS	93%	84%	84%	92%
BCPA_SSF[mono] vs. ENSEMBLE _CS	91%	82%	82%	91%
LCP Overlap vs. ENSEMBLE _CS	26%	2%	0%	1%

As shown in Figure 4.6 placement of least cost paths varies between the three resistance models. In particular, LCPs between patches located further apart showed less overlap as compared to paths between locations at smaller distances (see also supplement S4.11). The level of overlap between the LCPs of the three input models (BUFFER[invrev-mono], MAXENT[invrev-mono] and BCPA_SSF[mono]) were negatively correlated with spatial distance (Mantel R^2 = -0.53; $p < 0.001$). Further, we observed a significant decrease (Mantel R^2 = -0.46; $p < 0.001$) in the level of overlap between the combined LCPs of the these models and the LCPs derived from their ensemble model (ENSEMBLE_LCP] with increasing spatial distance (Figure 4.7). Overall, LCPs of different models overlapped up to 30% at short distances (Figure 4.7).

Finally, we observed only a low level of agreement between the conductance surface of ENSEMBLE_CS model and the sum of overlapping LCPs (Spearman correlation R_S = 0.27; see S4.12). However, buffered LCPs take up only a small area compared to the conductance surfaces and regions where all three LCPs overlapped made up 1% of the highest quantile of the ENSEMBLE_CS conductance surface (Table 4.5).

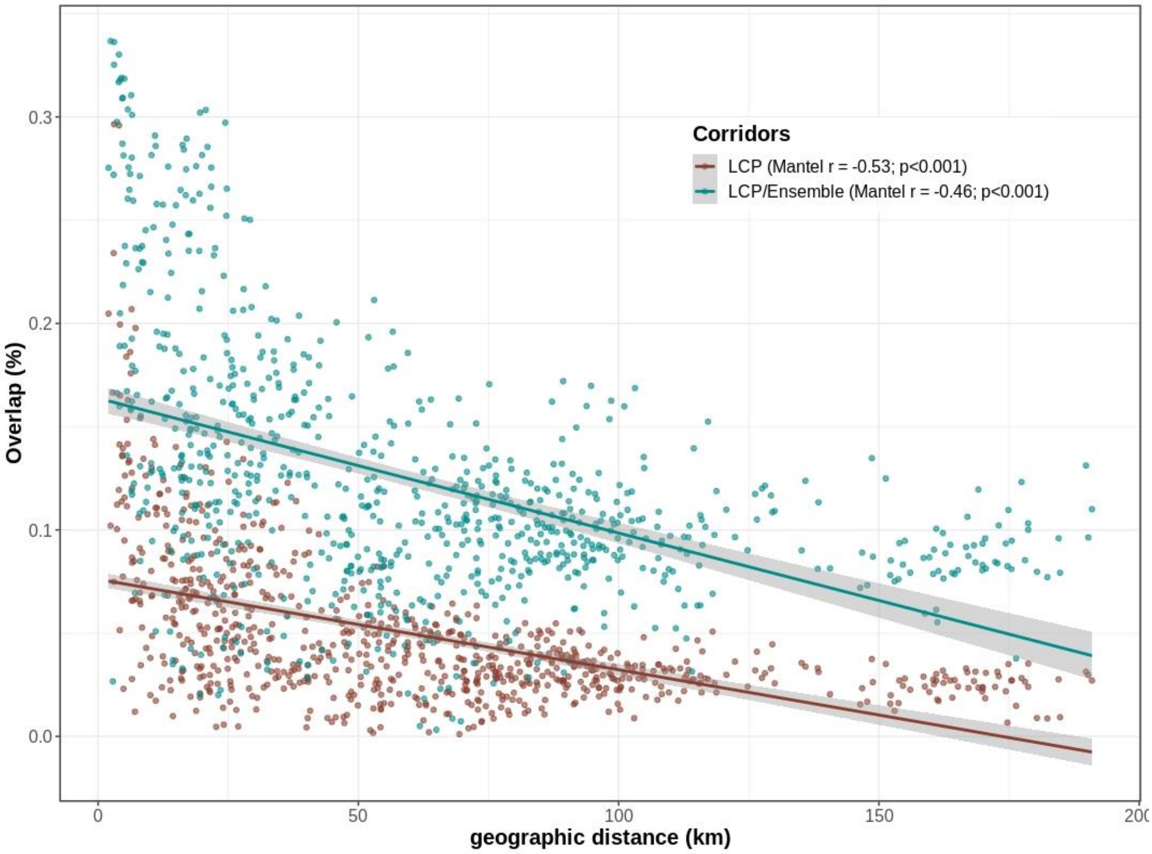


Figure 4.7 Correlation of geographic distance between habitat patches and level of overlap of LCPs derived from the highest ranked resistance models. Results shown for percent overlap between buffered LCPs of the three highest ranked models only (brown) and overlap between combined LCPs of the three models and the LCPs of their ensemble model (green).

Discussion

Benchmarking Resistance Models

Performance of resistance models heavily relies on how effective distances are quantified and how they are compared against genetic distances. Our results showed major differences between the two applied effective distance algorithms (circuitscape vs least-cost path) as well as the statistical approach for linking these effective distances with genetic distances (Mantel tests, MLPE and MRM). In contrast, the choice of the genetic distance parameter did not impact the results in our individual-based study design. We chose the PSA distance since it indicated higher correlations with effective distances but the multi-step model selection produced similar results for genetic-distance based on Rousset's a (data not shown here).

In general, we observed lower AICc values (MLPE) and higher correlations (MRM, Mantel test) for effective distances based on Circuitscape compared to LCP distance based models. Based on the multi-step model selection among the 27 resistance surfaces we did not find any indications that the ensemble approach clearly outperforms all individual resistance models. Still, the ensemble models were among the highest ranked models for both CS and LCP effective distances. Ensemble models are increasingly used in studies on species distribution to overcome limitations of single model algorithms and to delineate areas with high model agreement (Araújo and New 2007; Stohlgren et al. 2010). Therefore, corridors derived from such ensembles of multiple resistance models could potentially improve the shortcomings of relying only on single approaches and the assumptions behind them.

The answer to our second hypothesis that movement-informed resistance models outperform habitat-informed approaches in our study system was highly depending on the considered effective distance. Circuitscape distances derived from the selected habitat approach (MAXENT [mono]) best described the genetic distances in combination with IBD and IBB (Table 4.5). On the contrary, least-cost path distances based on the movement-informed resistance model (BCPA_SSF [mono]) performed best, again in combination with IBD and IBB. In most of the cases the expert-informed model (BUFFER [invrev-mono]) performed less well as the empirical approaches but here as well we observed differences between CS and LCP distances.

Generally, we found evidence for isolation by distance and barriers playing a key role in genetic differentiation of red deer in Schleswig-Holstein as both of the best-performing models for CS and LCP effective distances included the IBD and also IBB terms next to accounting for IBR. This confirms the results of a previous study on red deer populations in SH (Edelhoff et al. 2020; Chapter 2) which determined significant IBD based on population-based differentiation (F_{ST} values). The putative barriers tested for in our IBB hypothesis were the primary roads which are almost entirely fenced as well as the Kiel Canal. All of these have been shown to influence the structure of local red deer populations (STRUCTURE analysis in Chapter 2; Edelhoff et al. 2020).

The most significant cause for variation in resistance values was explained by the applied transformation functions (Figure 4.5). Overall, the best-performing models in our case indicated non-linear relationships between estimates of habitat suitability which confirms to results of other studies (Keeley et al. 2016; Zeller et al. 2017). In particular, models with monomolecular transformation showed higher correlations of effective distances with genetic distances.

We did not observe explicit differences in model performance when applying correlation based analyses. Overall, correlation coefficients based on either Mantel tests or multiple regression on distance matrices (MRM) showed only little variation with differences occurring at the third digit for some of the compared models (see Tables in supplement S4.8 and Table 4.4). Resistance values of the nine original input models were, to some extent, highly correlated (Table S4.6.2). This could ultimately lead to a high correlation of our tested hypotheses (expert-, habitat- and movement-informed models) which potentially limits the explanatory power of analyses relying solely on matrix correlations (Cushman and Landguth 2010a; Cushman et al. 2013; Zeller et al. 2016). Nevertheless, we assume that the Mantel test is a valid approach to select the highest-ranked resistance models within the same methodological group (expert-, habitat- and movement-informed) as performed in our first step of the model selection process (Cushman et al. 2006). In our case, differences in terms of model performance describing observed genetic distances was most prominent in AICc values derived from linear mixed models with MLPE. Such AIC-based model selection has been shown to provide an effective means with high accuracy for comparing multiple hypotheses on landscape resistance (Row et al. 2017; Shirk et al. 2018; Spear et al. 2015). However, the influence and non-consistency among the statistical approaches makes it difficult to benchmark the different methods for estimating functional connectivity. We therefore primarily focused here on the results of the MLPE models in the second part of our multi-step model selection analyses.

Comparing movement- and habitat-informed resistance models

The resistance model based on step-level resource selection using only potential dispersal movements was ranked highest among the three movement-informed models in our pre-selection (S4.4). This confirms other studies that already pointed out the importance of accounting for the actual dispersal process over general resource selection at the home range scale for estimating functional connectivity (Blazquez-Cabrera et al. 2016; Squires et al. 2013; Zeller et al. 2018). For example, the selected landscape variables and their coefficients (i.e., influence on selection probability) varied between the BCPA_SSF model in comparison to the highest-ranked, habitat-informed resistance model (MAXENT; Table 4.7). The latter represents more the primary habitat requirements of red deer without accounting for any movement processes (Vasudev et al. 2015). The functional scales of the most important landcover features, such as the amount of forest or urban (developed) areas, varied between the two models (Table 4.7). In summary, the step-level resource selection accounted for variables at a much smaller scale (100-200m) whereas the MAXENT model considered the same variables at a larger scale (1000m). This has an major impact on the evaluation of the influence of the landscape matrix on functional connectivity. Therefore, we observed the most prominent differences in the spatial prediction of these two input models in areas in-between the major habitat complexes (see supplement S4.13).

However, as pointed out earlier, depending on how effective distances are estimated, either the movement-informed or the habitat-informed model explained the observed genetic patterns better. Previous studies have indicated that wide-range connectivity is better represented in CS based distances (McRae and Beier 2007; Spear et al. 2015). Whereas others have shown that LCP distances are more suitable for deriving short range connectivity or modeling actual dispersal (Driezen et al. 2007; Sawyer et al. 2011; Zeller et al. 2018). Hence, we interpret our results that for short distances red deer are capable of moving through less suitable habitats (i.e. a hostile landscape matrix with high resistance) but for large scale connectivity and dispersal over long distances general habitat requirements need to be present, e.g. in small areas serving as stepping stones (Epps et al. 2007; Saura et al. 2014).

Overall, movement-informed resistance models are a promising approach for estimating connectivity but certainly have their limitations as well (Abrahms et al. 2017; Spear et al. 2010; Vasudev and Fletcher 2015). As in our case, telemetry studies, the primary source for movement data, are usually limited by the number of tagged (i.e. sampled) individuals. Next to that, sampled individuals are a major cause of observed variation in resource selection (Gillies et al. 2006; Wirsing and Heithaus 2014; Wittemyer et al. 2008). Further, observed movements most likely do not represent actual dispersal. This could be due to the age or sex of the sampled individuals in the data set. For example, in many species dispersal decisions are primarily found in young individuals (natal dispersal) and less likely to occur in adults (Elliot et al. 2014; Roffler et al. 2014).

Therefore, a study only tracking adult individuals may get misleading estimates in such movement models which might affect subsequent parameterization of resistance models. The movement data applied in our study was based on sampled individuals of various ages (see Table S4.1.1) and included only a marginal amount of dispersal movements. Individuals were mostly stationary (visual inspection; see Figure S4.5.2) with only a few excursions to areas outside of the core home ranges. Nevertheless, in our opinion, resource selection based on movement data from actual dispersal events is a key element for modeling functional landscape connectivity, even if overall sample size is low (see for example Zeller et al. 2017).

Table 4.7 Summary of the most important landcover variables included in the best performing resistance models. Since the applied methods are not directly comparable only the type of considered variables and their effect (i.e. slope of the coefficient) are presented here: binary (landcover present/not present), distance (to nearest feature of landcover) and the zone of influence [m] (e.g. based on a buffer or as the proportion of cover within a certain radius). Effects indicate whether the given landcover type increases (+), decreases (-) or has neutral influence (0) on landscape resistance.

Landcover	BUFFER		MAXENT		BCPA_SSF	
	Variable	Effect	Variable	Effect	Variable	Effect
Agriculture	binary	0	distance	0	*Not incl.*	*Not. Incl.*
Forest	200m	-	1000m	-	100m	-
Urban	100m - 1000m	+	1000m	+	200m	+
Roads	binary	+	distance 100m	0 +	100m	+
Water	binary	0	distance 100m	0 +	distance	0
Wetland/Heath	200m	-	distance	-	distance	0

Management Implications

The most relevant question for applying landscape genetics in conservation is the optimal placement (delineation) of corridors to facilitate or establish gene flow between populations in fragmented landscapes (Cushman et al. 2010; Koen et al. 2012; Rudnick et al. 2012). As many other studies before, our results show that deriving corridors either as conductance surfaces using Circuitscape or least-cost paths is highly depending on the applied resistance surface (Blazquez-Cabrera et al. 2016; Graves et al. 2014; McRae et al. 2016).

The results of our study indicated that only at shorter distances the placement of corridors is less impacted by the choice of resistance model since we observed significant decrease in the degree of overlapping LCPs with geographic distance (Figure 4.7). As discussed earlier, the ensemble models did not outperform other resistance models or improved the correlation between effective distances and genetic distances. Still, in terms of delineation of conservation corridors the ensemble models represented good summarizations of the other input models. For example, the conductance surface of the ENSEMBLE_CS model showed high correlation and overlap with the surfaces based on the three other models (habitat-, expert- and movement-informed) and to a certain degree with summarized LCPs of the three models as well. We thus do not want to discourage the application of ensembles if multiple sources of information are available to fit various types of resistance models based on habitat- or movement-information next to expert-opinion. Future research should focus on how this approach can be improved for example by applying weighting schemes to account for explanatory power of the individual models included in the ensemble (Araújo and New 2007).

Barriers such as primary roads play an important role in landscape connectivity for red deer in Schleswig-Holstein and have produced significant differentiation and restricted gene flow between local populations. In terms of landscape planning and conservation management, this calls for improvement by providing linkages or overpasses to mitigate the negative barrier effects (Beier et al. 2008; Corlatti et al. 2009; Epps et al. 2005). Optimal placement of these overpasses along the detected barriers should ideally also be derived from the corridors based on the ensemble models to ensure maximum efficiency (Epps et al. 2007; Sawyer et al. 2011).

Overall, the different models show unanimously that forests, the amount of urban (cultivated) areas as well as roads are the main factors influencing functional connectivity for red deer in Schleswig-Holstein (Table 4.7 and results in S4.3 to S4.5). Other landcover classes such as agriculture, water bodies or wetlands and heath play more of a subordinate or neutral role. The buffermap model (BUFFER) was the best performing resistance surface among the two expert-informed approaches. Although it is not directly comparable with the other approaches as it lacks any empirically derived coefficients it nevertheless shares many properties that are similar to the regression coefficients of the habitat- and movement-informed methods used in our analyses. Although the results are very specific to the study area, the BUFFER model could potentially be a useful tool for analyzing landscape connectivity in other parts of Germany where low genetic diversity and limited gene flow in red deer occurrences have recently been detected (Reiner and Willems 2019).

References

Abrahms, B., Sawyer, S. C., Jordan, N. R., McNutt, J. W., Wilson, A. M., and Brashares, J. S. (2017). Does wildlife resource selection accurately inform corridor conservation? *Journal of Applied Ecology, 54*(2), 412–422.

Adriaensen, F., Chardon, J. P., De Blust, G., Swinnen, E., Villalba, S., Gulinck, H., and Matthysen, E. (2003). The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning, 64*(4), 233–247.

Andersen, L. W., Fog, K., and Damgaard, C. (2004). Habitat fragmentation causes bottlenecks and inbreeding in the European tree frog (*Hyla arborea*). *Proceedings of the Royal Society B: Biological Sciences*, *271*(1545), 1293–1302.

Araújo, M. B., and New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, *22*(1), 42–47.

Arnold, T. W. (2010). Uninformative Parameters and Model Selection Using Akaike's Information Criterion. *Journal of Wildlife Management*, *74*(6), 1175–1178.

Avgar, T., Lele, S. R., Keim, J. L., and Boyce, M. S. (2017). Relative Selection Strength: Quantifying effect size in habitat- and step-selection inference. *Ecology and Evolution*, *7*(14), 5322–5330.

Baguette, M., Blanchet, S., Legrand, D., Stevens, V. M., and Turlure, C. (2013). Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews*, *88*(2), 310–326.

Baguette, M., and Van Dyck, H. (2007). Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. *Landscape Ecology*, *22*(8), 1117–1129.

Balkenhol, N., Holbrook, J. D., Onorato, D., Zager, P., White, C., and Waits, L. P. (2014). A multi-method approach for analyzing hierarchical genetic structures: A case study with cougars Puma concolor. *Ecography*, *37*(6), 552–563.

Balkenhol, N., Waits, L. P., and Dezzani, R. J. (2009). Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography*, *32*(5), 818–830.

Beier, P., Majka, D. R., and Spencer, W. D. (2008). Forks in the road: choices in procedures for designing wildland linkages. *Conservation Biology*, *22*(4), 836–851.

Beier, P., and Noss, R. F. (2008). Do Habitat Corridors Provide Connectivity? *Conservation Biology*, *12*(6), 1241–1252.

Bélisle, M. (2005). Measuring landscape connectivity: the challenge of behavioral landscape ecology. *Ecology, 86*(8), 1988–1995.

Betts, M. G., Gutzwiller, K. J., Smith, M. J., Robinson, W. D., and Hadley, A. S. (2015). Improving inferences about functional connectivity from animal translocation experiments. *Landscape Ecology*, *30*(4), 585-593.

Blazquez-Cabrera, S., Gaston, A., Beier, P., Garrote, G., Simon, M. A., and Saura, S. (2016). Influence of separating home range and dispersal movements on characterizing corridors and effective distances. *Landscape Ecology*, *31*, 2355–2366.

Borkowski, J., and Ukalska, J. (2008). Winter habitat use by red and roe deer in pine-dominated forest. *Forest Ecology and Management*, *255*(3–4), 468–475.

Bowcock, A., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J., and Cavalli-Sforza, L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature, 368* (6470), 455-457.

Boyce, M. S., Vernier, P. R., Nielsen, S. E., and Schmiegelow, F. K. . (2002). Evaluating resource selection functions. *Ecological Modelling*, *157*(2–3), 281–300.

Boyce, M. S., Mao, J. S., Merrill, E. H., Fortin, D., Turner, M. G., Fryxell, J., and Turchin, P. (2003). Scale and heterogeneity in habitat selection by elk in Yellowstone National Park. *Ecoscience*, *10*(4), 421–431.

Braunisch, V., Segelbacher, G., and Hirzel, A. H. (2010). Modelling functional landscape connectivity from genetic population structure: a new spatially explicit approach. *Molecular Ecology*, *19*(17), 3664–3678.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Broquet, T., Ray, N., Petit, E., Fryxell, J. M., and Burel, F. (2006). Genetic isolation by distance and landscape connectivity in the American marten (Martes americana). *Landscape Ecology*, *21*(6), 877–889.

Catchpole, E. A., Fan, Y., Morgan, B. J. T., Clutton-Brock, T. H., and Coulson, T. (2004). Sexual dimorphism, survival and dispersal in red deer. *Journal of Agricultural, Biological, and Environmental Statistics*, *9*(1), 1–26.

Chetkiewicz, C.-L. B., St. Clair, C. C., and Boyce, M. S. (2006). Corridors for Conservation: Integrating Pattern and Process. *Annual Review of Ecology, Evolution, and Systematics*, *37*(1), 317–342.

Clarke, R. T., Rothery, P., and Raybould, A. F. (2002). Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics*, *7*(3), 361–372.

Clevenger, A. P., Wierzchowski, J., Chruszcz, B., and Gunson, K. (2002). GIS-Generated, Expert-Based Models for Identifying Wildlife Habitat Linkages and Planning Mitigation Passages. *Conservation Biology*, *16*(2), 503–514.

Clutton-Brock, T. T. H., Guinness, F. F. E., Albon, S. D., and Albon. (1982). *Red Deer - Behavior and Ecology of Two Sexes*. University of Chicago press.

Corlatti, L., Hackländer, K., and Frey-Roos, F. (2009). Ability of wildlife overpasses to provide connectivity and prevent genetic isolation. *Conservation Biology : The Journal of the Society for Conservation Biology*, *23*(3), 548–556.

Coulon, A., Cosson, J., Angibault, J.-M., Cargnelutti, B., Galan, M., Morellet, N., Petit, E., Aulagnier S., and Hewison, A. (2004). Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual–based approach. *Molecular Ecology*, *13*, 2841–2850.

Craiu, R. V., Duchesne, T., Fortin, D., and Baillargeon, S. (2011). Conditional logistic regression with longitudinal follow-up and individual-level random coefficients: A stable and efficient two-step estimation method. *Journal of Computational and Graphical Statistics*, *20*(3), 767–784.

Crooks, K., and Sanjayan, M. (2006). *Connectivity conservation*. Cambridge University Press.

Cushman, S. A., McKelvey, K. S., Hayden, J., and Schwartz, M. K. (2006). Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *The American Naturalist*, *168*(4), 486–499.

Cushman, S. A., and Landguth, E. L. (2010a). Spurious correlations and inference in landscape genetics. *Molecular Ecology*, *19*(17), 3592–3602.

Cushman, S. A., and Landguth, E. L. (2010b). Scale dependent inference in landscape genetics. *Landscape Ecology*, *25*(6), 967–979.

Cushman, S. A., and Lewis, J. S. (2010). Movement behavior explains genetic differentiation in American black bears. *Landscape Ecology*, *25*(10), 1613–1625.

Cushman, S. A., Shirk, A., and Landguth, E. L. (2011). Separating the effects of habitat area, fragmentation and matrix resistance on genetic differentiation in complex landscapes. *Landscape Ecology*, *27*(3), 369–380.

Cushman, S. A., Wasserman, T. N., Landguth, E. L., and Shirk, A. J. (2013). Re-evaluating causal modeling with mantel tests in landscape genetics. *Diversity*, *5*(1), 51–72.

Cushman, S. A., Chase, M., and Griffin, C. (2010). Mapping Landscape Resistance to Identify Corridors and Barriers for Elephant Movement in Southern Africa. In *Spatial complexity, informatics, and wildlife conservation* (pp. 349-367). Springer, Tokyo.

Davies, Z. G., and Pullin, A. S. (2007). Are hedgerows effective corridors between fragments of woodland habitat? An evidence-based approach. *Landscape Ecology*, *22*(3), 333–351.

Diniz, M. F., Cushman, S. A., Machado, R. B., & Júnior, P. D. M. (2020). Landscape connectivity modeling from the perspective of animal dispersal. *Landscape Ecology*, *35*(1), 41-58.

Driezen, K., Adriaensen, F., Rondinini, C., Doncaster, C. P., and Matthysen, E. (2007). Evaluating least-cost model predictions with empirical dispersal data: A case-study using radiotracking data of hedgehogs (*Erinaceus europaeus*). *Ecological Modelling*, *209*(2–4), 314–322.

Duchesne, T., Fortin, D., and Courbin, N. (2010). Mixed conditional logistic regression for habitat selection studies. *The Journal of Animal Ecology*, *79*(3), 548–555.

Edelhoff, H., Fickel, J., Zachos, F. E., Epps, C. W., and Balkenhol, N. (2020). Genetic analysis of red deer (*Cervus elaphus*) administrative management units in a human-dominated landscape. *Conservation Genetics*, *21*, 261–276.

Elliot, N. B., Cushman, S. A., Macdonald, D. W., and Loveridge, A. J. (2014). The devil is in the dispersers: Predictions of landscape connectivity change with demography.

Journal of Applied Ecology, *51*(5), 1169–1178.

Engler, J. O., Balkenhol, N., Filz, K. J., Habel, J. C., and Rödder, D. (2014). Comparative landscape genetics of three closely related sympatric Hesperid butterflies with diverging ecological traits. *PLoS ONE*, *9*(9), e106526.

Epps, C. W., Wehausen, J. J. D., Bleich, V. C., Torres, S. S. G., and Brashares, J. S. J. (2007). Optimizing dispersal and corridor models using landscape genetics. *Journal of Applied Ecology*, *44*(4), 714–724.

Epps, C. W., Palsboll, P. J., Wehausen, J. D., Roderick, G. K., Ramey, R. R., and McCullough, D. R. (2005). Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters*, *8*(10), 1029–1038.

Excoffier, L., Laval, G., and Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, *1*, 47-50.

Fahrig, L. (2003). Effects of Habitat Fragmentation on Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 487–515.

Fahrig, L. (2007). Non-optimal animal movement in human-altered landscapes. *Functional Ecology*, *21*(6), 1003–1015.

Fahrig, L., and Merriam, G. (1994). Conservation of Fragmented Populations. *Conservation Biology*, *8*(1), 50–59.

Flather, C. H., and Bevers, M. (2002). Patchy reaction-diffusion and population abundance: the relative importance of habitat amount and arrangement. *The American Naturalist*, *159*(1), 40–56.

Fletcher, R. J., Maxwell, C. W., Andrews, J. E., and Helmey-Hartman, W. L. (2012). Signal detection theory clarifies the concept of perceptual range and its relevance to landscape connectivity. *Landscape Ecology*, *28*(1), 57–67.

Fortin, D., Beyer, H., Boyce, M., Smith, D., Duchesne, T., and Mao, J. (2005). Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park. *Ecology*, *86*(5), 1320–1330.

Franklin, J. (2009). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Frantz, A. C., Bertouille, S., Eloy, M. C., Licoppe, A., Chaumont, F., and Flamand, M. C. (2012). Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (Cervus elaphus), but not wild boars (Sus scrofa). *Molecular Ecology*, *21*(14), 3445–3457.

Garroway, C. J., Bowman, J., and Wilson, P. J. (2011). Using a genetic network to parameterize a landscape resistance surface for fishers, Martes pennanti. *Molecular Ecology*, *20*(19), 3978–3988.

Gillies, C. S., Hebblewhite, M., Nielsen, S. E., Krawchuk, M. A., Aldridge, C. L., Frair, J. L., Saher, D. J., Stevens, C. E., and Jerde, C. L. (2006). Application of random effects to the study of resource selection by animals. *Journal of Animal Ecology*, *75*(4), 887–898.

Goslee, S. C., and Urban, D. L. (2007). The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *Journal of Statistical Software*, *22*(7).

Graves, T., Chandler, R. B., Royle, J. A., Beier, P., and Kendall, K. C. (2014). Estimating landscape resistance to dispersal. *Landscape Ecology*, *29*(7), 1201–1211.

Grenouillet, G., Buisson, L., Casajus, N., and Lek, S. (2011). Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography, 34*(1), 9–17.

Guisan, A., and Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2–3), 147–186.

Gurarie, E., Andrews, R. D., and Laidre, K. L. (2009). A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, *12*(5), 395–408.

Gurarie, E., Bracis, C., Delgado, M., Meckley, T. D., Kojola, I., and Wagner, C. M. (2016). What is the animal doing? Tools for exploring behavioral structure in animal movements. *Journal of Animal Ecology*. *85*(1), 69-84.

Haddad, N. M., Holt, R. D., Fletcher, R. J., Loreau, M., and Clobert, J. (2016). Connecting models, data, and concepts to understand fragmentation’s ecosystem-wide effects. *Ecography*, *125*(3), 336–342.

Hanski, I. (1998). Metapopulation dynamics. *Nature, 396*, 41–49.

Hanski, I. and Ovaskainen, O. (2003). Metapopulation theory for fragmented landscapes. *Theoretical Population Biology*, *64*(1), 119–127.

Hardy, O., and Vekemans, X. (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, *2*(4), 618–620.

Hartl, G. B., Zachos, F., and Nadlinger, K. (2003). Genetic diversity in European red deer (*Cervus elaphus* L.): anthropogenic influences on natural populations. *Comptes Rendus Biologies*, *326*, 37–42.

Hilty, J., Lidicker, W., and Merenlender, A.M. (2012). *Corridor ecology: the science and practice of linking landscapes for biodiversity conservation*. Island Press.

Holderegger, R., and Wagner, H. H. (2006). A brief guide to Landscape Genetics. *Landscape Ecology*, *21*(6), 793–796.

Hutchison, D. W., and Templeton, A. R. (1999). Correlation of Pairwise Genetic and Geographic Distance Measures: Inferring the Relative Influences of Gene Flow and Drift on the Distribution of Genetic Variability. *Evolution*, *53*(6), 1898-1914.

Jacobs, S., Burkhard, B., Van Daele, T., Staes, J., and Schneiders, A. (2014). ‘The Matrix Reloaded’: A review of expert knowledge use for mapping ecosystem services. *Ecological Modelling*, *295*, 21–30.

Jarnemo, A. (2007). Seasonal migration of male red deer (*Cervus elaphus*) in southern Sweden and consequences for management. *European Journal of Wildlife Research*, *54*(2), 327–333.

Johnson, D. H. (1980). The Comparison of Usage and Availability Measurements for Evaluating Resource Preference. *Ecology*, *61*(1), 65–71.

Kamler, J., Jedrzejewski, W., and Jedrzejewska, B. (2008). Home ranges of red deer in a European old-growth forest. *The American Midland Naturalist*, *159*(1), 75–82.

Keeley, A. T. H., Beier, P., and Gagnon, J. W. (2016). Estimating landscape resistance from habitat suitability: effects of data source and nonlinearities. *Landscape*

Ecology, *31*(9), 2151–2162.

Keller, D., Van Strien, M. J., and Holderegger, R. (2012). Do landscape barriers affect functional connectivity of populations of an endangered damselfly? *Freshwater Biology*, *57*(7), 1373–1384.

Keyghobadi, N. (2007). The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology*, *85*(10), 1049–1064.

Keyghobadi, N., Roland, J., and Strobeck, C. (2005). Genetic differentiation and gene flow among populations of the alpine butterfly, *Parnassius smintheus*, vary with landscape connectivity. *Molecular Ecology*, *14*(7), 1897–1909.

Kindlmann, P., and Burel, F. (2008). Connectivity measures: a review. *Landscape Ecology*, *23*, 879–890.

Koen, E. L., Bowman, J., and Walpole, A. A. (2012). The effect of cost surface parameterization on landscape resistance estimates. *Molecular Ecology Resources*, *12*(4), 686–696.

Lande, U. S., Loe, L. E., Skjærli, O. J., Meisingset, E. L., and Mysterud, A. (2013). The effect of agricultural land use practice on habitat selection of red deer. *European Journal of Wildlife Research, 60(1), 69-76.*

Legendre, P., and Fortin, M.-J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, *10*(5), 831–844.

Maiorano, L., Boitani, L., Chiaverini, L., and Ciucci, P. (2017). Uncertainties in the identification of potential dispersal corridors: The importance of behaviour, sex, and algorithm. *Basic and Applied Ecology*, *21*, 66–75.

Manel, S., and Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology and Evolution*, *28*(10), 614–621.

Manel, S., Schwartz, M. K., Luikart, G., and Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, *18*(4), 189–197.

Manly, B.F.J., McDonald, L.L., and Thomas, D.L. (1993). Intro-duction to resource selection studies. In *Resource Selection by Animals: Statistical Design and Analysis for Field Studies* (pp. 1–13). Chapman & Hall, London, UK

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2), 209–220.

Mateo Sánchez, M. C., Cushman, S. A., and Saura, S. (2013). Scale dependence in habitat selection: the case of the endangered brown bear (Ursus arctos) in the Cantabrian Range (NW Spain). *International Journal of Geographical Information Science*, *28*(8), 1531-1546.

Mateo-Sánchez, M. C., Balkenhol, N., Cushman, S., Pérez, T., Domínguez, A., and Saura, S. (2015). A comparative framework to infer landscape effects on population genetic structure: are habitat suitability models effective in explaining gene flow? *Landscape Ecology*, *30*(8), 1405–1420.

McGarigal, K., Wan, H. Y., Zeller, K. A., Timm, B. C., and Cushman, S. A. (2016). Multi-scale habitat selection modeling: a review and outlook. *Landscape Ecology*, *31*(6), 1157-1160.

McRae, B.H. (2006). Isolation by resistance. *Evolution*, *60*(8), 1551–1561.

McRae, B. H., and Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America, 104*(50), 19885–19890.

McRae, B., Dickson, B., Keitt, T., and Shah, V. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, *89*, 2712–2724.

McRae, B. H., Shah, V., and Edelman, A. (2016). Circuitscape: modeling landscape connectivity to promote conservation and human health. *The Nature Conservancy,* 14.

Meißner, M., Reinecke, H., Zachos, F., Corsmann, M., Ritter, T., and Wölfel, H. (2008). *Der Rothirsch in Schleswig-Holstein*. Göttingen.

Milanesi, P., Holderegger, R., Caniglia, R., Fabbri, E., Galaverni, M., and Randi, E. (2016a). Expert - based versus habitat - suitability models to develop resistance surfaces in landscape genetics. *Oecologia*, *183*(1), 67–79.

Milanesi, P., Holderegger, R., Caniglia, R., Fabbri, E., and Randi, E. (2016b). Different habitat suitability models yield different least-cost path distances for landscape genetic analysis. *Basic and Applied Ecology*, *17*(1), 61–71.

Milner, J. M., Bonenfant, C., Mysterud, A., Gaillard, J.-M., Csányi, S., and Stenseth, N. C. (2006). Temporal and spatial development of red deer harvesting in Europe: biological and cultural factors. *Journal of Applied Ecology*, *43*(4), 721–734.

Morelle, K., Bunnefeld, N., Lejeune, P., and Oswald, S. A. (2017). From animal tracks to fine-scale movement modes: a straightforward approach for identifying multiple spatial movement patterns. *Methods in Ecology and Evolution*, *8*(11), 1488–1498.

Naidoo, R., Kilian, J. W., Du Preez, P., Beytell, P., Aschenborn, O., Taylor, R. D., and Stuart-Hill, G. (2018). Evaluating the effectiveness of local- and regional-scale wildlife corridors using quantitative metrics of functional connectivity. *Biological Conservation*, *217*, 96–103.

Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142.

Neteler, M., Bowman, M. H., Landa, M., and Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling and Software*, *31*, 124–130.

O'Connell, A., Nichols, J., and Kāranth, K. (2010). *Camera traps in animal ecology*: *methods and analyses*. Springer Science & Business Media.

Opdam, P. (1991). Metapopulation theory and habitat fragmentation: a review of holarctic breeding bird studies. *Landscape Ecology*, *5*(2), 93–106.

Peterman, W. E. (2014). *ResistanceGA : An R package for the optimization of resistance surfaces using genetic algorithms*.

Phillips, S. J., Dudik, M., and Schapire, R. E. (2004). A Maximum Entropy Approach to Species Distribution Modeling. *Proceedings of the 21st International Conference on Machine Learning*, 655–662.

Poor, E. E., Loucks, C., Jakes, A., and Urban, D. L. (2012). Comparing Habitat Suitability and Connectivity Modeling Methods for Conserving Pronghorn Migrations. *PLoS*

ONE, *7*(11), e49390.

Proctor, M. F., McLellan, B. N., Strobeck, C., and Barclay, R. M. R. (2005). Genetic analysis reveals demographic fragmentation of grizzly bears yielding vulnerably small populations. *Proceedings. Biological Sciences / The Royal Society*, *272*(1579), 2409–2416.

R Core Team. (2017). *R: A language and environment for statistical computing*. Retrieved from http://www.r-project.org/

Rayfield, B., Fortin, M.-J., and Fall, A. (2009). The sensitivity of least-cost habitat graphs to relative cost surface values. *Landscape Ecology*, *25*(4), 519–532.

Razgour, O. (2015). Beyond species distribution modeling: A landscape genetics approach to investigating range shifts under future climate change. *Ecological Informatics*, *30*, 250–256.

Razgour, O., Rebelo, H., Puechmaille, S. J., Juste, J., Ibáñez, C., Kiefer, A., Burke, T., Dawson, D. A., and Jones, G. (2014). Scale-dependent effects of landscape variables on gene flow and population structure in bats. *Diversity and Distributions*, *20*(10), 1173–1185.

Reding, D. M., Cushman, S. A., Gosselink, T. E., and Clark, W. R. (2013). Linking movement behavior and fine-scale genetic structure to model landscape connectivity for bobcats (Lynx rufus). *Landscape Ecology*, *28*(3), 471–486.

Reed, G. C., Litvaitis, J. A., Callahan, C., Carroll, R. P., Litvaitis, M. K., and Broman, D. J. A. (2016). Modeling landscape connectivity for bobcats using expert-opinion and empirically derived models: how well do they work? *Animal Conservation*, *20*(4), 308–320.

Reif, A., and Achtziger, R. (2004). Gebüsche, Hecken, Waldmäntel, Feldgehölze (Strauchformationen). In *Handbuch Naturschutz und Landschaftspflege* (pp. 1–45).

Reinecke, H., Meißner, M., Zachos, F., and Herzog, S. (2013). *Sicherung genetischer Diversität beim Rothirsch in der Kulturlandschaft.*

Reiner, G., and Willems, H. (2019). *Sicherung der genetischen Vielfalt beim hessischen Rotwild als Beitrag zur Biodiversität*.

Ricketts, T. H. (2001). The matrix matters: Effective isolation in fragmented landscapes. *American Naturalist*, *158*(1), 87–99.

Riitters, K. H., Wickham, J. D., and Wade, T. G. (2009). An indicator of forest dynamics using a shifting landscape mosaic. *Ecological Indicators*, *9*(1), 107–117.

Roffler, G. H., Talbot, S. L., Luikart, G., Sage, G. K., Pilgrim, K. L., Adams, L. G., and Schwartz, M. K. (2014). Lack of sex-biased dispersal promotes fine-scale genetic structure in alpine ungulates. *Conservation Genetics*, *15*(4), 837–851.

Roever, C. L., Beyer, H. L., Chase, M. J., and van Aarde, R. J. (2013). The pitfalls of ignoring behaviour when quantifying habitat selection. *Diversity and Distributions*, *20*(3), 322–333.

Rousset, F. (2000). Genetic differentiation between individuals. *Journal of Evolutionary Biology*, *13*(1), 58–62.

Rousset, François. (2008). GENEPOP’007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106.

Row, J. R., Knick, S. T., Oyler-McCance, S. J., Lougheed, S. C., and Fedy, B. C. (2017). Developing approaches for linear mixed modeling in landscape genetics through landscape-directed dispersal simulations. *Ecology and Evolution*, *7*(11), 3751–3761.

Rudnick, D. A., Ryan, S. J., Beier, P., Cushman, S. A., Dieffenbach, F., Epps, C. W., Gerber, L.R., Hartter, J. N., Jenness, J. S., Kintsch, J., Merenlender, A. M., Perkl, R. M., Periosi, D. V., and Trombulak, S. C. (2012). The role of landscape connectivity in planning and implementing conservation and restoration priorities. *Issues in Ecology*, (16), 1–23.

Ruiz-González, A., Gurrutxaga, M., Cushman, S. A., Madeira, M. J., Randi, E., and Gómez-Moliner, B. J. (2014). Landscape Genetics for the Empirical Assessment of Resistance Surfaces: The European Pine Marten (Martes martes) as a Target-Species of a Regional Ecological Network. *PloS One*, *9*(10), e110552.

Saura, S., Bodin, Ö., and Fortin, M. J. (2014). EDITOR'S CHOICE: Stepping stones are crucial for species' long-distance dispersal and range expansion through habitat networks. *Journal of Applied Ecology*, *51*(1), 171–182.

Sawyer, S. C., Epps, C. W., and Brashares, J. S. (2011). Placing linkages among fragmented habitats: do least-cost models reflect how animals use landscapes? *Journal of Applied Ecology*, *48*(3), 668–678.

Segelbacher, G., Cushman, S. A., Epperson, B. K., Fortin, M.-J., Francois, O., Hardy, O. J., Holderegger, R., Taberlet, P., Waits, L. P., and Manel, S. (2010). Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*, *11*(2), 375–385.

Shafer, A. B. A., Northrup, J. M., White, K. S., Boyce, M. S., Côté, S. D., and Coltman, D. W. (2012). Habitat selection predicts genetic relatedness in an alpine ungulate. *Ecology*, *93*(6), 1317–1329.

Shirk, A. J., Landguth, E. L., and Cushman, S. A. (2017). A comparison of individual-based genetic distance metrics for landscape genetics. *Molecular Ecology Resources*, *17*(6), 1308–1317.

Shirk, A. J., Wallin, D. O., Cushman, S. A., Rice, C. G., and Warheit, K. I. (2010). Inferring landscape effects on gene flow: a new model selection framework. *Molecular Ecology*, *19*(17), 3603–3619.

Shirk, A. J., Landguth, E. L., and Cushman, S. A. (2018). A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Molecular Ecology Resources*, *18*(1), 55–67.

Signer, J., Fieberg, J., and Avgar, T. (2017). Estimating utilization distributions from fitted step-selection functions. *Ecosphere*, *8*(4), e01771.

Singleton, P. H., Gaines, W. L., and Lehmkuhl, J. F. (2004). Landscape permeability for grizzly bear movements in Washington and southwestern British Columbia. *Ursus*, *15*(1), 90–103.

Skog, A., Zachos, F. E., Rueness, E. K., Feulner, P. G. D., Mysterud, A., Langvatn, R., Lorenzini, R., Hmwe, S. S., Lehoczky, I., Hartl, G. B., Stenseth, N. C., and Jakobsen, K. S. (2009). Phylogeography of red deer (*Cervus elaphus*) in Europe. *Journal of Biogeography*, *36*(1), 66–77.

Soille, P., and Vogt, P. (2009). Morphological segmentation of binary patterns. *Pattern Recognition Letters*, *30*(4), 456–459.

Spear, S. F., Balkenhol, N., Fortin, M.-J., McRae, B. H., and Scribner, K. (2010). Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. *Molecular Ecology*, *19*(17), 3576–3591.

Spear, S. F., Cushman, S. A., and McRae, B. H. (2015). Resistance Surface Modeling in Landscape Genetics. In *Landscape Genetics* (pp. 129–148).

Squires, J. R., DeCesare, N. J., Olson, L. E., Kolbe, J. A., Hebblewhite, M., and Parks, S. A. (2013). Combining resource selection and movement behavior to predict corridors for Canada lynx at their southern range periphery. *Biological Conservation*, *157*, 187–195.

Stevens, V. M., Verkenne, C., Vandewoestijne, S., Wesselingh, R. A., and Baguette, M. (2006). Gene flow and functional connectivity in the natterjack toad. *Molecular Ecology*, *15*(9), 2333–2344.

Stevenson-Holt, C. D., Watts, K., Bellamy, C. C., Nevin, O. T., and Ramsey, A. D. (2014). Defining landscape resistance values in least-cost connectivity models for the invasive grey squirrel: a comparison of approaches using expert-opinion and habitat suitability modelling. *PloS One*, *9*(11), e112119.

Stohlgren, T. J., Ma, P., Kumar, S., Rocca, M., Morisette, J. T., Jarnevich, C. S., and Benson, N. (2010). Ensemble habitat mapping of invasive plant species. *Risk Analysis*, *30*(2), 224–235.

Storfer, A., Murphy, M. A., Spear, S. F., Holderegger, R., and Waits, L. P. (2010). Landscape genetics: where are we now? *Molecular Ecology*, *19*(17), 3496–3514.

Taylor, P. D., Fahrig, L., Henein, K., and Merriam, G. (1993). Connectivity is a vital element of landscape structure. *Oikos*, *68*(3), 571–573.

Templeton, A. R., Shaw, K., Routman, E., and Davis, S. K. (1990). The genetic consequences of habitat fragmentation. *Annals of the Missouri Botanical Garden*, 13-27.

Thurfjell, H., Ciuti, S., and Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, *2*(1), 4.

van Etten, J. (2017). R package gdistance: Distances and Routes on Geographical Grids. *Journal of Statistical Software*, *76*(13), 1-21.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535–538.

van Strien, M. J., Keller, D., and Holderegger, R. (2012). A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Molecular Ecology*, *21*(16), 4010–4023.

Vasudev, D., and Fletcher, R. J. (2015). Incorporating movement behavior into conservation prioritization in fragmented landscapes: An example of western hoolock gibbons in Garo Hills, India. *Biological Conservation*, *181*, 124–132.

Vasudev, D., Fletcher, R. J., Goswami, V. R., and Krishnadas, M. (2015). From dispersal constraints to landscape connectivity: lessons from species distribution modeling. *Ecography*, 38, 967-978.

Vogt, P., and Riitters, K. (2017). GuidosToolbox: universal digital image object analysis. *European Journal of Remote Sensing*, *50*(1), 352–361.

Wang, Y. H., Yang, K. C., Bridgman, C. L., and Lin, L. K. (2008). Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*, *23*(8), 989–1000.

Ward, A. I., Dendy, J., and Cowan, D. P. (2015). Mitigating impacts of roads on wildlife: an agenda for the conservation of priority European protected species in Great Britain. *European Journal of Wildlife Research*, , *61*(2), 199-211.

Wasserman, T. N., Cushman, S. A., Schwartz, M. K., and Wallin, D. O. (2010). Spatial scaling and multi-model inference in landscape genetics: *Martes americana* in northern Idaho. *Landscape Ecology*, *25*(10), 1601–1612.

Wasserman, T. N., Cushman, S. A., Littell, J. S., Shirk, A. J., and Landguth, E. L. (2013). Population connectivity and genetic diversity of American marten (*Martes americana*) in the United States northern Rocky Mountains in a climate change context. *Conservation Genetics*, *14*(2), 529–541.

Wickham, J. D., and Norton, D. J. (1994). Mapping and analyzing landscape patterns. *Landscape Ecology*, *9*(1), 7–23.

Wirsing, A. J., and Heithaus, M. R. (2014). Accounting for individual behavioural variation in studies of habitat selection. *The Journal of Animal Ecology*, *83*(2), 319-321.

With, K., and Crist, T. (1995). Critical thresholds in species' responses to landscape structure. *Ecology*, *76*(8), 2446–2459.

Wittemyer, G., Polansky, L., Douglas-Hamilton, I., and Getz, W. M. (2008). Disentangling the effects of forage, social rank, and risk on movement autocorrelation of elephants using Fourier and wavelet analyses. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19108–19113.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114–138.

Zachos, F. E., Althoff, C., Steynitz, Y. V., Eckert, I., and Hartl, G. B. (2007). Genetic analysis of an isolated red deer (Cervus elaphus) population showing signs of inbreeding depression. *European Journal of Wildlife Research*, *53*(1), 61–67.

Zachos, Frank E, Frantz, A. C., Kuehn, R., Bertouille, S., Colyn, M., Niedzialkowska, M., Pérez-González, J., Skog, A., Sprěm, N., and Flamand, M.-C. (2016). Genetic structure and effective population sizes in European red deer (Cervus elaphus) at a continental scale: insights from microsatellite DNA. *Journal of Heredity*, *107*(4), 318-326.

Zeller, K. A., McGarigal, K., and Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. *Landscape Ecology*, *27*(6), 777–797.

Zeller, K. A., McGarigal, K., Beier, P., Cushman, S. A., Vickers, T. W., and Boyce, W. M. (2014). Sensitivity of landscape resistance estimates based on point selection functions to scale and behavioral state: pumas as a case study. *Landscape Ecology*, *29*(3), 541–557

Zeller, K. A., McGarigal, K., Cushman, S. A., Beier, P., Vickers, T. W., and Boyce, W. M. (2015). Using step and path selection functions for estimating resistance to movement: pumas as a case study. *Landscape Ecology*, *31*(6), 1319–1335.

Zeller, K. A., Creech, T. G., Millette, K. L., Crowhurst, R. S., Long, R. A., Wagner, H. H., Balkenhol, N., and Landguth, E. L. (2016). Using simulations to evaluate Mantel-based methods for assessing landscape resistance to gene flow. *Ecology and Evolution*, *6*(12), 4115-4128.

Zeller, K. A., Vickers, T. W., Ernest, H. B., Boyce, W. M., Pollinger, J., and Ernest, H. (2017). Multi-level, multi-scale resource selection functions and resistance surfaces for conservation planning: Pumas as a case study. *PLOS ONE*, *12*(6), e0179570.

Zeller, K. A., Jennings, M. K., Vickers, T. W., Ernest, H. B., Cushman, S. A., and Boyce, W. M. (2018). Are all data types and connectivity models created equal? Validating common connectivity approaches with dispersal data. *Diversity and Distributions*, *24*(7), 868–879.

Supplement

S4.1: Telemetry Data

Table S4.1.1 Summary and overview of the GPS telemetry data for 20 red deer individuals from Schleswig-Holstein. For each individual the number of relocations (n), the mean interval between consecutive relocations measured in hours (interval), sex, age class as well as start and end date and total runtime in days are provided.

id	**n**	**Interval**	**sex**	**age**	**start**	**end**	**runtime**
a5504	2284	6.51	m	young	2009-02-27	2010-11-09	620
a5506	1036	6.52	m	young	2009-04-05	2010-01-11	281
c5174	4573	4.65	f	young	2009-04-01	2011-09-04	886
c5175	3261	5.20	m	young	2009-02-19	2011-01-26	706
c5176	4879	4.29	m	young	2009-03-31	2011-08-20	872
c5178	5633	4.19	f	young	2010-02-16	2012-10-26	983
c5181	5178	4.25	m	midage	2008-04-03	2010-10-07	917
c5185	4313	4.39	f	old	2010-02-03	2012-04-01	788
c5186	2790	4.73	f	midage	2009-11-16	2011-05-21	551
c5187	1091	4.53	m	midage	2010-02-28	2010-09-22	206
c5188	895	9.13	m	midage	2009-06-21	2010-05-28	341
c5498	618	5.39	m	young	2010-03-07	2010-07-23	138
c5500	2380	5.04	m	midage	2009-05-13	2010-09-24	499
c5502	5940	3.81	m	young	2010-03-30	2012-10-27	942
c5503	3835	3.67	m	young	2010-03-24	2011-10-31	586
c5506	3156	5.27	f	midage	2010-03-05	2012-01-27	693
c5507	3381	4.64	f	midage	2009-11-16	2011-09-01	654
c5508	1689	4.34	m	midage	2010-04-03	2011-02-02	305
c5574	3042	5.31	m	young	2010-02-17	2011-12-22	673
c5581	1558	7.57	m	midage	2010-03-09	2011-07-14	492

S4.2: Genetic Data

The following 12 microsatellite markers were used:

- *INRA6* (Slate et al. 1998; Vaiman et al. 1994),
- *C143*, *C180*, *C229*, *T107*, *T115* (all tetranucleotide loci; Meredith et al. 2005),
- *HAUT14* (Thieven et al. 1995),
- *BM757*, *BM1818*, *ILSTS06* (Bishop et al. 1994),
- *CSSM14* (Moore et al. 1994),
- and *FSHB* (Moore et al. 1992).

S4.2.1 References

Bishop, M. D., Kappes, S. M., Keele, J. W., Stone, R. T., Sunden, S. L., Hawkins, G. A., Toldo, S. S., Fries, R., Grosz, M. D., and Yoo, J. (1994). A genetic linkage map for cattle. Genetics, 136(2), 619-639.

Meredith, E. P., Rodzen, J. A., Levine, K. F., and Banks, J. D. (2008). Characterization of an additional 14 microsatellite loci in California Elk (Cervus elaphus) for use in forensic and population applications. *Conservation Genetics*, *9*(6), 1715-1715.

Moore, S. S., Barendse, W., Berger, K. T., Armitage, S. M., and Hetzel, D. J. S. (1992). Bovine and ovine DNA microsatellites from the EMBL and GENBANK databases. *Animal Genetics*, *23*(5), 463-467.

Moore, S. S., Byrne, K., Berger, K. T., Barendse, W., McCarthy, F., Womack, J. E., & Hetzel, D. J. S. (1994). Characterization of 65 bovine microsatellites. *Mammalian Genome*, *5*(2), 84-90.

Slate, J., Coltman, D. W., Goodman, S. J., MacLean, I., Pemberton, J. M., and Williams, J. L. (1998). Bovine microsatellite loci are highly conserved in red deer (Cervus elaphus), sika deer (Cervus nippon) and Soay sheep (Ovis aries). *Animal Genetics*, *29*(4), 307-315.

Thieven, U., Harlizius, B., and Simon, D. (1995). Dinucleotide repeat polymorphism at the bovine HAUT1 and HAUT14 loci. *Animal Genetics*, *26*(2), 123–123.

Vaiman, D., Mercier, D., Moazami-Goudarzi, K., Eggen, A., Ciampolini, R., Lépingle, A., Martin, P., Levéziel, H., and Guérin, G. (1994). A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. *Mammalian Genome*, *5*(5), 288-297.

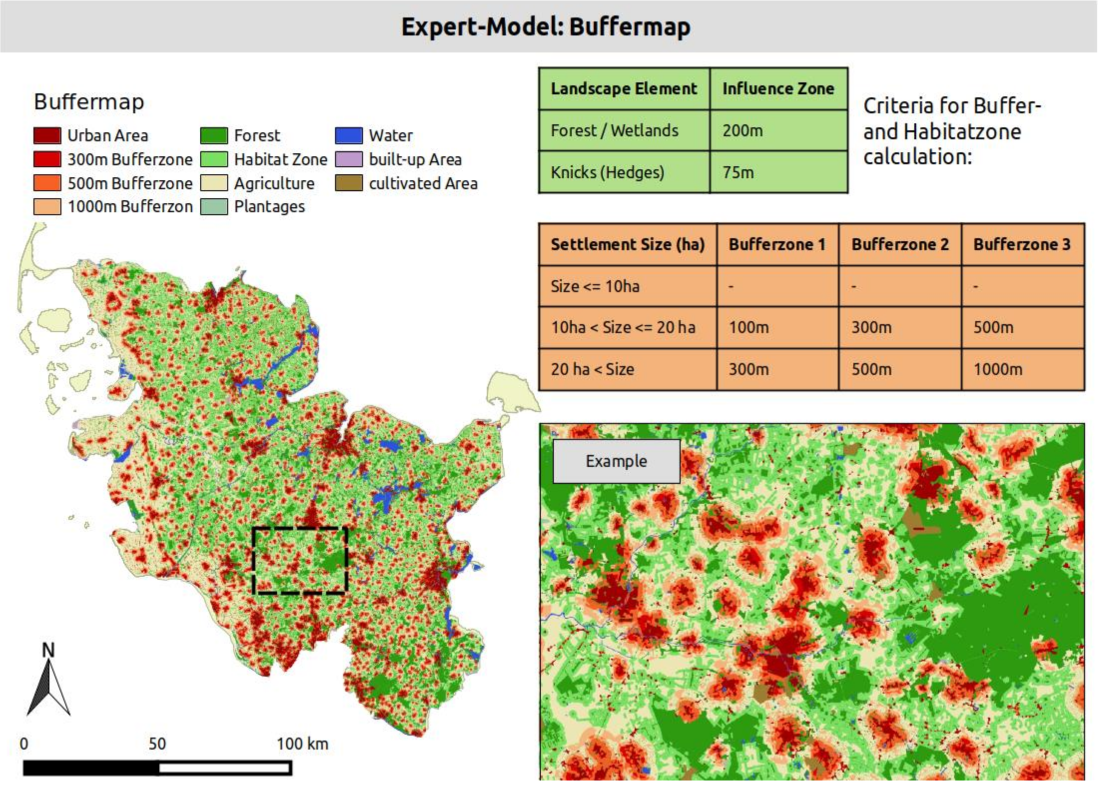
S4.3: Expert-informed Models

S4.3.1 Buffer Map

The first model was developed by wildlife biologists at the Institute for Wildlife Biology Göttingen Dresden e.V. as part of a long-term research project on red deer in Schleswig-Holstein (Meißner et al. 2008). The researchers performed a thorough survey including local expertise on red deer behavior and observed dispersal events to define potential effects of landscape features on red deer movements. The model distinguished between neutral, positive and negative landcover features. Additionally, the functional effects of these features were extended into the landscape with varying buffer sizes. The main assumption behind the designation of certain landscape features as positive or negative was based on how they likely influenced red deer demands on cover and security (e.g., as provided by forests) or its potential as a connecting landscape element.

Urban and cultivated areas as well as settlements of all sizes were considered as negative areas. However, the negative influence on to surrounding areas depended on the overall size of the single feature (patch unit). Small settlements (size <= 10 ha) were not assigned a buffer zone. Intermediate settlements and villages (10ha > size <= 20ha) exhibited three buffer zones of decreasing negative influence (100m, 300m, and 500m). Large urban and cultivated areas such as cities (size > 20ha) were also assigned three buffers of negative influence but of larger extend (300m, 500m and 1000m). Roads were assigned high values of resistance but with no buffer as the experts argued that there is limited evidence for radiating disturbance and also information on the variation of amount of traffic was not available (Meißner et al. 2008).

Forested areas make up the most prominent parts of the red deer distribution areas in Schleswig-Holstein. Therefore, they were considered a key feature with positive value for red deer. Next to forests other landcover types with high levels of vegetation and structure (cover) such as wetlands, heath and moor areas were also considered as positive areas for dispersal. Linear features such as mound hedges (Knicks; Meißner et al. 2008; Reif and Achtziger 2004) were also assigned positive values. Agricultural areas were interpreted as having a neutral effect on red deer dispersal. In case of overlapping areas of positive and negative buffers the negative buffer gets into a lower category (less resistant) but is the dominating buffer for that grid cell.



Landscape Element	Influence Zone
Forest / Wetlands	200m
Knicks (Hedges)	75m

Settlement Size (ha)	Bufferzone 1	Bufferzone 2	Bufferzone 3
Size <= 10ha	-	-	-
10ha < Size <= 20 ha	100m	300m	500m
20 ha < Size	300m	500m	1000m

Figure S4.3.0.1 Depiction of the input and development of the buffermap. Different categories of landcover and their zone of influence are shown on the left. Tables for sizes of the different buffer zones are shown on the right. The inlet provides an example of a smaller region of the study area to demonstrate how the zone of influence varies with varying size and type of the landcover (especially settlements and forested areas).

Originally, the model was developed to determine the potential connectivity of the landscape and for local management in order to find optimal locations for wildlife passages or determine potential conflicts with new constructions of roads. Here, we used the buffer map as an input for different resistance models. In order to assign resistance values to the different landscape features and their buffer zones we ranked them to values between 0.1 (low resistance) and 1.0 (high resistance). Overview provided in Table S4.3.1.

Table S4.3.1 Overview of resistance values assigned to different landcover classes in the buffermap. Defined classes and buffer sizes are based on the expert model developed by Meißner et al. (2008).

Landcover Class	**Resistance**
Habtiat (forest, hedgerows, heith and moors)	0.1
Habitat Influence Zone	0.2
Agriculture	0.3
Water	0.8
Plantages	0.8
Cultivated-Areas	0.9
Build-up Areas	1.0
Urban (buffer 100m)	1.0
Urban (buffer 300m)	0.9
Urban (buffer 500m)	0.8
Urban (buffer 1000m)	0.7

S4.3.2 Landscape Mosaic

The idea behind this approach is a situation where only limited knowledge about habitat requirements for a certain species are available. In this scenario a landcover model could at least be classified in the three categories of potentially positive (natural), neutral and negative (developed) areas. The landscape mosaic model accounts for a gradient between this three exclusive classifications and estimates which of the three categories is dominant in a certain part of the landscape or delineates areas of homogeneous and intermixed categories (Riitters et al. 2009).

In order to apply this method we first divided the landscape of our study area into three categories: natural, neutral, and developed areas. The division into these three categories was based on our assumptions on habitat requirements by red deer. By this, we mimic a very simplistic approach based only on basic landcover classifications which can be quickly derived from expert knowledge or literature review. Overall, this method can be a helpful tool to develop models of landscape resistance.

All patches of forest, mound hedges (Knicks), wetland and heath or moorland were classified as natural areas. Urban, settlements and cultivated areas as well as all types of primary roads were defined as developed areas. The remaining agricultural parts of the landscape made up the neutral classification.

We used the landscape mosaic analysis (Riitters et al. 2009; Wickham and Norton 1994) implemented in the *GUIDOS* software (Vogt and Riitters 2017) to assign each grid cell of our landscape raster a new value accounting also for the categories of the neighboring grid cells. The analysis weighs the amount of each category in the surrounding cells and defines new clusters based on a trigonometric gradient of influence of the natural, neutral, or developed category (see Figure S4.3.2).

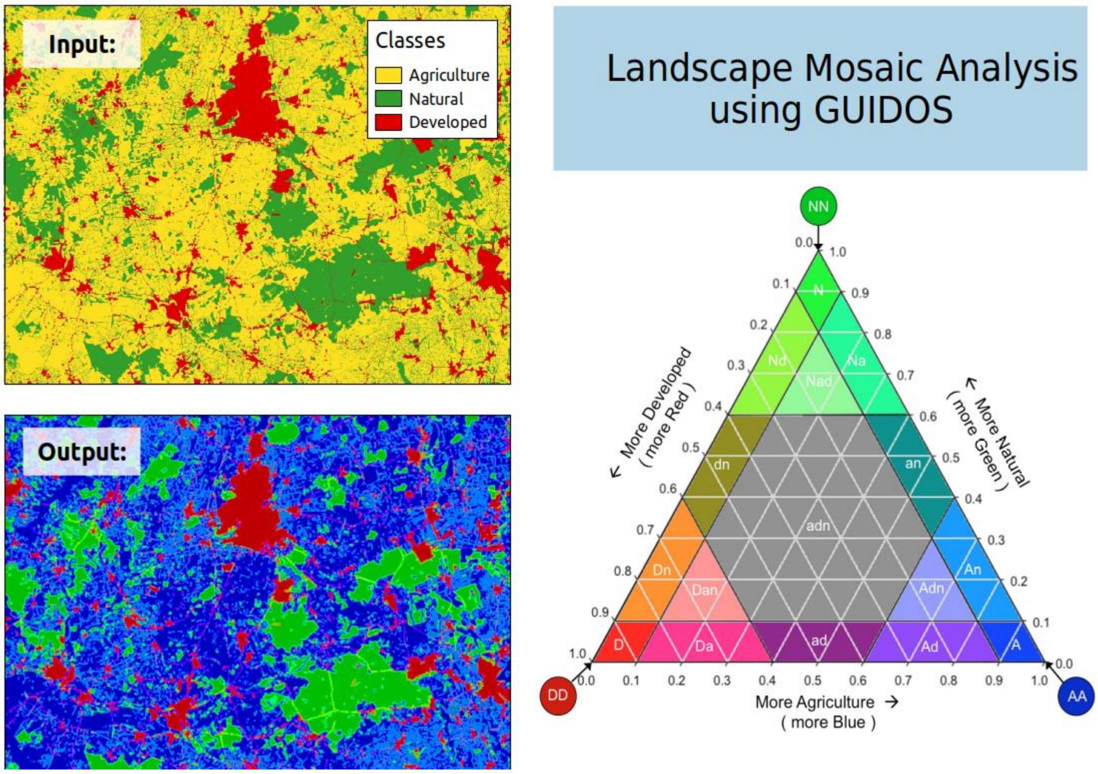


Figure S4.3.2 Explanation of the landscape mosaic approach: First, three categories of landcover (natural, neutral/agricultural, and developed) are defined (upper left).The landscape mosaic algorithm implemented in the GUIDOS software then uses a trigonometric analysis(right hand) to determine for each grid cell the dominating landcover category accounting for the composition of the neighboring cells. Based on this either just one, two or an intermix of all three categories are present in the surrounding area (output). Maps show small subregion of the study area for demonstrative purposes.

In the next step, we used these mosaic clusters as a foundation for a linear weighting scheme (Clevenger et al. 2002). Resistance values of each mosaic cluster were derived from linearly weighted resistance values of each category (natural = 1, neutral/agriculture = 5, developed = 10). Weights were based on the amount of natural, neutral and developed category within the cluster (Table S4.3.2). The final model was based on the sum of the weighted resistance values.

*Table S4.3.2 Weighting scheme to assign resistance values to different landscape mosaic categories. For each Mosaic Cluster the proportion of natural, agriculture (neutral) and developed landcover types were calculated. Natural areas were assigned a resistance value (cost) of 1, developed a cost of 10 and agricultural areas a cost of 5. The proportions of each category were used to weigh the resistance values and subsequently the total sum of costs was derived. Example calculation for first row: 0.05*1+0.9*5+0.05*10 = 5.05.*

Mosaic Cluster	**Natural**	**Agriculture**	**Developed**	**Cost_N**	**Cost_A**	**Cost_D**	**Cost_Sum**
A	0.050	0.900	0.050	1	5	10	5.050
D	0.050	0.050	0.900	1	5	10	9.300
N	0.900	0.050	0.050	1	5	10	1.650
Ad	0.000	0.750	0.250	1	5	10	6.250
An	0.250	0.750	0.000	1	5	10	4.000
Dn	0.250	0.000	0.750	1	5	10	7.750
Da	0.000	0.250	0.750	1	5	10	8.750
Na	0.750	0.250	0.000	1	5	10	2.000
Nd	0.750	0.000	0.250	1	5	10	3.250
Adn	0.125	0.750	0.125	1	5	10	5.125
Dan	0.125	0.125	0.750	1	5	10	8.250
Nad	0.750	0.125	0.125	1	5	10	2.625
ad	0.500	0.000	0.500	1	5	10	5.500
an	0.500	0.500	0.000	1	5	10	3.000
dn	0.500	0.000	0.500	1	5	10	5.500
adn	0.330	0.330	0.330	1	5	10	5.280
NN	1.000	0.000	0.000	1	5	10	1.000
AA	0.000	1.000	0.000	1	5	10	5.000
DD	0.000	0.000	1.000	1	5	10	10.000

S4.3.3 References

Clevenger, A. P., Wierzchowski, J., Chruszcz, B., and Gunson, K. (2002). GIS-Generated, Expert-Based Models for Identifying Wildlife Habitat Linkages and Planning Mitigation Passages. *Conservation Biology*, *16*(2), 503–514.

Meißner, M., Reinecke, H., Zachos, F., Corsmann, M., Ritter, T., and Wölfel, H. (2008). *Der Rothirsch in Schleswig-Holstein*. Göttingen.

Reif, A., and Achtziger, R. (2004). Gebüsche, Hecken, Waldmäntel, Feldgehölze (Strauchformationen). In *Handbuch Naturschutz und Landschaftspflege* (pp. 1–45).

Riitters, K., Vogt, P., Soille, P., and Estreguil, C. (2009). Landscape patterns from mathematical morphology on maps with contagion. *Landscape Ecology, 24*(5), 699–709.

Vogt, P., and Riitters, K. (2017). GuidosToolbox: universal digital image object analysis. *European Journal of Remote Sensing*, *50*(1), 352–361.

Wickham, J. D., and Norton, D. J. (1994). Mapping and analyzing landscape patterns. *Landscape Ecology*, *9*(1), 7–23.

S4.4: Habitat-informed Models

S4.4.1 Subsampling of Relocations

Presence points for red deer in Schleswig-Holstein were simulated by drawing random samples from the GPS relocations since other information on presence (occurrence) or true absence of red deer was not available for our study. We subsampled the data for all individuals using a grid-based (“fishnet”) approach: a grid with a cell size of 500 by 500m was placed over all relocations. For each grid cell containing one or multiple relocations only a single relocation was selected as a presence point. In the case of multiple relocations from the same or different individuals being present in the same grid cell the final relocation was chosen randomly. The process was repeated for ten times resulting in ten different sets of pseudo presence points used for our habitat suitability models. We fitted both, presence-only (MAXENT; Phillips et al. 2004) and presence-absence algorithms: generalized linear mixed (GLM) and random forest (RMF; Breiman 2001) models. For the latter we also simulated 1400 pseudo absence points for each of the ten runs. A buffer of 1000m was placed around all relocation points and excluded from the study area prior generating the random points to assure that they were placed outside the core areas of the red deer distribution.

Prior fitting the multi-scale multivariate habitat models we selected landcover-based variables using the following two steps:

1. for raster grids describing the proportion of a landcover type the most relevant radius (scale) was determined based on AIC (for the GLM) and AUC (for MAXENT and RMF; Hijmans 2012) values of univariate models (see also Zeller et al. 2018).
2. we accounted for covariation among considered variables and chose only landcover variables with absolute correlation values below 0.6 (based on Spearman's rank). If two variables showed correlations with $|r^2| \geq 0.6$ we only included the variable with the lowest AIC or AUC value of the univariate model respectively.

The remaining variables were applied in a multivariate, multi-scale model.The ten subsets of pseudo presence/absence points were used to run each model algorithm for ten times. Subsequently we used the fitted coefficients of each run to produce spatial predictions of habitat suitability (Guisan and Thuiller 2005). The mean across all ten runs was used as the final model output for the considered algorithm. Additionally, an ensemble model (Araújo and New 2007) was calculated using the mean of the three final habitat suitability models (referred to as SDM_Ensemble).

The inverse of the model predictions was used to produce models of landscape resistance based on habitat suitability (Keeley et al. 2016). For each habitat suitability model the inverse value was transformed into resistance values between 0 and 100 using a linear, a monomolecular and an inverse-reverse monomolecular transformation (Peterman et al. 2014; Zeller et al. 2018). In total we compared 12 different resistance surfaces based on habitat suitability.

S4.4.2 Results

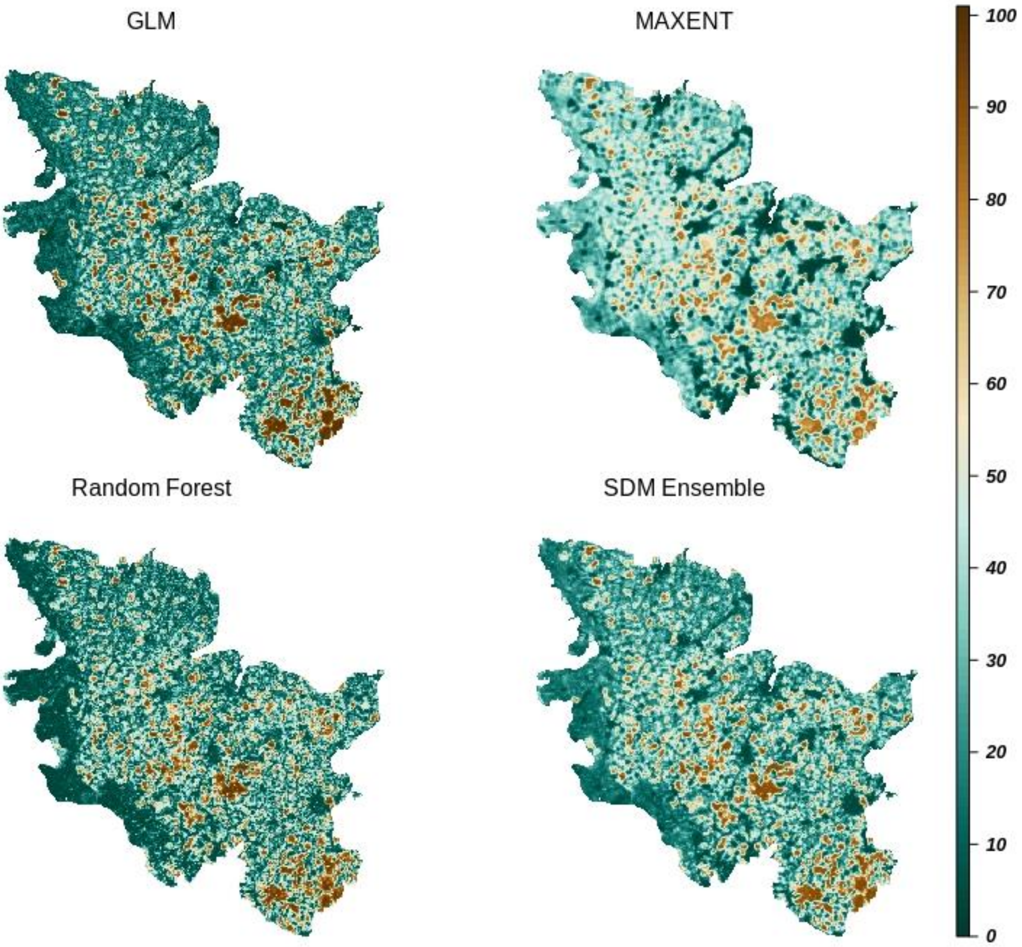


Figure S4.4.1 Spatial predictions of habitat suitability based on three different algorithms as well as their combination (mean value) as an ensemble model.

Table S4.4.1 Model coefficients of selected variables included in each of the three different habitat suitability models.

Landcover	**Random Forest**			**GLM**			**MAXENT**		
	type	**coeff (mean)**	**sd**	**type**	**coeff (mean)**	**sd**	**type**	**coeff (mean)**	**Sd**
agriculture	distance	0.001	0.000	distance	0.00	0.00	distance	0.03	0.05
forest (linear)	prop (100m)	-0.509	0.152	prop (100m)	2.10	0.81	prop (1000m)	-0.01	0.04
forest	prop (1000m)	0.928	0.021	prop (1000m)	6.81	0.42	prop (1000m)	3.01	0.13
Knicks (hedgerows)	distance	0.000	0.000	distance	0.00	0.00	prop (200m)	0.18	0.16
roads	distance	0.000	0.000	distance	0.00	0.00	prop (1000m)	-1.21	0.18
	prop (100m)	-2.680	0.161	prop (100m)	-6.15	1.86	not selected		
urban	distance	0.001	0.000	distance	0.00	0.00	prop (1000m)	-4.76	0.63
urban (linear)	prop (100m)	-4.095	0.309	prop (100m)	-13.50	2.93	not selected		
water	distance	0.000	0.000	distance	0.00	0.00	distance	0.00	0.00
	prop (100m)	-0.515	0.040	prop (100m)	-2.55	0.60	prop (1000m)	-3.08	0.55
wetland/heath	Prop (1000m)	0.766	0.105	prop (1000m)	7.42	0.99	distance	-1.28	0.27

S4.4.3 References

Araújo, M. B., and New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, *22*(1), 42–47.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Guisan, A., and Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009.

Hijmans, R. J. (2012). Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology*, *93*(3), 679–688.

Keeley, A. T. H., Beier, P., and Gagnon, J. W. (2016). Estimating landscape resistance from habitat suitability: effects of data source and nonlinearities. *Landscape Ecology*, *31*(9), 2151–2162.

Peterman, W. E., Connette, G. M., Semlitsch, R. D., and Eggert, L. S. (2014). Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Molecular Ecology*, *23*(10), 2402–2413.

Phillips, S. J., Dudik, M., and Schapire, R. E. (2004). A Maximum Entropy Approach to Species Distribution Modeling. *Proceedings of the 21st International Conference on Machine Learning*, 655–662.

Zeller, K. A., Jennings, M. K., Vickers, T. W., Ernest, H. B., Cushman, S. A., and Boyce, W. M. (2018). Are all data types and connectivity models created equal? Validating common connectivity approaches with dispersal data. *Diversity and Distributions*, *24*(7), 868–879.

S4.5: Movement-informed Models

We applied three different selection functions using the telemetry data of the 20 red deer individuals. In the following the different procedures are described with more specific detail. All analyses were performed using the *amt* R package (Signer et al. 2019).

S4.5.1 Resource Selection Analysis

We used a point-selection framework (Zeller et al. 2012) to estimate resource selection at the home range scale (third order RSF; Johnson 1980; Manly et al. 1993). We placed a minimum convex polygon (MCP) around the relocations of each individual and used this as a proxy for the utilized home range (Harris et al. 1990). Within each home range we produced twice the number of observed relocations as random points. Subsequently we extracted landscape variables for all observed relocations (used) and random points (available). Used variables were compared against available variables with a logistic regression model (Manly et al. 1993).

Prior fitting a final, multi-scale and multivariate model we determined the characteristic scale of each landscape variable. Univariate models were fitted for each percentage variable (i.e., percent cover in radius) at different scales (radii) and compared using AIC values (Zeller et al. 2014). The final scale for each type of landcover was selected based on the lowest AIC value. Additionally, we tested for covariation and in case of two variables showing a Spearman correlation $|r^2|$ above 0.6 we only kept the variable with the lower AIC value. All of the remaining variables served as input for our final multi-scale and multivariate model.

We used a two-step procedure to model resource selection while also accounting for individual variation arising from different sample sizes and levels of selectivity among the 20 individuals the data set was based on (Craiu et al. 2011; Murtaugh 2007). Therefore, we fitted the multivariate logistic regression model based on the final set of variables for each of the individuals separately. Subsequently, we calculated the mean of the 20 regression coefficients of each input variable to derive a population-level coefficient and applied a t-test to determine if it was significantly different from zero (Hosmer and Lemeshow 2004; Squires et al. 2013).

In order to spatially predict the probability of use only the significant coefficients were plugged in the point-selection function: $\widehat{w}(x) = \exp\,(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p\, x_p)$ (Johnson 1980). The exponential selection function is fitted with the population-level coefficients (β_1 to β_p) and the respective landscape variables (x_1 to x_p). The resulting selection scores among all grid cells of the landscape model were rescaled to values ranging from zero to one using a linear stretch procedure (DeCesare et al. 2016; Johnson et al. 2004). Finally, we used the inverse of these values (1 - selection score) to derive the final resistance surface based on point-level resource selection (Squires et al. 2013; Zeller et al. 2017).

S4.5.2 Step-Selection Analysis

In a second approach we again used all relocations of the 20 individuals. However, this time we estimated selection of landscape variables at the step-level using so called step-selection functions (SSF; Thurfjell et al. 2014; Zeller et al. 2012). For each individual the distribution of turning angles and step-lengths were sampled separately (Signer et al. 2019). Every observed step was compared to 10 random steps in used-availability study design. Landcover variables at used steps were compared against variables at available steps using conditional logistic regression (Thurfjell et al. 2014) models implemented in the *survival* R package (Therneau and Grambsch 2013).

Prior fitting a final, multi-scale and multivariate model we determined the characteristic scale of each landscape variable. Univariate models were fitted for each percentage variable (i.e., percent cover in radius) at different scales (radii) and compared using AIC values (Zeller et al. 2014). The final scale for each type of landcover was selected based on the lowest AIC value. Additionally, we tested for covariation and in case of two variables showing a Spearman correlation $|r^2|$ above 0.6 we only kept the variable with the lower AIC value. All of the remaining variables served as input for our final multi-scale and multivariate model.

We used a two-step procedure to model resource selection while also accounting for individual variation arising from different sample sizes and levels of selectivity among the 20 individuals the data set was based on (Craiu et al. 2011; Murtaugh 2007). Therefore, we fitted a multivariate, conditional logistic regression model based on the final set of variables for each of the individuals separately. Subsequently, we calculated the mean of the 20 regression coefficients of each input variable to derive a population-level coefficient and applied a t-test to determine if it was significantly different from zero (Hosmer and Lemeshow 2004; Squires et al. 2013).

In order to spatially predict the probability of use only the significant coefficients were plugged in the step-selection function $\widehat{w}(x) = \exp\,(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p\, x_p)$ (Johnson, 1980; Thurfjell et al. 2014). The exponential selection function is fitted with the population-level coefficients (β_1 to β_p) and the respective landscape variables (x_1 to x_p). The resulting selection scores among all grid cells of the landscape model were rescaled to values ranging from zero to one using a linear stretch procedure (DeCesare et al. 2016; Johnson et al. 2004). Finally, we used the inverse of these values (1 - selection score) to derive the final resistance surface based on path-level resource selection (Squires et al. 2013; Zeller et al. 2017).

S4.5.3 Behavioral Change Point Analysis and Step-Selection Function

In this approach we first selected movement patterns that could potentially be interpreted as dispersal behavior. For this we applied a behavioral change point analysis (BCPA, Gurarie et al. 2009) to each individual movement trajectory. We chose the BCPA because it is also applicable for movement data with irregular sampling intervals between relocations (Edelhoff et al. 2015; Chapter 3). The BCPA was based on the resulting time series of net-squared displacement values (NSD, Calenge et al. 2009). NSD measures (Figure S5.3.1) the squared displacement between the first and a current relocation of the trajectory and is commonly applied to characterize dispersal movements or migration patterns (Bunnefeld et al. 2011; Edelhoff et al. 2015; Chapter 3).

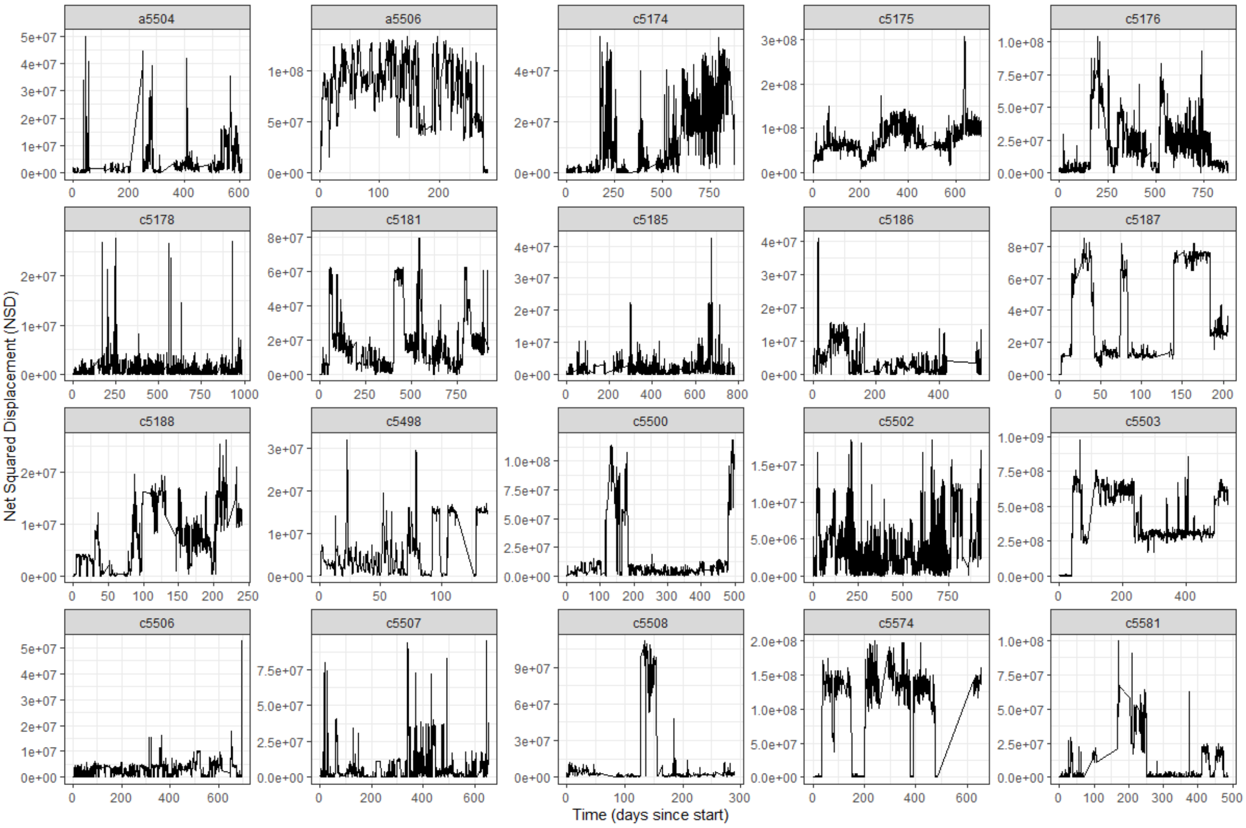


Figure S4.5.1 Time-series of the net-squared displacement values for each individual. The time in days since the start of the GPS tagging is plotted on the x-axis.

The potential change-points in movement behavior estimated by the BCPA were then used to cut out segments of different movement patterns within each individual trajectory. In order to be able to interpret the underlying behaviors we used a set of path-characteristics to describe the observed patterns, in particular we calculated the sinuosity (i.e., straightness), mean NSD, as well as the cumulative distance and total displacement of each segment (Benhamou 2004).

We then used a cluster analysis two separate the into two distinctive groups based on the multiple path-characteristics (see Zhang et al. 2015) for another example). Among the two clusters we chose the one with the highest mean of NSD values and interpreted the contained path-segments as potential dispersal or "dispersal-like" movement behavior. On average 12% of the relocations of an individual were assigned to the potential dispersal cluster (min = 2%, max= 30%). The spatial distribution of the regular and dispersal relocations are plotted in Figure S5.2. Subsequently, we used the same modeling procedure for step-selection as described above only using the steps from the potential dispersal cluster this time. The final result was another resistance surface based on this subset of relocations determined by a behavioral change point analysis.

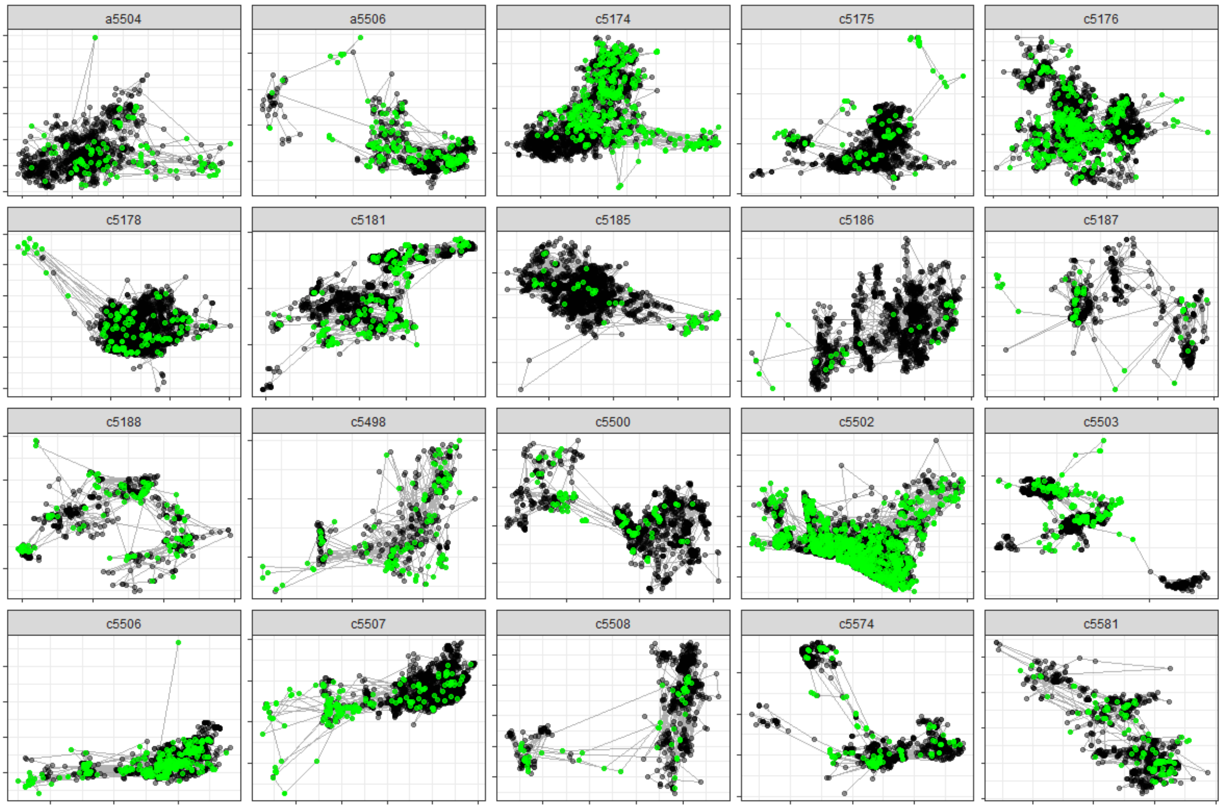


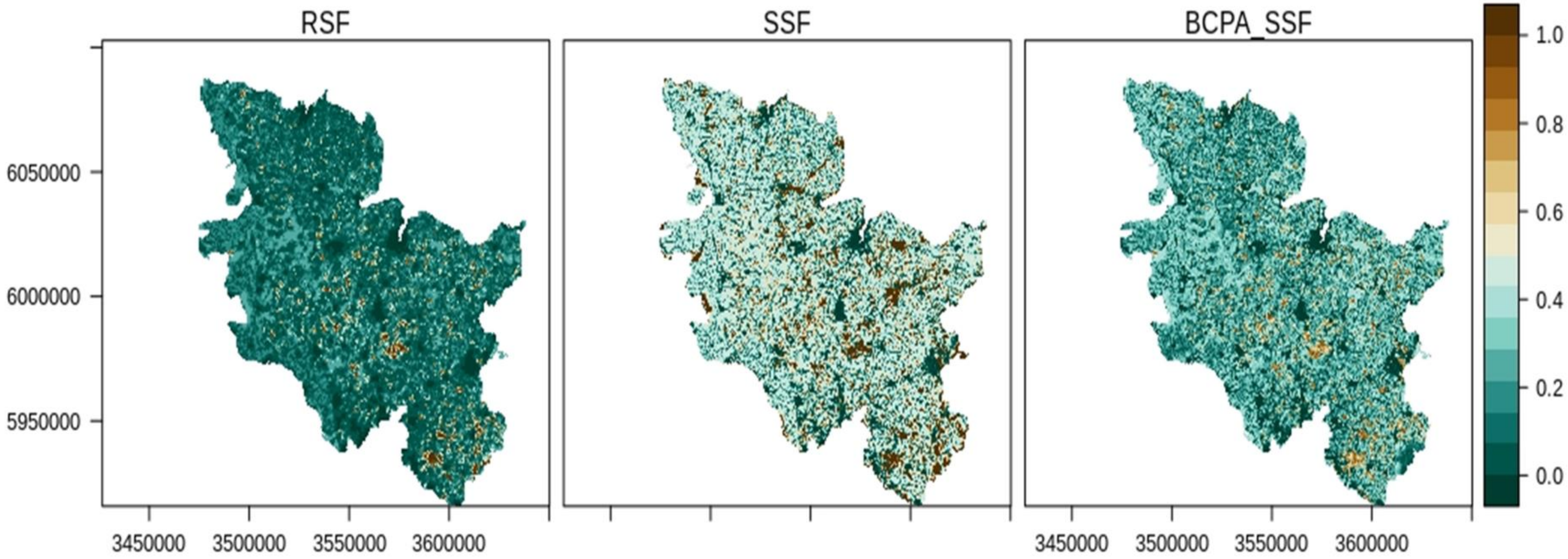
Figure S4.5.2 GPS relocations for all 20 red deer individuals. Green points indicate relocations of the potential dispersal cluster. Black points represent relocations from regular movement behavior. The amount of relocations of potential dispersal behavior varied substantially.

S4.5.4 Result

Table S4.5.1 Shows the results of the population-level coefficients (averaged across the 20 individuals) derived from regression models comparing used versus available landscape variables at the point-level (RSF) or the step-level (SSF, BCPA_SSF). Because of the variable pre-selection process for each approach separately, variables included in the final models were not consistent.

Landcover	**RSF**				**SSF**				**BCPA_SSF**			
	Type	mean	sd	sign	type	mean	sd	sign	type	mean	sd	sign
agriculture					prop. (100m)	-0.77	0.12	0.00				
forest (in)	prop. (1000m)	-21.18	20.77	0.32	prop. (100m)	-1.53	2.82	0.59	prop. (100m)	2.07	1.03	0.06
forest	prop. (100m)	1.44	0.25	0.00					prop. (100m)	0.94	0.09	0.00
hedge rows	prop. (100m)	-4.04	1.82	0.04					prop. (100m)	-2.32	0.74	0.01
roads	prop. (200m)	-12.72	2.72	0.00	prop. (100m)	-10.89	1.31	0.00	dist	0.00	0.00	0.86
									prop. (100m)	-8.78	1.70	0.00
urban	prop. (1000m)	-10.50	3.44	0.01	dist	0.00	0.00	0.87	prop. (200m)	-11.48	2.24	0.00
					prop. (200m)	-8.58	1.41	0.00				
urban (lin)	prop. (200m)	-32.08	5.39	0.00								
water	prop. (1000m)	21.67	30.22	0.48	dist	0.00	0.00	0.14	dist	0.00	0.00	0.02
					prop. (100m)	-0.34	0.59	0.58				
wetland	dist	0.00	0.00	0.08	dist	0.00	0.00	0.06	dist	0.00	0.00	0.05
	prop. (100m)	-106.47	58.35	0.08	prop. (100m)	-0.05	0.30	0.86				

Figure S4.5.3 presents the spatial prediction of the fitted selection functions derived at the point-level (RSF) and step-level (SSF, BCPA_SSF). Values range between zero and one whereas values closer to one indicate a higher probability of being selected for movement by an red deer individual.



S4.5.5 References

Benhamou, S. (2004). How to reliably estimate the tortuosity of an animal’s path: straightness, sinuosity, or fractal dimension? *Journal of Theoretical Biology*, *229*(2), 209–220.

Bunnefeld, N., Börger, L., van Moorter, B., Rolandsen, C. M., Dettki, H., Solberg, E. J., and Ericsson, G. (2011). A model-driven approach to quantify migration patterns: individual, regional and yearly differences. *The Journal of Animal Ecology*, *80*(2), 466–476.

Calenge, C., Dray, S., and Royer-Carenzi, M. (2009). The concept of animals’ trajectories from a data analysis perspective. *Ecological Informatics*, *4*(1), 34–41.

Craiu, R. V., Duchesne, T., Fortin, D., and Baillargeon, S. (2011). Conditional logistic regression with longitudinal follow-up and individual-level random coefficients: A stable and efficient two-step estimation method. *Journal of Computational and Graphical Statistics*, *20*(3), 767–784.

DeCesare, N. J., Hebblewhite, M., Lukacs, P. M., and Hervieux, D. (2016). Evaluating sources of censoring and truncation in telemetry-based survival data. *Journal of Wildlife Management*, *80*(1), 138–148.

Edelhoff, H., Signer, J., and Balkenhol, N. (2015). Path segmentation for beginners: an overview of current methods for detecting changes in animal movement data. *Movement Ecology*, *4*(1), 21.

Gurarie, E., Andrews, R. D., and Laidre, K. L. (2009). A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, *12*(5), 395–408.

Harris, S., Cresswell, W., Forde, P., Trewhella, W., Wollard, T., and Wray, S. (1990). Home-range analysis using radio-tracking data–a review of problems and techniques particularly as applied to the study of mammals. *Mammal Review*, *20*(2–3), 97–123.

Hosmer, D., and Lemeshow, S. (2004). *Applied logistic regression*. John Wiley and Sons.

Johnson, C. J., Seip, D. R., and Boyce, M. S. (2004). A quantitative approach to conservation planning: Using resource selection functions to map the distribution of mountain caribou at multiple spatial scales. *Journal of Applied Ecology*, *41*(2), 238–251.

Johnson, D. H. (1980). The Comparison of Usage and Availability Measurements for Evaluating Resource Preference. *Ecology*, *61*(1), 65–71.

Manly, B.F.J., McDonald, L.L., and Thomas, D.L. (1993). Intro-duction to resource selection studies. In *Resource Selection by Animals: Statistical Design and Analysis for Field Studies* (pp. 1–13). Chapman & Hall, London, UK

Murtaugh, P. A. (2007). Simplicity and complexity in ecological data analysis. *Ecology*, *88*(1), 56–62.

Signer, J., Fieberg, J., and Avgar, T. (2019). Animal movement tools (amt): R package for managing tracking data and conducting habitat selection analyses. *Ecology and Evolution*, *9*(2), 880–890.

Squires, J. R., DeCesare, N. J., Olson, L. E., Kolbe, J. A., Hebblewhite, M., and Parks, S.

A. (2013). Combining resource selection and movement behavior to predict corridors for Canada lynx at their southern range periphery. *Biological Conservation*, *157*, 187–195.

Therneau, T. M., and Grambsch, P. M. (2000). Testing proportional hazards. In *Modeling survival data: extending the Cox model* (pp. 127-152). Springer, New York, NY.

Thurfjell, H., Ciuti, S., and Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, *2*(1), 4.

Zeller, K. A., McGarigal, K., Beier, P., Cushman, S. A., Vickers, T. W., and Boyce, W. M. (2014). Sensitivity of landscape resistance estimates based on point selection functions to scale and behavioral state: pumas as a case study. *Landscape Ecology*, *29*(3), 541–557.

Zeller, K. A., McGarigal, K., and Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. *Landscape Ecology*, *27*(6), 777–797.

Zeller, K. A., Vickers, T. W., Ernest, H. B., Boyce, W. M., Pollinger, J., and Ernest, H. (2017). Multi-level, multi-scale resource selection functions and resistance surfaces for conservation planning: Pumas as a case study. *PLOS ONE*, *12*(6), e0179570.

Zhang, J., O'Reilly, K. M., Perry, G. L. W., Taylor, G. A., and Dennis, T. E. (2015). Extending the Functionality of Behavioural Change-Point Analysis with k-Means Clustering: A Case Study with the Little Penguin (Eudyptula minor). *Plos One*, *10*(4), e0122811.

S4.6: Variance and Correlation of Resistance Surfaces

Table S4.6.1 Results of generalized linear models fitted to explain observed resistance values depending on either the transformation function, the original input model (SDM, MAXENT, etc.), or the methodological category (type).

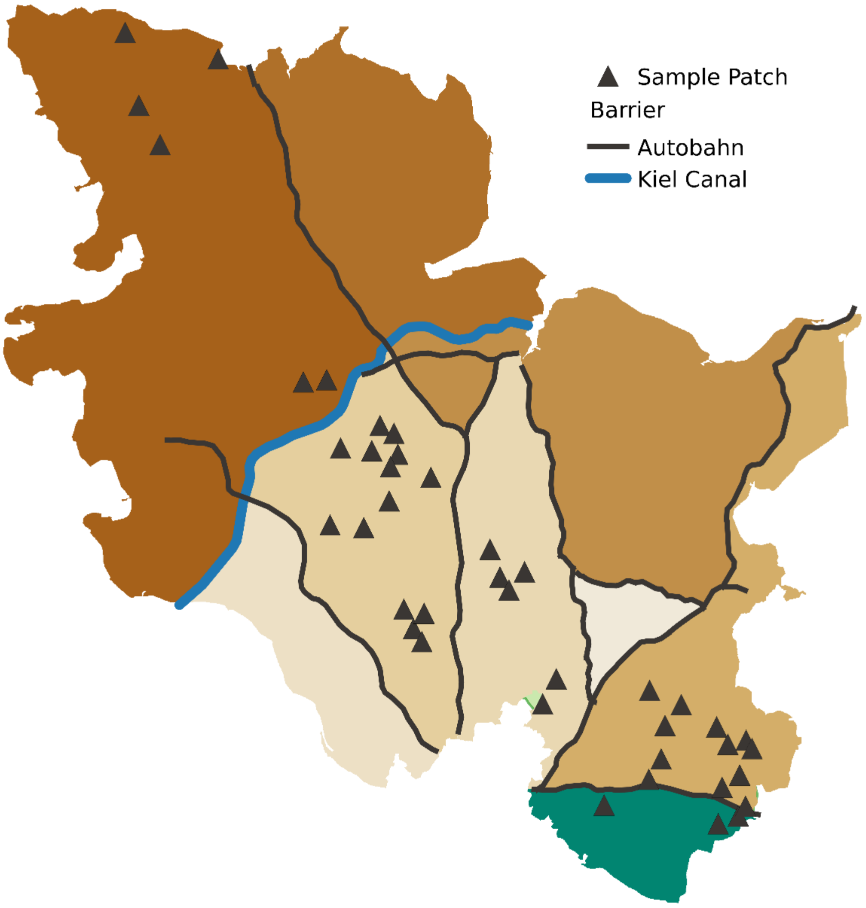
Variable	**AIC**	**Deviance expl (D^2)**
Transformation	239488.7	0.456945
Model	252215.4	0.109733
Type (Method)	253139.6	0.076731

Table S4.6.2 Spearman correlation of the sampled resistance values (1000 points) of each of the nine input models (results shown for all linear transformations).

	BCPA.SSF	**BUFFER**	**GLM**	**LS.MOSAIC**	**MAXENT**	**RANDOM.FOR**	**RSF**	**SDM.ENS**	**SSF**
BCPA.SSF		0.58	0.51	0.54	0.45	0.54	**0.82**	0.55	**0.90**
BUFFER	0.58		**0.63**	**0.69**	**0.67**	**0.63**	**0.66**	**0.71**	**0.65**
GLM	0.51	**0.63**		**0.60**	**0.70**	**0.91**	0.56	**0.95**	**0.62**
LS.MOSAIC	0.54	**0.69**	**0.60**		0.45	**0.63**	0.44	**0.62**	**0.65**
MAXENT	0.45	**0.67**	**0.70**	0.45		**0.67**	**0.63**	**0.85**	0.50
RANDOM.FOR	0.54	**0.63**	**0.91**	**0.63**	**0.67**		0.54	**0.94**	**0.66**
RSF	**0.82**	**0.66**	0.56	0.44	**0.63**	0.54		**0.63**	**0.78**
SDM.ENS	0.55	**0.71**	**0.95**	**0.62**	**0.85**	**0.94**	**0.63**		**0.65**
SSF	**0.90**	**0.65**	**0.62**	**0.65**	0.50	**0.66**	**0.78**	**0.65**	

S4.7: Isolation by Barrier

Figure S4.7.1 Map showing the putative barriers considered in our isolation by barrier hypothesis. Distance values between sampled patches increased with number of barriers (canal, primary roads) in-between them .



S4.8: Mantel tests

Table S4.8.1 Results of Mantel correlation tests for all resistance surfaces using effective distance based on Circuitscape analysis and genetic distance derived from the proportion of shared alleles.

Resistance surface	**Mantel r**	**llim.2.5%**	**ulim.97.5%**
buffer [invrev-mono]	0.246	0.235	0.259
ensemble [cs]	0.241	0.229	0.256
bcpa_ssf [mono]	0.241	0.228	0.255
maxent [mono]	0.236	0.222	0.251
maxent [lin]	0.230	0.216	0.244
ls_mosaic [invrev-mono]	0.229	0.217	0.243
maxent [invrev-mono]	0.226	0.211	0.242
bcpa_ssf [lin]	0.221	0.209	0.237
ssf [invrev-mono]	0.221	0.207	0.234
buffer [mono]	0.216	0.203	0.231
sdm_ensemble [lin]	0.212	0.198	0.227
random_forest [mono]	0.211	0.198	0.224
sdm_ens [mono]	0.211	0.197	0.227
ssf [mono]	0.211	0.198	0.227
rsf [mono]	0.210	0.195	0.225
bcpa_ssf [invrev-mono]	0.208	0.194	0.221
ssf [lin]	0.207	0.194	0.222
ls_mosaic [mono]	0.207	0.193	0.223
glm [mono]	0.205	0.189	0.218
rsf [invrev-mono]	0.204	0.190	0.220
sdm_ens [invrev-mono]	0.203	0.191	0.218
random_forest [lin]	0.202	0.188	0.217
rsf [lin]	0.201	0.186	0.217
glm [lin]	0.196	0.181	0.213
glm [invrev-mono]	0.192	0.175	0.207
random_forest [invrev-mono]	0.189	0.174	0.204
ls_mosaic [mono]	-0.039	-0.055	-0.023

Table S4.8.2 Results of Mantel correlation tests for all resistance surfaces using effective distance based on circuitscape analysis and genetic distance derived from Rousset's a.

Resistance surface	**Mantel r**	**llim.2.5%**	**ulim.97.5%**
buffer [invrev-mono]	0.219	0.205	0.233
ensemble [cs]	0.218	0.203	0.234
maxent [mono]	0.218	0.204	0.234
bcpa_ssf [mono]	0.215	0.200	0.230
ls_mosaic [invrev-mono]	0.212	0.196	0.226
maxent [lin]	0.211	0.196	0.228
maxent [invrev-mono]	0.207	0.192	0.222
buffer [mono]	0.203	0.188	0.221
ssf [invrev-mono]	0.203	0.186	0.219
rsf [mono]	0.199	0.183	0.218
ssf [mono]	0.199	0.182	0.216
sdm_ens [mono]	0.198	0.182	0.215
sdm_ens [lin]	0.198	0.180	0.214
ls_mosaic [mono]	0.197	0.181	0.215
random_forest [mono]	0.196	0.180	0.213
ssf [lin]	0.195	0.180	0.215
glm [mono]	0.193	0.177	0.210
rsf [invrev-mono]	0.193	0.177	0.214
bcpa_ssf [lin]	0.192	0.178	0.208
rsf [lin]	0.192	0.174	0.210
sdm_ens [invrev-mono]	0.191	0.176	0.208
random_forest	0.190	0.174	0.207
glm [lin]	0.187	0.168	0.205
glm [invrev-mono]	0.183	0.166	0.200
random_forest [invrev-mono]	0.180	0.164	0.200
bcpa_ssf [invrev-mono]	0.177	0.161	0.193
ls_mosaic [mono]	-0.002	-0.023	0.017

Table S4.8.3 Results of Mantel correlation tests for all resistance surfaces using effective distance based on least-cost analysis and genetic distance derived from the proportion of shared alleles.

Resistance surface	**Mantel r**	**llim.2.5%**	**ulim.97.5%**
buffer [invrev-mono]	0.216	0.203	0.229
maxent [invrev-mono]	0.212	0.199	0.226
bcpa [mono]	0.208	0.193	0.223
ensemble [lcp]	0.208	0.196	0.222
maxent [lin]	0.208	0.194	0.223
ssf [lin]	0.207	0.192	0.220
maxent [mono]	0.207	0.191	0.222
ls_mosaic [lin]	0.206	0.192	0.219
bcpa_ssf [lin]	0.206	0.191	0.220
ssf [mono]	0.205	0.191	0.220
ls_mosaic [invrev-mono]	0.205	0.192	0.218
ssf [invrev-mono]	0.200	0.187	0.215
glm [mono]	0.199	0.186	0.215
sdm_ensemble [lin]	0.198	0.184	0.212
ls_mosaic [mono]	0.198	0.185	0.214
rsf [lin]	0.198	0.184	0.212
sdm_ens [mono]	0.198	0.182	0.213
bcpa [invrev-mono]	0.197	0.183	0.213
rsf [mono]	0.197	0.181	0.211
random_forest [mono]	0.196	0.182	0.210
buffer [mono]	0.194	0.180	0.209
random_forest [lin]	0.194	0.179	0.208
glm [lin]	0.194	0.181	0.208
rsf [invrev-mono]	0.193	0.178	0.209
sdm_ens [invrev-mono]	0.192	0.177	0.206
random_forest [invrev-mono]	0.184	0.171	0.201
glm [invrev-mono]	0.182	0.168	0.200

Table S4.8.4 Results of Mantel correlation tests for all resistance surfaces using effective distance based on least-cost analysis and genetic distance derived from Rousset's a.

Resistance surface	**Mantel r**	**llim.2.5%**	**ulim.97.5%**
buffer [invrev-mono]	0.191	0.177	0.209
maxent [invrev-mono]	0.187	0.171	0.203
ssf [lin]	0.186	0.169	0.204
ssf [mono]	0.185	0.169	0.201
maxent [lin]	0.185	0.168	0.199
bcpa_ssf [mono]	0.185	0.171	0.202
maxent [mono]	0.184	0.169	0.203
ensemble [lcp]	0.183	0.167	0.200
ls_mosaic [lin]	0.183	0.166	0.199
bcpa_ssf [lin]	0.182	0.166	0.198
ls_mosaic [invrev-mono]	0.182	0.165	0.202
glm [mono]	0.180	0.163	0.198
rsf [lin]	0.180	0.164	0.198
rsf [mono]	0.180	0.163	0.198
ssf [invrev-mono]	0.180	0.165	0.196
sdm_ensemble [mono]	0.179	0.164	0.198
sdm_ens [lin]	0.178	0.161	0.196
ls_mosaic [mono]	0.178	0.163	0.198
random_forest [mono]	0.178	0.162	0.194
bcpa_ssf [invrev-mono]	0.177	0.161	0.199
buffer [mono]	0.177	0.161	0.197
glm [lin]	0.176	0.158	0.193
random_forest [lin]	0.175	0.158	0.193
rsf [invrev-mono]	0.174	0.156	0.193
sdm_ens [invrev-mono]	0.172	0.153	0.189
random_forest [invrev-mono]	0.168	0.151	0.186
glm [invrev-mono]	0.166	0.148	0.186

S4.9: Ensemble Models

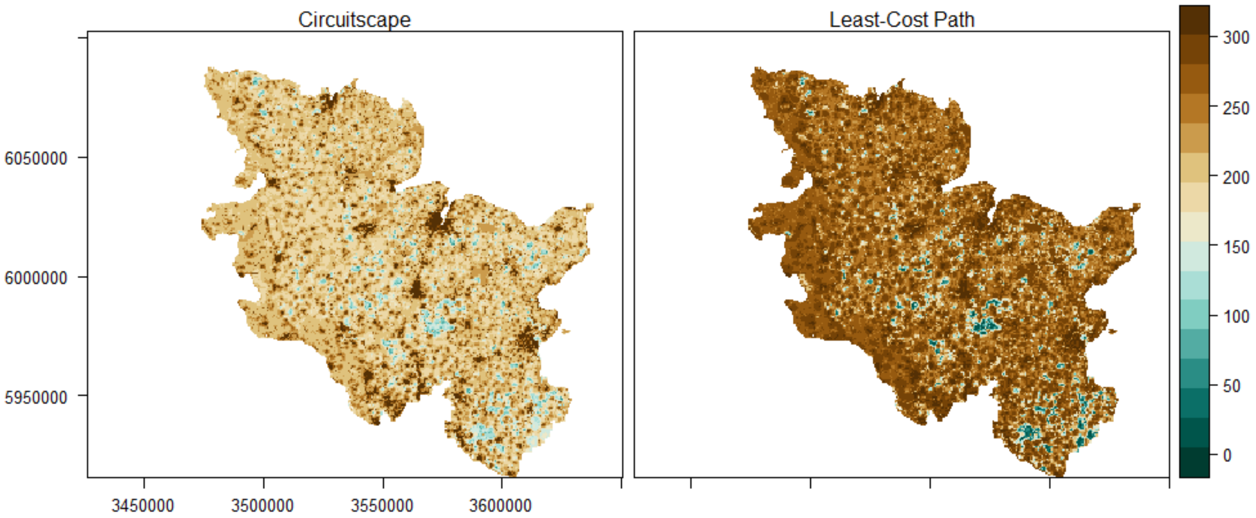


Figure S4.9.1 Maps depicting the two ensemble models based on the three highest-ranked models identified using the Circuitscape distance and least-cost path distance..

S4.10: Conductance Surfaces

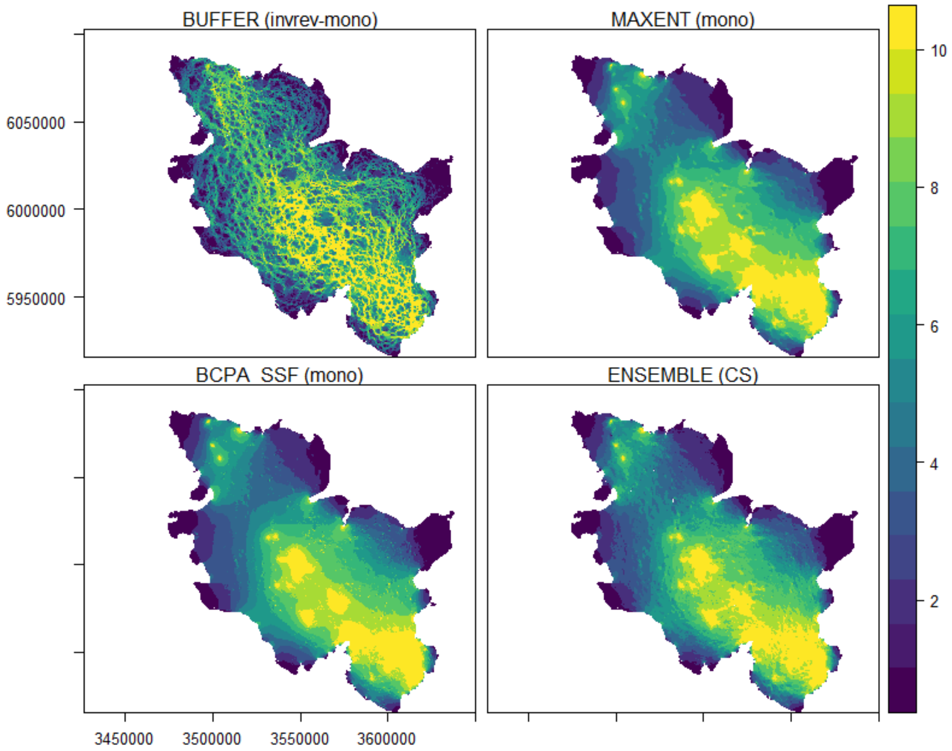


Figure S4.10.1 Cumulative conductance surfaces derived with Circuitscape. Results shown for the three selected models based on expert-opinion (BUFFER), habitat suitability (MAXENT) and dispersal movements (BCPA_SSF) as well as the ensemble of all three models.

S4.11: Least-Cost Paths

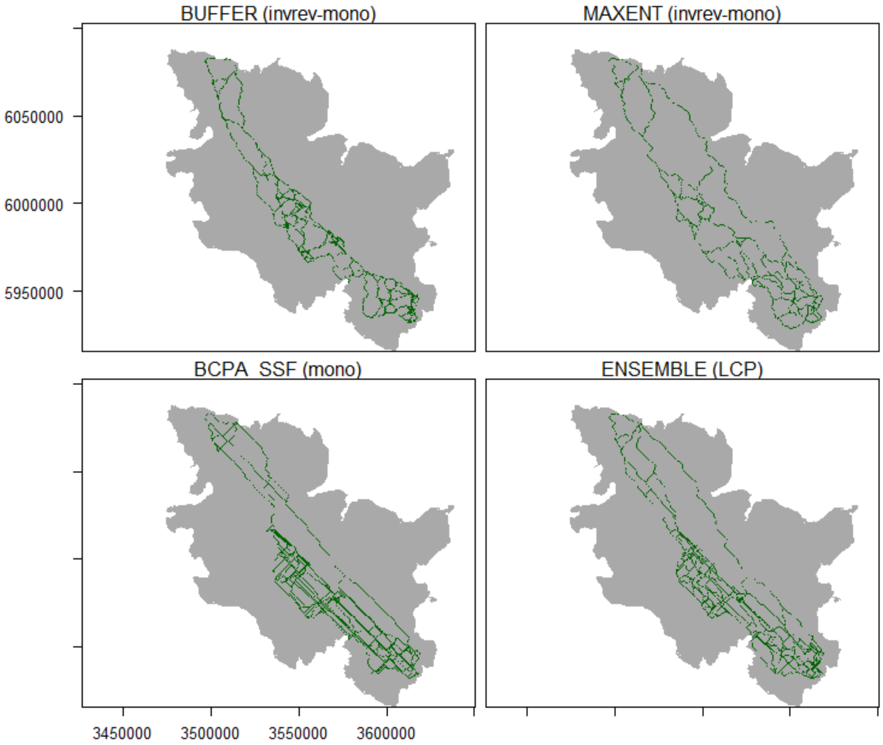


Figure S4.11.1 Least-cost paths for the three highest ranked resistance models based on expert-opinion (BUFFER), habitat suitability (MAXENT) and a movement-analysis (BCPA_SSF) as well as their ensemble model.

S4.12: Corridor Overlap

Table S4.12.1 Pairwise coefficients based on Spearman rank correlation between the Circuitscape conductance surfaces of the selected models based on expert-knowledge (BUFFER[invrev-mono]), habitat suitability (MAXENT[mono]) and movement data (BCPA_SSF[mono]), as well as their ensemble (ENSEMBLE_CS). Additionally, conductance surfaces were correlated to the raster grid summarizing the number of overlapping LCPs.

	EXPERT	**HABITAT**	**MOVEMENT**	**ENSEMBLE**	**LCP Overlap**
EXPERT	1.00	0.68	0.67	0.71	0.27
HABITAT	0.68	1.00	0.98	0.97	0.27
MOVEMENT	0.67	0.98	1.00	0.97	0.26
ENSEMBLE	0.71	0.97	0.97	1.00	0.27
LCP Overlap	0.27	0.27	0.26	0.27	1.00

S4.13: Comparing Movement and Habitat Models

We visually inspected the differences between the best performing movement-informed model (BCPA_SSF) and habitat-informed model (MAXENT). The spatial predictions of the step-selection function and habitat suitability respectively (see S4.4.2 and S4.5.4) were considered here. We first calculated the quantiles of each model and then subtracted the MAXENT model from the BCPA_SSF model. Results are shown in the following figure.

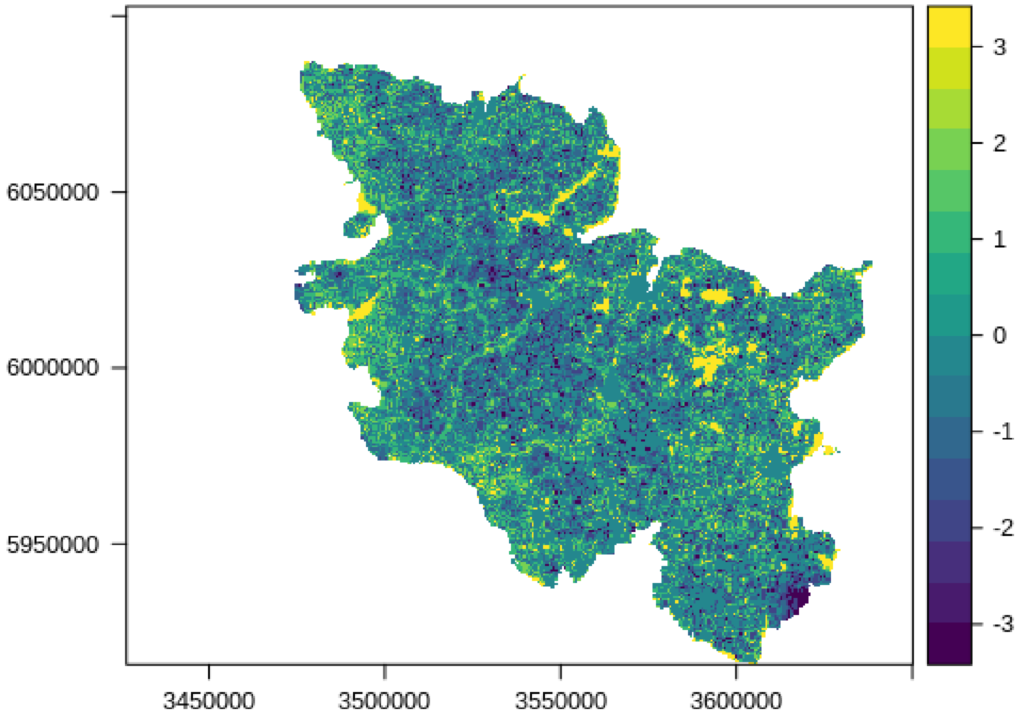


Figure S4.13.1 Differences between quantiles of the spatial predictions of the best performing movement-informed model and the highest ranked habitat-informed model: p(BCPA_SSF) – p(MAXENT). Positive values indicate that the BCPA_SSF predicted higher probability of selection as compared to the habitat suitability derived from MAXENT: p(BCPA_SSF) > p(MAXENT). A value of zero delineates areas where quantiles of the two models are equal: p(BCPA_SSF) = p(MAXENT). Negative values are the result of higher quantiles of suitability based on MAXENT as compared to BCPA_SSF resource selection: p(MAXENT) > p(BCPA_SSF).

General Discussion

The aim of this thesis was to uncover the genetic consequences of fragmentation and other restrictions primarily caused by human activities to dispersal and gene flow of red deer in Northern Germany. For this, I utilized both, information derived from genetic data as well as movement data, and applied a landscape genetic framework to study different possibilities for estimating functional connectivity.

Except for a few studies on detection of barriers to gene flow (Coulon et al. 2008; Frantz et al. 2012; Kuehn et al. 2007) large ungulates such as red deer in Europe are underrepresented in landscape genetic studies. In order to benchmark the extensive toolbox available to landscape genetics and to evaluate the effects of landscape fragmentation I used red deer in Schleswig-Holstein as an example to demonstrate the methodological and conceptual differences of multiple approaches for estimating functional connectivity.

In summary, the presented thesis provides important insights for applied conservation of wildlife and planning of corridors. Furthermore, the potential applications of landscape genetics and movement ecology to assess landscape connectivity are presented. For example, movement ecology provides an integral framework to explore the potential factors shaping the movements of organisms and the ecological consequences of these movements such as gene flow (Nathan 2008).

Effects of landscape fragmentation and other human-related restrictions

As shown in the second chapter of this thesis, genetic data has the great potential to illustrate the effects of landscape fragmentation but also of various other anthropogenic restrictions on wildlife populations. Genetic data provided valuable insight on the hierarchical structure of local populations which indicated multiple causes for the observed patterns of differentiation, some were based on the historic development of the populations (e.g., introduction of individuals from other parts of Europe) others were due to barriers such as primary roads or the Kiel Canal which separated once connected populations. Overall, genetic diversity was comparably low and even populations within reachable distances were asymmetrically connected, with high genetic exchange among some local populations and reduced connectivity of others. However, current delineation of red deer management units does not account for the observed levels of connectivity or isolation among them. In addition, the results indicated that population densities in neighboring management units also effect the level of genetic diversity within local populations. Therefore, densities of red deer should be taken more into consideration by local management. In particular, detailed information on the size and age- as well as sex-structure of local populations is necessary to gain a better understanding of the processes driving dispersal and gene flow (Draheim et al. 2016; Tucker et al. 2017).

Another important observation was that some red deer individuals dispersed from Denmark to an established population just north of the Kiel canal. This implies that dispersal over longer distances through the fragmented landscape is still possible and gives hope that in the future such processes can be revoked or promoted by protecting dispersal corridors and building overpasses to link these corridors across barriers. In conclusion, a better understanding on how landscape composition either impedes or facilitates effective dispersal of red deer is needed to address the observed loss of genetic diversity.

Assessing landscape connectivity and incorporating movement data

Throughout this thesis several methods have been applied for assessing landscape connectivity with focus on red deer in Schleswig-Holstein. Utilizing an extensive dataset comprising high resolution landcover data as well as telemetry relocations, several resistance models were fitted based on expert-knowledge, habitat suitability and movement analyses. As described in the first and fourth chapter all of the methodological approaches have their pros and cons (Spear et al. 2015). Overall, landscape genetics provided a valuable framework for objectively comparing these different models of landscape connectivity (Cushman et al. 2006; Storfer et al. 2010). As pointed out in multiple recent studies accounting for dispersal-specific movements and habitat selection during these movements turned out to be a promising approach and for some species even outperform other commonly applied approaches for estimating landscape connectivity (Roffler et al. 2016; Vasudev and Fletcher 2015; Zeller et al. 2018). Based on these indications I compared the habitat selected during potential dispersal movements of red deer with other regularly used methods for inferring primary habitat requirements. Thus, in order to identify potential dispersal movements from regular movements within established home ranges, the application of a path-segmentation method was necessary.

Identifying changes in movement behavior

The third chapter of this thesis presented an extensive review on the different methods for identifying behavioral patterns from movement data and showed how they can be applied to answer research questions within the movement ecology paradigm. Movement data provides highly relevant information not only for landscape genetics but also for studying the ecological consequences of animal space use and movement in general (Nathan et al. 2008). For example gaining a better understanding on how behavioral patterns and their interaction with the environment shape resource selection and thus the distribution of individuals or populations in the landscape (Boyce et al. 2003; Fleming et al. 2014; Roever et al. 2013). As shown in the fourth chapter, path-segmentation can also be an appropriate tool for identifying potential dispersal movements within a given dataset in order to estimate resource selection during these movements and infer functional connectivity for a given study species.

Overall, researchers are capable now of gaining much more insight on species movement behavior and with much more detail (Gurarie et al. 2016). In particular, with the growing availability of high resolution movement data and constant improvement of the applied tagging devices (Cagnacci et al. 2010). The provided overview on path-segmentation methods should therefore help to utilize the available date to its full potential and propagate this kind of analyses or spark new ideas for research in movement ecology (e.g., Hansen et al. 2019)

Benchmarking resistance models for red deer

Benchmarking the multitude of methods available for estimating functional connectivity with special focus on red deer was not straight forward. As described in the fourth chapter, performance of the different methodological approaches was also depending on the derived effective distances and applied statistical procedures to link them with genetic distances. Nevertheless, the findings indicated that for large scale connectivity the habitat-informed resistance model performed best as compared to the movement-informed model which best explained local scale connectivity derived from least-cost paths. This confirmed the hypothesis that at least for short range dispersal movements, red deer are capable of moving through less suitable habitats which has been observed in many other species as well (Abrahms et al. 2017; Zeller et al. 2018).

Although, overall performance of ensemble models was not significantly better in terms of describing the observed genetic patterns, I want to argue that if multiple data formats (telemetry data, presence data) are available an ensemble approach could very much overcome and compensate for drawbacks of each of the different analyses (Araújo and New 2007). In particular, for deriving conservation corridors from the resulting resistance surface the ensemble approach could most likely be more representative of the different processes driving connectivity at the local as well as the large scale.

The results of the benchmark analysis are especially relevant for other regions of Germany where similar issues of low genetic diversity as well as decreased gene flow in red deer populations have recently been observed (Reiner and Willems 2019). Although, detailed movement data such as GPS telemetry relocations are not always directly available at least a presence-based habitat model (Phillips et al. 2004) or even the expert-informed approach like the applied buffer map could be combined in an ensemble to infer landscape connectivity. However, transferring the results from Northern Germany to other regions has to be done with caution as habitat requirements and other factors driving the dispersal process could vary within the different localities.

Outlook

A major focus of this thesis was on the effects of the landscape matrix on dispersal and functional connectivity. However, the dispersal process consists of three important stages: the first being the decision by an animal to leave its current home range (emigration), the second stage of traversing through the landscape matrix (i.e., actual dispersal) and finally the third stage is to establish a new home range (immigration). In particular, the first and last stage are driven by local variables such as habitat quality, resource availability and population size (Pflüger and Balkenhol 2014). As discussed in chapter two local densities also seems to play an important role in genetic diversity and differentiation of red deer populations. Therefore, the potential effects of different local variables (e.g., habitat availability, population density, sex- and age-structure) need to be addressed in future research to obtain a complete picture on the processes driving effective dispersal of red deer in the study area (Pflüger and Balkenhol 2014). For example, individual-based simulations could be applied to estimate the potential effects of these variables and infer dispersal probabilities between populations which in turn could be compared to the observed genetic distances (Anadón et al. 2012; Fordham et al. 2014; Hoban 2014; Schumaker et al. 2014). Information on local densities and sex-ratio could be derived from spatial explicit capture-recapture using fecal samples (Borchers 2012; Royle et al. 2013). This would have the advantage of also gaining new genetic insight (e.g., diversity, gene flow) at the same time.

As shown in the fourth chapter, the availability of suitable habitats play an important role for large scale connectivity. The highest-ranked habitat model (MAXENT) could be utilized to delineate major areas of primary habitat for red deer (Franklin 2009). Subsequently, network analyses could be applied to test the significance of these habitat patches (i.e., stepping stones) on overall connectivity depending on their size and spatial arrangement (Bodin and Saura 2010; Rubio et al. 2014; Saura et al. 2014).

Finally, I expect the outcomes of this thesis to contribute to a better understanding on the reliability of available tools in a landscape genetic context and to provide evidence-based outcomes from the perspective of an important indicator species that might improve connectivity measures for environmental management from a regional to a state-wide scale. Overall, providing conservation areas and corridors as well as overpasses that facilitate exchange are a very important first step for reestablishing the great mobility potential of red deer and most likely leading the way for many other species. After the establishment of such mitigation measures a genetic reanalysis of the red deer populations should be performed in the future in order to give insight on the actual success of these measures and test if management goals (e.g., increasing gene flow) were accomplished. Although, this will take time since deer individuals have next to disperse also reproduce successfully at new established ranges.

References

Coulon, A., Morellet, N., Goulard, M., Cargnelutti, B., Angibault, J.-M., and Hewison, A. J. M. (2008). Inferring the effects of landscape structure on roe deer (Capreolus capreolus) movements using a step selection function. *Landscape Ecology*, *23*(5), 603–614.

Frantz, A. C., Bertouille, S., Eloy, M. C., Licoppe, A., Chaumont, F., and Flamand, M. C. (2012). Comparative landscape genetic analyses show a Belgian motorway to be a gene flow barrier for red deer (Cervus elaphus), but not wild boars (Sus scrofa). *Molecular Ecology*, *21*(14), 3445–3457.

Kuehn, R., Hindenlang, K. E., Holzgang, O., Senn, J., Stoeckle, B., and Sperisen, C. (2007). Genetic effect of transportation infrastructure on roe deer populations (Capreolus capreolus). *The Journal of Heredity*, *98*(1), 13–22.

Nathan, R., Getz, W. M., Revilla, E., Holoyak, M., Kadmon, R., Saltz, D., and Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, *105*(49), 19052–19059.

Abrahms, B., Sawyer, S. C., Jordan, N. R., McNutt, J. W., Wilson, A. M., and Brashares, J. S. (2017). Does wildlife resource selection accurately inform corridor conservation? *Journal of Applied Ecology*, *54*(2), 412–422.

Anadón, J. D., Wiegand, T., and Giménez, A. (2012). Individual-based movement models reveals sex-biased effects of landscape fragmentation on animal movement. *Ecosphere*, *37*.

Araújo, M. B., and New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, *22*, 42–47.

Bodin, Ö., and Saura, S. (2010). Ranking individual habitat patches as connectivity providers: Integrating network analysis and patch removal experiments. *Ecological Modelling*, *221*(19), 2393–2405.

Borchers, D. (2012). A non-technical overview of spatially explicit capture-recapture models. *Journal of Ornithology*, *152*, 435–444.

Boyce, M. S., Mao, J. S., Merrill, E. H., Fortin, D., Turner, M. G., Fryxell, J., and Turchin, P. (2003). Scale and heterogeneity in habitat selection by elk in Yellowstone National Park. *Ecoscience*, *10*(4), 421–431.

Cagnacci, F., Boitani, L., Powell, R. A., and Boyce, M. S. (2010). Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society of London Series B - Biological Sciences*, *365*(1550), 2157–2162.

Cushman, S., McKelvey, K., Hayden, J., and Schwartz, M. K. (2006). Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *The American Naturalist*, *168*(4), 486–499.

Draheim, H. M., Moore, J. A., Etter, D., Winterstein, S. R., and Scribner, K. T. (2016). Detecting black bear source – sink dynamics using individual-based genetic graphs. *Proceedings of the Royal Society B*, *283*, 2016002.

Fleming, C. H., Calabrese, J. M., Mueller, T., Olson, K. A., Leimgruber, P., and Fagan, W. F. (2014). From Fine-Scale Foraging to Home Ranges: A Semivariance Approach to Identifying Movement Modes across Spatiotemporal Scales. *The American Naturalist*, *183*(5), E154-67.

Fordham, D. A., Shoemaker, K. T., Schumaker, N. H., Akçakaya, H. R., Clisby, N., and Brook, B. W. (2014). How interactions between animal movement and landscape processes modify local range dynamics and extinction risk. *Biology Letters*, *10*, 20140198.

Franklin, J. (2009). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Gurarie, E., Bracis, C., Delgado, M., Meckley, T. D., Kojola, I., and Wagner, C. M. (2016). What is the animal doing? Tools for exploring behavioral structure in animal movements. *Journal of Animal Ecology*, 85, 69-84.

Hansen, N. A., Driscoll, D. A., Michael, D. R., and Lindenmayer, D. B. (2019). Movement patterns of an arboreal gecko in fragmented agricultural landscapes reveal matrix avoidance. *Animal Conservation*, *23*(1), 48-59.

Hoban, S. (2014). An overview of the utility of population simulation software in molecular ecology. *Molecular Ecology*, *23*(10), 2383–2401.

Nathan, R. (2008). An emerging movement ecology paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19052–19059.

Pflüger, F. J., and Balkenhol, N. (2014). A plea for simultaneously considering matrix quality and local environmental conditions when analysing landscape impacts on effective dispersal. *Molecular Ecology*, *23*(9), 2146–2156.

Phillips, S. J., Dudik, M., and Schapire, R. E. (2004). A Maximum Entropy Approach to Species Distribution Modeling. *Proceedings of the 21st International Conference on Machine Learning*, 655–662.

Reiner, G., and Willems, H. (2019). *Sicherung der genetischen Vielfalt beim hessischen*

Rotwild als Beitrag zur Biodiversität.

Roever, C. L., Beyer, H. L., Chase, M. J., and van Aarde, R. J. (2013). The pitfalls of ignoring behaviour when quantifying habitat selection. *Diversity and Distributions*, *20*(3), 322–333.

Roffler, G. H., Schwartz, M. K., Pilgrim, K. L., Talbot, S. L., Sage, G. K., Adams, L. G., and Luikart, G. (2016). Identification of landscape features influencing gene flow: How useful are habitat selection models? *Evolutionary Applications*, *9*(6), 805–817.

Royle, J. A., Chandler, R. B., Sollmann, R., and Gardner, B. (2013). *Spatial Capture-Recapture*. Academic Press.

Rubio, L., Bodin, Ö., Brotons, L., and Saura, S. (2014). Connectivity conservation priorities for individual patches evaluated in the present landscape: how durable and effective are they in the long term? *Ecography, 38*(8), 782-791.

Saura, S., Bodin, Ö., and Fortin, M. J. (2014). EDITOR’S CHOICE: Stepping stones are crucial for species’ long-distance dispersal and range expansion through habitat networks. *Journal of Applied Ecology*, *51*(1), 171–182.

Schumaker, N. H., Brookes, A., Dunk, J. R., Woodbridge, B., Heinrichs, J. A., Lawler, J. J., Carroll, C., and LaPlante, D. (2014). Mapping sources, sinks, and connectivity using a simulation model of northern spotted owls. *Landscape Ecology*, *29*(4), 579–592.

Spear, S. F., Cushman, S. A., and McRae, B. H. (2015). Resistance Surface Modeling in Landscape Genetics. In *Landscape Genetics* (pp. 129–148).

Storfer, A., Murphy, M. A, Spear, S. F., Holderegger, R., and Waits, L. P. (2010). Landscape genetics: where are we now? *Molecular Ecology*, *19*(17), 3496–3514.

Tucker, J. M., Allendorf, F. W., Truex, R. L., and Schwartz, M. K. (2017). Sex-biased dispersal and spatial heterogeneity affect landscape resistance to gene flow in fisher. *Ecosphere*, *8*(6), e01839.

Vasudev, D., and Fletcher, R. J. (2015). Incorporating movement behavior into conservation prioritization in fragmented landscapes: An example of western hoolock gibbons in Garo Hills, India. *Biological Conservation*, *181*, 124–132.

Zeller, K. A., Jennings, M. K., Vickers, T. W., Ernest, H. B., Cushman, S. A., and Boyce, W. M. (2018). Are all data types and connectivity models created equal? Validating common connectivity approaches with dispersal data. *Diversity and Distributions*, *24*(7), 868–879.

Acknowledgements

This thesis was funded by the scholarship program of the German Federal Environmental Foundation (Deutsche Bundesstiftung Umwelt, DBU). The majority of the applied data, in particular the telemetry data, were originally obtained during the pilot study "Sicherung genetischer Diversität beim Rothirsch in der Kulturlandschaft" financed by the Federal Office for Agriculture and Food (Bundesanstalt für Landwirtschaft und Ernährung, BLE; project number: 07BM010).

This dissertation would not have been possible without the many inspiring discussions and exchange of ideas with my peers and colleagues. A lot of people have crossed my path in the last seven years and were involved in the presented research one way or another.

In particular, I would like to thank:

- my first supervisor Prof. Dr. Niko Balkenhol, who gave me the opportunity to be part of his department and research team from the very beginning. I am very grateful for his patience and longtime support and guidance - I think we both learned a lot along the way.
- Dr. Christian Roos and Prof. Dr. Kerstin Wiegand for kindly serving as referees and co-referees.
- The Co-Authors of my publications for their input and pushing our research forward.
- The Fulbright Foundation for giving me the opportunity to take my research abroad and learn new perspectives.
- Marcus Meißner and the employees of IWGD for giving me the opportunity to be part of their long term research project, providing me with their valuable data and knowledge on red deer in Schleswig-Holstein.
- Dr. Johannes Signer who was there from the very beginning of our journey at the Department of Wildlife Sciences and for his R support.

- My former colleagues, office mates and students at the department: Femke Pflüger, Laura Richter, Kathrin Mayer, Scott Appleby, Maarten Hoffman, Ludmila Osipova, and Camilla Tusche. Thank you for all the good times we had and hopefully continue to have.
- Horst Reinecke for enlightened talks and discussions not only about statistics, wildlife research and management but also the important things in life (football).
- My new colleagues at the LWF for giving me the chance and especially time to finish my thesis most notably Dr. Wibke Peters and Alois Zollner.
- Katharina Westekemper and Poldi for the relaxing afternoon walks, many thoughtful conservations and all the help on the way to the finish line.
- Jan Engler, for his guidance and support, the good friendship, and mentorship until the very end of this thesis.
- Kirsten Weingarth and Dominik Dachs for the good friendship, the helpful support along the way (especially at the very end), wonderful times at the "Dachsbau" and many more to come.
- Last but not least my parents, for inspiring me, and for their ongoing support, believing in me and making my academic career in wildlife ecology possible.

Thesis Declaration

I hereby confirm that I have written this doctoral thesis independently, that I have not used other sources or facilities other than the ones mentioned, that I have not used unauthorized assistance and that I have not submitted this thesis previously in any form for another degree at any university or institution.

Hendrik Edelhoff

Freising, 31.12.2019