Developing Quantitative Methods for Movement Data

by

Jed A. Long
B.Sc., University of Guelph, 2006
M.Sc., University of Victoria, 2009

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Geography

Developing Quantitative Methods for Movement Data

by

Jed A. Long
B.Sc., University of Guelph, 2006
M.Sc., University of Victoria, 2009

Supervisory Committee

Dr. Trisalyn A. Nelson, Supervisor
(Department of Geography, University of Victoria)

Dr. Rosaline Canessa, Departmental Member
(Department of Geography, University of Victoria)

Dr. Farouk S. Nathoo, Outside Member
(Department of Mathematics & Statistics, University of Victoria)

Supervisory Committee

Dr. Trisalyn A. Nelson, Supervisor
(Department of Geography, University of Victoria)

Dr. Rosaline Canessa, Departmental Member
(Department of Geography, University of Victoria)

Dr. Farouk S. Nathoo, Outside Member
(Department of Mathematics & Statistics, University of Victoria)

ABSTRACT

Scientists are now able to collect ubiquitous data on individual-level movement at increasingly fine spatial and temporal resolutions. Despite this surge in data availability, methods for extracting relevant information about spatial-temporal movement patterns remain limited in scope and sophistication. The objective of this PhD research is to develop novel quantitative approaches for analyzing spatial-temporal patterns in modern movement datasets. A review of the state-of-the-art in quantitative movement analysis identifies the current breadth of available methods, while highlighting key limitations and fragmentation in the literature across multiple disciplines. Existing theory from the geographical literature, namely time geography is applied to a novel application wildlife movement ecology (termed the PPA home range), in an attempt to expose these ideas to wildlife researchers. The PPA home range method has several advantages over existing methods, most notably its ability to identify omission and commission error in existing home range techniques. Next, an advance to time geography theory is proposed for incorporating object kinetics (i.e., velocity and acceleration) into a probabilistic movement model termed kinetic-based probabilistic time geography. Kinetic-based probabilistic time geography provides a more accurate model for predicting object movement when object kinetics are relevant (e.g., with fast moving vehicles, or athletes). A novel method (termed the

DI index) for quantifying dynamic interactions between moving objects is presented, focusing specifically on examining cohesive movement behaviour. The DI index is advantageous over existing dynamic interaction measures in that it is computed at the local level, facilitating a finer treatment of interactive movement behaviour. The DI index is then contrasted with seven alternative measures of dynamic interaction to examine the effectiveness of each at identifying expected and unexpected interactive behaviour, at a range of sampling resolutions, in the context of wildlife movement ecology. The results highlight the value of the DI index, especially as a local level index, capable of identifying variable and infrequent interactions in pairs of moving objects. In summary, this dissertation contributes to the rapidly expanding body of quantitative movement research by providing: 1) a cross-disciplinary methodological review, 2) expanding the application of core time geography theory to wildlife ecology, 3) advancing time geographic theory in development of kinetic-based probabilistic time geography, 4) developing a novel index (the DI index) for measuring inter-object interactions, and 5) examining the effectiveness of available dynamic interaction measures, and their sensitivity across sampling resolutions, in the context of wildlife ecology.

Contents

List of Tables

List of Figures

Acknowledgements

First and foremost, I would like to thank my supervisor Trisalyn Nelson. Over the last 6 years you have been the driving force behind my research, providing the best possible environment to conduct my postgraduate studies. By pushing me to improve and develop, most notably in my writing, you have allowed me to reach some lofty goals and achievements. My time in SPAR at UVic has been a truly rewarding experience, and one that I know I will look back at fondly for the rest of my life.

To my committee members, Rosaline Canessa and Farouk Nathoo, I am grateful for your support and guidance.

To my co-authors, Trisalyn Nelson, Farouk Nathoo, thank-you for all your hard work and timely responses in the generation of manuscripts and research discussion. To Stephen Webb thank you for making the initial contact and the opportunity to work with you and the deer telemetry data you and Ken Gee have collected.

Thanks to Gord Stenhouse and the Foothills Research Institute for graciously providing the grizzly bear data.

A big shout-out to the University of Victoria Ultimate Frisbee Club for their on-going participation in various sports-related GPS-tracking experiments, you guys rule.

Over the last six years I have been fortunate to have some great SPAR lab-mates who have been a fun-loving group and have made coming into work that much more enjoyable. Colin Robertson, Carson Farmer, Mary Smulders, Ben Stewart, Nick Gralewicz, Jessica Fitterer, Keith Holmes, Mathieu Bourbonais, Shanley Thompson, Cesar and Liliana Perez, and so many more; without you guys I probably would have never finished, and I certainly would have had a lot less fun.

My amazing family, friends, and teammates have provided me with a positive atmosphere, providing a level of comfort that has allowed me to thrive in my academic life. To mom and dad, I love you and thanks for being that foundation and source of unwavering support for both me and my brothers.

Finally, to my wonderful wife Gina, I love you! Thanks for being my best friend, partner, sounding board, proof-reader, presentation tester, complaint manager, financial advisor, and overall life-long companion. You are the best distraction I could ever have, and I look forward to embarking on our new adventure together.

Co-Authorship Statement

All the manuscripts contained in this dissertation (Chapters 2-6) were co-authored, and the following outlines each of the authors' contributions, as well as that of the doctoral candidate. An appropriate reference conferring the publishing status of each chapter is provided.

Chapter 2

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

JL designed the review, performed the literature search, and prepared the manuscript. TN aided in the preparation of the manuscript with comments, edits, and advice on content and structure.

Chapter 3

Long, J.A., Nelson, T.A. (2012) Time geography and wildlife home range delineation. *Journal of Wildlife Management*. 76(2): 407-413.

JL developed the PPA home range method, performed the analysis, and prepared the manuscript. TN aided in the preparation of the manuscript with comments and edits, and advice on home range analysis.

Chapter 4

Long, J.A., Nelson, T.A., Nathoo, F.S. (In Review) Towards a kinetic-based probabilistic time geography. *International Journal of Geographical Information Science*. Submitted January 24, 2013.

JL designed the model, performed the analysis, and prepared the manuscript. TN aided in the preparation of the manuscript with comments, edits, and advice on linkages to GIScience. FN provided assistance on statistical background and model development, and aided in preparation of the manuscript with comments and edits.

Chapter 5

Long, J.A., Nelson, T.A. (2013) Measuring dynamic interaction in movement data. *Transactions in GIS*. 17(1): 62-77.

JL developed the DI index, collected the ultimate frisbee data, performed all analysis, and prepared the manuscript. TN obtained the grizzly bear data from an outside source, and aided in preparation of the manuscript with comments and edits.

Chapter 6

Long, J.A., Nelson, T.A., Webb, S.L., Gee, K.L. (In Review) A critical examination of indices of dynamic interaction for wildlife telemetry studies. *Journal of Animal Ecology*. Submitted April 8, 2013.

JL designed the study, performed all analysis, and prepared the manuscript. TN aided in preparation of the manuscript with comments and edits. SW contributed to study design and research questions, assisted in manuscript preparation through comments, edits, and interpretation, and collected the GPS telemetry data. KG assisted in manuscript preparation with comments and edits, and made available the GPS telemetry data.

Chapter 1

Introduction

1.1 Introduction

The manner by which individuals (i.e., people, their vehicles, or even wildlife) move about their environment provides an illuminating view of individual preferences, motivations, and behaviour. While early movement research focused on the identification of broader-scale movement patterns (e.g., Lenntorp, 1976; Van Winkle, 1975), contemporary research has primarily focused on an altogether different aspect of movement, examining the finer-scale geometry of movement problems relating to daily behaviour (e.g., Kwan, 1998; Neutens *et al.*, 2010) or movement within constrained environments (e.g., tourism parks, Shoval and Isaacson, 2007). Movement data, defined as the recursive collection of the space-time location of an individual, is now ubiquitously collected, allowing scientists to perform increasingly detailed analysis of movement patterns (Laube *et al.*, 2007). In fact, movement tracking technologies have become a part of everyday modern culture and our now routinely integrated into devices such as cell-phones, watches, and vehicle navigation systems (Raper *et al.*, 2007).

Fine-scale movement data, such as collected via GPS, are now being used in a variety of research applications for addressing many unique types of questions, originating from a host of disciplines. In human research, movement data are being used, for example, to study human exposure to chemicals (Elgethun *et al.*, 2003), investigate daily activity levels (Rodriguez *et al.*, 2005), monitor patients with cognitive diseases (Miskelly, 2005; Shoval *et al.*, 2008), and investigate gender related differences in daily routines (Kwan, 1999). Transportation research has included the study of traffic patterns (Verhein and Chawla, 2006), management of fleet services (Mintsis

et al., 2004), and optimizing emergency response teams (Derekenaris *et al.*, 2001). Perhaps the largest body of research exists in the study of wildlife, as their cryptic nature necessitated the use of mechanical tracking technologies early-on. With wildlife, movement data have been used to study migration patterns (Sawyer *et al.*, 2005), foraging behavior (Hebblewhite *et al.*, 2008), relationships with habitat variables (Mosnier *et al.*, 2003), and inter- and intra-species interactions (Bandeira de Melo *et al.*, 2007; Merrill *et al.*, 2010). Beyond these more applied examples, disciplines such as geography, computer and database science, sport science, and robotics are also using and developing methods for movement data in varying capacities.

While considerable focus has been placed on the collection and viewing of movement data, there remain relatively few options for more formal quantitative analysis (Wolfer *et al.*, 2001; Laube *et al.*, 2007). New and online geovisualization platforms, such as Google Earth, provide simple tools for animated viewing of movement paths (Goodchild, 2008), and given the considerable illustrative value associated with movement data, geovisualization remains at the forefront of movement research (e.g., Kraak, 2003; Andrienko and Andrienko, 2007; Pultar *et al.*, 2010). As such, the study of movement continues to be one of the latest examples of the data-rich information-poor age where geographic science currently resides. The challenge moving forward lies in deriving meaningful quantitative techniques for extracting information on the patterns observed in increasingly detailed records of object movement stemming from modern tracking systems.

1.1.1 Characteristics of Modern Movement Data

Movement is a continuous process that occurs through both space and time. A moving object traverses a path through the space-time continuum. Movement data are made up of discrete points recorded from an objects movement path, thus representing only samples (often termed fixes) extracted from the true path. In most cases, samples are recorded at regular sampling intervals (in either space or time), but can also occur at irregular intervals (Calenge *et al.*, 2009). Early tracking systems recorded fixes at relatively coarse intervals, for example obtaining only 1 fix every two weeks to monitor caribou movements (Bradshaw *et al.*, 1995). Modern tracking systems now facilitate increasingly fine sampling intervals (e.g., GPS tracking of athletes with sub-second fix frequencies, Coutts and Duffield, 2010) representing movement data as a (near) continuous movement path (Laube *et al.*, 2007).

Originally, collecting data for studying movement was a labor intensive task, requiring manual tracking of subjects (Weiss and Boutourline, 1963), or relying on detailed travel diaries (Hartenstein and Iblher, 1967). These studies often focused on the activities associated with movement, and less so on the spatial patterns of movement. With time, scientists replaced manual tracking and travel diaries with novel tracking technologies for recording movement data (Cochran and Lord, 1963; Tester and Heezen, 1965; Craighead, 1982; Wolf *et al.*, 2001). Early quantitative tracking technologies (e.g., radio or satellite telemetry, early GPS) where characterized by considerable spatial uncertainty, making their use problematic in more rigorous spatial analyses. In May 2000 the United States military removed selective availability the means by which GPS accuracy was intentionally degraded. GPS based position estimates immediately improved from an error range of approximately 100 metres to only a few metres. Wildlife scientists immediately recognized the potential of GPS for wildlife movement research as it provided the ability to collect space-time location estimates with increased accuracy, precision, and reliability (Hansen and Riggs, 2008). Geographers, as well, realized the potential of GPS, and began replacing travel diaries with GPS tracking devices in human movement studies in order to generate more reliable, and objectively collected movement data (Wolf *et al.*, 2001).

GPS technologies have revolutionized the study of movement in two ways. First, GPS provides a more accurate spatial representation of the movement path than any previously developed tracking system. That is, the spatial error in GPS studies is minimal (< 10 m) and, for most movement applications, is usually negligible in the context of the overall movement path extent. Second, and perhaps more importantly, GPS technology allows location estimates to be recorded with high frequency. Thus, while other systems provide location estimates with coarse sampling intervals (e.g., every hour); GPS systems can facilitate fine sampling intervals (e.g., every second). In the absence of other limiting factors (most notably battery life) modern GPS tracking systems can provide a (near) continuous representation of movement, with minimal spatial error (Laube *et al.*, 2007). The change from coarser tracking methods to fine-scale GPS movement tracking has happened quickly, and thus researchers have struggled with developing suitable techniques for analyzing fine-scale GPS-based movement data, often reverting to existing methods ill-suited for fine-scale GPS data (see Kie *et al.*, 2010).

1.1.2 Scope of Movement Analysis

The study of movement has evolved from within geography (Hägerstrand, 1970), but also in other disciplines such as wildlife ecology (Turchin, 1998), transportation (Miller, 1991), and computer and database science (Güting and Schneider, 2005; Verhein and Chawla, 2008). The diversity of disciplines interested in studying movement has led to the divergence of techniques used for collecting, managing, and especially for analyzing movement data. In some cases, different disciplines have come up with a host of unique methods for examining essentially the same problems (for example dynamic interaction measures by wildlife ecologists, and path similarity measures in the movement data mining literature). Furthermore, the language used to describe movement data and concepts is rarely consistent, even within disciplines, which can make comparison across techniques unnecessarily challenging.

The fragmentation of the literature on movement analysis makes it difficult to assess the state-of-the science, but also makes it challenging for those from each discipline to stay current with the latest approaches. Geography, as a discipline, is routinely faced with this dilemma, as geographers commonly take a multi-disciplinary approach to problem solving, drawing on theory and methods from a host of related disciplines (Johnston, 2003). As a geographer, I have tried to take a multi-disciplinary approach to movement research, examining concepts and ideas that have relevance in a host of disciplines and applications. Further, I have attempted to make my research accessible to various disciplines outside geography who are interested in similar components of movement analysis. Specifically, I have tried to expose fundamental geographic theory on movement, namely time geography, to other disciplines, such as wildlife ecology.

Research on moving objects has surged in recent years, primarily due to the development of tracking technologies, such as GPS (Miller, 2010). In the study of human movement, emphasis has been placed on the discovery of recurring patterns, from the often large databases of GPS paths. Of particular interest has been the identification of similar movement patterns or clusters of paths, representing popular movement thoroughfares and accumulation points (Verhein and Chawla, 2006).

These same technologies used in the study of human movement by geographers have been paramount to the growth of wildlife movement ecology and related research (Tomkiewicz *et al.*, 2010). Wildlife ecologists have focused on measures of space use – termed home range (Burt, 1943), or utilization distribution (Van Winkle,

1975), in order to map where animals go. Home ranges and utilization distributions are routinely linked to underlying environmental variables to study wildlife habitat composition and preferences (Aebischer *et al.*, 1993). Also, when multiple animals (sometimes even from different species) are tracked simultaneously it is of interest to study the manner in which these individuals interact (Kernohan *et al.*, 2001). Studying wildlife interactions can provide novel insight on complex social behaviour, such as mating, and familial structure (Kenward *et al.*, 1993).

Movement data, despite being best described as following a path, is most commonly made up of a series of spatial points, each depicting the spatial location of the object at a moment in time. Thus, movement data are often commonly viewed statically (e.g., in a geographic information system – GIS) as a collection of spatial points. It seems natural, given the efficiency with which movement data are stored and displayed in a GIS, that the tools and techniques commonly used in traditional spatial analyses have been modified for movement applications. For example, many methods commonly used in the analysis of spatial point patterns (e.g., k-nearest neighbor analysis, kernel density estimation) are routinely implemented with movement data stored as points (Worton, 1989; Kenward *et al.*, 1993; Güting *et al.*, 2010). However, movement is a dynamic space-time process, and movement data (a set of recorded points sampled from movement path) are collected sequentially, implying an ordered connection between spatial locations through time. In the strictest statistical sense, this means that movement data are often in violation of the assumptions of spatial independence associated with classic spatial statistical tests that are often applied to movement data. This is due to the fact that movement data suffer from the inherent serial autocorrelation of location fixes (Swihart and Slade, 1985; Dray *et al.*, 2010). More fundamentally, movement data present an altogether different set of problems requiring unique solutions explicitly tailored to movement applications (Calenge *et al.*, 2009).

1.1.3 Objectives and Content Overview

The goal of this research is to advance methods for quantitative analysis of modern movement data. Given the current lack of analytical tools for movement data, one of the primary goals of my PhD research is to develop quantitative analysis methods tailored specifically to fine-scale movement datasets. In my research, I have tried to focus on methods that explicitly consider the space-time properties of movement as

a path by specifically incorporating the ordered sequencing of records. Finally, my hope was to expose other disciplines (especially wildlife ecology) to existing movement theory and research currently present in the geographical literature.

As a first step it is essential to assess the state-of-the art for movement research. Given that quantitative movement analysis is the domain of many research applications (e.g., geography, wildlife ecology, computer science), the literature on movement analysis is especially fragmented. Thus, while reviews on discipline-specific components of movement analysis are available (e.g., Laver and Kelly, 2008; Güting and Mamoulis, 2011) there exists no comprehensive review of quantitative movement analysis. Chapter 2 provides such a review of quantitative movement methods, within which I have identified seven different categories of methods related to movement analysis problems common between various disciplines. As part of this review, I also discuss the challenges routinely faced in movement analysis, for example the role of scale (granularity, analysis scale, spatial scale) in movement research. From my assessment of the current state of movement research I propose several avenues for future research with emerging movement datasets. Specifically, I am interested in the development of methods related to two underlying themes in movement analysis: 1) Hägerstrand's time geography, and 2) inter-object interactions.

In Chapter 3, I provide a novel application of Hägerstrand's time geography for use in analyzing wildlife space use. The method I have developed, termed the potential path area home range, represents an alternative approach to delineating wildlife home ranges, one that explicitly considers the space-time constraints in wildlife movement. Further, this method serves as a complementary view to existing methods, capable of identifying commission and omission errors often prevalent in other methods. By utilizing Hägerstrand's time geographic framework, the potential path area home range explicitly considers the temporal dimension in the analysis of wildlife space use, a key component absent in many existing methods.

With Chapter 4 I further develop the construction of quantitative movement theory and mathematical definitions (e.g., Miller, 2005) associated with Hägerstrand's time geography. Recently, two extensions to Hägerstrand's time geography have been developed, probabilistic time geography (Winter, 2009; Winter and Yin, 2010, 2011), which models movement probabilities within time geographic structures; and kinetic time geography, (Kuijpers *et al.*, 2011) which alters the shape of time geographic structures to accommodate object kinetics. In Chapter 4, I develop a kinetic-based probabilistic time geographic model that generalizes the work of Winter and Yin

(2010) in order to account for object kinetics, as in Kuijpers *et al.* (2011). Kinetic-based probabilistic time geography represents a primarily theoretical advancement, but one that has many interesting potential applications in movement analysis, for example with location based services.

In Chapter 5, I present a novel method for quantifying inter-object interactions, termed the dynamic interaction (DI) index. The DI index extends an existing approach for quantifying movement interactions (Shirabe, 2006) by considering interaction in movement speed and direction independently, but more importantly through the development of an appropriate local-level statistic. The DI index is demonstrated using simulated examples, but also wildlife (grizzly bear), and athlete (ultimate frisbee) movement case-studies.

In Chapter 6, I examine the suite of available methods for quantifying dynamic interaction patterns (including the DI index from Chapter 5) in the context of wildlife movement analysis. First, I review the suite of available techniques for quantifying dynamic interaction patterns in wildlife movement data. Then, using a simulation study, I examine the effectiveness of each technique at correctly measuring expected vs. unexpected dynamic interaction patterns, at three different scales. I follow, with an example using empirical GPS movement data of white tailed deer, in which I further justify the use of the DI index developed in Chapter 5.

Finally, with Chapter 7, I present some concluding remarks. I summarize my developments and contributions to the state of movement research in geography and wildlife ecology. Specifically, I discuss the role of the methods and theory I have developed in furthering the science of movement analysis. As well, I lay down several courses for further research, which build upon the ideas and methods I have proposed here. In future works, I hope to address many of these research areas in order continue exploring suitable methods for quantitative movement analysis in geography, wildlife ecology, and other interesting applications.

Chapter 2

A Review of Quantitative Methods for Movement Data

Abstract

The collection, visualization, and analysis of movement data is at the forefront of geographic information science research. Movement data are generally collected by recording an object's spatial location (e.g., XY coordinates) at discrete time intervals. Methods for extracting useful information, for example space-time patterns, from these increasingly large and detailed datasets have lagged behind the technology for generating them. In this article we review existing quantitative methods for analyzing movement data. The objective of this article is to provide a synthesis of the existing literature on quantitative analysis of movement data while identifying those techniques that have merit with novel datasets. Seven classes of methods are identified: 1) time geography, 2) path descriptors, 3) similarity indices, 4) pattern and cluster methods, 5) individual-group dynamics, 6) spatial field methods, and 7) spatial range methods. Challenges routinely faced in quantitative analysis of movement data include difficulties with handling space and time attributes together, representing time in GIS, and using classic statistical testing procedures with space-time movement data. Areas for future research include: investigating equivalent distance comparisons in space and time, measuring interactions between moving objects, development of predictive frameworks for movement data, integrating movement data with existing geographic layers, and incorporating theory from time geography into movement models. In conclusion, quantitative analysis of movement data is an active

research area with tremendous opportunity for new developments and methods.

2.1 Introduction

The study of movement in geographic information science (GISci) has followed a similar trajectory to the discipline of geography, whereby early work relied heavily on qualitative methods. In the 1960's and 70's the discipline of geography experienced a quantitative revolution whereby theory and methods were developed for explaining how place and space could be modeled as quantitative entities. The quantitative revolution produced developments in statistical methods designed specifically for spatial data, for instance spatial autocorrelation measures (Cliff and Ord, 1973). Only later in the quantitative revolution did theoretical frameworks for quantitative analysis of movement emerge most notably Hägerstrand's 1970 time geography. As the quantitative revolution in geography sputtered in the late 1970's (Johnston, 1997) Hägerstrand's ideas were primarily used as context for examining human behavior (e.g., Parkes and Thrift, 1975; Pred, 1981), rather than as an analytical toolkit for quantitative research. An exception is the work of Lenntorp (1976) and Burns (1979), which represent seminal pieces using time geography in quantitative analysis.

In the 1990's, triggered by the development of geographic information systems (GIS), quantitative analysis again moved to the forefront of the geographic literature (Sheppard, 2001). The term geographic information science (GISci) was coined to refer collectively to the science behind the collection, storage, representation, and analysis of geographic datasets (Goodchild, 1992). The term amalgamated those interested in the study of geographic information including geographers, computer scientists, and statisticians. As technologies for recording the paths of moving objects have evolved (e.g., video, cell-phone, and GPS tracking) contemporary GIScientists have found new opportunities for quantitative analysis using time geography with GISci (e.g., Miller, 1991; Kwan, 1998). Other quantitative methods for analyzing movement have stemmed from geography's strong legacy in spatial point pattern analysis (e.g., Gao *et al.*, 2010), as movement data are commonly represented by a sequence of points. Computational geometry has played a leading role in recent advances in analyzing movement data (e.g., Laube *et al.*, 2005). As well, methods for representing movement data using areal data formats, for example polygons (Downs and Horner, 2009) or fields (Downs, 2010), remain ongoing research areas. The study of movement is of interest in many applications outside of GISci, for example wildlife

ecology (Nathan *et al.*, 2008), urban planning (Drewe, 2005), and military applications (Wells, 1981). Further, the study of movement has a long history in physics. Even Hägerstrand's time geography was strongly influenced by the ideas of physicists from the early 20th century (Rose, 1977; Hallin, 1991). For example, the diagram of the space-time cone from time geography can be clearly related to the past and future light-cones used in Einstein's relativity.

Movement is a complex process that operates through both space and time. Representing the temporal dimension in geographic studies has presented a challenge for GISci to move beyond static (map-based) representations of space (Chrisman, 1998; Laube *et al.*, 2007). Despite notable advances at incorporating temporal dynamics in GISci (e.g., Pultar *et al.*, 2010), integrating the study of space and time remains at the forefront of GISci research, as evidenced by the special symposium on space-time integration in GISci at the 2011 annual meeting of the Association of American Geographers. How to effectively integrate time into the quantitative analysis of movement, specifically movement data stored in a GIS, is at the core of this review.

The growth of spatial methods for quantitative analysis of movement data has been facilitated by developments in movement databases that now provide efficient methods for storing, indexing, and querying movement data (Güting and Schneider, 2005). Despite the large body of existing literature on the topic of moving object databases, it remains an active area of research as new tools (e.g., Güting *et al.*, 2010) and applications (e.g., Jensen *et al.*, 2010) continue to develop. Data visualization methods have developed alongside these readily available movement databases; in GISci this practice is termed geovisualization (Dykes, 2005). Given the sheer volume of data often contained in movement databases, geovisualization can be a powerful tool for identifying patterns in movement databases – a process referred to as visual analytics (Thomas and Cook, 2005). A complete treatment of either of these topics is beyond the scope of this review, and we restrict the contents of this review to, as the title suggests, those methods for analyzing movement data that are quantitative in nature. We would point those interested in more information on movement databases to the comprehensive book by Güting and Schneider 2005 and a recent special issue on data management for mobile services (VLDB Journal, 20(5), Güting and Mamoulis, 2011). For those interested in more information on visual analytics for movement data we refer readers to Andrienko *et al.* (2007), and to the special issue from IJGIS entitled geospatial visual analytics: focus on time (IJGIS, 24(10), Andrienko *et al.*, 2011).

The objective of this review is to provide an unbiased evaluation of the usefulness and shortcomings of existing quantitative methods for movement data, while highlighting techniques that have particular merit with emerging movement datasets. Challenges to the development and application of quantitative methods with movement data are identified in an attempt to locate avenues for future research. An outline of this article is as follows; section 2.2 contains a brief introduction to the properties of movement data, and how movement data is typically represented in a GIS. In section 2.3 we review the existing literature on quantitative analysis of movement data separated into seven classes of methods: 1) time geography, 2) path descriptors, 3) similarity indices, 4) pattern and cluster methods, 5) individual-group dynamics, 6) spatial field methods, and 7) spatial range methods. With section 2.4 we provide a discussion of the challenges routinely faced in GISci when analyzing movement data and, what we feel are, some future directions for quantitative movement analysis. Lastly, we close with some conclusions.

2.2 Movement Data

Movement is a continuous process that operates in both the spatial and temporal domains. Movement data are used to represent the continuous process of movement for geographical analysis. Due to existing geospatial data collection and storage techniques, movement data are most commonly represented as a collection of spatial point objects with time stored as an attribute. A more formal definition of movement data is the collection M_t of $t = 1...n$ ordered records each comprised of the triple $[ID, S, T]$, where ID is a unique object identifier, S are spatial coordinates, and T a sequential (non-duplicated) time-stamp (Hornsby and Egenhofer, 2002). A number of terms are used synonymously for movement data (see Table 2.1); here we use the term path to represent the ordered sequence of records portraying individual/object movement, the term fix when discussing a single record from a path, and the term movement database to describe a collection of paths. The term movement data is used in broader contexts when discussing the study of movement, to refer generally to fixes, paths, and movement databases.

While movement data have historically been collected using a variety of techniques, most current acquisition schemes use some form of wireless sensor (e.g., GPS, cellular phone records, radio telemetry). Calenge *et al.* (2009) identify two types of sampling commonly employed in the collection of movement data regular and irregu-

Table 2.1: Terms used synonymously for describing movement data.

Description	**Term**	**Synonomous Terms**
A single record of object movement (of the form $< ID, S, T >$).	Movement Fix (M_t)	point, observation, relocation
A sequence of ordered records in time depicting the movement of a single object.	Movement Path (M^a)	space-time path (Hägerstrand, 1970), trip-chain (Kondo and Kitamura, 1987), geospatial lifeline (Mark, 1998), trajectory, trace, track
A collection of records depicting the movements of many objects or the same object at different occasions, potentially including attribute information.	Movement Database	moving objects database (Güting and Schneider, 2005)

lar. Regular paths are those where fixes are acquired at an even temporal interval, for example recording one fix per minute. Irregular paths are those where fixes are acquired at unequal temporal intervals, for example paths collected from cell phone call records. The term granularity is used to refer to the resolution of a path (Hornsby and Egenhofer, 2002). Finer granularities are associated with frequent sampling intervals, and provide a detailed representation of movement. Conversely, coarser granularities correspond to sparse sampling and less-detailed representation of movement. Technological developments now facilitate finer sampling intervals in movement paths (e.g., 1 fix / second), and movement data can be used to represent a (near) continuous movement path (Laube *et al.*, 2007). However, these sensor-specific sampling designs may not be suitable for all analysis questions, requiring the use of re-sampling (up- or down-sampling) to fit a given research need (see Turchin, 1998; Hornsby and Egenhofer, 2002, for a more thorough discussion of changing granularity).

Spaccapietra *et al.* (2008) present an alternative view of movement data granularity, defining a path as consisting of stops and moves separating a path into periods of movement and stationary behavior. This conforms with the event-based model for movement data outlined by Stewart Hornsby and Cole (2007) which contrasts with the coordinate-based representation of movement typically employed. An event based model for movement data still allows for the detection of movement patterns, but with focus placed on combinations or sequences of events that identify a specific behavior, such as an exodus of objects out of a zone or region (Stewart Hornsby and

Cole, 2007). Further, event based models allow for enriching movement data with the geographic information associated with events, for instance if events are related to spatial regions the attributes of each region.

2.3 Review of Methods

This section contains a review of quantitative analysis methods that exist within seven areas of movement research; 1) time geography, 2) path descriptors, 3) path similarity indices, 4) pattern and cluster methods, 5) individual-group dynamics, 6) spatial field methods, and 7) spatial range methods. We emphasize techniques we feel have particular merit for analysis with novel and emerging movement datasets.

2.3.1 Time Geography

The concept of time geography was first presented in the 1960's and 1970's by Torsten Hägerstrand at the Research Group for Process and System Analysis in Human Geography at the University of Lund, Sweden (Lenntorp, 1999). Time geography (Hägerstrand, 1970) represents a framework for investigating the constraints, such as an object's maximum travel speed, on movement in both the spatial and temporal dimensions. Hägerstrand expanded on the purely physical limitations of movement, identifying three other types of constraints: capability, coupling, and authority constraints. Capability constraints limit the activities of the individual because of their biological construction and abilities, for example the necessity to eat and sleep. Coupling constraints represent specific locations in space-time an individual must visit that limit movement possibilities. Authority constraints are opposite of coupling constraints, locations in space time an individual cannot visit, for example a mall after it has closed. Contemporaries expanded on Hägerstand's work providing both theoretical (Parkes and Thrift, 1975; Pred, 1981) and applied (Lenntorp, 1976; Burns, 1979) extensions. Originally, time geography was used solely to investigate the movement of humans, but has since been reformulated for use with transportation networks (Miller, 1991) and wildlife ecology (Baer and Butler, 2000).

Time geography uses volumes (Figure 2.1) capable of capturing the movement limits of an object. A 3-D space (often termed cube Kraak 2003, or aquarium, Kwan 2004), with two spatial axes representing geographic space and a third orthogonal axis for time, is used to develop time geography volumes. The space-time cone (Figure

2.1a) identifies the future movement possibilities of an object. A space-time prism (Figure 2.1b) is used to quantify movement possibilities between known start and end locations. The potential path area is the projection of the space-time prism onto geographic space (Figure 2.1c), and is a purely spatial measurement of movement capability. A path is used to portray the trajectory of movement through space-time. Bundling (Figure 2.1d) occurs when multiple paths coincide in space and time, for example taking the same bus to work. Typically, time geography is discussed qualitatively in terms of the aforementioned volumes, but Miller (2005) has provided mathematical definitions for time geography concepts that can be used in more rigorous quantitative analyses.

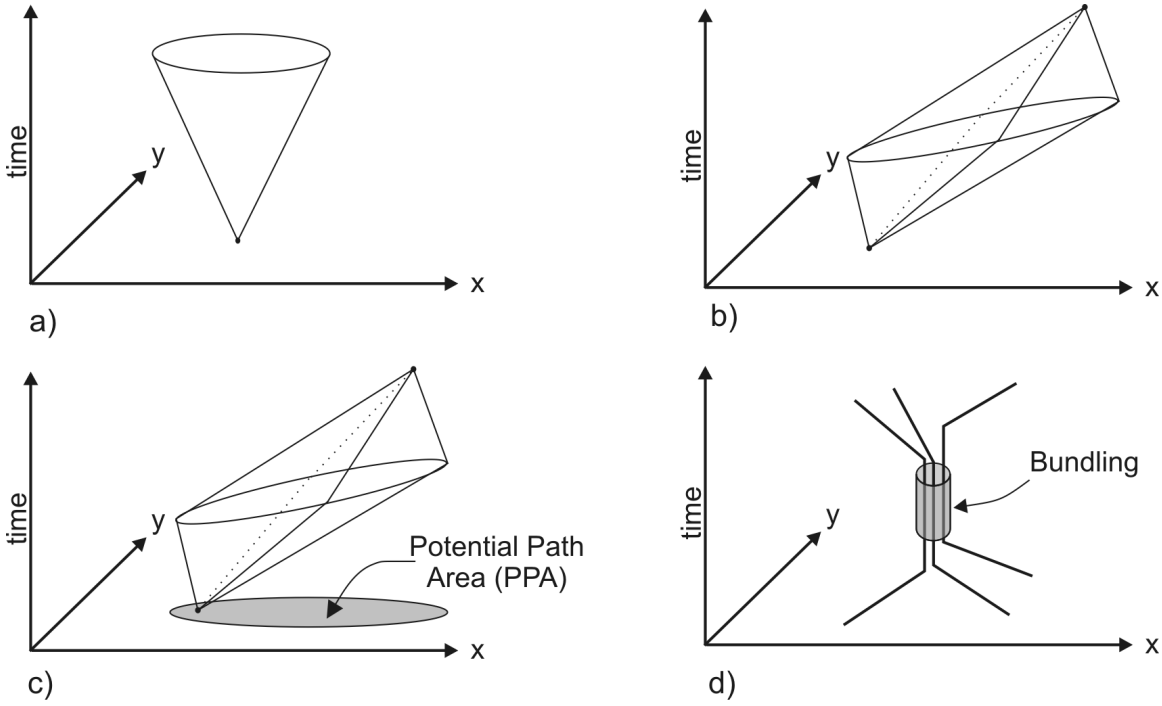

Figure 2.1: Volumes used in Hägerstrand's time geography: a) space-time cone, b) space-time prism, c) potential path area, and d) path bundling.

Recently, with advances in GISci and movement data, time geography is experiencing a resurgence (Miller, 2003). Lenntorp (1999) explains how time geography has reached 'the end of its beginning', suggesting that current and future research using GIS and novel movement datasets will present new and exciting developments in time geography. Examples include using time geography to investigate mobility data on a network (Miller and Wu, 2000), factoring in uncertainty (Neutens *et al.*, 2007), field-based time geography (Miller and Bridwell, 2009, further discussed in

section 2.3.6), and the development of a probabilistic time geography (Winter, 2009, further discussed in section 2.3.6).

Time geography represents a useful tool for quantitative analysis of movement as it contains a framework for measuring space-time bounds on movement. Movement models that fail to consider the constraints provided by space and time often result in misleading conclusions (Long and Nelson, 2012). Methods that explicitly consider time geography principles, even unknowingly (e.g., Yu and Kim, 2006), avoid such deceptions.

2.3.2 Path Descriptors

Path descriptors are measurements of path characteristics, for example velocity, acceleration, and turning azimuth. Typically path descriptors may be calculated at each point in a movement dataset, and can be scaled appropriately to represent interval or global averages. Dodge *et al.* (2008) categorize a number of path descriptors as primitive parameters, primary derivatives, or secondary derivatives based on simple measurements in space, time, and space-time (see Table 2.2). Ecologists routinely use simple path descriptors in the study of wildlife movement (Turchin, 1998). Measures of movement tortuosity have also been developed for the study of wildlife, for example path entropy (Claussen *et al.*, 1997), sinuosity (Benhamou, 2004), and fractal dimension (Dicke and Burrough, 1988). Related to these are stochastic movement models (i.e., models where fixes are obtained via random draws from distributions for movement displacement and turning angle) such as Lévy flights (Viswanathan *et al.*, 1996) and correlated random walks (Kareiva and Shigesada, 1983). When movement data are statistically fit to such models, interpretation of model parameters can provide useful quantitative inference.

2.3.3 Path Similarity Indices

Path similarity indices are routinely used to quantify the level of similarity between two movement trajectories. It is desirable for similarity indices to take the form of a metric distance function, as metric functions are able to distinguish objects on an interval scale of measurement (Sinha and Mark, 2005). A metric distance function (d) is one that computes a generalized scalar distance between two objects while satisfying the following four properties (Duda *et al.*, 2001):

Table 2.2: Parameters extractable from movement data sorted by dimension. After Table 1 from Dodge *et al.* (2008).

	Primitive	**Primary Derivatives**	**Secondary Derivatives**
Spatial (x, y)	position	distance, direction, spatial extent	spatial distribution, change of direction, sinuosity
Temporal (t)	instance, interval	duration, travel time	temporal distribution, change of acceleration
Spatial-temporal (x, y, t)	–	speed, velocity	acceleration, approaching rate

i. Non-negativity: $d(x, y) \leq 0$;

ii. Reflexivity (uniqueness): $d(x, y) = 0$, iff $x = y$;

iii. Symmetry: $d(x, y) = d(y, x)$;

iv . Triangle Inequality: $d(x, z) \leq d(x, y) + d(y, z)$

The simplest similarity metric is a Euclidean measurement. Sinha and Mark (2005) implement a time-weighted distance metric where spatial proximity (Euclidean) is weighted by its temporal duration. Sinha and Mark (2005) also present a modified version of the time-weighted distance metric for the situation where the two objects move over different time intervals. Because the time-weighting is based on the duration an object spends at a given spatial location, this index works best with movement data defined as a series of stops and moves such as suggested by Spaccapietra *et al.* (2008).

Yanagisawa *et al.* (2003) present an alternative Euclidean-based similarity index that focuses on the shape of the movement path by normalizing the spatial coordinates of a path to a common plane. Euclidean measurements in the normalized spatial plane are used to identify similarly shaped movement paths. Euclidean distance is appropriate for comparisons in the spatial or temporal domains. However, Euclidean measurements are limited when data are represented with different scales (spatial and temporal). That is, what is the temporal equivalent to a 1 km distance in space? Despite these limitations, Euclidean distance similarity indices are frequently

implemented by fixing either space or time and considering Euclidean distance in the other dimension, such as the above examples.

Other distance metrics may be more appropriate for assessing path similarities. The Hausdorff distance is a shape comparison metric commonly used to evaluate the similarity of two point sets (Huttenlocher *et al.*, 1993), which has also been used to measure the similarity of movement paths. Given two movement paths M_a and M_b, the Hausdorff distance is defined as:

$$H(M^a, M^b) = max(h(M^a, M^b), h(M^b, M^a)) \tag{2.1}$$

with

$$h(M^a, M^b) = \max_{t \in T}(\min_{s \in S}(d(M_t^a - M_s^b))) \tag{2.2}$$

where t and s are used to index fixes from M_a and M_b respectively, and d is a distance operator (e.g., Euclidean). Not originally designed for movement data, the Hausdorff distance performs poorly when analyzing movement paths as it fails to consider the ordering of points (Zhang *et al.*, 2006), and is sensitive to outliers and data noise (Shao *et al.*, 2010). As such, modified versions of the Hausdorff distance metric have been designed specifically for use with movement paths (e.g., Atev *et al.*, 2006; Shao *et al.*, 2010).

The Fréchet distance metric may be more appropriate as a path similarity index as it was initially designed for comparing polygonal curves. Formally the Fréchet distance for two movement paths M_a and M_b is defined as:

$$\delta_F(M^a, M^b) = \inf_{\alpha,\beta} \max_{t \in [0,1]} d(M^a(\alpha(t)), M^b(\beta(s))) \tag{2.3}$$

Where α (resp. β) is an arbitrary continuous non-decreasing function from $[0, 1]$ onto $[t_1...t_n]$ (resp. $[s_1...s_n]$) and d is a distance operator (Alt and Godau, 1995). In simple terms, the Fréchet distance measures the maximum distance apart of two coinciding movement paths. The Fréchet distance, is best conceptualized using the analogy of a person walking their dog, where no backwards movement is allowed. In the dog walking example, the Fréchet distance is the minimum length of the dog's leash. The discretized form of the Fréchet distance metric (Eiter and Mannila, 1994) is useful for its computation with movement data collected by discrete fixes, as described in section 2.2. In applications involving objects that move with the same temporal granularity this calculation is simply the maximum distance in space between any

pair of fixes taken at the same time. However, when object movement is recorded at differing temporal granularities or extents, the value of the Fréchet distance metric is through the use of the scaling functions (α, β) to measure similarity.

Vlachos *et al.* (2002) use longest common subsequences (LCSS), a method taken from time-series analysis, to identify similar movement paths. The LCSS is defined as the number of consecutive fixes from two (or more) paths ($M_a, M_b, \ldots$) that are within d spatial and τ temporal units of each other. This method can be extended to paths that move at a distance, using mapping function $f(M)$ to translate M_b onto a space equivalent to M_a. LCSS is advantageous as it is able to address issues relating movement paths taken at different temporal granularities and/or extents. LCSS is efficient even with paths that contain a significant amount of data noise. When outlying fixes are likely to influence the calculation of other similarity indices LCSS is advantageous as it is insensitive to extreme outliers. The disadvantage of the LCSS method is that it relies on the subjective definition of thresholds d and τ, and it fails the triangle inequality test (iv. above), and is therefore not a metric distance function.

Similarity indices have also been extended to objects moving along a network. For example, Hwang *et al.* (2005) calculate similarity using points-of-interest, such as major intersections. Movement paths are considered similar if they pass through the same points-of-interest in the same order. This index is not a metric distance function, but moves away from Euclidean based measurements which are inappropriate in a network scenario.

Recently, a new similarity method has been proposed by Dodge *et al.* (2012). Here, a movement path is separated into segments where specific movement parameter patterns (and derivatives of) are observed. In their example, velocity is the parameter of interest, and the metrics deviation from the mean and sinuosity are used to define movement parameter classes. For example, the letters A–D could be used to denote 4 unique movement parameter classes, and a path could then be represented as the sequence [ACBCACBDBDA]. To assess the similarity of two paths, a modified version of the edit distance (a string matching algorithm) is computed on the movement parameter class sequences. This method measures similarity in the selected movement parameters, rather than in the space-time geometry of the movement paths. As such, it may be more appropriate when similarity in various parameters, rather than space-time geometry is specifically of interest, for instance, in the study of hurricane path dynamics, as demonstrated by Dodge *et al.* (2012).

When objects interactively move with each other at a distance, they often exhibit correlated movement. Typically, similarity indices may identify such correlated movements by mapping the spatial coordinates of one path onto the spatial plane equivalent to the other. Alternatively, Shirabe (2006) presents a method for computing the correlation coefficient between two movement paths, each represented as a vector time-series. Consider a path M with $t = 1 \dots n$ fixes, then for $t = 2 \dots n$, $V = [M_t - M_{t-1}] = [v_t]$, is a vector time series of M. Given two two movement paths (M_v, M_w) represented as vector time-series V and W, the correlation coefficient is defined as:

$$r(V,W) = \frac{\sum_{t=1}^{n-1}(v_t - \bar{v}) \cdot (w_t - \bar{w})}{\sqrt{\sum_{t=1}^{n-1}|v_t - \bar{v}|^2}\sqrt{\sum_{t=1}^{n-1}|w_t - \bar{w}|^2}} \tag{2.4}$$

Where $\bar{v} = \frac{1}{n-1}\sum_{t=1}^{n-1} v_t$ (resp. $\bar{w}$) are mean coordinate vectors of (V, W). Note that a movement path of n fixes is comprised of $n-1$ movement vectors, this distinction we keep for consistency with other methods. The numerator in (2.4) is the covariance, which indicates how the two motions deviate together from their respective means (Shirabe, 2006). Geometrically, the dot product in the numerator is the product of vector lengths multiplied by the cosine of the angle between them, which can be interpreted as the similarity. The correlation index ranges from -1 to 1, identifying both negatively and positively correlated movements. Important to note is that this correlation coefficient relies on each movement's deviation from the respective mean, not the raw values of each observed movement. Relating correlations to a global mean can be advantageous in cases where two movements are correlated, but do not move in the same direction. The first drawback of the formulation in (2.4) is that we are unable to disentangle the effects of correlation in azimuth vs. magnitude of movements. A metric decomposed into each of these components would be advantageous in situations where such distinctions are necessary. A second drawback of equation (2.4) is that it requires that the fixes from each movement path be taken simultaneously in order to be valid, which is not always realistic. However, Shirabe (2006) does present an extension for modifying (2.4) to measure movement path correlations at a temporal lag.

2.3.4 Pattern and Cluster Methods

Many applications are interested in identifying broad spatial-temporal patterns from large movement databases (Benkert *et al.*, 2007; Palma *et al.*, 2008; Verhein and Chawla, 2008). For example, in the study of tourist behavior, often the goal is to identify places of interest that are frequently visited (e.g., Ahas *et al.*, 2007). Alternatively, studying commuter patterns typically involves the identification of intersections and routes being used by multiple individuals (Verhein and Chawla, 2006). In these situations, pattern and cluster methods are employed to identify similar movement behaviors or places of interest.

Early work on indexing and querying movement databases coming from the computer and database science literature (e.g., Güting *et al.*, 2000; Pfoser *et al.*, 2000) has been essential to the development of pattern and cluster methods. For instance, many methods for identifying patterns and clusters in large movement databases implement a simple spatial or temporal query (Erwig *et al.*, 1999). Alternatively, pattern or cluster methods may implement one of the aforementioned path similarity indices and perform pair-wise similarity computations over all permutations of stored movement paths. Paths identified as similar based on a query or similarity index may convey some movement pattern, or belong to the same cluster. The use of the term cluster' comes from methods for statistical analysis of spatial point patterns (Diggle, 2003), as many approaches used in point pattern analysis have been adopted for movement data. For example, both Gao *et al.* (2010) and Güting *et al.* (2010) describe methods for performing k-nearest neighbor queries in movement databases.

For the most part, the identification of patterns and clusters in large movement databases focus on one of space, time, or space-time. Methods that identify spatial clusters look at space first and time second, if at all (e.g., Benkert *et al.*, 2007). The simplest methods for detecting spatial clusters in movement databases generally require that fixes from individual paths be represented as spatial points. Other spatial methods look to define regions of interest (static or dynamic) and identify times at which movement fixes are clustered in these spaces (Giannotti *et al.*, 2007). Alternatively, temporal clusters look at time first and space second, (e.g., D'Auria *et al.*, 2005; Nanni and Pedreschi, 2006). Temporal clustering is enhanced (Palma *et al.*, 2008) when movement paths are represented by a sequence of stops (representing activities) and moves (Spaccapietra *et al.*, 2008).

Space-time approaches to identifying patterns and clusters strive to consider space

and time simultaneously. This is difficult, as previously mentioned, due to scaling differences between space and time. Most space-time approaches fail to properly scale space and time and degenerate to spatial clustering methods linked through time (e.g., Kalnis *et al.*, 2005). Such methods routinely consider the following problem: given p mobile objects, $M_i, i = 1 \ldots p$. Each M_i consists of n fixes taken at coinciding times $t = (1 \ldots n)$. A set of $\alpha (1 \leq \alpha \leq p)$ spatial clusters are identified at each time t (for example with multivariate clustering) using the spatial (x, y) coordinates of $M_i(t)$. In one example, Shoshany *et al.* (2007) link clusters through time using linear programming. In their example, moving objects M_i can switch between clusters, but all M_i must belong to a cluster, as well clusters can emerge or disappear over time. The appeal of this approach is that linear programming, frequently used in optimization research, can identify flows and trends in movement data clusters.

Spatial-temporal association rules (STAR) learning represents an algorithm-based method for discovering spatial-temporal patterns in movement databases (Verhein and Chawla, 2006, 2008). The patterns found by STAR methods are able to identify sources, sinks, and thoroughfares in large mobility databases. Verhein and Chawla (2008) demonstrate a STAR-miner software that implements their algorithm, and apply it to a caribou dataset. STAR patterns rely on pre-determined spatial units (termed regions) over which the algorithm is run. Unfortunately, the use of explicit spatial regions in their derivation means that STAR are especially sensitive to changes in the definition of regions (known as the modifiable areal unit problem - Openshaw, 1984).

Pattern and cluster methods for movement data have also drawn on existing methods from other applications. Shoval and Isaacson (2007) propose sequence alignment methods, originally used to analyze DNA, as a way to identify patterns in human travel behavior. With movement data, sequence alignment methods are able to identify groups of objects that follow a similar sequence of events (e.g., using an event based movement data representation, as in Stewart Hornsby and Cole, 2007). Shoval and Isaacson (2007) apply sequence alignment methods to tourist movement data and conclude that sequence alignment methods have potential for identifying patterns of spatial behavior in large movement databases. In another example, Eagle and Pentland (2009) introduce a method for discovering eigenbehaviors in movement databases. Eigenbehaviors represent trends or routines in individual movement data. Principal component analysis is used to identify the eigenbehaviors of each person in their dataset. In their example using the movements of people's daily routines,

three trends emerge: workday, weekend, and other behaviors. Increasingly complex questions could be addressed using the eigenbehavior method.

2.3.5 Individual-Group Dynamics

The term individual-group dynamics is used to classify a suite of methods that focus on individual object movement within the context of a larger group. This differs fundamentally from methods for identifying patterns and clusters in movement databases. Most current methods for investigating individual-group dynamics rely on computational algorithms capable of searching movement databases for specific, pre-defined patterns. These algorithms are often computationally demanding and inefficient (Gudmundsson *et al.*, 2007), and thus primarily used only in small, case-study examples.

Laube *et al.* (2004, 2005) provide the most comprehensive examination of individual group-dynamics. Their concept of relative motion (REMO) can be used to detect specific patterns (constancy, concurrence, and trend-setters) in groups of moving objects. Constancy represents when an object moves in the same direction for a number of consecutive fixes. An episode of concurrence occurs when multiple moving objects move in the same direction at the same time. Trend-setters are objects that move in a given direction ahead of a concurrence episode by a group of objects. Trend-setting is identified as the most interesting property, and examined in more detail using the sport of soccer as an example. Players that exhibit trend-setting behavior are able to better anticipate the movement of play. Their concept of trend-setting has been further developed for identifying leaders and followers in groups of moving objects, which is potentially useful for the analysis of wildlife movement data (Andersson *et al.*, 2008). Laube *et al.* (2005)'s REMO method uses only movement azimuths to determine relative motion. All other movement attributes, such as speed or distance, are ignored in their derivation. Thus, REMO is useful only in situations where a group of objects move with similar speeds and are contained in a relatable geographic space, such as the soccer example. Another disadvantage is that the REMO method relies on the definition of azimuthal breakpoints to define when objects are moving in a similar direction (e.g., East is between 45° and 135°). Due to their discreteness, these breakpoints can lead to misleading interpretations, for example when objects move in similar directions on either side of a breakpoint. Alternatively, Noyon *et al.* (2007) evaluate the relative movement of objects from the point-of-view of an observer

within the system. Using changes in relative inter-object distance and velocity, Noyon *et al.* (2007) identify relative behavior, for example collision avoidance. Furthermore, Noyon *et al.* (2007) suggest that such relative movement behavior also include other regions-of-interest such as lines and polygons, which they include in their derivation.

Another problem routinely encountered in the study of movement is the detection of flocks and convoys (e.g., groups of individuals that move as a cohesive unit). A flock (see Figure 2.2a) is defined as a group of at least m moving objects (M) contained within a circle of radius r over a minimum time interval - τ (Gudmundsson and van Kreveld, 2006; Benkert *et al.*, 2008). Alternatively, a convoy (see Figure 2.2b) is defined as a group of at least m moving objects (M) that are density connected at a distance d over a minimum time interval - τ (Jeung *et al.*, 2008). Density connected implies that there exists a sequence of segments connecting all points in the convoy, each segment with length $\leq d$. This definition of convoy relaxes the circular requirement of flocks affording flexibility in the shape and extent of convoys that can be identified, for example Canada geese forming their characteristic V-shape. Methods that look at flock/convoy behavior have obvious usefulness in the study of wildlife herds, but also in monitoring crowd dynamics at large events (Benkert *et al.*, 2008). Like space-time clustering, methods describing flocks or convoys build upon Hägerstrand's concept of bundling, identifying areas where objects move coincidentally in space-time. The fundamental difference between the identification of flocks or convoys and space-time cluster methods is that the definition of a flock or convoy explicitly considers the individual in relation to the group in its definition. That is, focus is placed on membership to a given group, with explicit consideration of minimum requirements for flock or convoy behavior (e.g., the parameters m and τ). Space-time cluster methods focus more on identifying broader patterns, typically from large movement databases, and generally rely on pair-wise comparisons of individual movement paths.

Recently, free space diagrams have been proposed for identifying single-file motion in movement databases (Buchin *et al.*, 2010). To conceptualize a free space diagram consider two movement paths (M_a and M_b), over the time intervals m and n respectively, where the trajectory between fixes is given by some linear or other model (e.g., Tremblay *et al.*, 2006). The functions ψ_a and ψ_b give the position of the objects a and b at time t. The free space diagram for a and b (following Buchin *et al.*, 2010) is

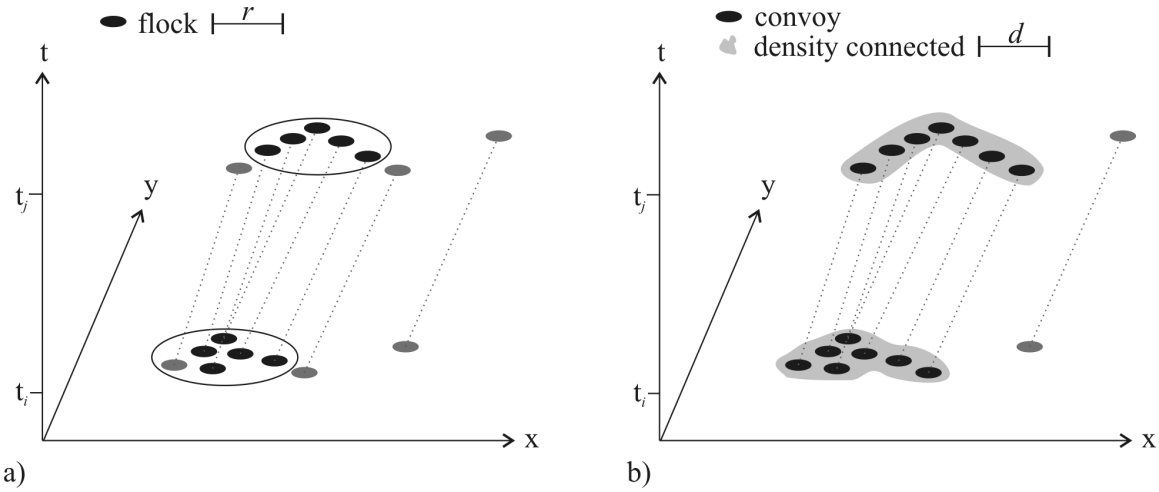

Figure 2.2: Comparison between definitions of a) flocks, and b) convoys. A flock requires objects be contained in a circle of radius – r, while a convoy is defined as those objects that are density connected at distance – d. Both methods require that objects be included in the group over a minimum time interval – τ.

given by:

$$F_\delta(M^a, M^b) = \{(t^a, t^b) \in [1, n] \times [1, m] : |\psi_a(t^a), \psi_b(t^b)| \leq \delta\} \tag{2.5}$$

which defines the set of all points in ψ_a and ψ_b that have a Euclidean distance below some threshold – δ. The map of F_δ describes a two dimensional space where the axes correspond to the two paths, and the free space is defined as anywhere along the paths where the distance between the two paths is below the threshold δ. Buchin *et al.* (2010) demonstrate a method for interpreting free-space diagrams capable of identifying single-file movement patterns in groups of moving objects. A criticism of this method is that it relies on a subjectively defined threshold – δ, to constrain the single-file movement process. Single-file motion has intuitive meaning, but is especially difficult to conceptualize geometrically. Methods that use Euclidean geometry to measure the spatial separation between leaders and followers (e.g., Andersson *et al.*, 2008) are inadequate for identifying single-file movement warranting the free-space diagram approach.

2.3.6 Spatial Field Methods

Often it is of interest to represent a movement path (or many movement paths) as a spatial field in order to identify areas in space (or space-time) that are more or less

frequently visited. Field based representations are especially useful for visualizing large quantities of movement data when maps become cluttered. As many other spatial datasets are stored as raster fields, a field-based representation of movement allows quantitative map comparisons to be performed in a GIS.

Most methods for representing movement data as spatial fields have evolved from those used to analyze spatial point patterns. When spatial point pattern methods are employed the temporal component of movement fixes is ignored. Spatial point pattern methods can be separated into quadrat or density based methods (Diggle, 2003). The simplest quadrat methods involve subdividing a study area into a regular grid and determining point densities within each cell (e.g., Dykes and Mountain, 2003; Hadjieleftheriou *et al.*, 2003). Cells with high point densities indicate spatial locations of high use. Hengl *et al.* (2008) proposes a quadrat based space-time density measure based on distance and velocity within each cell (2.6).

$$D_{xyt}(j) = \frac{\hat{d}_j}{\hat{v}_j} \tag{2.6}$$

Here $D_{xyt}(j)$ is the space-time density at cell j, $\hat{d}_j$ is the length of the movement path within cell j, and $\hat{v}_j$ is the average velocity of movement within cell j. For a single moving object the space-time density is simply interpreted as the duration of time the object spends within each cell. If calculated for a movement database of many objects, areas with higher space-time densities represent those where more objects spend more time, the opposite with low values (Hengl *et al.*, 2008). This approach has been extended for three-dimensional visualization, where density is related to the lengths of multiple paths in 3-D voxels defined by two spatial dimensions and a temporal dimension (Demšar and Virrantaus, 2010). Voxel densities are visualized in a space-time cube (aquarium), and can be used for exploratory analysis of large movement databases.

Density based methods in spatial point pattern analysis stem from bivariate probability models, where movement fixes represent sampled locations from a two-dimensional probability density function (Silverman, 1986). In the analysis of wildlife, density based models are frequently used to generate estimates of animal space use (also discussed in S3.7). Worton (1989) first applied kernel density estimation (KDE) to wildlife movement data to derive such a surface, termed a utilization distribution (Jennrich and Turner, 1969). In movement applications, KDE can be interpreted

as the intensity of space use based upon a collection of fixes. Calculation of KDE requires selection of a kernel shape and bandwidth parameter, with no consensus on the best way to do so (Hemson *et al.*, 2005; Kie *et al.*, 2010). Alternatively, Downs (2010) has proposed time geography's potential path area (see Figure 1) to replace the kernel shape and bandwidth parameter, representing a novel approach for integrating temporal constraints into KDE analysis. Downs (2010) replaces the traditional kernel function with one based on the potential path area (termed geo-ellipse – G) from time geography (2.7).

$$\hat{f}_t(x) = \frac{1}{(n-1)((t_E - t_s)v)^2} \tag{2.7}$$

The numerator in this function sums the distance between a given point x and the object's locations (M) at times i and j. The denominator is the maximum distance the object could have travelled in that time interval given its maximum velocity v. Others have seen the need to move away from continuous representations of space, and have developed KDE for networks (Borruso, 2008; Okabe *et al.*, 2009). Such analysis is more appropriate for depicting the movement of urban travelers as their movement is restricted to travel networks of roads, paths, and sidewalks.

Random walks and diffusion theory have also been used to model movement as a continuous spatial field. Horne *et al.* (2007) use Brownian bridges to model wildlife movement as a continuous probability surface. Between two consecutive mobility points the probability density that an object is at a given location at time t is defined using a bivariate normal probability density function. More recently, probabilistic time geography has been proposed (Winter, 2009), where a similar probability surface is based on discrete random walks in a cellular automata environment. Winter and Yin (2010) extend on the ideas of Winter (2009) to include directed movements. Random walks are used to derive a probability surface which explicitly considers the time geographic constraints on object movement, using a similarly defined bivariate normal probability surface. Both Winter and Yin (2010) and Horne *et al.* (2007) discuss the fact that determining movement probabilities based on random walks is limited when objects do not move randomly. Future work looking at probabilistic movement using other movement models (e.g., correlated random walks or on a network) is thus warranted for moving objects that can be modeled this way. Alternatively, Miller and Bridwell (2009) propose a field-based time geography. Field-based time geography uses movement cost surfaces in the calculation of time geography volumes. Movement possibilities are evaluated in a similar manner to Winter and Yin

(2010) but based on an underlying movement cost surface (e.g., as in least-cost path analysis in GIS, Douglas, 1994). This approach is advantageous in that it directly considers underlying variables impacting movement, however is limited in that an accurate cost surface must be derived.

Brillinger (2001); Brillinger *et al.* (2004) provide a unique approach for discovering patterns in movement data. Stochastic differential equations are used to model movement as a Markov process. The drift term in the stochastic movement model can be interpreted as a spatial velocity field and used for exploratory analysis. The spatial velocity field represents a potential function, whereby points of attraction and repulsion can be identified. Methods for statistical inference (e.g., jackknifing) can be used to identify statistically significant movement patterns within this velocity field (Brillinger *et al.*, 2002). Brillinger (2007) further applies this approach for analyzing the flow of play in soccer, where the spatial velocity field for ball movement is used to investigate a team's attack formation.

2.3.7 Spatial Range Methods

Spatial range can be broadly defined as the area (generally represented as a polygon) containing an object's movement. Measures of spatial range can be useful for examining object mobility and space use. Aspatial metrics, such as net displacement (Turchin, 1998), provide no information on the spatial distribution of movement, simply measuring distance, thus spatial measurements are warranted. Furthermore, researchers are often interested in intersections and/or differences in movement ranges (e.g., Righton and Mills, 2006). In such cases it is advantageous to represent point/line movement data in an areal format (e.g., as a polygon).

The practice of representing movement data using spatial polygons has been developed primarily by wildlife ecologists for studying wildlife home ranges (Burt, 1943), however, the concept of home range has also been applied to other subjects (e.g., children, Andrews, 1973). Spatial range methods typically rely on the geometric properties of movement data, for example the calculation of the minimum convex polygon, a common measure of wildlife home range (Laver and Kelly, 2008). Other geometric methods include harmonic mean (Dixon and Chapman, 1980), Voronoi polygons (Casaer *et al.*, 1999), and characteristic hull (Downs and Horner, 2009). It is also common to extract spatial range polygons from spatial field representations of movement (e.g., those from S3.6) by extracting polygon contours based on density.

For example, with KDE a 95% volume contour is frequently used to delineate wildlife home range, while a 50% volume contour is used to delineate core habitat areas (Laver and Kelly, 2008). These spatial range methods ignore temporal information stored in movement data and are likely to contain areas never visited by the object (commission error), and miss actually visited locations (omission error) (Sanderson, 1966).

Time geography volumes may also be used for generating spatial range estimates. Long and Nelson (2012) propose a spatial range method for wildlife movement data based on time geography's potential path area (Figure 2.1c). This method is capable of identifying omission and commission errors in other spatial range methods (Long and Nelson, 2012). Such time geographic analysis is commonly used to study accessibility in the context of human movement (Kwan, 1998). The value of the potential path area as a spatial range method is that it explicitly considers the temporal sequencing of movement data in a time geography context. Spatial range methods that consider the temporal component of movement data are advantageous over purely spatial methods (such as convex polygons) as they consider movement data as a sequence of spatial points taken through time, rather than as an arbitrary collection of spatial points.

2.4 Discussion

2.4.1 Time

The first and foremost challenge to the quantitative analysis of movement data is how to effectively characterize time. Despite having well-developed theory and tools for analyzing space, geographers and the GISci community have historically struggled with the temporal dimension (Peuquet, 1994). Time is a single, continuous dimension that can be portrayed as either monotonically linear or cyclical (Frank, 1998). If time is portrayed as linear, objects are not capable of re-visiting instances in time. If time is portrayed as cyclical, the beginning of a new cycle implies that time is reset to some initial state, thus revisiting is facilitated. For example, consider research on human daily routines; within each day time is treated linearly, but is reset at the beginning of each day signifying the start of a new cycle. Movement data collected over long periods may contain both linear and cyclical temporal patterns, confounding representation and analysis.

Theoretical constructs for including time in GIS have long been discussed (Langran

and Chrisman, 1988; Peuquet, 1994) but remain challenging. Some spatial datasets are easily represented at discrete time intervals in a GIS as different layers, for example land cover data in different years. This representation allows for vertical analysis through time using relatively simple map algebra (Mennis *et al.*, 2005). Vertical analysis through time is not straightforward with movement data, as objects move in both space and time and cannot be explicitly linked through the spatial dimension. Others have suggested the notion that geography's fetish for the static (Raper, 2002) may lie at the root of the time problem. In practice, researchers have begun to use a 3-D aquarium (drawing on Hägerstrand's ideas) for representing time in GIS, however this is principally a visualization tool (e.g., Kraak, 2003; Andrienko *et al.*, 2007; Shaw *et al.*, 2008). Dynamic views (i.e., animations) may overcome the drawbacks of static portrayals of movement, allowing more fluid representations of velocity and acceleration properties (Andrienko *et al.*, 2005). However, dynamic views are also visual-based, and lack potential for developing quantitative analyses.

The challenge has been finding appropriate ways to simultaneously represent the different scales of measurement for temporal and spatial attributes associated with movement. Consider that it is common to use measurements of time and space interchangeably in queries associated with movement from everyday life, for example if you were asked the question: how far is it from here to the grocery store? You might answer with "about 2 kilometers" or alternatively with "about a 5 minute drive". Here, a question of spatial distance associated with movement can be equivalently answered using a spatial measurement (2 km) or temporal measurement (5 minutes). This has led to alternative conceptualizations of movement where space and time can be represented using relationships that can scale from spatial to temporal measurements, and vice-versa (Parkes and Thrift, 1975). For example, travel can be considered as the consumption of physical distance through time (Forer, 1998). However in the previous scenario, you may have also answered with "about a 5 minute drive, depending on traffic". Alternatively, one might add that it depends on mode of transport (e.g., whether you walk or drive). This alternative view demonstrates the non-linear and dynamic relationship that exists between space and time which confounds the direct exchange of measurements of space and time (Forer, 1998). With movement data, time is often stored alongside spatial attributes (e.g., $[x, y, t]$), which naturally lends itself to Euclidean-type measurements in the space-time aquarium. However, as demonstrated, time is poorly represented by such direct physical measurements, because time cannot be represented as a linear function of space. As there is still

no consensus on the best way to represent time with movement data, research on how to effectively characterize space and time in movement data continues to require development.

Distance in space is easily computed using Euclidean (or other, such as network) measurements. Differences in time are generally measured using clock times. The conceptualization of a single space-time proximity measure remains one of the biggest hurdles with quantitative analysis of movement data. Moving forward it is imperative to go beyond simple Euclidean based measures, as time and space do not operate on equal scales (Peuquet, 2002). The Fréchet distance (Alt and Godau, 1995) is an example of a novel method for comparing the similarity of two movement paths that may prove useful in future analyses. Nearest neighbor computations (e.g., Gao *et al.*, 2010), most useful with movement data stored as points, may also provide avenues for exploration. Normalizing different data scales, common to other branches of quantitative analysis such as multivariate cluster analysis (Duda *et al.*, 2001), may be useful for comparing movement processes across scales and relates to work using fractals for describing movement datasets (Dicke and Burrough, 1988). Normalization, however, may mask scale specific patterns, and should be done with caution only when scale specific behavior is less-important. Fundamentally, space and time have different dimensions and require special consideration when analyzed together.

2.4.2 Scale

With any spatial analysis the selection of analysis level (scale) will influence the outcome of quantitative measures and the resulting inferences and conclusions (Dungan *et al.*, 2002). The study of scale and its impacts in spatial analysis remains a key topic in geographic studies. In the analysis of movement data Laube *et al.* (2007) identify four levels of analysis: instantaneous, interval, episodal, and global (Figure 2.3). The instantaneous ("local") level represents measures computed at any point along a movement path. Interval ("focal") level analysis takes the form of a moving temporal window, but may also use a moving spatial window. Episodal ("zonal") level analysis looks at specific partitions of movement data, often related to some known event. Most common is global level analysis, where a movement dataset is represented as a complete path, from beginning to end, as a single entity. While some methods are specifically designed for a given level of analysis others can be applied to various levels. Methods that can be applied at different analysis levels may

not scale from one level to the next, meaning results at a lower level may not sum to the global result, as is the case with some spatially local statistics (termed LISA – Anselin, 1995).

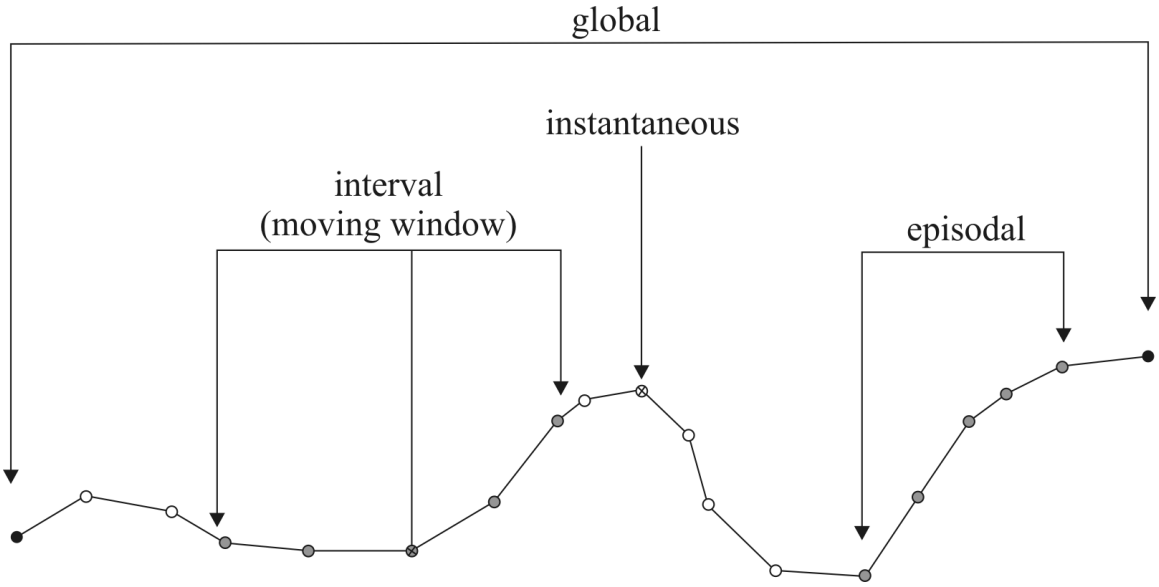

Figure 2.3: Four analysis levels for movement data: instantaneous, interval, episodal, and global. After Figure 2 from Laube *et al.* (2007).

Quantitative methods are also sensitive to changes in the temporal granularity at which movement data is represented (Laube and Purves, 2011). Methods for changing granularity can be used when process scale is explicitly known, however this is rarely the case. When movement data are over-sampled (i.e., too fine a granularity) data noise can mask broader-scale process signals. When movement data are under-sampled (i.e., too coarse a granularity) important movement events are missed, leading to incorrect parameter estimates. Some ecologists have suggested that movement data should not be sampled at even time intervals, but rather as a sequence of moves or steps relating to individual behavior (Wiens *et al.*, 1993; Turchin, 1998). This aligns with the view of (Spaccapietra *et al.*, 2008) that human movement data are best represented as a series of stops (representing activities, as in the event-based model of Stewart Hornsby and Cole, 2007) and moves. However, many developed methods tend to perform better when implemented with regularly sampled movement data (e.g., Downs and Horner, 2012). As the toolbox of methods for the quantitative of analysis of movement grows, it will be important to identify at what analysis level(s) and over which temporal granularities various methods are

appropriate.

As previously identified, and following from Laube *et al.* (2007) and Laube and Purves (2011), there are two fundamental issues of scale associated with movement analysis, that is, analysis level and temporal granularity. (Laube and Purves, 2011) suggest a third issue of scale may also exist, in that many approaches for movement analysis are tested only on small, idealized datasets, and do not perform as expected when carried out on larger, real-life datasets. As a result, many existing methods cannot be readily implemented in practical scenarios with large volumes of movement data. We take an alternative view on this issue. Testing of methods with smaller, idealized datasets limits the scope of movement analysis to realistic and manageable problem sets, which are in turn appropriate with subsets of a larger movement database. For example, the detection of trend-setters (Laube *et al.*, 2005) is only useful if there is some expectation about where, if observed, this pattern is meaningful. In applied research, one should be able to identify specific scenarios, within a larger movement database, where a given technique is appropriate. Once these specific scenarios are identified, for example using spatial-temporal queries, apply the technique of interest on this subset of the movement database. The result is a multi-tiered analysis, where a specified technique is only performed on smaller, appropriate subsets of the data. The goal being to break down larger movement datasets into pieces resembling the idealized scenarios upon which various techniques are useful.

2.4.3 Statistical Significance

Often, it is desirable to examine quantitative problems using a statistical lens, that is, to determine if some pattern is different than an expectation. For those less familiar with statistical inference in GISci, we point the reader to the text by O'Sullivan and Unwin (2010), which provides an introduction to these concepts. Spatial statistics often rely on the concept of complete spatial randomness (CSR) as an a priori assumption for assessing the statistical significance of observed spatial patterns (Cressie, 1993). With some types of spatial statistics (e.g., join counts, Cliff and Ord, 1981) the distributions for computing statistical tests are analytically derived. With other statistics, specifically most spatially local measures, simulation procedures are used to generate test distributions, making these statistics primarily exploratory (Boots, 2002).

Random walks have been suggested as being to movement data what CSR is

to spatial data (Winter and Yin, 2010). Two key methodological developments have included random movement in their derivation: Brownian bridge home ranges (Horne *et al.*, 2007) and probabilistic time geography (Winter and Yin, 2010). However, these two examples represent essentially the same problem: defining a probability surface for movement between two known locations in space-time. Authors of both methods concede that random movement is inappropriate for modeling objects that move non-randomly, but contend that it represents a necessary starting point.

The development of space-time statistics for movement is still in its infancy and lacks clear direction for future research. Some have taken alternative views on this problem, for example treating movement data as a bivariate time series using spatial coordinates as dependent variables (e.g. Jonsen *et al.*, 2003). Others have looked at geographic space first, often ignoring the temporal component altogether (e.g., Casaer *et al.*, 1999). Both approaches are limited as they do not consider movement as a dynamic process that is a function of both space and time. To adequately address the process of movement, novel statistical techniques must consider space and time simultaneously in their derivation. This will be challenging however, as inferential statistics are ill-suited to the multidimensional complexity of movement (Holly, 1978).

2.4.4 Emerging Trends in Quantitative Movement Analysis

Technological advances now facilitate real-time capture and analysis of movement data on both wildlife and humans. In wildlife applications, real-time data acquisition is providing opportunities for conservation and wildlife management. Dettki *et al.* (2004) implemented a real-time tracking system for moose in Sweden, where data on moose movements could be used to initiate the start-up and shut-down of forestry operations in seasonal moose ranges. This idea relates directly to recent work identifying the importance of timing in time geographic measures of space-time accessibility (Neutens *et al.*, 2010; Delafontaine *et al.*, 2011a). As the interface between wildlife and humans narrows, other potential applications exist for real-time tracking. Consider a problematic large carnivore (e.g., lion or bear) residing in a national park. Rather than relocating or exterminating this animal, a real-time tracking system could be used to monitor the animal's movements. Park managers could use this information to improve park safety and minimize human-animal conflicts through trail/site closures and surveillance efforts.

Further developments with real-time movement data will involve the creation of in-

creasingly sophisticated models for predicting future movement locations. The space-time cone from time geography (see Figure 1a) provides only the boundary for future movement possibilities (e.g., O'Sullivan *et al.*, 2000), factoring in the uneven distribution of future movement possibilities (e.g., Winter, 2009) provides more useful information for prediction. Future movement possibilities can be linked to contextual factors such as obstacles (Prager, 2007), underlying movement cost surfaces (Miller and Bridwell, 2009), and object kinetics (Kuijpers *et al.*, 2011). Further developments towards probabilistically predicting future movements based on contextual factors will provide researchers and analysts with powerful tools for linking real-time movement data with other data sources.

With human movement data a new field that is gaining momentum focuses on leveraging real-time location data in everyday applications: location based services (Raper *et al.*, 2007). Location based services have developed coincidentally with the availability of location-aware devices (e.g., GPS enabled cell-phones and handheld devices), which are now integral to people's daily routines (Kumar and Stokkeland, 2003). However, given the revealing nature of personal movement data, concerns over the privacy and ownership rights of personal movement information continue to surface (e.g., Dobson and Fisher, 2003). With location based services, the fundamental goal is to tailor individual applications, services, and marketing to a user's real-time location (Raper *et al.*, 2007). For example, methods for predicting future movements based on contextual factors, when applied in a real-time application, could provide increased functionality and improve user experiences with location based services. As methods for analyzing real-time movement data emerge, their development in conjunction with applications from location based services should be conducted in order to facilitate their adoption in this field.

With the development of technologies for acquiring movement data, the ability to capture finely grained movement data has increased substantially. Opportunities exist for investigating properties of movement previously not feasible with coarser grained movement data. For example, investigating velocities, accelerations, and the role of momentum in moving objects is an area of opportunity. Current research is developing methods for incorporating physical kinetics (based on object velocity and acceleration) into the calculation of time geography volumes, such as those from Figure 2.1 (Kuijpers *et al.*, 2011). Another avenue for future work is the development of a probabilistic time geographic framework, such as by Winter (2009), that considers the influence of kinetics into the calculation of future movement probabilities.

Methods for investigating interactions between individuals in groups of moving objects continue to develop, but remain limited in overall scope and sophistication. Laube *et al.* (2005)'s relative motion concept can identify trendsetters, but uses only movement azimuth in its derivation. Others have developed other ways to identify specific types of interactions between moving individuals (e.g., Andersson *et al.*, 2008; Buchin *et al.*, 2010). As our ability to characterize these patterns grows, it may be more useful to investigate methods for quantifying the strength of interactions that occur in movement databases. That is, can we measure how interactive are the movements of two individuals. The work of Shirabe (2006) provides a necessary starting point for this research which could be further investigated in light of this problem. Further, it may be necessary to examine outside factors influencing the levels of interaction between individuals (e.g., barriers and obstacles represented as lines/polygons, Noyon *et al.*, 2007). Subsequently, how to accommodate other data sources into models for measuring individual level interactions in movement data remains an open research problem.

With time geography, Hägerstrand provided a theoretical context for looking at the constraints of object movement. Contemporary geographers continue to expand on time geographic concepts incorporating a range of ideas into time geographic theory (e.g., Winter, 2009; Miller and Bridwell, 2009; Delafontaine *et al.*, 2011b). As discussed by Lenntorp (1999), Hägerstrand's time geography represents a set of conceptual and methodological building blocks for use in analyzing and understanding movement as a process. As the quantitative toolkit for analyzing movement continues to grow and develop, those methods including theory and ideas from time geography in their derivation will have increased value in a broader range of applications.

Other theoretical frameworks have also been successfully implemented in movement research. For example, the idea that movement is motivated by an underlying field (e.g., Brillinger, 2001) suggests that forces of attraction and repulsion may influence movements. Such points of attraction, for example in wildlife, may be used to investigate central place foraging theory (Orians and Pearson, 1979). Markovian models have also been used to demonstrate how movement operates as a diffusion process (e.g., Skellam, 1951). Diffusion, originally used to describe random dispersal of organisms, can also be related to crowd dynamics in humans (Batty *et al.*, 2003). The use of theoretical constructs in quantitative methods, such as the aforementioned examples, demonstrates thoughtful development of ideas that in the end are easier to interpret for both the reader and analyst.

It has been suggested that movement methods must consider the "geography behind trajectories" (Bogorny *et al.*, 2009) in order to understand the geographic processes affecting observed movement patterns. Movement analysis is no longer limited by available data, but rather by the tools required to manage and analyze movement databases in more efficient and sophisticated ways (Miller, 2010). Thus, the continued development of methods capable of integrating increasingly large and complex movement databases with available spatial and temporal layers is warranted. With such analysis, the goal is to identify relationships between movement patterns and underlying spatial and/or temporal variables. Data mining work is beginning to enrich movement data with underlying geographic datasets (Alvares *et al.*, 2007; Bogorny *et al.*, 2009). Quantitative methods for movement data must be further developed to consider underlying geographic variables in order for movement to be understood as a function of the environment. Similarly, novel movement datasets are emerging where attribute data are recorded along with spatial and temporal records (e.g., $[ID, S, T, A]$, where A represents some attribute data). For example, wildlife tracking systems are being equipped with devices, such as cameras (Hunter *et al.*, 2005), that simultaneously record information alongside movement fixes. The inclusion of attributes with movement fixes can be termed marked movement data, comparable to the term marked point pattern in the spatial statistics literature (Cressie, 1993). Inclusion of attributes (numerical or categorical) alongside spatial locations in movement data represents an area of opportunity for advanced analysis in the movement-attribute space, as existing methods are not designed for marked movement data.

2.5 Conclusions

Novel movement datasets are not only becoming readily available they are changing how data on movement processes are captured. Traditionally, movement data have been collected as samples taken at coarse temporal granularities. Coarsely collected movement data represents movement discretely and with considerable uncertainty between sampled points. More recently, movement data are being collected at extremely fine temporal granularities, such as 5 fixes/second with athletes. Finely grained movement data represents a (near) continuous form of movement data which contains minimal uncertainty in space-time location. Not only are existing methods ill-suited for finely grained movement data, but the types of questions being asked must also be revisited to consider that uncertainty between consecutive fixes is neg-

ligible.

Within GIS data formats, there is a clear lack of appropriate structures for handling movement data. Those interested in purely visualizing movement data have circumvented these problems by generating independent platforms for visualizations (Andrienko *et al.*, 2005). However, the development of quantitative methods is still hindered by difficulties representing the temporal domain within GIS. The development of geospatial data formats exclusively for movement data will invigorate future research into quantitative methods for movement.

There is a clear need for novel quantitative methods for extracting information and generating knowledge from ever-expanding movement datasets (Wolfer *et al.*, 2001; Laube *et al.*, 2007). Most existing methods can be classified as data mining algorithms, which are used to identify and categorize trends in movement databases, based on some a priori notion about movement. Emerging problems investigate more complex patterns and relationships contained in movement datasets, such as the identification of flocking behavior (Benkert *et al.*, 2008). Methods that are able to quantify interactions between individuals (Laube *et al.*, 2005), and with environmental variables (Patterson *et al.*, 2009) in movement databases will be increasingly relevant in more sophisticated movement analyses. Movement models capable of quantifying relationships between moving objects and dynamic features in the environment (e.g., traffic conditions) are justified in order to measure the significance of events or changes on object movement.

Chapter 3

Time Geography and Wildlife Home Range Delineation

3.1 Abstract

We introduce a new technique for delineating animal home ranges that is relatively simple and intuitive: the potential path area (PPA) home range. PPA home ranges are based on existing theory from time geography, where an animal's movement is constrained by known locations in space-time (i.e., n telemetry points) and a measure of mobility (e.g., maximum velocity). Using the formulation we provide, PPA home ranges can be easily implemented in a Geographic Information System (GIS). The advantage of the PPA home range is the explicit consideration of temporal limitations on animal movement. In discussion, we identify the PPA home range as a stand-alone measure of animal home range or as a way to augment existing home range techniques. Future developments are highlighted in the context of the usefulness of time geography for wildlife movement analysis. To facilitate the adoption of this technique we provide a tool for implementing this method.

3.2 Introduction

Animal home ranges are used to study many aspects of wildlife ecology including habitat selection (Aebischer *et al.*, 1993), territorial overlap (Righton and Mills, 2006), and movement impacts of offspring status (Smulders, 2009). Home ranges often serve as the primary spatial unit for wildlife research and represent the area to which an

animal confines its normal movement (Burt, 1943). Wildlife telemetry data, typically collected with radio or Global Positioning System (GPS) collars, provide a collection of space-time locations for an animal. Telemetry data are commonly converted to home ranges to identify spatial patterns in animal movement and answer specific research questions.

To derive animal home ranges, wildlife scientists have used existing methods in geometric topology and spatial smoothing to transform a set of telemetry points into a polygon animal home range. The two most common methods for computing animal home ranges are the minimum convex polygon (MCP), and kernel density estimation (KDE; Laver and Kelly, 2008). MCP continues to be used extensively in wildlife movement analysis (Laver and Kelly, 2008) despite considerable drawbacks, such as sensitivity to sampling intensity and outliers, convex assumption, and inclusion of large, unused interior areas (Worton, 1987; Powell, 2000; Börger *et al.*, 2006). The prevalence of MCP is likely owing to its ease of implementation in common Geographic Information System (GIS) platforms and that it requires no input parameters. Kernel density estimation (KDE) has been influential in home range analysis since its introduction by Worton (1989). KDE remains contentious in animal movement analysis owing to issues with selecting an appropriate kernel bandwidth (Hemson *et al.*, 2005; Kie *et al.*, 2010), which can significantly affect results (Worton, 1989). Unfortunately, KDE-based home ranges can be misleading when telemetry points are irregularly shaped (Downs and Horner, 2008) or when animals occupy patchy environments (Mitchell and Powell, 2008). A number of other lesser-used methods also exist (e.g., harmonic mean, Dixon and Chapman 1980; local nearest-neighbor convex hull, Getz and Wilmers 2004; Brownian bridge, Horne *et al.* 2007; characteristic hull, Downs and Horner 2009), but have yet to become widely adopted.

The objective of this article is to demonstrate a new approach for integrating time attributes accompanying telemetry data when calculating animal home ranges. Drawing on concepts from time geography (Hägerstrand, 1970), we develop a new approach for computing animal home ranges that explicitly considers the temporal constraints of animal movement. Time is largely ignored in existing home range techniques, and used primarily for separating data into temporal groups such as seasons (Nielsen *et al.*, 2003). We discuss the value of this method in context of existing home range research, including existing examples moving towards a time-geographic approach.

3.3 Methods

3.3.1 Background: Time Geography

Time determines bounds on an object's movement in space (Parkes and Thrift, 1975). With time geography (Hägerstrand, 1970), these constraints are represented as volumes containing all accessible locations in a 3-dimensional space-time continuum consisting of geographic coordinates x and y and time (t) (frequently termed the space-time cube, Kraak 2003; or space-time aquarium, Kwan and Lee 2004). If both starting and end points are known (as with a collection of telemetry fixes) then the space-time prism represents the set of all accessible locations to the object during that movement segment (Figure 3.1). The projection of the space-time prism onto the geographic plane is termed the potential path area (PPA), and represents all locations accessible to an object given its start and end points and assumed maximum rate of travel (Figure 3.1). An object's maximum traveling velocity affects the extent of these volumes into geographic space.

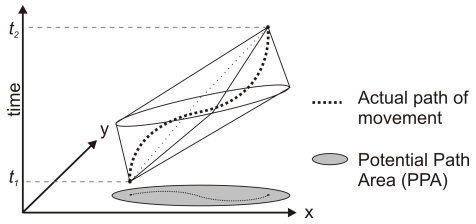

Figure 3.1: Diagram of Hägerstrand's (1970) time geography. The space-time prism contains the set of all locations accessible to an individual given telemetry fixes at t_1 and t_2, and a velocity parameter (v_{max}). The projection of the space-time prism onto the geographical plane is called the potential path area (PPA), used here for delineating wildlife home ranges.

3.3.2 Potential Path Area (PPA): A New Measure of Animal Home Range

We will focus on potential uses of PPA in wildlife movement analysis, specifically the calculation of a PPA animal home range. The PPA represents the set of all accessible locations between 2 known locations in space and time (Miller, 2005). Geometrically, the PPA is an ellipse with focal points located at 2 known locations, the origin and destination. The spatial extent of the PPA depends on the animal's maximum velocity

(v_{max}), which may be explicitly known or empirically estimated from the data.

Visually, conceptualizing the creation of a PPA ellipse is best done using the pins-and-string method (Figure 3.2a). Consider placing pins at the known start (i) and end (j) locations of an animal movement segment. A single string is then tied to each point, connecting the 2 pins. The length of the string is D_{max}, representing the maximum distance the animal can travel given its maximum velocity (v_{max}) and the time difference between points i and j (Δt).

$$D_{max} = v_{max} \times \Delta t \tag{3.1}$$

The PPA ellipse is drawn by moving a pencil around the 2 points, but inside of the string, keeping the string tight at all times. Any point located along or within the PPA ellipse is reachable by the animal during this movement segment.

Mathematically, given that in unconstrained space PPA is an ordinary ellipse, we can derive PPA using parameters of an ellipse related to animal movement in time and space. We define v_{max} and Δt as above, the maximum velocity of the animal and the time difference between known telemetry locations i and j. We define a PPA ellipse using 4 parameters: a center point, a major axis, a minor axis, and a rotation angle (Figure 3.2b). We calculate the center point as the midway point between the spatial (x, y) coordinates of telemetry points i and j. We define the major axis (a) as:

$$\begin{aligned} a &= D_{max} \\ &= v_{max} \times \Delta t \end{aligned} \tag{3.2}$$

With this we can define the minor axis (b) as:

$$b = \sqrt{a^2 - d^2} \tag{3.3}$$

Where d is the Euclidean distance between points i and j. Rotation angle (R_θ) is the angle the ellipse is rotated from the horizontal, and defined using x and y coordinates of telemetry points i and j:

$$R_\theta = \tan^{-1}\left(\frac{y_j - y_i}{x_j - x_i}\right) \tag{3.4}$$

Using these parameters, we can generate the PPA ellipse for any pair of known loca-

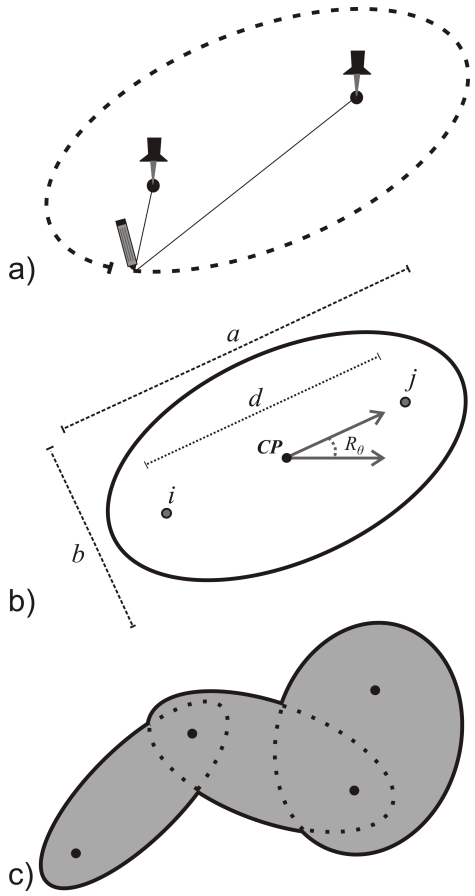

Figure 3.2: a) Pins-and-strings method for generating potential path area (PPA) ellipses. The length of the string is equal to the longest distance the animal could travel (D_{max}) given parameter v_{max} and the time difference between points. b) Geometric properties of a PPA ellipse with telemetry points i and j. CP is the center point and d is the Euclidean distance between points i and j; a and b are lengths of the major and minor axis respectively; and R_θ is the rotation angle. c) Computation of the PPA home range involves combining multiple $(n-1)$ PPA ellipses.

tions in space-time.

A PPA home range can be computed by generating PPA ellipses for a set of animal locations. A telemetry dataset of n recordings requires calculation of $n-1$ PPA ellipses, which are combined to produce the PPA home range (Figure 3.2c). Formally this is defined as the union of $n-1$ PPA ellipses such that:

$$PPA_{HR} = \cup[PPA_{i,i+1}],\ i \text{ in } \{1,\ldots,n-1\} \tag{3.5}$$

The mathematical formulation of this method (represented by equations (3.1) through

(3.5)) is easily implemented in a GIS.

3.3.3 Estimating v_{max}

The PPA home range method requires a single input parameter v_{max} that has obvious biological connotations and in some cases may be explicitly known based on a fine understanding of an organism's mobility. This parameter could be related to an organism's maximum velocity. For example, cheetahs have a maximum speed of up to 120 km/h (Sharp, 1997); however it is unreasonable to expect a cheetah to maintain that speed over longer intervals, characteristic of telemetry datasets. It is more useful to compute the maximum distance a cheetah could cover in 30 minutes and derive v_{max} from this. In practice, v_{max} should relate biologically to the temporal frequency of recordings.

In many cases however, a biologically reasonable estimate of v_{max} will not be explicitly known and a researcher will be required to estimate it from the data. For each pair of consecutive relocation fixes we can compute the segment velocity (v_i) by:

$$v_i = \frac{d_i}{t_i} \tag{3.6}$$

where d_i is the distance and t_i the time difference between consecutive fixes. Computing v_i for all $n-1$ segments will provide a distribution of v values which can be used to generate estimates for v_{max}. The simplest would be to take $\max(v_i)$ – the maximum observed velocity as v_{max}; however, this is problematic as it produces a straight-line (degenerative ellipse) between any consecutive pair of fixes that have this maximum value. A more robust approach is to estimate a value for v_{max} based on the ordered distribution of the v_i. Following Robson and Whitlock (1964) an estimate of v_{max} could take the form:

$$\hat{v}_{max} = v_m + (v_m - v_{m-1}) \tag{3.7}$$

where v_i are in ascending order such that $v_1 < v_2 < \ldots < v_{m-1} < v_m$ and $m = n - 1$. This estimate for v_{max} has an approximate $100(1-\alpha)\%$ upper confidence limit given by:

$$U_{Lim}(v_{max}) = v_m + \frac{(1-\alpha)(v_m - v_{m-1})}{\alpha} \tag{3.8}$$

Cooke (1979) and van der Watt (1980) have extended the work of Robson and Whitlock (1964) deriving estimates with lower mean squared errors and smaller confidence

intervals, at the cost of added complexity. In the case where $v_m = v_{m-1}$, the result from (3.7) will equal $\max(v_i)$ and cause degenerate ellipses to be produced for pairs of consecutive points that have this maximum value. The method of van der Watt (1980) is advantageous as it avoids the problem of degenerate ellipses through careful selection of the parameter k in the equation:

$$\hat{v}_{max} = \left(\frac{k+2}{k+1}\right) v_m - \left(\frac{1}{k+1}\right) v_{m-k} \tag{3.9}$$

where $1 < k < m$ representing the k^{th} ordered value of v_i. This estimate for v_{max} has an approximate $100(1-\alpha)\%$ upper confidence limit given by:

$$U_{Lim}(v_{max}) = v_m + \left(\frac{1}{(1-\alpha^{1/k})} - 1\right)^{-1} (v_m - v_{m-k}) \tag{3.10}$$

In the previously stated problem scenario where $v_m = v_{m-1}$ it would be useful to take k to be the largest value such that $v_{m-k} < v_m$. In general (3.9) has been shown to be an improved estimator of v_{max} over (3.7) (van der Watt, 1980); however, it requires that the researcher select an appropriate value for k. Alternatively, a more conservative analysis could use the upper confidence interval limits (e.g., (3.8) or (3.10)) as an estimator for v_{max}.

For demonstration, we simulate an animal trajectory using a correlated random walk ($n = 2000$). Using these data as a surrogate for animal movement data, we calculate animal home range using two common, existing techniques (MCP and KDE) and the new PPA home range approach. We used the (Robson and Whitlock, 1964) method given by (3.7) for estimating the v_{max} parameter from the data. The temporal sampling interval of telemetry fixes is known to influence output home range size and shape using MCP (Börger *et al.*, 2006) and KDE (Downs and Horner, 2008), but also will influence the PPA home range. To demonstrate this effect, we re-sampled our simulated animal trajectory using only 25% ($n = 500$) of the points and re-estimated the v_{max} parameter using (3.7).

3.4 Results

Estimated values for the v_{max} parameter were 10.6 for the original correlated random walk dataset ($n = 2000$), and 6.8 for the re-sampled dataset ($n = 500$). Maps of

output home ranges using the original correlated random walk (Figure 3.3 a-c) and re-sampled dataset (Fig. 3.3 d-f) are given to view differences and similarities between MCP, KDE, and PPA. Areas of output home ranges for the original correlated random walk are: MCP – 53,739, KDE – 25,045, PPA – 37,232. For the re-sampled dataset home range areas are: MCP – 52,391, KDE – 29,110, PPA – 60,658. Coarsening of the sampling frequency (through re-sampling) caused MCP area to decrease, while both KDE and PPA increased in area.

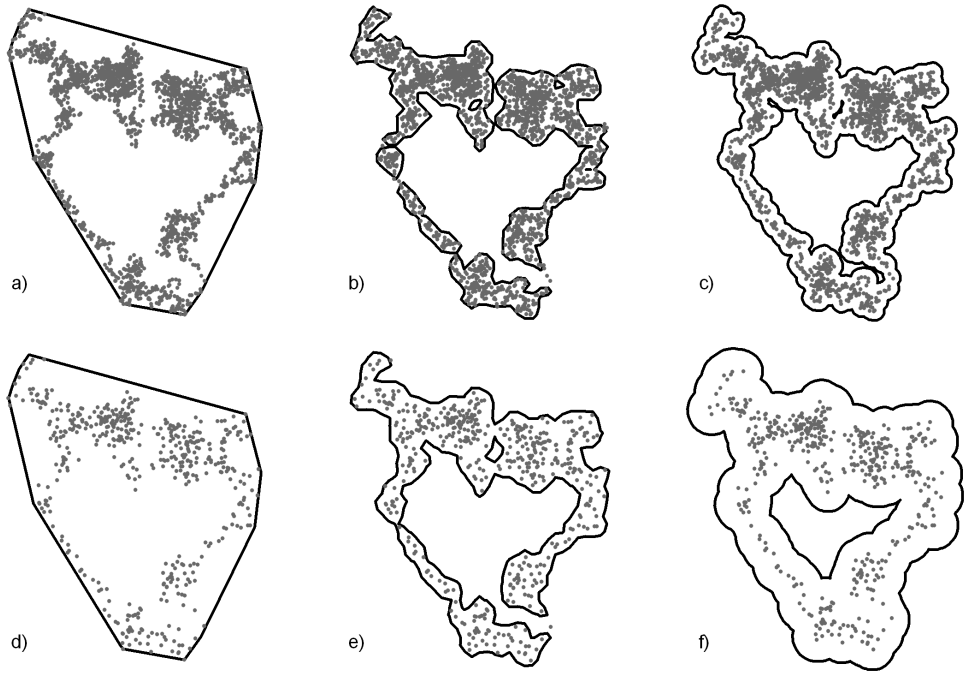

Figure 3.3: Home range polygons for a simulated dataset with $n = 2000$ (top) re-sampled to $n = 500$ (bottom) using minimum convex polygon (MCP; a, d), kernel density estimation (KDE; b, e), and potential path area (PPA; c, f).

3.5 Discussion

In this example, the effect of changing sampling frequency had minimal effect on home range computed using MCP (Figure 3.3 a, d), however this will not always be the case (Börger *et al.*, 2006). With KDE, fewer points lead to increased uncertainty in the bandwidth selection process, resulting in a wider bandwidth selection, and in general a larger output home range. With the PPA home range method uncertainty is a function of the time between consecutive known locations, rather than the number of

points. As a result, PPA home ranges are comprised of fewer, larger ellipses to account for uncertainty in animal location between consecutive known points, and produce larger home range estimates. We suggest that PPA home ranges be employed only when telemetry data are collected using a relatively short sampling interval (e.g., dense GPS telemetry data). In these situations, uncertainty between consecutive fixes will be relatively low. In cases where the temporal duration between fixes is substantially longer (e.g., with most VHF collars), the ellipses produced by the PPA algorithm will be large, resulting in significant overestimations of home range size. We withhold from specifying an absolute threshold on sparse telemetry data where the PPA method should not be used as it will be dependant on both the species (e.g., large vs. small mammal) and application (seasonal home range vs. migratory behavior). Comparison of the PPA home range with existing methods (e.g., KDE and MCP) should provide information as to whether or not the PPA approach is appropriate with a given dataset (see Figure 3.4 and the accompanying discussion below).

The conceptual and computational simplicity of the PPA home range may be its greatest asset. We define the PPA home range simply as: given a set of sampled locations (telemetry points) the PPA home range contains all locations in geographic space that the animal could have visited. The PPA method can be easily implemented in a GIS and requires only 1 input parameter, maximum travelling velocity – v_{max}, that can be derived using biological knowledge or estimated directly from the data (e.g., using (3.7) or (3.9)). If telemetry data are categorized into distinct behavioral segments (e.g., Jonsen *et al.*, 2005; Gurarie *et al.*, 2009) where differing v_{max} would be expected, PPA home range analysis could be further enhanced.

Despite the intuitive structure of ideas from time geography, they are largely absent from wildlife movement research. In one example, Baer and Butler (2000) use time geographic theory for modeling wildlife movement building upon Hägerstrand's (1970) concept of bundling, representing animals congregating in space-time. Regions where bundling occurs can be used to identify specific ecological activity in groups of animals (e.g., locating scarce resources). Wentz *et al.* (2003) implement time geographic constraints for animal movement, interpolating between sampled telemetry locations to model movement paths. Time geography volumes are used by Wentz *et al.* (2003) to constrain random walks between sampled locations. More recently, Downs (2010) presents a novel approach for incorporating time geographic principles, specifically the potential path area (termed geo-ellipse), into kernel density estima-

tion. Downs (2010) uses the geo-ellipse in place of a circular kernel in the density estimation. Several advantages of this approach are identified, such as replacing subjective selection of kernel bandwidth by an objective parameter maximum travelling velocity. Time geographic kernel density estimation assigns zero density to regions outside of the PPA home range, creating a utilization distribution density allocated only to accessible regions.

Wildlife do not use the space within their home range evenly motivating use of an intensity surface (utilization distribution) to analyze animal space use (Jennrich and Turner, 1969). Utilization distributions more adequately portray patterns of space use within wildlife home ranges and provide more reliable estimates of overlap and/or fidelity compared with discrete home range methods (Fieberg and Kochanny, 2005). However, these advantages come at the cost of added complexity in deriving the utilization distribution with many researchers continuing to use discrete measures of home range over utilization distributions in analysis due to their simplicity (Laver and Kelly, 2008). Kernel density estimation remains the most popular method for computing utilization distributions despite considerable drawbacks with newer (temporally dense) telemetry data (Hemson *et al.*, 2005; Kie *et al.*, 2010). Horne *et al.* (2007) propose the Brownian bridge approach for computing the utilization distribution. A Brownian bridge simply defines the probability density a random walk passes through a location given the known start and end points. Like the PPA home range, with the Brownian bridge approach telemetry data are analyzed using pairs of consecutive telemetry fixes. This method relies on a variance parameter (σ_m) that is difficult to interpret but can be estimated from the data using an optimization algorithm. The PPA method is essentially the discrete equivalent of the Brownian bridge approach, but with simple, intuitive, and easy to estimate parameters that can be straightforwardly computed in a GIS. Getz and Wilmers (2004) propose the use of overlapping local convex hulls to generate a utilization distribution. A similar approach could be adopted with PPA ellipses to generate a utilization distribution based on the areas under overlapping ellipses. The derivation of an overlap-based utilization distribution for PPA ellipses remains an area for future investigation.

Wildlife researchers now routinely collect temporally dense telemetry data using sophisticated tracking technologies (e.g., GPS; Tomkiewicz *et al.*, 2010). Such temporally dense telemetry data provide a more detailed and informative view of animal movement. Given continued advancements in technology in the future it is likely that we will be analyzing (near) continuous animal trajectories. This improved representa-

tion of animal movement necessarily results in highly autocorrelated movement data. Much attention has been given to the problems autocorrelated telemetry data pose with traditional methods for studying wildlife movement (Swihart and Slade, 1985; Otis and White, 1999; Fieberg *et al.*, 2010). Many existing methods, developed for use with temporally sparse telemetry data, are ill equipped for dense telemetry data. The PPA home range method is advantageous with temporally-dense telemetry data, as it is capable of including rich temporal information into the derivation of home range. With few exceptions (e.g., Horne *et al.*, 2007) existing home range techniques ignore rich temporal information contained in telemetry datasets. Including temporal information in analysis is beneficial as points are no longer considered independent observations, but rather as a sequence of recordings taken over time.

Certain land cover types (e.g., dense forest; Rempel *et al.*, 1995) can interfere with locating technologies resulting in missing recordings. Missing data points are problematic in subsequent analysis as bias towards specific cover types can occur (Frair *et al.*, 2004). By explicitly considering the temporal sequencing of points, PPA home ranges adjust for missing telemetry recordings by way of a larger Δt value in these areas, providing an unbiased estimator of home range.

Commission errors (locations included in the home range but never visited) and omission errors (locations visited but not included in the home range) are important properties of output home range polygons that require careful consideration (Sanderson, 1966). All home range methods short of a direct trace of an animal's movement path will include commission errors. Omission errors occur with most methods, but can be avoided by substantially overestimating home range size. This is equivalent to selecting an overly large bandwidth with KDE. Substantial overestimation limits utility for wildlife research as the signature of animal behavior is masked. The PPA home range method can be used in tandem with other methods to examine commission and omission errors. Consider a simple comparison, by intersecting the PPA home range with commonly employed home range techniques MCP and KDE (Figure 3.4). The PPA home range represents the largest spatial unit such that no omission error occurs, due to explicit consideration of the time geography constraints on animal movement. Potential omission errors are then easily represented as those areas included in the PPA home range, but not in other techniques. Areas not included in the PPA home range but included in other methods can be considered inaccessible regions and an unnecessary source of commission error. With MCP, potential omission errors are likely to occur near edges of MCP home ranges. Due to the con-

vex assumption, MCP home ranges almost always include inaccessible areas as well (Powell, 2000). KDE home range polygons are not guaranteed to include all sampled telemetry points, therefore explicitly known errors of omission may exist.

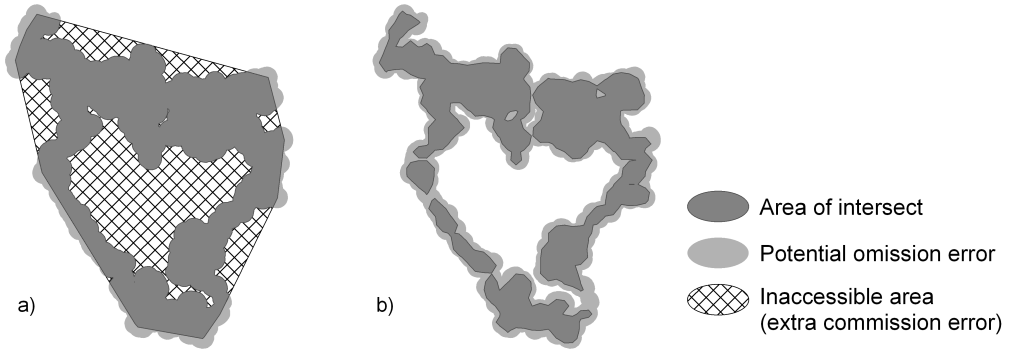

Figure 3.4: Intersections between a) minimum convex polygon (MCP) and potential path area (PPA) and b) kernel density estimation (KDE) and PPA (for $n = 2000$); demonstrating how PPA home ranges can be used to augment existing techniques by identifying omission errors and inaccessible areas.

All measures of home range are indirect and based on specific properties of the telemetry data from which they are derived. Most existing methods use only the spatial properties of telemetry data represented as points. The PPA method provides a complementary view that not only considers spatial information but also temporal information. Using the demonstrated intersection technique, omission errors and inaccessible regions (unnecessary commission error) using existing home range methods can be mapped and quantified. This represents a significant contribution towards home range analysis that carefully considers these types of errors as has been previously suggested (Sanderson, 1966). Often studies employ multiple methods when delineating wildlife home ranges to evaluate a range of possibilities (e.g., Righton and Mills, 2006). The PPA home range should be included in such studies as it can be used to augment other techniques by providing information on omission and commission errors.

In this derivation of PPA home range, all geographical space is considered equally navigable. In reality, environmental factors (e.g., topography, land cover, water bodies) influence an animal's ability to traverse the landscape. As well, external factors such as inter- and intra-species competition (Schwartz *et al.*, 2010), and habitat requirements (Sawyer *et al.*, 2007), motivate wildlife movement, and subsequent home range delineations. Optimally, PPA home ranges would be based on the time ge-

ography constraints across an unequal surface (see Miller and Bridwell, 2009), that considers competition, habitat, topography, and barriers to wildlife movement. Future work should investigate combining available environmental datasets into animal specific movement cost surfaces. Movement cost surfaces could then be integrated into time geographic analysis to compute more realistic PPA home ranges. However, incorporating movement cost surfaces may take away from the attractiveness of time geography methods due to added complexity.

3.6 Management Implications

We have presented a new technique for deriving animal home ranges that is simple and intuitive, but also designed specifically for use with emerging temporally-dense telemetry datasets, such as those now routinely collected with GPS collars. However, we suggest the PPA approach not be adopted with temporally-coarser telemetry data (e.g., VHF collars) as it can lead to overestimation of home range size and misleading interpretations. The PPA home range can be used as a stand-alone measure of animal home range, or to augment existing techniques by identifying potential omission errors and inaccessible areas making it flexible for use with both novel and existing analyses. When performing PPA home range analysis the method for obtaining the v_{max} parameter (e.g., through biological reasoning or by one of the estimation approaches we provide) along with the parameter value should be explicitly stated, as it will influence the resulting home range area. To those wishing to implement the PPA home range technique in their own research we have provided access to a tool for implementing the PPA home range. For more information please go to: `http://www.geog.uvic.ca/spar/tools.html`.

Chapter 4

Towards a Kinetic-Based Probabilistic Time Geography

Abstract

Time geography represents a powerful framework for quantitative analysis of individual movement. Time geography effectively delineates the space-time boundaries of possible individual movement by characterizing movement constraints. The goal of this paper is to synchronize two new ideas, probabilistic time geography and kinetic-based time geography, to develop a more realistic set of movement constraints that consider movement probabilities related to object kinetics. Using random-walk theory, the existing probabilistic time geography model characterizes movement probabilities for the space-time cone using a normal distribution. The normal distribution has a symmetric probability density function and is an appropriate model in the absence of skewness which we relate to object kinetics. Moving away from a symmetric distribution for movement probabilities, we propose the use of the skew-normal distribution to model kinetic-based movement probabilities, where the degree and direction of skewness is related to movement direction and speed. Following a description of our model, we use a set of case-studies to demonstrate the skew-normal model: a random walk, a correlated random walk, wildlife data, cyclist data, and athlete movement data. Our results show that for objects characterized by random movement behavior the existing model performs well, but for object movement with kinetic properties (e.g., athletes), the kinetic-based model provides a substantial improvement. Future work will look to extend the kinetic-based probabilistic framework to the space-time

prism.

4.1 Introduction

Over the past decade there has been rekindled interest in using ideas from Hägerstrand's 1970 time geography (Figure 4.1) in quantitative geographic analysis (Lenntorp, 1999; Miller, 2003). This resurgence is largely due to availability of movement data, obtained using various methods for tracking individual level movements. Concepts from time geography are now routinely used as an analytic framework for quantitative movement analysis (Lenntorp, 1999). Supported by recent developments presenting rigorous mathematical definitions for time geography (Miller, 2005), increasingly sophisticated quantitative analyses of movement data are emerging. For example, Delafontaine *et al.* (2011b) have introduced algorithms for incorporating physical barriers and obstacles into quantitative time geographic analysis.

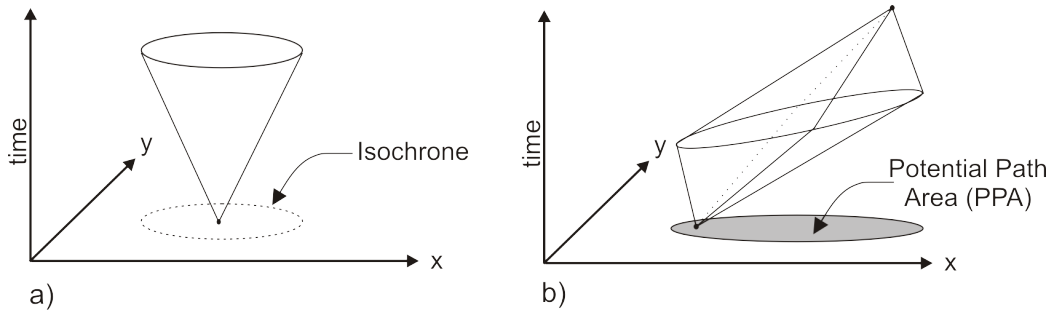

Figure 4.1: Structures originating from Hägerstrand's time geography. a) space-time cone, along with an isochrone – a line of equal movement possibility in the future. b) space-time prism, along with the potential path area – the projection of the prism onto the spatial plane.

Object kinetics, defined by an object's current speed and direction of movement, along with acceleration, can similarly influence movement opportunities defined by time geography (Kuijpers *et al.*, 2011). For instance, in classical time geography, movement boundaries are calculated with the unrealistic expectation that an object can make instantaneous changes in velocity. With object kinetics (and other physical constraints) ignored, time geographic structures (i.e., the space-time cone and space-time prism) substantially overestimate movement opportunities. Kuijpers *et al.* (2011) have quantified the influence of object kinetics (from velocity and acceleration) on time geographic boundaries, termed kinetic-based time geography. Consid-

eration of object kinetics provides a more realistic representation of time geography's boundaries, as kinetic-based time geography will exclude locations in space-time not accessible based on an individuals kinetic movement abilities.

As a quantitative framework, time geography (and kinetic-based time geography) is used to characterize the space-time boundaries on object movement, delineating locations in space and time as either accessible or not. Such a binary definition (i.e., accessible, not accessible) of time geography does not account for unequal movement probabilities within time geographic structures (e.g., those in Figure 4.1). Unequal movement probabilities are a result of locations and paths that are more likely to be visited than others, for instance due to shorter, more direct movement routes.

Several approaches have been proposed to model movement probabilities within time geographic volumes (Miller and Bridwell, 2009; Winter, 2009; Downs, 2010), determining, for instance, the probability an object will be found at a given location in space and time. Such a model for modeling movement probabilities is termed probabilistic time geography, which quantifies variation in movement probabilities in time geography (Winter and Yin, 2010, 2011). With the current probabilistic models, calculations typically assume random movement (i.e., random walks), resulting in the use of a bivariate normal distribution for modeling potential movements in space. A random movement assumption has been used extensively in wildlife movement models, especially with coarser tracking intervals (Turchin, 1998; Codling *et al.*, 2008). Assuming random movement is a limitation, as most objects move non-randomly with directed, linear movements and often revisit specific locations with regularity (Gonzalez *et al.*, 2008).

Kuijpers *et al.* (2011) identify several lingering questions in terms of kinetic-based time geography, the first of which is quantifying unequal movement probabilities in kinetic time geography structures, much like probabilistic time geography. The objective of this research is to develop a model for quantifying movement probabilities for kinetic-based time geography, which we term kinetic-based probabilistic time geography. We generalize the model for probabilistic time geography, proposed by Winter and Yin (2010, 2011), to account for object kinetics. The skew-normal distribution is proposed in place of the normal distribution used in Winter & Yin to model future movement probabilities in the space-time cone building upon previous attempts at factoring object kinetics into movement uncertainty models (Prager and Yu, 2005) and interpolation algorithms (Yu and Kim, 2006).

The chapter is organized as follows. We introduce and develop the skew-normal

model for kinetic-based probabilistic time geography in section 4.2, followed by a short discussion of the model. Section 4.3 outlines a case study, with five different datasets (a random walk, a correlated random walk, wildlife data, cyclist data, and athlete data), used to compare the kinetic-based model against the existing probabilistic time geography model from Winter & Yin. In section 4.4, we discuss case-study results and model limitations, followed by some potential applications of kinetic-based probabilistic time geography. Finally, with section 4.5, we conclude with remarks on the impact of this work along with some areas for future research.

4.2 Kinetic-Based Probabilistic Time Geography

4.2.1 Model Derivation – One Dimension (Linear Movement)

We will first demonstrate the concept using the 1-Dimensional situation (i.e., an object moving along a straight line), where an object at a moment in time (t), located at point x_t, moves with some velocity (v_t). As in traditional time geography, the object has a maximum travelling velocity parameter (v_{max}). The goal of a kinetic-based probabilistic time geography model is to derive future movement probabilities at time $t + \Delta t$ that include consideration of object kinetics, defined as a function of its current velocity v_t (see Figure 4.2).

To be a candidate for kinetic-based probabilistic time geography, the model should satisfy three general characteristics in order to relate to object movement. First, the candidate model should revert back to the normal model proposed by Winter and Yin (2010, 2011) in the absence of kinetic properties. Reducing to the normal model in the absence of initial kinetics seems reasonable, as movement in any direction should be equally probable. Second, the shift in the probability mass should be proportional to the objects current velocity (as demonstrated in Figure 4.2). Here, interpretation of the initial kinetic properties may differ based on application, allowing flexibility in model development. Finally, the mode of the resultant distribution should be identifiable. The mode of the resulting distribution relates clearly to the most probable location of future movement, which can be used as an expectation in more formal analysis and model goodness-of-fit testing.

A candidate model that satisfies the aforementioned properties, is the skew-normal distribution (Azzalini, 1985) which we propose as a generalization of the normal probability density function (pdf) from Winter and Yin (2010, 2011). Thus in order

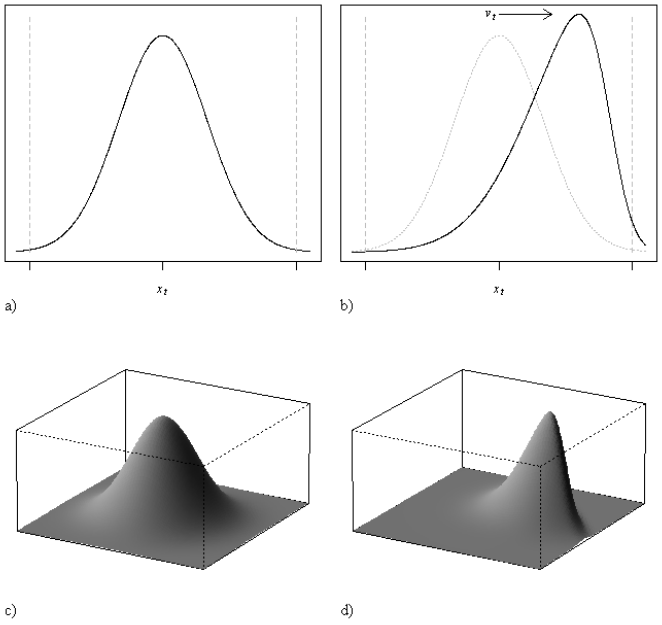

Figure 4.2: a) Probabilistic time-geography for an object moving in a single dimension; b) Incorporating object kinetics (e.g., v_t); c) and d) Extension of a) and b) to two-dimensions: the spatial plane.

to develop the model in the one-dimensional case, we are interested in modeling the movement possibilities of an object in a probabilistic fashion using a univariate skew-normal pdf (denoted SN_1) which takes the following form (Azzalini, 1985):

$$f(x) = \frac{2}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left(\alpha\left(\frac{x-\xi}{\omega}\right)\right) \tag{4.1}$$

Where the functions are the pdf and cumulative distribution function respectively of the standard normal distribution. The SN_1 model requires the selection of three parameters that govern the location ($\xi \in \mathbb{R}$), scale ($\omega \in \mathbb{R}^+$) and shape ($\alpha \in \mathbb{R}$) of the SN_1 pdf. Given its form, (equation (4.1)) can be expressed alternatively as:

$$f(x) = \frac{1}{\omega\pi}e^{-\frac{(x-\xi)^2}{2\omega^2}}\int_{-\infty}^{\alpha\left(\frac{x-\xi}{\omega}\right)} e^{-\frac{t^2}{2}}dt \tag{4.2}$$

Following Azzalini (1985) and Arellano-Valle and Azzalini (2008), an alternative parameterization may be used to represent the pdf in terms of the first three moments of the distribution (i.e., mean – μ, variance – σ^2, and skewness – γ) with respect

to ξ, ω, and α. Using measurable object movement properties, and some existing theory from probabilistic time geography (Winter and Yin, 2010, 2011) we will build a probabilistic model for object movement that considers object kinetics using the SN_1 distribution. To do so we will work with the alternative parameterization (Azzalini, 1985), to relate known movement properties to SN_1 parameters. A system composed of three non-linear equations (in ξ, ω, and α) is used to derive a realistic SN_1 parameterization to probabilistically define object movement possibilities that incorporates kinetics. As will be seen, it is advantageous to investigate the three alternate parameters in reverse order starting first with γ.

The third moment (γ) of a SN_1 can be related to the shape parameter (α) directly by:

$$\gamma = \frac{4-\pi}{2} \frac{\left(\delta\sqrt{2/\pi}\right)^3}{\left(1-2\delta^2/\pi\right)^{3/2}}, \text{ where } \delta = \frac{\alpha}{\sqrt{1+\alpha^2}} \tag{4.3}$$

We wish to restrict γ to $[-1, 1]$ as the maximum theoretical skewness is ~ 1, obtained by setting $\delta = 1$ in (4.3). Further, the goal is to relate to the kinetic properties of the object, which will vary depending on the object type and context (Prager and Yu, 2005). We propose a model where the skewness of the SN_1 (modeled via parameter γ), is calculated from the kinetic properties of the object directly, and is relative to the object's maximum velocity. A simple formulation for γ which satisfies the above conditions is the ratio of v_t to v_{max}.

$$\gamma = -\frac{v_t}{v_{max}} \tag{4.4}$$

The negative sign in (4.4) reflects the fact that if initial velocity is in the positive direction, the direction of the skewness is negative (i.e., if v_t is positive the bulk of the distribution should be in the positive direction). By substituting (4.4) into (4.3) one can solve for the shape parameter α, which will have a unique, real-valued solution.

The second moment (σ^2) can be expressed in terms of the shape parameter (α – which has already been identified) and the scale parameter (ω) by:

$$\sigma^2 = \omega^2\left(1 - \frac{2\delta^2}{\pi}\right), \text{ where } \delta = \frac{\alpha}{\sqrt{1+\alpha^2}} \tag{4.5}$$

We are motivated to use what has already been shown from probabilistic time geogra-

phy (Winter and Yin, 2010, 2011) to relate the variance of the SN_1 to time geography properties. Winter and Yin (2010) suggest that the variance of a normal pdf relates directly to the maximum extent of the space-time cone volume (i.e., $v_{max} \times \Delta t$) through the simple idea that at its maximum extent, the pdf is zero. Following Winter and Yin (2010) we can approximate that the pdf is 0 at 3σ (i.e., by definition 99.7 % of the normal pdf volume is within three standard deviations of the mean). We adopt an identical assumption for use with the SN_1 pdf; that is:

$$3\sigma = v_{max} \times \Delta t \tag{4.6}$$

By substituting the solved values for σ (4.6), and α (4.3), into (4.5) one can obtain a quadratic equation in terms of ω. Since the scale parameter (ω) is strictly positive, of interest is the positive solution. This leaves only the remaining parameter (ξ) to identify.

Unfortunately, the first moment (μ) of a SN_1 is not very meaningful in the context of object movement. However, the mode of a SN_1 (denoted as $\hat{\mu}$) can be used to model the most probable location of future movement. For a SN_1 pdf $\hat{\mu}$ is not available in analytic form, but can be found by solving for the root of the first derivative of the SN_1 pdf (Gupta and Gupta, 2004), that is we must solve:

$$f(x)' = \frac{d}{dx}\left[\frac{2}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left(\alpha\left(\frac{x-\xi}{\omega}\right)\right)\right] = 0 \tag{4.7}$$

The SN_1 pdf is unimodal and therefore (4.7) possesses a single, unique root. Unfortunately, (4.7) cannot be easily represented in an analytical form, requiring the use of numerical methods to obtain the root. It is intuitive enough to visualize the most probable location of future movement[1] occurring at the mode (e.g., Figure 4.1b). We propose a simple model where $\hat{\mu}$ is a function of the objects current location (x_t), current velocity (v_t), and the time difference into the future (Δt).

$$\hat{\mu} = x_t + v_t \times \Delta t \tag{4.8}$$

More sophisticated formulations for $\hat{\mu}$ may be warranted that consider the ratio of v_t to v_{max}, the magnitude of Δt, and the objects acceleration. By substituting $\hat{\mu}$ (obtained from (4.8)) for x into (4.7), along with the previously computed values for

[1]Here we assume, as in physics, that moving objects tend to continue their motion unless acted on by other forces. That is, it is *most* probable that the object does not change speed or direction

ω and α, one can obtain a function for the single remaining unknown – ξ, which can be solved using numerical methods.

In summary, using known values for x_t, v_t, Δt, and v_{max}, we derive a system of three non-linear equations to solve for SN_1 parameters α, ω, and ξ using the following steps.

1. Substitute (4.4) into (4.3) in order to explicitly solve for the shape parameter – α.

2. Substitute α and (4.6) into (4.5) and solve for the scale parameter – ω, where $\omega > 0$.

3. Substitute values for α and ω, along with the computed value for $\hat{\mu}$ from (4.8) into (4.7), to solve for the location parameter – ξ.

Recall that in step 3 this procedure requires that (4.7) be solved numerically as it is not analytically tractable. Solving of the above system of non-linear equations is done in the mathematical software Maple (Maplesoft, Waterloo, Ontario). The resulting values for parameters ξ, ω, and α can be used to model the future movement possibilities for the object based on the SN_1 model. We have used the '`sn`' package available in R (R Development Core Team, 2012) to build and sample from skew-normal distributions.

4.2.2 Extending the Model – Two Dimensions (Spatial Movements)

Extension of the univariate model to two dimensions for application with movement data recorded in the spatial plane (i.e., with x, y coordinates) requires the consideration of several key properties. When an object exhibits kinetics, this movement is associated with a direction in the spatial plane. Consider this direction to be the axis-of-movement (AoM), and thus there is an associated axis perpendicular to the movement (A+M). In practice it may be useful to examine movement based on these two axes using rotations of the natural (x, y) coordinates (Figure 4.3). These two newly defined axes (AoM and A+M) are useful properties for developing and comparing candidate models.

Again, we consider the three basic characteristics required for candidate models, as suggested for the univariate case, that is: 1) if no initial velocity exists, the model

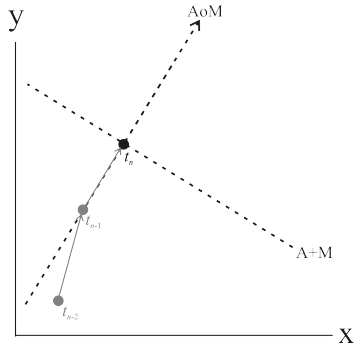

Figure 4.3: Diagram showing how axis of movement (AoM) and axis perpendicular to movement (A+M) can be interpreted from a movement dataset.

should reduce back to the normal model from Winter and Yin (2010, 2011); 2) the initial velocity is proportional to the shift in the probability mass; and 3) the mode of the model distribution is identifiable. In the two dimensional case, we consider two alternative properties of candidate models. Let $f_m(s)$ be the function describing the movement probability surface across space (s) for model m. First, the model should exhibit reflectional symmetry about the AoM; defined as:

$$f_m(s) = f_m(r_{AoM}(s)) \tag{4.9}$$

where r_{AoM} signifies a reflection along the line defined by the AoM. For most objects, moving in unconstrained space, turning left and right are equally probable. For objects moving along a network turning probabilities may favor left or right turns in specific scenarios.

The second consideration is the structure of the resulting distribution. This consideration arose after experimentation with multiple candidate models that seemed reasonable, but exhibited differing resultant shape characteristics. We can examine the structure of fm(s) to examine symmetry, but also discuss how well the shape of fm(s) aligns with boundaries proposed by Kuijpers *et al.* (2011). The multivariate skew-normal distribution (Azzalini and Dalla Valle, 1996) offers a potentially useful model for modeling future movement probabilities in the spatial plane (i.e., bivariate skew-normal model Figure 4.4a). The bivariate skew-normal uses the same three parameters as the univariate skew-normal, replacing scalar values by their multidimensional vector/matrix alternatives, where ξ is a location vector, ω is a scale/covariance matrix, and α is a skewness vector. Again using the proposed alternate parameterization (Arellano-Valle and Azzalini, 2008) one could attempt to relate these parameters

to the moments of the bivariate skew-normal distribution, similar to the univariate case. However, parameterizing the bivariate skew-normal is extremely difficult. Recall that we used numerical methods to solve for ξ in the univariate case, which become intractable for the bivariate situation. Further, in the bivariate case the scale/covariance matrix induces assymetries into the model by interrelating the scale and skewness parameters (Arellano-Valle and Azzalini, 2008) and therefore would not satisfy the reflectional symmetry property we desire. The absence of symmetry suggests that the bivariate skew-normal distribution may not be useful in this particular application. A seemingly logical alternative would be to model the movement of the object as two independent SN_1 distributions (Figure 4.4b), one for movement in the X direction and one for the Y direction. However, when the magnitude of $v_x \neq v_y$ this form of a model also introduces similar unwanted assymetries in $f_m(s)$, and therefore does not satisfy the reflectional symmetry condition.

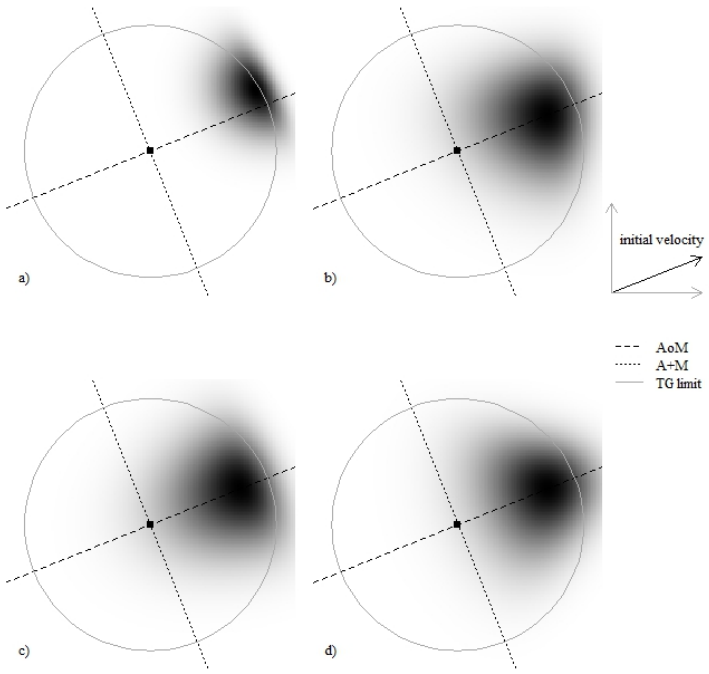

Figure 4.4: Output probability surfaces, termed $f_m(s)$, for candidate models for predicting future movement possibilities in spatial (2-dimensional) movement applications. a) bivariate skew-normal, b) two univariate skew-normals, aligned with the x- and y-axis, c) univariate skew-normal aligned with the AoM, normal aligned with the A+M, and d) two univariate skew-normals, each aligned at 45° to the AoM, constructed as in Figure 4.3b.

Given our success at implementing the univariate skew-normal, a potential bivariate skew-normal model would include the product of a skew-normal aligned with the AoM and a normal model aligned with the A+M (Figure 4.4c). The selection of the normal for the A+M is to satisfy the symmetry requirement, although any symmetric distribution could be accommodated here. The use of the normal distribution here however ensures that we satisfy criterion 2); that the model reduces to that of Winter & Yin in the absence of initial kinetics.

Alternatively, we propose the use of two univariate skew-normal distributions aligned at 45° of either side of the AoM (Figure 4.4d). The motivation for choosing this formulation is based on repeated experimentation with two independent SN_1 distributions. Based on this orientation it can be shown that object kinetics (along the AoM) can be decomposed into two equal and orthogonal vectors along these corresponding axes. Given an object located at the origin with an initial velocity in direction θ (i.e., θ from the horizontal axis) it is trivial to compute the rotated coordinate system (see Figure 4.5). Under this rotated coordinate system, object kinetics will be identical in the rotated axis (x' and y') and computed by:

$$v_{x'} = v_{y'} = \sqrt{\frac{v^2}{2}} \tag{4.10}$$

Where x' and y' are the rotated coordinates for two orthogonal axes taken to be 45° from the AoM. Based on this model we can construct a bivariate skew-normal as the product of two identical univariate skew-normal distribution aligned at 45° from the AoM. As can be seen in Figure 4.4d, this model accommodates all of the requirements of a candidate model.

Unlike with random movement where a strong foundation of theory exists for using the bivariate normal distribution for modeling future movement probabilities (e.g., Pearson, 1905; Skellam, 1951), no general theory exists for deriving future movement probabilities for kinetic movements. Thus, we chose to further evaluate only the rotated skew normal model, based on qualitative assessment and initial data-driven comparisons between models. Based on our observations and trials we found the rotated skew-normal model provided better alignment with the kinetic time geographic boundaries (Kuijpers *et al.*, 2011, , see also Figure 4.6), but also showed better agreement with movement data based on initial tests. However, the skew-normal / normal model and rotated skew-normal models generate rather similar $f_m(s)$ surfaces (i.e., Figure 4.4 c and d), and a more thorough investigation of the differences between the

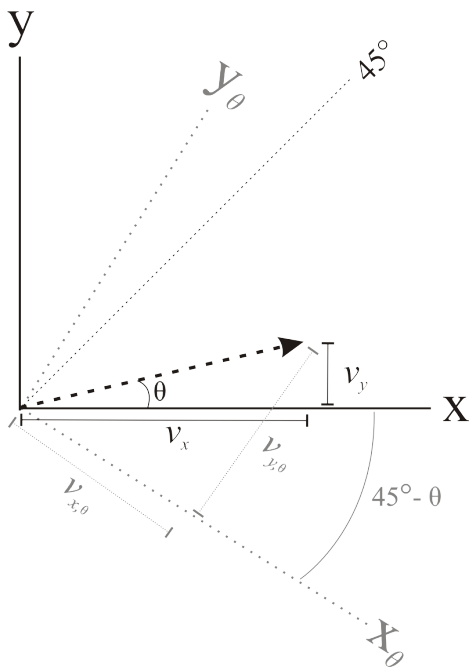

Figure 4.5: Diagram showing how a rotated coordinate system set up at 45° angles from the AoM can be used to decompose a movement vector into two orthogonal velocities of equal magnitude ($v_{x,\theta}$ and $v_{y,\theta}$).

models is warranted. From here forward, the rotated SN model (with two axis at 45° from the AoM - Figure 4.4d) will be implemented and referred to as the SN-model.

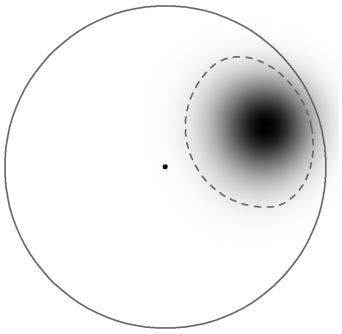

Figure 4.6: Comparison of proposed SN model (probability surface in greyscale) with kinetic time geographic boundaries (dashed line) defined by Kuijpers *et al.* (2011). The classic time geographic boundary (large grey circle) is shown for comparison.

4.2.3 Model Discussion

The model we have proposed is impacted by the assumptions necessarily made to solve the system of equations associated with the skew-normal parameters (i.e., equations (4.3) to (4.8)). The first assumption is that the skewness parameter (γ) is proportional to the ratio of the objects current velocity – v_t to v_{max} (i.e., in (4.4)). Setting the skewness to the ratio of v_t to v_{max} is logical as this bounds γ on $[-1, 1]$, which is the natural range for this parameter. However, the relationship between object kinetics (defined here using the ratio of v_t to v_{max}) and γ may be non-linear and alternative definitions of (4.4) may be warranted provided they maintain γ on the range $[-1, 1]$. For instance, here we ignore the effect of acceleration (Kuijpers *et al.*, 2011) in our kinetic definition.

The second assumption we make is on the variance parameter in (4.6), where existing theory from Winter and Yin (2010, 2011) suggests that at three standard deviations the pdf should equal the time geographic boundary of movement (i.e., $v_{max} \times \Delta t$). Use of this formulation for σ^2 means that in the absence of kinetics the model reverts to that proposed by Winter & Yin, a property of the model we intended to maintain. For moving objects, an increased kinetic effect will result from an object moving with a faster relative initial velocity, or a finer sampling interval. Using the definition in (4.5) and keeping variance constant, it can be shown that at higher levels of observed skewness (equating to increased kinetics) the scale parameter (ω), which roughly describes the width of the skew-normal distribution, will be smaller in magnitude then with a lower kinetic effect. A smaller width associated with increased kinetics is a positive result in light of what we might expect with movement situations (i.e., lesser movement opportunities with increased kinetics – Kuijpers *et al.*, 2011) and further evidence that the proposed model is suited to movement applications. A lingering issue with the Winter & Yin model is probability surfaces defined beyond the physical limits imposed by time geography. Winter and Yin (2011) use the classical time geographic boundaries in order to truncate the model distribution. Similarly, here it would be appropriate to truncate the SN model surfaces using the kinetic boundaries defined by Kuijpers *et al.* (2011, see also Figure 4.6).

The final assumption we make in the model is given by (4.8). Here we assume that the most likely location of future movement (the mode of the resulting two-dimensional surface) is at the location (Δt into the future) associated with unchanging speed and direction by the object. By assuming that moving objects are most likely

to maintain both speed and direction, the SN model is founded on fundamental rules from motion-based physics (i.e., Newton's first law of motion). This assumption is also apparent in models used to match movement data (e.g., GPS traces) to road networks (e.g., Krumm *et al.*, 2007). However, the assumption that movement speed and direction are most likely to be constant and unchanging may not hold as Δt increases (e.g., in (4.8)), however there may be some psychological factor that suggests this relationship is approximately true. In ecology, the tendency of organisms to continue moving in the same direction is termed persistence (Othmer *et al.*, 1988). In most cases this is unlikely to be related to physical kinetics, but rather other underlying motivations, such as migratory phases, or habitat requirements. It may be useful to consider a persistence-based definition of kinetics in ecological examples to more appropriately factor in these types of kinetic properties of wildlife movements. This would allow kinetic-based ideas from time geography to be included with more coarsely collected wildlife movement datasets (i.e., those with sampling intervals of minutes to hours).

4.3 Case Study

4.3.1 Data

We have attempted to evaluate the proposed SN model using a combination of simulated and real-world movement datasets (Table 4.1; Figure 4.7). The first dataset is a random walk. Similar random models have been suggested by early ecologists as null models for organism movement (Skellam, 1951). The second dataset is a correlated random walk. Correlated random walk models are considered one of the best models for the movements of wildlife (Kareiva and Shigesada, 1983; Turchin, 1998), and commonly used to simulate movement data for method testing (Nams, 2005; Börger *et al.*, 2008; Long and Nelson, 2012). The first real dataset used is Caribou data tracking the movement of a single caribou across northern British Columbia over the course of a single year. Location fixes were obtained at a sampling interval of $\Delta t = 4$ hour, using a VHF telemetry system resulting in minimal missing fixes. The second real dataset is a GPS track of a commuter cyclist. Cycling data were recorded using a commercial, handheld GPS set to a sampling interval of $\Delta t = 5$ seconds. The final real dataset is generated using sport-specific GPS units (GPSports, Fyshwick, Australia) from athletes participating in an ultimate frisbee game. Here GPS relocations

of an athlete are collected at a sampling interval of 5 Hz ($\Delta t = 0.2$ s), representing an extremely detailed dataset on individual movement. This sports data has been previously explored in Long and Nelson (2013a), in the context of measuring dynamic interactions in player movements.

Table 4.1: Two simulated and three real-world datasets used to evaluate the existing probabilistic time geography model with the proposed kinetic-based probabilistic time geography model.

Dataset	**Type**	n	Δt	v_{max}	**Comments**
RW	Simulated	1000	–	3.7	`simm.brown()` function in R package `adehabitatLT`, $h = 1$.
CRW	Simulated	1000	–	3.8	`simm.crw()` function in R package `adehabitatLT`, h= 1, $r = 0.8$.
Caribou	Real	1772	4 hr	0.77 m/s	Caribou tracked via satellite VHF telemetry during 2000.
Cyclist	Real	247	5 s	13.5 m/s	Movements of the first author while cycling; tracked using a commercial GPS.
Athlete	Real	288	0.2 s	7.6 m/s	Ultimate frisbee player, over a 1 minute interval of a training match. Collected using a sport-specific GPS device.

Note: v_{max} estimated from the data following Long and Nelson (2012).

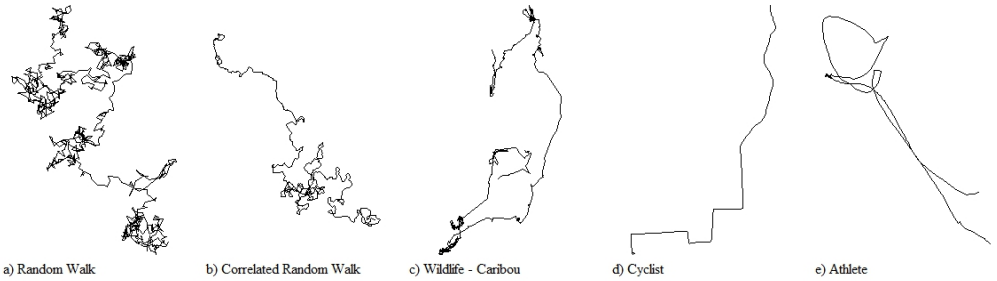

Figure 4.7: Five example datasets used in evaluating the SN model against the Winter & Yin model; see Table 1 for more details on each dataset.

4.3.2 Methods

Model Set-up

Movement datasets can be used for examining predictive movement models (such as kinetic-based probabilistic time geography) by attempting to predict successive movement fixes based on the previous fixes. In order to do this, we compare the observed location of each fix with the modeled probabilities obtained from either the Winter & Yin model or the SN model. That is, for each fix i in a trajectory we compute two probability surfaces $f_{W\&Y}(x)$ and $f_{SN}(x)$ (e.g., figure 4.6) that can be used to predict, probabilistically, future movement locations (i.e., fix $i+1$). We extract the observed fix probability from both the Winter & Yin and SN models, along with the maximum observed probability in order to evaluate the two models.

Model Evaluation

It is useful to evaluate the predictive ability of the model by examining how well the predictive model aligns with the observed movement data. Typically, one would use a measure of, in this case spatial, distance (e.g., |observed - expected|) to quantify this agreement. Given that we use the mode explicitly in our derivation of skew-normal models (which is not necessarily equivalent to the expected value) we suggest some alternative measures of model agreement.

When one model is a special case of another, as in our situation where the normal model is a special case of the SN-model, the likelihood ratio of the two models can be used as a comparison statistic (Kalbfleisch, 1985). Owing to its mathematical properties, the natural logarithm of this ratio, termed the log-likelihood ratio is routinely implemented:

$$\Lambda_i = \ln \frac{p_a(x_i)}{p_b(x_i)} \tag{4.11}$$

here Λ_i is the log-likelihood ratio for observation i, and $p(x_i)$ is the modeled probability (for model a or b) at observation i. Positive values favor the model a, while negative values favor the model b, values near 0 signify that both models perform equally. In our examples, model a is the normal model from Winter and Yin (2010, 2011) and model b is the skew-normal (SN) model incorporating object kinetics. As a result, $\Lambda_i < 0$ indicate the SN model provided a better fit to the data, while $\Lambda_i > 0$ indicate the Winter & Yin model demonstrates better agreement. We plot the Λ_i of a particular movement dataset as a time-series to examine temporal trends in model

differences and report the mean values ($\bar{\Lambda}_i$). Further, a global measure of agreement, the log-likelihood ratio statistic, can be computed as:

$$\mathrm{LLR} = -2 \sum \Lambda_i \tag{4.12}$$

where LLR is the log-likelihood ratio statistic, which is approximated by a chi-square distribution, with degrees of freedom (d.f.) equal to the difference in the free parameters in model a and b. In our case, model a is the Winter & Yin model and contains 1 free parameter; while model b is the SN model and contains 3 free parameters. Therefore, the d.f. for the LLR test statistic is $3 - 1 = 2$. We use LLR to test for whether the use of the more complex SN model provides a significant improvement (with $\alpha = 0.01$) over the Winter & Yin model.

To further examine the agreement of the models with the data, we define a statistic that compares the observed probability for movement i as a ratio of the maximum modeled probability (the mode of the predictive surface). Termed the predictive probability, the statistic takes the form:

$$PP_{k,i} = \frac{p_k(x_i)}{p_k(\hat{\mu}_i)} \tag{4.13}$$

Where $PP_{k,i}$ is the predictive probability of the k^{th} model for observation i. The numerator is simply the observed probability from the model at observation i. This value is then taken as a ratio of the observed maximum probability (expected value – or mode) of the model, denoted which is used here to appropriately scale values. The ratio defined by (4.11) can be thought of as a performance measure of the model at each data point, with values closer to 1 signifying that the data and model show good agreement, while values near 0 suggest the model and data are not well aligned. The mean values (PP_k) are reported for each model and a pairwise t-test (with $\alpha = 0.01$) was used to examine whether the evaluative measure ($PP_{k,i}$) differs significantly between the two models.

With these five datasets, we have differing expectations of SN model performance when compared with the existing Winter & Yin model. Given that the model of Winter & Yin is based on random walks, we expect the Winter & Yin model to perform better with the random walk dataset. With the correlated random walk dataset we might expect the SN model provide a better agreement, although decreasing the correlation parameter (r – see Table 4.1) could change this outcome as the correlated

random walk would exhibit more random-like behavior. Similarly, wildlife movements are commonly modeled as variations of correlated random walks. We expect that at a relatively coarse sampling interval ($\Delta t = 4$ hr) we will see similar results with the wildlife data as with the correlated random walk. In the cyclist example, we expect that the directed and linear nature of cyclist movement will favor the SN model. Further, the kinetic effect is likely dependent on the sampling interval chosen (here $\Delta t = 5$ s), and further decreasing the sampling interval would initiate an even greater kinetic effect. Finally, with the athlete movement data we expect the SN model to outperform the existing model due to the relatively high influence of kinetics in athlete movement and the extremely fine sampling interval ($\Delta t = 0.2$ s).

4.3.3 Results

As expected, for the random walk dataset, the probabilistic time geography model from Winter and Yin (2010, 2011) performed better based on both evaluative tests. The log-likelihood ratio plot (Figure 4.8a) demonstrates the unpredictable nature of a random walk, with both models outperforming the other in some cases, but on average the normal model of Winter & Yin seems to provide better agreement ($\bar{\Lambda}_i = 0.424$), further supported by the non-significant LLR. For the random walk dataset, the predictive probability of the SN model ($PP_k = 0.456$) is lower than the Winter & Yin model ($PP_k = 0.599$), a difference that is highly significant (Table 4.2). However, both values are relatively low, which suggests that neither model is particularly adept at predicting successive locations of this particular random walk dataset.

Table 4.2: Results for each of the five example datasets comparing the SN model against the existing model of Winter & Yin.

	$\bar{\Lambda}_i$	**LLR**	PP_k **(SN)**	PP_k **(W&Y)**	**Diff.**
RW	-0.184	-848	0.456	0.599	-0.143*
CRW	0.0779	359*	0.682	0.618	0.0642*
Caribou	-0.00643	-52.4	0.958	0.973	-0.0151*
Cyclist	0.381	430*	0.945	0.548	0.396*
Athlete	0.304	401*	0.949	0.650	0.299*

*denotes significant value ($p < 0.01$)

Similarly, as expected with the correlated random walk, the SN model outperformed the model of Winter & Yin using both visual and statistical tests. As can

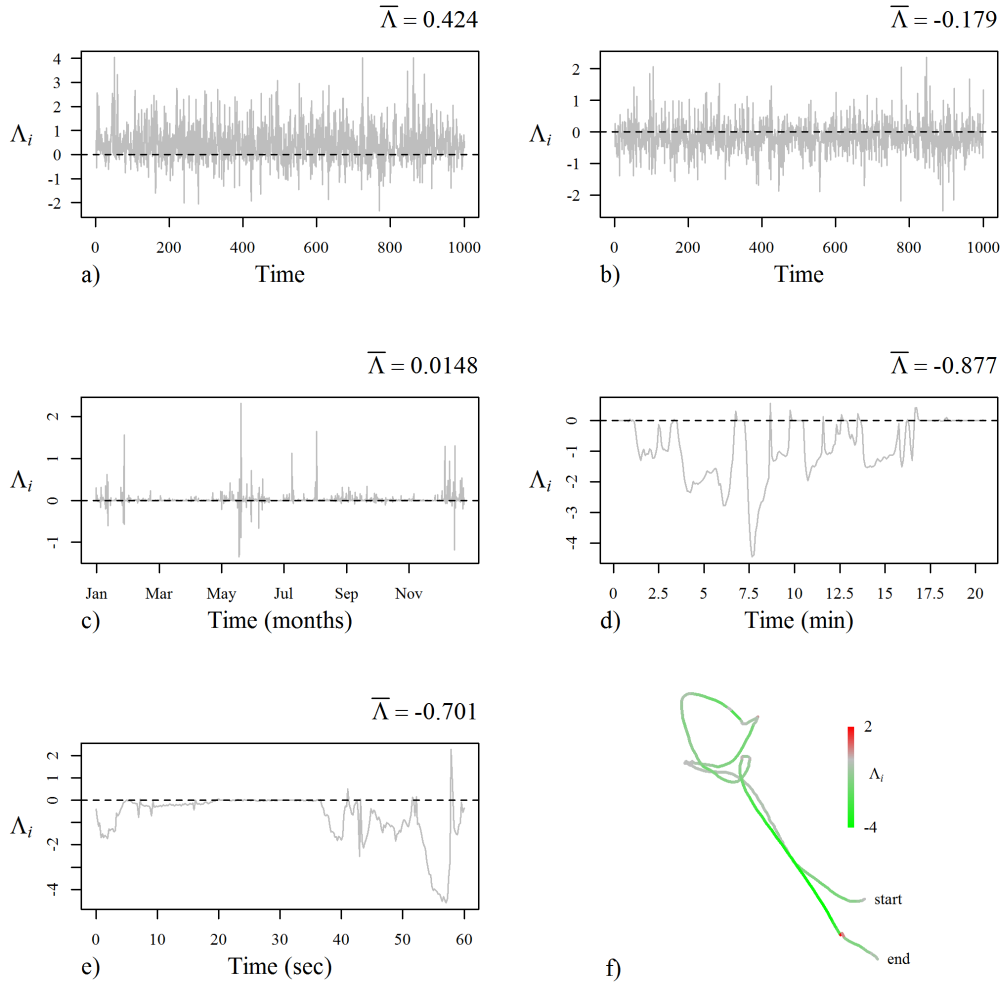

Figure 4.8: Λ_i results for each of the five sample datasets: a) RW, b) CRW, c) Caribou, d) Cyclist, and e) Athlete. As an example, a map f), of the Λ_i values associated with the athlete movement dataset can be used to visualize in which parts of the movement trajectory the SN model outperforms the Winter & Yin model (and vice versa). Values for $\Lambda_i > 0$ indicate where the Winter & Yin model agrees better with the movement data, while values for $\Lambda_i < 0$ indicate where the SN model shows better agreement.

be seen in the log-likelihood ratio plot (Figure 4.8b), the difference between the two models in the correlated random walk is similar but opposite than the random walk ($\bar{\Lambda}_i = -0.179$). A significant LLR = 359 suggests that the SN outperforms the Winter & Yin model. With the correlated random walk, the predictive probability ($PP_k = 0.682$) of the SN model is higher than the Winter & Yin model ($PP_k = 0.618$), a difference that is significant (Table 4.2).

Plotting the Λ_i from the caribou dataset (Figure 4.8c) demonstrates that, for the most part, the Λ_i values are near 0. With the caribou example $\bar{\Lambda}_i = 0.0148$, an indication that the model from Winter & Yin slightly outperforms the SN model. However, during specific intervals the SN model outperforms the Winter & Yin model (e.g., the interval occurring in late May). These periods correspond with more active caribou movements associated with annual migration phases. When the caribou is making extensive movements the kinetic-based model may be superior, but during low movement phases the two models perform similarly. The LLR indicates that indeed there is no significant advantage of choosing the more complex SN-model over the simpler model of Winter & Yin. The test comparing the $PP_{k,i}$ of each model for the caribou dataset revealed that, on average, the SN model has lower predictive probability ($PP_k = 0.958$) than the Winter & Yin model ($PP_k = 0.973$) for this dataset, a small difference, but one that is still significant (Table 4.2).

From the plot of the Λ_i for the cyclist dataset (Figure 4.8) it is clear that during specific intervals the SN model demonstrates better agreement (negative Λ_i values). However, at other instances the two models perform identically (i.e., when $\Lambda_i = 0$). Here the cyclist has stopped moving, and in the absence of kinetics the two models are equivalent, thus $\Lambda = 0$. With the cyclist dataset, $\bar{\Lambda}_i = -0.877$, which suggests that the SN model outperforms the Winter & Yin model, further supported by the significant LLR = 430. A $PP_k = 0.945$ was observed with the SN model, while a much lower $PP_k = 0.548$, was found with the Winter & Yin model, a difference again found to be highly significant (Table 4.2).

The results from the athlete dataset are similar to those from the cyclist dataset. During specific mobile periods the SN model shows better agreement, while during other periods (of stationary behaviour) the two models are similar (Figure 4.8e). The fact that $\bar{\Lambda}_i = -0.701$ again suggests that the SN model demonstrates better agreement with this dataset, supported by a highly significant LLR = 401. The predictive probability test confirms this observation with $PP_k = 0.949$ for the SN model and $PP_k = 0.650$ for the Winter & Yin model, a highly significant difference (Table 4.2).

4.4 Discussion

We have used two simulated examples along with three real-world datasets to demonstrate the usefulness of the kinetic-based approach to modeling future movement pos-

sibilities in a time-geographic framework. From these examples it is clear that with applications involving a relatively high kinetic effect (i.e., fast moving objects, with finely sampled movement data), the SN model for probabilistic time geographic proposed here is a far more useful predictor of future movement probabilities than the existing definition based solely on random movement.

As discussed by Kuijpers *et al.* (2011), ignoring object kinetics may be reasonable when estimating broad-scale patterns from finely sampled movement data; for example, when looking at long-term transportation trends. Similarly with coarsely sampled movement data the physical kinetics of movement will not be relevant. For instance, data collected by legacy radio-tracking systems of wildlife use sampling frequencies in the order of hours to days. However, the development of kinetic-based time geography has clear merit in applications where object kinetics are relevant in the construction of time geographic volumes. Such applications include the analysis of finely sampled wildlife movement data (Cagnacci *et al.*, 2010), human powered movements, such as by athletes (this chapter), and the movements of vessels such as ships and airplanes (Knighton and Claramunt, 2001), as well, the role of kinetics can be clearly demonstrated when examining automobile trajectories (Yu and Kim, 2006).

Wildlife tracking systems are now being equipped with real-time data transfer mechanisms in order to monitor wildlife movements in real-time (Urbano *et al.*, 2010). These systems can be used to guide management strategies (e.g., forest harvesting) in important conservation areas based on the location of wildlife (Dettki *et al.*, 2004). Kinetic-based probabilistic time geography could be used to improve movement predictions and guide conservation strategies by identifying, probabilistically, specific areas of concern. Video tracking systems are also commonly used to derive movement data of multiple target objects in a fixed spatial domain (e.g., athletes on a playing surface, Liu *et al.*, 2009; Lu *et al.*, 2009). With video tracking, movement trajectories are often interrupted by visual occlusions, and a single trajectory will become divided into numerous segments (Liu *et al.*, 2009; Lu *et al.*, 2011). A kinetic-based probabilistic time geography could be useful as a model for connecting trajectory segments in multi-object video tracking systems. Another area of spatial research that is rapidly expanding is the development of location based services (Raper *et al.*, 2007). Location based services leverage a client's location through a location aware device (e.g., GPS embedded in a cell-phone) in order to tailor services to clients based on location. Popular examples include restaurant locating or real-time navigation ap-

plications on a smart-phone. In such applications, kinetic-based probabilistic time geography could improve spatial locating or preference selection by incorporating the motion of the client, especially if they are travelling in a fast moving vehicle such as a car.

For many movement applications researchers are interested in extracting patterns from datasets where movement is confined to a (known) travel network (e.g., Miller and Wu, 2000). In these situations the spatial domain cannot be represented as an open two-dimensional plane, but rather as a set of connected network links that facilitate essentially one-dimensional movement within the spatial plane. Turns can occur along network links, but primarily at nodes, where movement may proceed in one of multiple directions. The framework we have introduced for modeling kinetic-based movement probabilities can still hold in this situation (e.g., Figure 4.9). Along network links the univariate skew-normal formulation can be used in lieu of the two-dimensional model. At network nodes, the probability density beyond the node can be divided between the available links based on individual node turning probabilities that may reflect pre-determined preferred route choices, and even turning times. A hybrid one-dimensional model draws on the calculations already being used in network analysis algorithms for computing travel times along street networks.

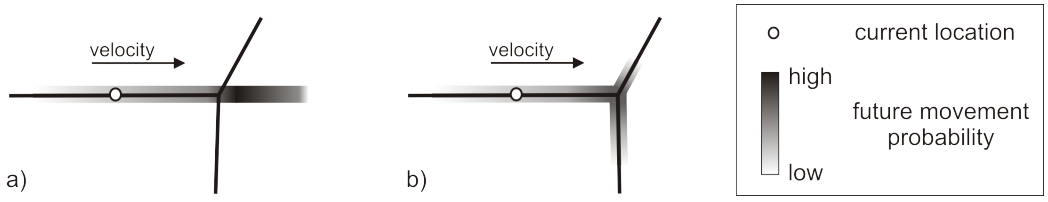

Figure 4.9: Example of how a hybrid one-dimensional model for kinetic-based probabilities could be applied on a network. a) Kinetic probabilities derived for a moving object along a network link; modeled probabilities extend along the current link, but go beyond the node. b) Turning incorporated at the node, with probability of right turn > left turn.

Other models for modeling movement probabilities in time geography also exist. Miller and Bridwell (2009) propose field-based time geography where movement probabilities are defined using a movement cost surface. Field-based time geography represents the combination of time geography theory with common GIS operations, (i.e., those used in least-cost path analysis, Douglas, 1994). Downs (2010) has introduced ideas from time geography into kernel density estimation (commonly used in the study of wildlife movement). Downs replaces the traditional circular kernel (e.g.,

Gaussian, quartic) with the potential path area from time geography (Figure 4.1b) and computes a density surface representing the probability an object visits a given location (termed a utilization distribution). Time geographic kernel density estimation can be used to define the interior structure of the potential path area, and has since been extended to work with network-based applications (Downs and Horner, 2012).

How to model movement probabilities has also been examined in the context of wildlife movement ecology. Horne *et al.* (2007) have derived a similar probabilistic surface to the Winter & Yin model based on the notion of a Brownian bridge (random walks connected by two end points). Benhamou (2011) has suggested that biased random bridges represent a more suitable model for such movement and has developed a biased random bridge movement model. Both the Brownian bridge and biased random bridge utilize a bimodal distribution for modeling movement probabilities between two fixed locations, which effectively models movement probabilities within the potential path area. Winter and Yin (2010) model movement locations in the space-time prism where movement probabilities are the result of unimodal distributions computed for slices of the space-time prism. Computing the integral (over time) of the Winter and Yin (2010) model would produce a surface for comparison with the Brownian bridge and biased random bridge, providing novel insight on the differences and similarities between these approaches.

4.5 Conclusion

We quantify movement probabilities for the space-time cone from time geography using a formulation that incorporates object kinetics. Quantifying the interior structure of time geography volumes is currently an area of active research with different methods relying on various underlying assumptions. The kinetic-based approach we describe is useful for studying movement data at fine temporal granularities, or where kinetic properties (physical or otherwise) are expected, but may not be appropriate with coarser temporal granularities or slow moving objects. A time geography that incorporates movement kinetics, both in the calculation of volume boundaries as in (Kuijpers *et al.*, 2011), and in the interior structure of those volumes as we describe here, will provide a more powerful, and realistic model for studying object movement when kinetic properties are inherent. Future endeavors will involve extending the kinetic-based probabilistic time geography model to the space-time prism, necessary

for evaluating movement datasets where fixes are most appropriately represented as start and end anchor points of prisms. Further, we hope to investigate ways for examining intersection probabilities with the kinetic-based model, similar to those proposed by Winter and Yin (2010, 2011), which will allow more sophisticated time geographic questions to be investigated.

Chapter 5

Measuring Dynamic Interaction in Movement Data

5.1 Abstract

The emergence of technologies capable of storing detailed records of object locations has presented scientists and researchers with a wealth of data on object movement. Yet analytical methods for investigating more advanced research questions from such detailed movement datasets remain limited in scope and sophistication. Recent advances in the study of movement data has focused on characterizing types of dynamic interactions, such as single-file motion, while little progress has been made on quantifying the degree of such interactions. In this article, we introduce a new method for measuring dynamic interactions (termed DI) between pairs of moving objects. Simulated movement datasets are used to compare DI with an existing correlation statistic. Two applied examples, team sports and wildlife, are used to further demonstrate the value of the DI approach. The DI method is advantageous in that it measures interaction in both movement direction (termed azimuth) and displacement. As well, the DI approach can be applied at local, interval, episodal, and global levels of analysis. However the DI method is limited to situations where movements of two objects are recorded at simultaneous points in time. In conclusion, DI quantifies the level of dynamic interaction between two moving objects, allowing for more thorough investigation of processes affecting interactive moving objects.

5.2 Introduction

The study of individual movement has entered a new era whereby researchers from various fields can benefit from fine resolution object movement data. Technical developments associated with location aware technologies, such as GPS, are transforming representations of movement. Despite improvements in spatially explicit movement datasets, the scope and sophistication of research questions are limited by a lack of methods and analysis (Wolfer *et al.*, 2001). Laube *et al.* (2007) suggest that within geography, reliance of geographic information systems (GIS) and spatial statistics on 2-dimensional representations may be limiting the development of more complex analyses of movement, while disciplines outside of geography may be unaware of the power of spatial (and space-time) analysis. To optimally utilize new movement datasets, analytical techniques capable of addressing more advanced research questions are required.

Recently, the identification and measurement of *dynamic interactions* between moving objects has become an active area of research, likely owing to readily available fine granularity movement data. Dynamic interaction, a term from the wildlife ecology literature, can be defined as the way the movements of two individuals are related (Macdonald *et al.*, 1980) or as inter-dependency in the movements of two individuals (Doncaster, 1990). Alternatively, the terms association (Stenhouse *et al.*, 2005), relative motion (Laube *et al.*, 2005), and correlation (Shirabe, 2006) have been used to refer to dynamic interactions between moving objects in other examples. All of these terms refer to the same general idea: identifying of how the movements of one individual are related to another. Recent work on dynamic interactions has focused on methods for identifying dynamic interaction patterns defined a priori (for example single file motion, Buchin *et al.* 2010; or chasing behavior, de Lucca Siqueira and Bogorny 2011). However limited work exists on quantifying the strength of dynamic interactions present in movement data. With this in mind we are motivated to investigate methods for measuring the strength of dynamic interactions when there is an expectation that such behavior occurs. This approach differs from recent developments in movement analysis which focus on identifying patterns, defined *a priori*, from large movement databases.

The objective of this work is to extend a previously developed statistic (Shirabe, 2006) to a measure capable of quantifying the degree of dynamic interaction between moving objects. The new method (termed DI) measures dynamic interaction in co-

incidental movement segments, that is, it requires movement data of two individuals recorded simultaneously. The DI method is separable into components measuring dynamic interaction in movement direction (azimuth) and movement distance (displacement), termed DI_θ and DI_d respectively. Further, DI is appropriate with the four analysis levels (local, interval, episodal, and global – see Figure 5.1) identified by Laube *et al.* (2007) with the beneficial property of local values (denoted here using lower-case di) that aggregate to the interval, episodal and global values. Lastly, DI is derived in a way to allow for a time-lagged approach, but also extensions including time- and distance-based weighting schemes.

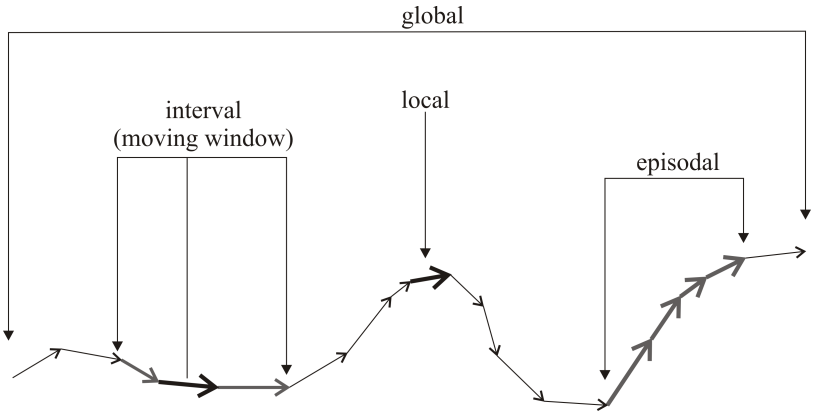

Figure 5.1: Diagram of four analysis levels used in movement data analysis (after Figure 2 in Laube *et al.*, 2007). Local level statistics are calculated for each individual movement segment. Interval level analysis computes a running average statistic using a moving window. Episodal level analysis computes the statistic over a selected 'episode' or period of the dataset. Global level analysis computes the statistic over the entire movement path.

5.3 Related Work

This research is motivated by an existing technique (Shirabe, 2006) for measuring the strength of dynamic interactions (termed correlations) present in movement data. The use of the term correlation by Shirabe stems from the fact that the statistic takes the form of a Pearson product-moment correlation coefficient. Consider two moving objects M_a and M_b, whose spatial coordinates (x, y) are recorded coincidentally at discrete times $t = 1...n$, termed fixes. Now consider for any M with $t = 2...n$, $V = [M_t - M_{t-1}] = [v_t]$, is a vector time series of M with $n - 1$ vector segments.

A correlation statistic for movement data defined this way takes the form (Shirabe, 2006):

$$r\left(V^b\right) = \frac{\sum_{t=1}^{n-1}(v_t^a - \bar{v}^a)\cdot(v_t^b - \bar{v}^b)}{\sqrt{\sum_{t=1}^{n-1}|v_t^a - \bar{v}^a|^2}\sqrt{\sum_{t=1}^{n-1}|v_t^b - \bar{v}^b|^2}} \tag{5.1}$$

where $\hat{v} = \frac{1}{n-1}\sum_{t=1}^{n-1} v_t$ are mean coordinate vectors of V. The correlation statistic (r) is defined over the interval $[-1, 1]$ with a score of 1 being perfect positive correlation and a score of -1 perfect negative correlation, with 0 denoting no correlation.

The statistic – r, could be advanced in three ways. First, it is dependant on the mean vector of each path, and thus measures correlations in movement deviations from their respective means. The statistic, r, cannot be used for testing direct interactions between two moving objects unless their corresponding mean vectors are identical or near identical. An improved statistic would not rely on this overall mean value. Second, r is unable to disentangle the effects of correlations in movement azimuth and distance, while being sensitive to both. Decomposing such a statistic into components based on movement direction (termed azimuth) and distance (displacement) would be beneficial, as it would allow interactions in these two independent components of movement to be analyzed separately. A third improvement would be a statistic that measures the interaction of each individual movement segment (i.e., local level – Laube *et al.*, 2007). By definition, r produces a single resulting value for the entire path (i.e., global level – Laube *et al.*, 2007). When movement patterns are characterized by periods of interactive and non-interactive behavior, or varying levels of interactive behavior, a local level statistic will allow a finer treatment of dynamic interactions.

Measurements of dynamic interaction in movement data have also been developed by wildlife researchers interested in a finer understanding of wildlife movement processes. The types of interactions studied in wildlife are classified as either static or dynamic interactions (Doncaster, 1990; Macdonald *et al.*, 1980). Static interaction relates to how two individuals use space coincidentally, while dynamic interaction reflects how the movements of two individuals are related, for example attraction (Macdonald *et al.*, 1980). Typically, measures of dynamic interaction summarize the proximity of simultaneous movement points. Doncaster (1990) introduced one such measure of dynamic interaction based on the variance/covariance matrix of the spatial

coordinates of simultaneous wildlife telemetry fixes; others have used Euclidean distance as an indicator of interaction (Bandeira de Melo *et al.*, 2007; Stenhouse *et al.*, 2005). Stenhouse *et al.* (2005) further investigated dynamic interaction in grizzly bears (termed associations) by measuring dynamic interaction in movement direction (azimuth – θ) defined as:

$$f_t\left(\theta_t^a, \theta_t^b\right) = \frac{\left|-(\theta_t^a - \theta_t^b) - 180\right|}{180} \tag{5.2}$$

Equation (5.2) ranges from 0 – 1, with values of 1 when direction of movements is identical and zero when completely opposite (i.e., at 180°).

Measuring dynamic interactions in moving object databases is also directly related to a larger body of literature on identifying similar movement trajectories (Sinha and Mark, 2005; Vlachos *et al.*, 2002; Yanagisawa *et al.*, 2003). Similarity indices are commonly employed as a first-step for identifying broader patterns or for detecting clusters in larger movement databases (Benkert *et al.*, 2008; Gao *et al.*, 2010). Moving object pairs that are highly interactive could also be said to follow a similar trajectory in many of these applications, and the methods for detecting dynamic interactions in movement data could be useful for detecting similar movement trajectories.

Recently, many new techniques have been developed for categorizing various dynamic interaction patterns commonly found in movement data. Laube *et al.* (2005) developed a method for detecting RElative MOtion (REMO) classes based upon interpreting patterns of movement direction in groups of moving objects. For example, trend-setting, when one object moves with anticipation of the movement of others, is identifiable using the REMO approach. Noyon *et al.* (2007) use changes in inter-object distance and velocity to identify relative behavior such as collision avoidance. Benkert *et al.* (2008) present an algorithm for finding flock patterns in movement databases; which tests whether a group of moving objects are contained in a circle radius r over a given time interval. The study of flocking behavior is useful in the study of wildlife and crowd dynamics (Batty *et al.*, 2003). Buchin *et al.* (2010) have developed a method for identifying single-file motion in groups of moving objects. Single-file motion is detected using free-space diagrams, derived from the Fréchet distance metric for comparing polygonal curves (Alt and Godau, 1995). Related to single-file motion is the detection of chasing behavior, identifiable using the algorithm proposed by de Lucca Siqueira and Bogorny (2011). The methods mentioned above are capable of identifying specific types of dynamic interactions in movement data as

defined a priori. However, such methods are unable to quantify the strength of dynamic interactions present, thus motivating the development of quantitative measures of dynamic interaction.

5.4 Derivation

In developing a measure of dynamic interaction we consider the rather optimal data situation (as in Shirabe, 2006) where two moving objects' (M_a and M_b) spatial coordinates (x, y) are recorded coincidentally at discrete times $t = 1...n$, termed fixes. For any M with $t = 2...n$, $V = [M_t - M_{t-1}] = [v_t]$, is a vector time series of M with $n-1$ vector segments. For each movement segment define two fundamental properties: direction (θ), termed azimuth, and length (d), termed displacement. Azimuth (θ) is the angle between a movement segment and a constant axis, most commonly the horizontal axis (Figure 2a). Displacement (d) is the Euclidean distance between two consecutive fixes in a movement segment (Figure 5.2a). We are interested in deriving a measure of dynamic interaction that separately quantifies interactions in azimuth and displacement (Figure 5.2b-e).

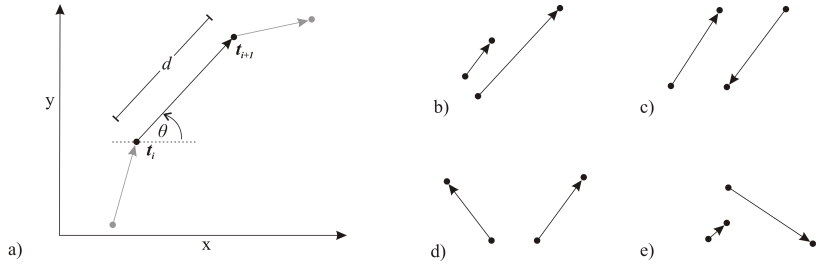

Figure 5.2: a) Diagram of movement properties azimuth (θ) and displacement (d). Examples of movement segments that exhibit: b) positive interaction in θ and low interaction in d; c) negative interaction in θ and high interaction in d; d) no interaction in θ and high interaction in d; and e) no interaction in θ or d.

5.4.1 Azimuth – θ

To investigate the interaction in movement azimuths we take the cosine of the angle between them. This is simply calculated as:

$$\mathrm{di}_\theta = f_t\left(\theta_t^a, \theta_t^b\right) = \cos(\theta_t^a - \theta_t^b) \tag{5.3}$$

where θ_t is the angle of movement at time-step t. Here f_t has a range of $[-1, 1]$ as desired. The function f is 1 when movement segments have the same orientation, 0 when movement segments are perpendicular, and -1 when in complete opposing directions. In practice if either object (or both) do not move (5.3) is undefined, because θ_t is undefined. Thus, we must consider two alternative scenarios; first if one object moves and one remains stationary, and second if both objects remain stationary. Here we make the assumption that if one moves and the other remains stationary the two objects exhibit no directional interaction, and if both are stationary they are positively interactive. Considering these two alternative scenarios, a complete definition for (5.3) is:

$$f_t\left(\theta_t^a, \theta_t^b\right) = \begin{cases} 0, & \text{if one of } \theta_t^a \text{ or } \theta_t^b \text{undefined} \\ 1, & \text{if both } \theta_t^a \text{ and } \theta_t^b \text{undefined} \\ \cos(\theta_t^a - \theta_t^b), & \text{otherwise} \end{cases} \tag{5.4}$$

5.4.2 Displacement $- d$

Interaction in movement displacement could be measured using a variety of functions. However, it is desirable to have the function (g_t) fall in the range of $[-1, 1]$, where a value of 0 represents no interaction and 1 positive interaction. Note there is no consideration of negative interaction in displacement. Using this definition g_t can be thought of as a scaling function to f_t, and maintains the statistic on the range $[-1, 1]$. We propose the following function for g_t:

$$\text{di}_d = g_t\left(d_t^a, d_t^b\right) = 1 - \left(\frac{|d_t^a - d_t^b|}{d_t^a + d_t^b}\right)^{\alpha} \tag{5.5}$$

where $|\cdot|$ is the absolute value operator, and α is a scaling parameter defaulting to 1. The function approaches zero when $d_t^a >>> d_t^b$ or vice-versa, and is 1 when $d_t^a = d_t^b$. The effect of the scaling parameter (α) on the function g_t is demonstrated in Figure 5.3. Parameter α can be adjusted to place stricter or looser requirements on similarity in displacement denoting interaction. As α is increased larger differences in displacement are still considered as positively interactive. A closer examination of (5.5) reveals that it is undefined when $d_t^a + d_t^b = 0$, (i.e., both objects are stationary). If we consider both objects remaining stationary as positive interaction, a more robust

definition of (5.5) is:

$$g_t\left(d_t^a, d_t^b\right) = \begin{cases} 1, & \text{if } d_t^a + d_t^b = 0 \\ 1 - \left(\frac{|d_t^a - d_t^b|}{d_t^a + d_t^b}\right)^{\alpha}, & \text{otherwise} \end{cases} \tag{5.6}$$

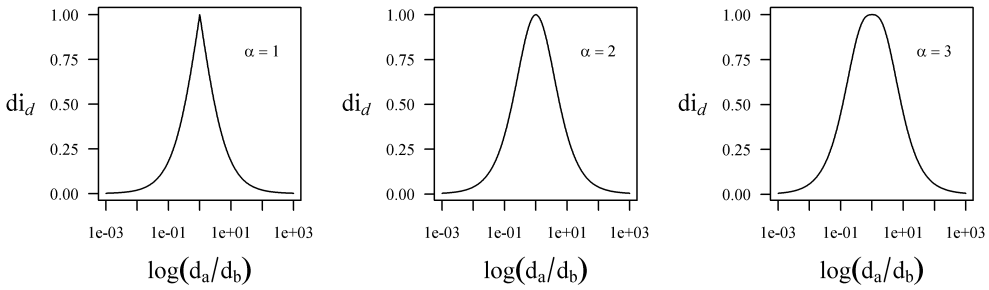

Figure 5.3: Relationship between $\log(d_a/d_b)$ and di_d, for values of $\alpha = 1, 2, 3$.

Thus, for two corresponding movement segments, a measure of dynamic interaction is the product between the azimuthal term (f_t) and displacement term (g_t):

$$\text{di} = \text{di}_\theta \times \text{di}_d = f_t\left(\theta_t^a, \theta_t^b\right) \times g_t\left(d_t^a, d_t^b\right) \tag{5.7}$$

We are motivated to use the functions f_t and g_t to provide a statistic that covers the range $[-1, 1]$ as was done in Shirabe (2006). Positive values of di_t correspond to cohesive or positively interactive movements, while negative values can be interpreted as repulsion or opposing movements. Values near zero should be interpreted as having no interaction.

The di statistics measure dynamic interaction based on similarity in azimuth (θ) and displacement (d) of simultaneous movement segments but do not account for the proximity of moving objects. Thus, di represents a similarity index taken in a normalized plane (i.e., the distance between the two objects has no impact on the resulting value). We are motivated to use this type of formulation as the spatial proximity required for dynamic interaction to occur is application specific. It is up to the analyst to decide if two moving objects maintain a requisite proximity for dynamic interaction to occur, then such interaction can be measured using di. In cases where actual spatial contact is required, for example when identifying points-of-interest in

large movement databases (e.g., Benkert *et al.*, 2007), the di method should not be employed.

We have made assumptions in the equations for di_θ and di_d regarding how to analyze dynamic interactions when objects do not move (i.e., θ is undefined and $d = 0$). In certain cases interpretation of these situations will be clear, for example, if one object stops moving, does the other? However in practice, many applications may not facilitate such straight-forward interpretation. For example, when studying urban travelers does stopping at a red-light signify a change to dynamic interaction even if they will eventually go straight? In light of these concerns, these assumptions can be modified to accommodate different situations that may arise in various movement scenarios to fit a given application.

5.4.3 Global analysis

A global version of the di statistic can be used to measure the overall interaction in a set of movement segments. First, it is useful to recognize that we can identify global interaction in azimuth or displacement individually by summing the interaction values for each individual segment and dividing by the number of segments. This form of a global DI gives equal weight to each segment. .

$$\text{DI}_\theta(V^a, V^b) = \frac{1}{n-1} \sum_{t=1}^{n-1} \text{di}_\theta \tag{5.8}$$

$$\text{DI}_d(V^a, V^b) = \frac{1}{n-1} \sum_{t=1}^{n-1} \text{di}_d \tag{5.9}$$

A global measure of overall dynamic interaction DI can also be derived.

$$\text{DI} = \sum_{t=1}^{n-1} \text{di}_\theta \times \text{di}_d \tag{5.10}$$

It is important to note that in the local version $di = di_\theta \times di_d$, but with the global statistic, due to summation rules, $DI \neq DI_\theta \times DI_d$. This can make interpretation of global values of DI less straightforward than with local values. However, if we were to alternatively define the global version as $DI = DI_\theta \times DI_d$, then the equation defined by (5.10) would no longer hold. Thus, interpretation of DI values is best done separately for each component (i.e., DI, DI_θ, and DI_d).

The global formulation is also appropriate for interval and episodal levels of analysis. Here we simply replace n with some interval or episode length $\hat{n}$, where $\hat{n} < n$. This type of analysis can be illuminating when analyzing interactions in larger movement datasets, where varying levels of dynamic interaction may occur at different points in the movement paths.

5.4.4 Time- and Distance-based Weighting

In instances where the sampling interval of the n fixes is unequal it is desirable to scale the statistic based on the temporal duration of each movement segment. In practice, this would give more weight to segments of longer duration and less weight to shorter segments. Temporal weighting may also be used to account for missing fixes, common to GPS-based tracking data. Let Δ_t correspond to the temporal duration of segment t, where T is the total duration of the entire movement path. Then a time weighted version of (5.10) is defined as:

$$\mathrm{DI} = \sum_{t=1}^{n-1} \frac{\Delta_t}{T} \mathrm{di}_\theta \times \mathrm{di}_d \tag{5.11}$$

Viewed in light of the uncertainty associated with movement data, this form of temporal weighting may be counter-intuitive. That is, it may be logical to assign weights inversely proportional to the duration between fixes; lower weights to segments with higher uncertainty (i.e., more time between fixes) and higher weights to segments with higher certainty or finer space-time resolution. Similarly, we can define a distance-based weighting scheme for (5.10) where movements with larger displacement have increased weight in calculation of the statistic. Varying distance-based weights could be used when dynamic interactions of a specific movement behavior are of interest. For example in the study of wildlife long directed movements are often interspersed with shorter random movements distinguishing migratory and foraging behavior (Turchin, 1998). Distance weighting could be used to tailor the measurement of dynamic interactions to either of migratory or foraging behaviors in this case. A possible distance-based weighting scheme would be the average displacement of two segments: $d_t^{avg} = (d_t^a + d_t^b)/2$, and $\sum d_t^{avg} = D$. Based on the average displacement a

distance-weighted version of (5.10) is defined as:

$$\mathrm{DI} = \sum_{t=1}^{n-1} \frac{d_t^{avg}}{D} \mathrm{di}_\theta \times \mathrm{di}_d \tag{5.12}$$

However, the average displacement of two objects movement segments is misleading when one object has a large displacement and the other has a small displacement. Thus, other distance measures are worth investigating for alternative distance-based weighting schemes, keeping in mind that the sum of the weights should equal one. The equations (5.11) and (5.12) can be combined to provide a time- and distance-based weighting scheme. It is important to note that time- and distance-based weighting is really only useful when interpreting global results when there is benefit to assigning segments weights based on duration or distance.

Another interesting extension to studying correlations in movement paths is when movements interact with a temporal lag, for example when trend-setting occurs, as described by Laube *et al.* (2005). The DI statistic can be modified to evaluate dynamic interactions at a temporal lag. To measure dynamic interactions at a temporal lag, select a time lag – k, where k is generally taken to be a multiple of the fix interval (i.e., if fixes are taken at even intervals the time between consecutive fixes). Then we can, alternatively define di_θ and di_d as:

$$\mathrm{di}_\theta = f_t(\theta_t^a, \theta_{t+k}^b) \tag{5.13}$$

$$\mathrm{di}_d = g_t(d_t^a, d_{t+k}^b) \tag{5.14}$$

The global statistics (DI, DI_θ, DI_d) can be computed as before, using the time lagged versions of di_θ (5.13) and di_d (5.14).

5.5 Data

5.5.1 Simulated Data

Six simulated data sets are used to highlight the utility of the DI statistic and the benefit of extensions it makes to r (Shirabe 2006). A single random walk ($n = 10$) is used to generate a movement path that is the bases for the simulation examples. We used manual permutations to the spatial coordinates of the original random walk

to produce five new movement paths that represent five unique dynamic interaction scenarios (Table 5.1). The first scenario simulates two objects moving with strong-positive dynamic interaction. The second scenario uses the same two paths as the first scenario, but one is rotated at 45°, simulating strong interaction in displacement, and low interaction in azimuth. The third scenario simulates positive interaction in azimuth and no interaction in displacement. The fourth scenario simulates negative interaction in azimuth and no interaction in displacement. The fifth scenario simulates no interaction in azimuth and strong interaction in displacement. The sixth scenario uses a second independent random walk to simulate random interactions between two moving objects.

Table 5.1: Simulated movement scenarios, depicting different types of dynamic interactions, used to examine the differences between the new interaction statistic (DI) and an existing method (r).

Scenario	**Azimuth (θ)**	**Displacement (d)**
1	positive interaction	interaction
2	positive interaction (rotated 45°)	interaction
3	positive interaction	no interaction
4	negative interaction	no interaction
5	no interaction	interaction
6	random	random

5.5.2 Athletes – Ultimate Frisbee

In team sports players (objects) movements are expected to be highly interactive. Often a defending player is tasked with "covering" an offensive player, and their movements are in reaction to that offensive player. In the sport of ultimate frisbee, offensive players move about the field in an attempt to get open for a pass from their teammates. Defending players cover them, in an attempt to intercept or dissuade passes from being completed. As such, in ultimate frisbee the movements of an offensive player and their defender are highly interactive. We used 5 Hz sports-specific GPS devices (GPSports, Fyshwick, Australia) to monitor the movements of two ultimate frisbee players over a one minute segment during a training game. In this example, the two players cover each other for the entirety of the one minute period. A total of $n = 276$ GPS locations (out of a possible 300) were simultaneously recorded. Most of the missing locations occur when the players are relatively stationary. At 5

GPS locations per second this represents an extremely detailed movement dataset, appropriate for investigating the intricate movements of athletes.

5.5.3 Grizzly Bears in Alberta, Canada

To further demonstrate DI, we investigate a previously published dataset containing GPS telemetry locations of a number of grizzly bears in Alberta, Canada (Stenhouse *et al.*, 2005). Stenhouse *et al.* (2005) revealed that various bear combinations showed evidence of dynamic interaction during different seasons, in particular male-female interactions were strongest during spring when mating activity occurs. To demonstrate DI, we examine one specific male-female bear combination that exhibited a relatively strong association during the mating season (male (G006) and female (G010) – see Figure 4 in Stenhouse *et al.* 2005). Grizzly bear GPS collars were programmed to obtain a location fix every four hours, however missing entries are frequent. As a result, only 112 simultaneous GPS fixes were obtained for the two bears during period from May 28, 2000 to July 08, 2000. In this example, we incorporate time-based weighting in order to account for unevenness in fix intervals (ranging from 4 hours to over 6 days).

5.6 Results

5.6.1 Simulated Data

Using the six simulated datasets we compared global values for DI, DI_θ, and DI_d with Shirabe's (2006) r statistic (Figure 5.4) to reveal both the similarities and differences between these two methods. In scenario 1, where both movements are highly interactive in both displacement and azimuth, DI and r are very similar. In scenario 2 DI and r are similar, however using the DI method we can identify that interaction is higher in displacement ($\text{DI}_d = 0.977$), and lower in azimuth ($\text{DI}_\theta = 0.664$). In contrast, scenario 3 reveals a situation where DI and r exhibit substantially different results. Using DI_θ and DI_d we can further examine the nature of the interaction in both azimuth and displacement, in this case $\text{DI}_d = 0.287$ and $\text{DI}_\theta = 0.992$. High DI_θ independent of DI_d could be useful in measuring interactive movement patterns via different modes of transportation (e.g., walking vs. biking), or scale independent movement behavior in wildlife. Scenario 4 demonstrates an example where negative

dynamic interaction is present (i.e., repulsion). In this case, DI is small and negative (DI = -0.278) due to low interaction in displacement ($DI_d = 0.280$), while r_{xy} is large and negative ($r = -0.805$). Scenario 5, shows the case where low DI is a function of low interaction in azimuth ($DI_\theta = -0.095$), despite having a strong level of interaction in movement displacement ($DI_d = 0.979$), while $r_{xy} = -0.532$. Measurement of high vs. low DI_d independent of DI_θ could be used in behavior analysis to identify objects with similar diurnal activity patterns (i.e., temporal patterns of long and short movements). In Scenario 6, both DI and r show values near 0, as would be expected from two independent random motions. It is interesting to note that $DI_d = 0.649$ is relatively high in this example, as the random walks used identical parameters for their displacement distributions.

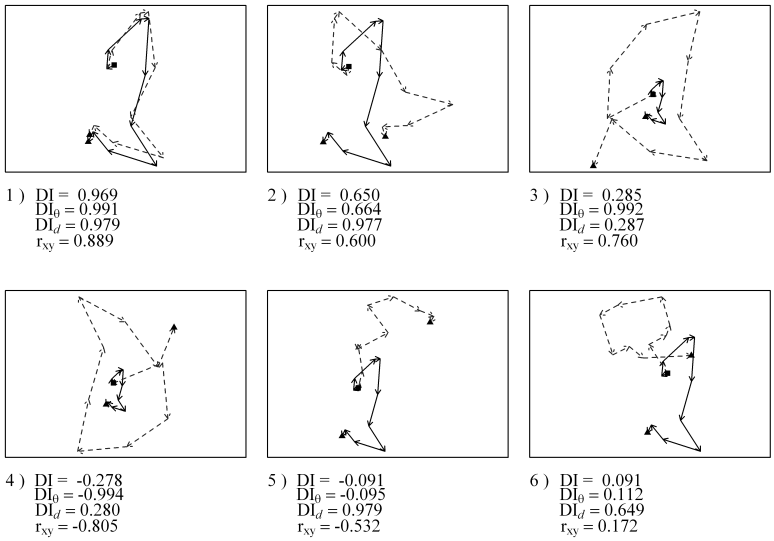

Figure 5.4: Results from global analysis of 6 simulated example scenarios, comparing the new DI method with the Shirabe (2006) correlation statistic – r. Original path is solid and black, while the path in dashed grey portrays variations based on six simulated scenarios (see Table 5.1).

5.6.2 Athletes – Ultimate Frisbee

In the Ultimate Frisbee example, the two players positively interact in movement azimuth ($DI_\theta = 0.682$) and movement displacement ($DI_d = 0.730$). The global statistic shows that a substantial level of interaction exists between the two athletes (DI = 0.572). Local analysis enables the identification of times/locations where the athletes exhibit more or less interactive movements (Figure 5.5). In the ultimate frisbee ex-

ample, local analysis is more informative than the global measure, as the movement path consists of many (shorter) movement segments. Maps of local di can be combined with a time-series graph of di, di_θ, and di_d related to times/locations during the game where the defending player did a poor job covering the offensive player. We use episodal level analysis to segregate the movement paths into episodes of high vs. low interaction in order to further investigate the interactive behavior of these two athletes. For example, from 0 – 20 and 38 – 40 seconds (highlighted in blue in Figure 5.5), high and positive di values suggest the defending player is providing good defensive coverage (for these two episodes DI = 0.757). While from 20 - 38 seconds (highlighted in red in Figure 5.5) di values are much lower, an indication of less interactive movement and poor defensive coverage (for this episode DI = 0.122).

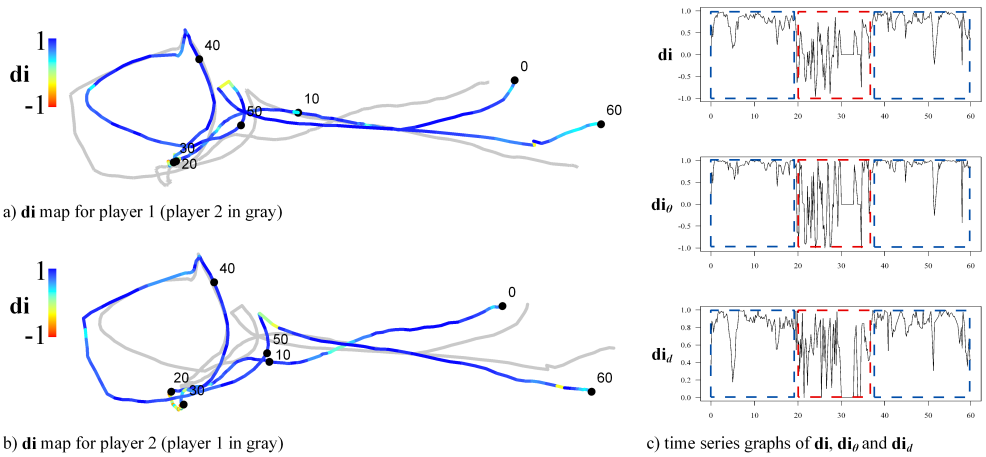

Figure 5.5: Local analysis showing maps of di values for a) player 1, and b) player 2, from the ultimate frisbee example. c) time series graphs of di, di_θ, and di_d can be used to identify periods of high and low dynamic interaction. Highlighted in blue in the time series graphs (c) are periods where player 1 does a good job covering player 2 (DI = 0.757). Highlighted in red is a period where the player 1 does a poorer job covering player 2 (DI = 0.122).

5.6.3 Grizzly Bears in Alberta, Canada

In the grizzly bear example it was revealed that the male and female bears showed substantial interaction (DI = 0.578) over the 42 day period from May 28, 2000 to July 8, 2000, using time-based weighting (see equation (5.11)) to account for missing fixes. Similarly, time weighted results for azimuth (DI_θ = 0.663) and displacement

($DI_d = 0.731$) reveal that both azimuth and displacement were strongly related during this period. Local analysis revealed that the strong interaction seen with the global results was a function of highly cohesive movements during the middle of June, while at the beginning of June the two animals show little interaction (see Figure 5.6). Again we perform analysis at the episodal level for separate periods identified visually from the local analysis as having low and high dynamic interaction (low interaction: May 28 – June 09; high interaction: June 09 – 29). The period of high interactions has a time-weighted DI = 0.492, while the period of low interaction has a time-weighted DI = 0.029. Highly interactive behavior by mating grizzly bears is common in this region, as males will attempt to confine female movements to a 'mating area' (Hamer and Herrero, 1990). Interpretation of maps and graphs of di facilitates the identification of where and when such behavior occurs.

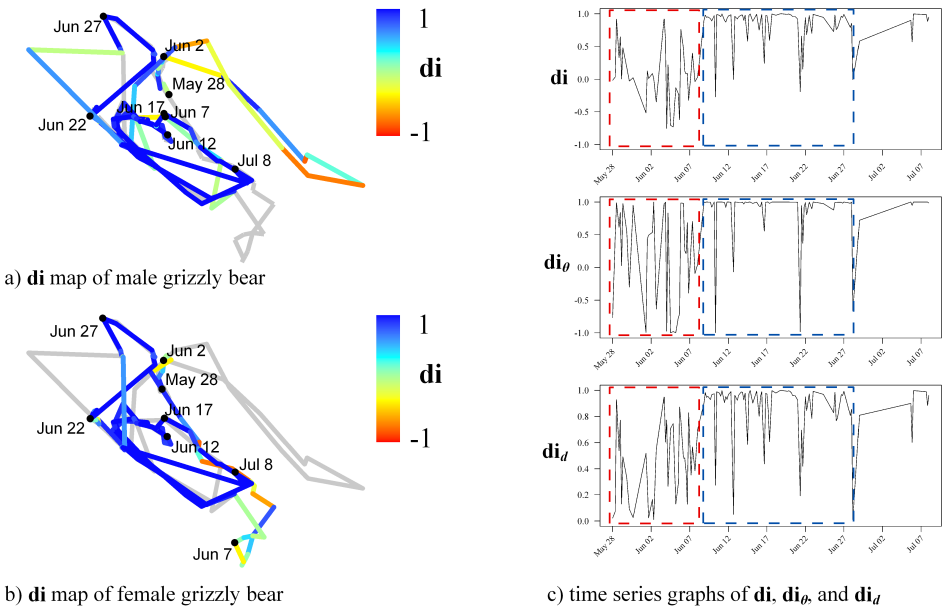

Figure 5.6: Local analysis showing maps of di values for a) the male grizzly bear (G006), and b) the female grizzly bear (G010), from the grizzly bear example. c) time series graphs of di, di_θ, and di_d can be used to identify periods of high and low dynamic interaction. Highlighted in red in the time series graphs (c) is a period where the bears exhibit low dynamic interaction (DI = 0.029). Highlighted in blue is period where the bears exhibit strong dynamic interaction (DI = 0.492), in this example indicative of mating behavior.

5.7 Discussion

DI has three fundamental advantages over an existing method (Shirabe, 2006) for measuring interactions (termed correlations) in movement data. First, the existing method follows a traditional correlation coefficient structure and is thus dependent on the mean vector of a movement vector time series. In most cases, this mean movement vector will have little relevance in the context of the analysis. However, in cases where interactions are expected to occur relative to some mean movement trajectory, the method from Shirabe (2006) is still advantageous. For instance, two objects moving radially from a point (at some acute angle) may exhibit dynamic interaction (e.g., Figure 4a in Shirabe, 2006). Second, DI is decomposed into components measuring interaction in movement azimuth and displacement. This property enables analysts to identify situations where movements are related in one component but not the other. For example, in scenario 3, DI_d is low, however strong interaction is present in DI_θ, indicating that the objects move with similar azimuths but not displacements, a conclusion not discernable from the r_{xy} statistic. Lastly, the di statistics we have developed are calculated independently for each simultaneous movement segment. The di values can be mapped and analyzed in a time-series fashion providing a local level analysis. Local analysis reveals spatial-temporal information about locations of increased or decreased interaction along the movement trajectory. Furthermore, the local level statistics (di, di_θ, and di_d) are easily aggregated to coarser levels of analysis (interval, episodal, and global).

Other research areas where measuring movement interactions could provide new and unique insight include transportation, human-activity, and other wildlife and sporting examples. In transportation applications measuring interactions in large movement databases could be used for generating information on commuter behavior. Examples from human-activity research where interactions are important include tourist behavior (e.g., Shoval and Isaacson, 2007) or crowd dynamics (Batty *et al.*, 2003). With wildlife movement data, the detection of interactions is important in the study of resource selection (Millspaugh *et al.*, 1998) and social behavior (Bandeira de Melo *et al.*, 2007; Kenward *et al.*, 1993), but also for examining offspring dependency, and inter-/intra-species behavior. Finally, a number of sporting examples exist where measuring movement interactions could provide new and unique insight including soccer, American football, and ice hockey.

We use simulated movement data to highlight the advantages of DI over an exist-

ing method in a small set of specific scenarios designed to show the range of dynamic interactions present in movement data. When two movements are highly interactive (e.g., scenario 1) both methods successfully identify the high level of dynamic interaction. Also, when two movements show opposing or repulsive movements (e.g., scenario 4) both methods are able to identify this behavior. The value of the DI method is demonstrated in scenarios 3, 4, and 5, where interactions in either azimuth or displacement are coupled with no interaction in the other component. This type of analysis may be useful, for example, when object movement is dependent on a temporal factor. For instance, many wildlife species are active only at specific times of the day and remain dormant during other periods. Measuring positive dynamic interactions in displacement, irrespective of azimuth, may be useful in identifying whether or not different species or individuals operate with similar circadian cycles (Merrill and Mech, 2003).

The example from athletes playing ultimate frisbee demonstrates the value of measuring dynamic interactions at the local and episodal levels of analysis. Local and episodal analysis revealed periods of varying degrees of dynamic interaction, which can be related to player performance (i.e., how well the defensive player was able to cover the offensive player). In many team sports, player evaluation has traditionally been conducted by human observers. More recently, data driven analyses have become common in the evaluation of players in team sports (e.g., Fearnhead and Taylor, 2011). When a player's movement can be directly related to specific abilities, for instance the soccer example in (Laube *et al.*, 2005), the measurement of dynamic interactions, using the DI method can enhance player evaluation using novel sport-specific movement datasets.

The DI method we have developed requires that movement locations be recorded simultaneously. Such a tidy form of movement data (i.e., where objects locations are recorded simultaneously) may not always be available, limiting the ability to implement this method. In such cases, path interpolation methods (e.g., Tremblay *et al.*, 2006) could be used to estimate the locations of one object at coinciding times. Similarly, in many applications the assumption that movement data are collected at a regular interval is not satisfied (e.g., with movement data collected using cell-phone records). This is also the case in many wildlife telemetry studies where missing fixes are common. In the grizzly bear example, we demonstrate the value of temporal weighting the DI statistic to account for uneven sampling intervals. Further, we highlighted how local and episodal analyses can provide unique and valuable insights into

the nature of dynamic interactions present in movement datasets. Local analysis reveals the times and locations of dynamic interactions not discernable from global level statistics. When comparing male and female grizzly bears, the dynamic interactions were likely due to mating behavior. This example demonstrates the value of quantifying dynamic interactions in wildlife movement datasets, as they can be related directly to specific social activities.

When movement data are collected at too fine a granularity, the movement process (e.g., dynamic interaction) can be masked by data noise (termed over-sampling, Turchin, 1998). In these cases, down-sampling can be used to reduce data redundancy in the movement path and improve the process signal to noise ratio. The DI statistics can then be computed on the re-sampled movement dataset, as another form of interval and/or episodal analysis (e.g., Laube *et al.*, 2007). Variations of this procedure at different interval and episodal scales can lead to increasingly complex and cross-scale investigations of dynamic interactions in moving object datasets. Recently, Laube and Purves (2011) have discussed the impact that movement data granularity (i.e., sampling resolution) has on metrics used to quantify and describe movement trajectories (e.g., mean speed). The DI method is similarly impacted by the granularity at which movement data are represented. For example, at a coarse granularity objects may exhibit positive dynamic interactions, while at a fine granularity their movements may show negative dynamic interaction (see Figure 5.7). Both the granularity at which the data are represented and analysis level selected will impact the results and subsequent interpretation of DI. One of Laube and Purves (2011) main recommendations is that movement data analysis be conducted across a range of scales (granularities and analysis levels) to correctly understand observed patterns.

5.8 Conclusions

Movement data are being collected for a variety of research agendas involving the study of humans, their vehicles, and wildlife. Central to analyzing movement data is the measurement of dynamic interactions between pairs of moving objects. We have developed a new statistic (DI) for measuring dynamic interactions in discrete movement data (e.g., with a GPS). The basic properties of movement segments – azimuth and displacement, are used to detect dynamic interactions in azimuth, displacement, and overall movement. The DI method can be applied at four analysis levels (local, interval, episodal, and global – Laube *et al.*, 2007) associated with move-

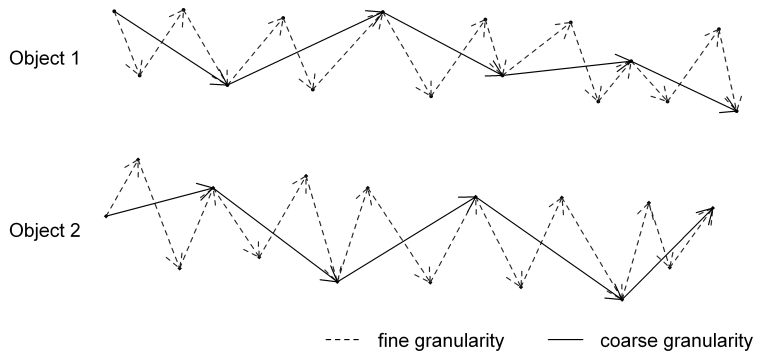

Figure 5.7: A pair of moving objects that exhibit negative dynamic interaction when analyzed at a fine granularity (dashed line, DI = -0.47) but positive dynamic interaction when analyzed at a coarser granularity (solid line, DI = 0.49). This example illustrates how changes in data granularity can impact results and interpretation of DI.

ment data, and results can be aggregated across analysis levels. We introduce both time- and distance-based weighting schemes that can be useful in specific situations. The measurement of dynamic interactions at a temporal-lag, an example of trend-setting (Laube *et al.*, 2005), can be easily incorporated. Like many spatial analysis techniques the DI method is impacted by the granularity at which movement data is represented. A detailed investigation of cross-scale effects is warranted to provide a better understanding of how the measurement of dynamic interaction is impacted by changing data granularities.

In some situations the nature of movement interactions will not simply involve two moving objects, but rather involve two moving objects impacted by a third. Consider the grizzly bear example; the bears exhibit varying levels of dynamic interaction over the course of the time period. The level of interaction is likely affected by their position relative to the location of other objects, including other bears, roads, or sources of attraction or repulsion (i.e., food or danger). Future research will develop approaches for measuring third-party interactions, whereby pairs of moving objects interact with respect to a third stationary or moving object. To those wishing to measure dynamic interactions with their own applications we have developed code for implementing DI in the statistical software package R (R Development Core Team, 2012), for more information please visit: `http://www.geog.uvic.ca/spar/tools.html`

Chapter 6

A Critical Examination of Indices of Dynamic Interaction for Wildlife Telemetry Studies

Abstract

Wildlife scientists continue to be interested in studying ways to quantify how the movements of animals are inter-dependent – dynamic interaction. While a number of applied studies of dynamic interaction exist, little is known about the comparative effectiveness and applicability of available methods used for quantifying interactions among animals. We highlight the formulation, implementation, and interpretation of the suite of eight currently available indices of dynamic interaction. Point- and path-based approaches are contrasted to demonstrate differences between methods and underlying assumptions on telemetry data. Correlated random walks were simulated at a range of sampling resolutions to generate scenarios of unexpected or expected dynamic interaction behaviour. We evaluate the effectiveness of each index at identifying differing types of interactive behaviours at each sampling resolution. Each index is then applied to an empirical telemetry dataset of three white-tailed deer (*Odocoileus virginianus*) dyads. Results from the simulated data show that three indices of dynamic interaction reliant on statistical testing procedures are susceptible to Type I error, an effect that increases with fine sampling resolutions. In the white-tailed deer examples, a recently developed index for quantifying local-level cohesive movement behaviour (the di index) provides revealing information on the presence of

infrequent and varying interactions in space and time. Point-based approaches implemented with finely sampled telemetry data over-estimate the presence of interactions (Type I errors). Indices producing only a single global statistic (7 of the 8 indices we tested) are unable to quantify infrequent and varying interactions through time. The quantification of infrequent and variable interactive behaviour has important implications for the spread of disease, and the prevalence of social behaviour in wildlife. Guidelines are presented to inform researchers wishing to study dynamic interaction patterns in their own telemetry datasets. Finally, we make openly available a set of tools, in the statistical software R, for computing each index of dynamic interaction presented herein.

6.1 Introduction

The development of GPS tracking technologies is revolutionizing wildlife movement and behaviour research (Cagnacci *et al.*, 2010) and has led to increased interest in the study of interactions among individual animals, for example, mating behaviour (Stenhouse *et al.*, 2005) and predator-prey dynamics (Eriksen *et al.*, 2008). Interactive behaviour can be characterized as either static or dynamic (Macdonald *et al.*, 1980). Static interaction can be defined simply as the joint-space use between two individuals (Kernohan *et al.*, 2001), and is typically measured by an index of home range overlap or volume of intersection using utilization distributions (Millspaugh *et al.*, 2004). Alternatively, dynamic interaction refers to how the movements of two individuals are related (Macdonald *et al.*, 1980), or as the inter-dependency in the movement of two individuals (Doncaster, 1990). Tests for dynamic interaction can be used to examine attraction or avoidance behaviour (Doncaster, 1990), simultaneous joint-space use (Minta, 1992), or cohesiveness in the movements of two individuals (Long and Nelson, 2013a). A number of techniques for studying dynamic interactions have been developed and adopted widely in wildlife telemetry studies (see Table 6.1), but little is known about the effectiveness of each at identifying true dynamic interaction patterns. Similarly, it is difficult for researchers to compare results among existing methods because most applied studies typically implement only a single method.

Novel tracking technologies (e.g., GPS, Argos, Platform terminal transmitters (PTTs), global satellite Iridium systems, etc.) are changing the manner in which wildlife telemetry data are recorded and transmitted. A wide variety of systems can be programmed to collect telemetry fixes at variable sampling resolutions (Tomkiewicz

Table 6.1: Selected examples of applications involving the study of dynamic interactions using wildlife telemetry data.

Index	**Species**	**Data**	**Study Objective**	**Citation**
Prox	white-tailed deer (*Odocoileus virginianus*)	VHF	Parturition	Bertrand *et al.* (1996)
–	maned wolves (*Chrysocyon brachyurus*)	GPS	Familial bonds	Bandeira de Melo *et al.* (2007)
Ca	maned wolves (*Chrysocyon brachyurus*)	VHF	Inter- & intra-sex behaviour	de Almeida Jácomo *et al.* (2009)
–	red wolves (*Canis rufus*)	GPS	Sociality, group behaviour	Karlin and Chadwick (2011)
Don	wood mice (*Apodemus sylvaticus*)	VHF	Mating	Tew and Macdonald (1994)
–	raccoons (*Procyon lotor*)	VHF	Philopatry	Gehrt and Fritzell (1998)
–	coyote (*Canis latrans*)	VHF	Inter- & intra-sex behaviour	Chamberlain *et al.* (2000)
–	badgers (*Meles meles*)	VHF	Sociality, group behaviour	Böhm *et al.* (2008)
Cs	lynx (*Lynx canadensis*)	VHF	Inter- & intra-sex behaviour	Poole (1995)
–	red & grey squirrels (*Sciurus vulgaris & carolinensis*)	VHF	Interspecific	Kenward and Hodder (1998)
–	wolves & moose (*Canis lupus & Alces alces*)	GPS	Interspecific	Eriksen *et al.* (2008)
Lixn	badgers (*Taxidea taxus*)	VHF	Inter- & intra-sex behaviour	Minta (1993)
–	grizzly bear (*Ursus arctos*)	VHF	Inter- & intra-sex behaviour	Mace and Waller (1997)
–	mountain lion (*Puma concolor*)	VHF	Familial bonds	Nicholson *et al.* (2011)
HAI	coyote (*Canis latrans*)	VHF	Management and control	Bromley and Gese (2001)
–	coyote (*Canis latrans*)	VHF	Habitat relations	Atwood and Weeks (2003)
Cr	caribou (*Rangifer tarandus*)	VHF	Sociality, group behaviour	Shirabe (2006)
DI	grizzly bear (*Ursus arctos*)	GPS	Mating	Long and Nelson (2013a)

et al., 2010). Sampling resolutions (the frequency at which telemetry fixes are collected) previously unattainable are now routinely implemented in modern tracking studies. However, many studies still employ coarser resolution telemetry systems (e.g., VHF telemetry), which have a lower unit cost in order to monitor a greater number of individuals (Girard *et al.*, 2006). In addition to the wide range of technologies, study objectives typically dictate sample size, sampling resolution, and study design, all of

which may influence the ways in which data are analyzed and interpreted. Therefore, researchers are often left with the difficult task of identifying which, of a suite of available analytical techniques, are appropriate for meeting study objectives.

The overall goal of this study is to demonstrate the effectiveness of eight available indices for measuring dynamic interaction common to wildlife telemetry data using both simulated and empirical data. Our objectives were to: 1) review the formulation and interpretation of each of the eight indices, 2) detect expected vs. unexpected dynamic interaction using simulated data at varying sampling resolutions, 3) evaluate each method using empirical data collected on white-tailed deer fitted with GPS collars, 4) highlight the advantages and disadvantages of each approach, and 5) provide guidance on the selection, use, and interpretation of dynamic interaction indices common to analysis of wildlife telemetry data. Areas of future research are discussed to encourage the development of additional tools and algorithms that can be used in association with dynamic interaction analysis. Last, we provide a tool for the R statistical computing environment that allows researchers to implement each of the eight indices of dynamic interaction presented herein.

6.2 Indices of Dynamic Interaction

Indices of dynamic interaction can be broadly categorized as point-based or path-based, based on how they represent telemetry data (as points, or as connected segments paths). Clear conceptual differences in the calculation and interpretation of the eight dynamic indices are apparent from their individual formulation (see below and Table 6.2), but also between point- and path-based approaches. Point-based indices typically examine attraction/avoidance behaviour, while path-based indices look at cohesive movement behaviour. Of the eight currently available indices of dynamic interaction, six are point-based, while two are path-based (Table 6.2). The terminology and notation used for describing telemetry data and concepts relating to measurement of dynamic interaction is introduced in Table 6.3.

6.2.1 Proximity analysis

Dynamic interaction is most simply quantified as a measure of nearness in space. Researchers have used proximity analysis (Prox) to understand the frequency at which

Table 6.2: Eight indices of dynamic interaction for wildlife telemetry data. Refer to Table 3 for terminology. In all indices, except for Lixn, simultaneous fixes ($T_{\alpha\beta}$) are determined using a temporal threshold (t_c) and d_c is a threshold distance for proximal fixes ($S_{\alpha\beta}$).

Index	**Reference**	**Sig. Test**	**Data**	**Tests**	**Interpretation**
Prox	–	–	Point	Ratio of $ST_{\alpha\beta}$ fixes to $T_{\alpha\beta}$ fixes, based on d_c.	Prox can be interpreted much like Ca (see below) and is similarly based on d_c.
Ca	Cole (1949)	–	Point	The proportion of all fixes that are $ST_{\alpha\beta}$ based on d_c.	Ca > 0.5 - attraction; Ca < 0.5 - no association.
Don	Doncaster (1990)	χ^2 test	Point	If the distribution of distances of $T_{\alpha\beta}$ is different than the distances of permutations of all fixes.	Based on the contingency table and a χ^2 test looks for significant attraction in $ST_{\alpha\beta}$ for a given d_c.
Lixn	Minta (1992)	χ^2 test	Point	The simultaneity of usage of the shared area of each home range.	Lixn > 0 shared use is simultaneous (attraction); Lixn < 0, shared use is solitary (avoidance); Lixn ~ 0 shared use is random.
Cs	Kenward *et al.* (1993)	Wilcoxon signed-rank test	Point	For differences between distances of $T_{\alpha\beta}$ and distances of permutations of all fixes.	Cs ~ 1 - attraction; Cs ~ -1 - avoidance.
HAI	Atwood and Weeks (2003)	–	Point	Number of $ST_{\alpha\beta}$ fixes within the shared area of the home range against solitary use of shared area.	HAI ~ 1 - attraction; HAI ~ 0 - avoidance.
Cr	Shirabe (2006)	–	Path	Correlation of movement segments tested against respective path means. Identical to Pearson correlation statistic (r).	Cr ~ 1 - positive correlation (cohesion); Cr ~ -1 - negative correlation (opposition); Cr ~ 0 – no correlation (random).
DI	Long and Nelson (2013a)	–	Path	Cohesion in individual movement segments (global and local), with respect to distance and direction.	DI ~ 1 cohesive movement; DI ~ -1 - opposing movement; DI ~ 0 - random movement.

Table 6.3: Terminology and notation used for describing telemetry data and dynamic interaction methods.

Term	Explanation
α or β	Individuals of a dyad (telemetry data)
dyad	Pair of individuals (α and β)
fix	A telemetry record (spatial location and time stamp)
t_c	Time threshold
d_c	Distance threshold
$\mathrm{T}_{\alpha\beta}$	Temporally simultaneous fixes based on t_c
$\mathrm{S}_{\alpha\beta}$	Spatially proximal fixes based on d_c
$\mathrm{ST}_{\alpha\beta}$	Spatially proximal and temporally simultaneous fixes based on d_c and t_c
v_t, w_t	Movement segment, vector connecting two consecutive fixes
$\hat{V}$, $\hat{W}$	Mean movement segment for an entire path

two individuals are near each other. The simplest such index is the proximity rate:

$$\mathrm{Prox} = \frac{ST_{\alpha\beta}}{T_{\alpha\beta}} \tag{6.1}$$

where Prox measures the proportion of simultaneous fixes (defined using t_c) that are spatially proximal (based on threshold d_c). The value of implementing Prox is that it is easily interpreted and indicates whether or not two individuals spend time in close contact (Table 6.2). Prox has been used as an indicator of attraction between individuals (e.g., Bertrand *et al.*, 1996), and as an estimate of contact rates, which is useful when studying disease spread dynamics (Baker and Harris, 2000). The requirement of a distance threshold is both advantageous and problematic. In some cases, such as with contact rates, a biologically motivated spatial distance may be used. However, in many cases, it will be chosen subjectively by the researcher, owing to previously used thresholds in the literature, or some other property of the data such as error or sampling interval. In these cases, the subjectivity of the chosen threshold will impact the results as various choices for d_c will change Prox results.

6.2.2 Coefficient of association

The coefficient of association (Ca) was first introduced by Cole (1949) for measuring interspecific associations in field samples and has since been identified as a potential measure of dynamic interaction in wildlife telemetry data (Bauman, 1998). Coefficient

of association is calculated as:

$$\text{Ca} = \frac{2AB}{A+B} \tag{6.2}$$

where AB is the number of $\text{ST}_{\alpha\beta}$ fixes and A (resp. B) is the total number of all fixes in α(resp. β). Ca is similar to Prox, only Ca measures the rate of all fixes that are $\text{ST}_{\alpha\beta}$, not just the simultaneous fixes; thus Ca is measuring the same phenomenon as Prox. Typically, $\text{Ca} > 0.5$ indicates attraction, while $\text{Ca} < 0.5$ indicates no association (Kernohan *et al.*, 2001, Table 6.2). Like Prox, Ca is a useful indicator of attraction and contact rates, as defined by the threshold d_c. However, it is similarly affected by the subjectivity with which d_c is determined.

6.2.3 Coefficient of sociality

The coefficient of sociality (Cs) was proposed by (Kenward *et al.*, 1993) as an alternative measure of attraction using the raw distances between fixes, rather than a user-defined threshold. The formulation of Cs is a variant of Jacobs' index (Jacobs, 1974), a metric originally proposed for measuring food selection by wildlife. Cs is calculated as:

$$\text{Cs} = \frac{d_E - d_O}{d_E + d_O} \tag{6.3}$$

where d_O is the mean spatial distance between $\text{T}_{\alpha\beta}$ fixes and d_E is the expected mean distance, based on n^2 permutations of the $\text{T}_{\alpha\beta}$ fixes. The statistic is symmetric (on $[-1, 1]$) where positive values suggest attraction while negative values suggest avoidance. A Wilcoxon signed-rank test can be used to examine the significance of the resulting Cs value (Table 6.2). A major limitation is determining a reasonable expectation of mean distance to test against. Generally, the n^2 permutations of all $\text{T}_{\alpha\beta}$ telemetry fixes are used, but this can be misleading especially in the absence of static interaction. Others have suggested that d_E can be determined via simulations using, for example, correlated random walks (Miller, 2012).

6.2.4 Doncaster's non-parametric test

Doncaster (1990) proposed a non-parametric test for interaction (Don) by examining the separations between the n $\text{T}_{\alpha\beta}$ fixes and the unpaired $n^2 - n$ permutations of the $\text{T}_{\alpha\beta}$ fixes, and is analogous to the Knox test for space-time clustering (Knox, 1964). The cumulative distribution of the $\text{T}_{\alpha\beta}$ fix distances can be compared graphically with the cumulative distribution of the $n^2 - n$ permutated distances. For example,

Don is useful for determining a suitable distance threshold (d_c) by identifying where the $T_{\alpha\beta}$ plot is below the expected line based on the permutations. Upon selecting a suitable d_c value, a contingency table can be constructed, identifying the number of $T_{\alpha\beta}$ and non-$T_{\alpha\beta}$ (termed 'unpaired') fix distances that are above and below the threshold d_c. A χ^2 test (with 1 d.f.) or a binomial test can be used to examine the statistical significance of the counts of $T_{\alpha\beta}$ and non-$T_{\alpha\beta}$ distances above and below d_c (Table 6.2). A modified version of Don replaces the expectations derived from the $n^2 - n$ permutations of the $T_{\alpha\beta}$ fixes with a simulation procedure based on correlated random walks (White and Harris, 1994).

6.2.5 Minta's test for spatial and temporal interaction

Minta (1992) introduced three statistics (L_{AA}, L_{BB}, and Lixn) for examining the spatial and temporal interactions between two individuals. The Minta statistics require the calculation of individual home ranges a priori, which are then divided (using a spatial intersection) into three areas: 1) belonging to α only, 2) belonging to β only, and 3) shared by α and β (often termed the overlap zone; see Table 6.3 for notations). The first two statistics (L_{AA} and L_{BB}) represent spatial interaction statistics, and examine how each individual utilizes the shared area. The number of fixes contained in each area (i.e., α's area, β's area, and the shared area) are tested against expectations representing the probability of finding the animal in each of these areas. Expectation probabilities can be derived by using either the proportions of all fixes contained in each area, or the overlap area percentages (see Minta, 1992). Essentially, L_{AA} (L_{BB} respectively) tests how the individual uses the independent vs. shared area of its home range. L_{AA} (L_{BB}) ranges from -1 to 1 with negative values indicating spatial avoidance of the shared area, positive values indicating spatial attraction to the shared area, and values near 0 indicating random use of the shared area.

A more suitable dynamic interaction index is the Lixn statistic, which can be interpreted as a temporal interaction index. Lixn is derived using the same expectation probabilities as with L_{AA} (and L_{BB}). The Lixn statistic is a function of the ratio of simultaneous use and avoidance of the shared area to that of the solitary use of the shared area, and is calculated by:

$$\text{Lixn} = log\left(\frac{\left(\frac{n_{\alpha\beta}}{p_{\alpha\beta}} + \frac{n_{00}}{p_{00}}\right)}{\left(\frac{n_{\alpha 0}}{p_{\alpha 0}} + \frac{n_{0\beta}}{p_{0\beta}}\right)}\right) \tag{6.4}$$

where n represents the number of observed fixes, p is the expectation probability, and the subscripts α and β signify each individual's presence in the shared area, while the subscript 0 signifies absence from the shared area. Thus, Lixn measures the simultaneous use of the shared area. Positive Lixn suggest simultaneous use of shared area (attraction), while negative values indicate solitary use of the shared area (avoidance). Lixn near 0 indicates indifference or random use of the shared area. The Lixn statistic can be tested for significance using a χ^2 test with 1 d.f. from the contingency table of observed and expected frequencies of use of the shared area (Table 6.2).

6.2.6 Half-weight association index

The half-weight association index (HAI - Brotherton *et al.*, 1997) represents a companion test to the Minta (1992) Lixn temporal interaction statistic (Atwood and Weeks, 2003). As in Lixn, HAI is based on the shared area between the two individual home ranges (the overlap zone). The HAI statistic is however a more localized approach, focusing only on those fixes contained in the shared area (or only on the area of static interaction). HAI is calculated as:

$$\mathrm{HAI} = \frac{n}{n + \frac{x+y}{2}} \tag{6.5}$$

where n is the number of $\mathrm{ST}_{\alpha\beta}$ fixes in the shared area, and x and y are the number of solitary fixes (for α and β, respectively) within the shared area. Values near 1 indicate attraction (within the shared area) and values near 0 indicate avoidance (within the shared area; Table 6.2). HAI is computed identically to Ca, but only for those fixes contained in the shared area of the home range. Thus, for a dyad consisting of two individuals with identical home ranges, Ca = HAI.

6.2.7 Correlation index

Shirabe (2006) introduced a correlation index (Cr) for analyzing movement data, which can be considered a type of path-based measure of dynamic interaction. The Cr index takes the form of a Pearson product-moment correlation statistic for multivariate data (in this case bivariate in the two spatial dimensions, x and y). With Cr, movement data are represented as time-series with vectors corresponding to movement segments that connect consecutive fixes (see Table 6.3). Cr measures differences

in corresponding segments with respect to overall path means to determine the correlation structure of the data. Cr is calculated as:

$$\mathrm{Cr} = \frac{\sum_{t=1}^{n-1}(v_t^a - \bar{v}^a) \cdot (v_t^b - \bar{v}^b)}{\sqrt{\sum_{t=1}^{n-1}|v_t^a - \bar{v}^a|^2}\sqrt{\sum_{t=1}^{n-1}|v_t^b - \bar{v}^b|^2}} \tag{6.6}$$

where v_t and w_t represent movement segments (for α and β) corresponding to time t, and $\bar{v}$ and $\bar{w}$ (means of path vectors). Cr is interpreted similarly to other correlation statistics; values range from -1 to 1 where positive values indicate stronger correlation (cohesive movement), negative values indicate negative correlation (opposing movement), and values near 0 indicate random movement with respect to the other individual (Table 6.2).

6.2.8 Dynamic interaction index

Recently, Long and Nelson (2013a) introduced a dynamic interaction index (DI) based on path-based methods, similar to earlier attempts of Shirabe (2006). The DI index attempts to measure the cohesiveness of corresponding movement segments. DI is constructed as the mean of a localized version (termed di). Here, di is calculated as

$$\mathrm{di} = \left(1 - \frac{\left|d_t^\alpha - d_t^\beta\right|}{d_t^\alpha + d_t^\beta}\right) \times \cos\left(\theta_t^\alpha - \theta_t^\beta\right) \tag{6.7}$$

whereas DI is calculated by

$$\mathrm{DI} = \frac{1}{n-1}\sum_{t=1}^{n-1}\mathrm{di} \tag{6.8}$$

where d is displacement and θ (the direction) of segment t for individual α or β. The localized di is simply the product of terms measuring cohesiveness in displacement and direction for each corresponding segment. Thus, di is capable of separately measuring cohesiveness in the distance and direction components (di_d and di_θ respectively), which can be averaged into global statistics (see Long and Nelson, 2013a, for more details). Temporal trends in di can be used to identify periods of cohesive, opposing, and random movement within a dyad. The DI approach, as in Cr, measures cohe-

siveness irrespective of proximity between corresponding movement segments (Table 6.2). Thus, to justify DI analysis, the researcher is required to have some a priori expectation of cohesive movement, perhaps from one of the distance-based measures of dynamic interaction (e.g., Prox or Ca), a measure of static interaction (e.g., home range overlap), or simultaneous capture (e.g., familial groups).

6.3 Testing Indices

6.3.1 Simulated Data

In order to test methods of dynamic interaction, we simulated situations where dynamic interaction would be expected vs. unexpected. Generating movement dyads with unexpected dynamic interaction can be accomplished by simulating two independent correlated random walks (CRW) (White and Harris, 1994; Miller, 2012). In the unexpected scenarios, the origin of the second independent CRW was chosen to be a random location within the bounding box of the first CRW (see Figure 6.1a). Unfortunately, no algorithm has been developed for generating pairs of trajectories where dynamic interaction would be expected. Therefore, we propose the following procedure for simulating trajectories where dynamic interaction is expected. First, simulate a single CRW. Second, generate a second trajectory from the first by inducing a small amount of random noise to the spatial coordinates of each telemetry fix. This produces a relatively strong level of dynamic interaction over the entire path (dependent on the variance of the noise). A more realistic scenario is one where the amount of dynamic interaction varies along the trajectory. To generate a scenario where dynamic interaction varies along the trajectory, we separated the original trajectory into two intervals (of which the length varied between 30-70% of the entire trajectory). In the first interval, we simulated strong dynamic interaction created via the random noise procedure, and in the second interval, we simulated unexpected dynamic interaction using an independent CRW (Figure 6.1b). In all cases, the parameters for the simulated CRW (h - step length parameter, r - turning angle correlation parameter, and n - number of fixes) were selected randomly from the values $h = 1 - 5$, $r = 0 - 0.7$, and $n = 1000 - 2000$. In total, 1000 simulated scenarios were created to serve as a testing dataset, each containing two simulated dyads, one with expected dynamic interaction and one with unexpected dynamic interaction. In this context, Type I error will refer to when an index identifies dynamic interaction in an

unexpected scenario (false positive), and Type II error when an index fails to identify dynamic interaction in an expected scenario (false negative).

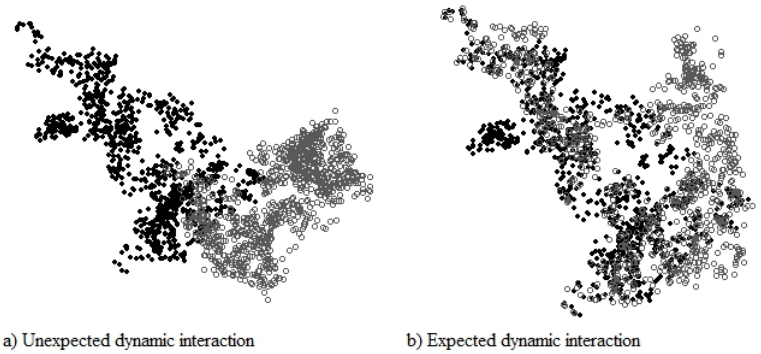

Figure 6.1: Depictions of two simulated scenarios (dyads), each containing two individuals where dynamic interaction would be unexpected (a), and where dynamic interaction would be expected (b).

In order to examine the effect of varying sampling resolutions on indices of dynamic interaction, we systematically down-sampled the simulated trajectories representing high resolution telemetry data, to two coarser levels (medium resolution – 25% of the original fixes, and coarse resolution – 10%, of the original fixes). This resulted in 1000 scenarios at each of three sampling resolutions (i.e., high, medium, coarse), each containing one dyad with expected dynamic interaction, and the other with unexpected dynamic interaction.

6.3.2 Empirical Data: White-tailed Deer GPS Telemetry Data

We collected data on white-tailed deer from two study areas in south-central Oklahoma, USA. Study site 1 was 1,214 ha in size, and was surrounded by a 15-strand, high-tensile electric fence, thus restricting movement across property boundaries (2.5-m tall; Webb *et al.*, 2009). Study site 2 was 1,861 ha and consisted of 5-strand barbed-wire fences, which allowed deer to cross property boundaries unrestricted. Vegetation on both study sites was consistent with that of the Cross Timbers and Prairies ecoregion (Gee *et al.*, 1994). On both study sites, we captured deer during JanuaryMarch (1998-2005 on study site 1; 2010-2012 on study site 2) using modified drop-net systems (Gee *et al.*, 1999). We sedated deer using intramuscular injections of telazol (4.4 mg/kg) and xylazine (2.2 mg/kg), and used yohimbine as an antagonist at 0.125 mg/kg. We fitted deer with GPS collars (ATS G2000 remote-release

collars; Advanced Telemetry Systems, Inc., Isanti, MN) programmed to collected 1 fix every 15 (study site 1) or 30 minutes (study site 2). On study site 1, collars were capable of collecting data for ~ 3 months, and on study site 2, collars collected data for ~ 6 months. All capture, handling, and marking procedures were consistent with the guidelines of the American Society of Mammalogists (Gannon *et al.*, 2007) and were approved by permit from the Oklahoma Department of Wildlife Conservation.

Three pairs (i.e., dyads) of white-tailed deer were used to further demonstrate the performance of each index of dynamic interaction using empirical telemetry data. We began with high-resolution GPS data recorded at a 30 min sampling resolution[1], but systematically resampled data to reflect coarser fix intervals (i.e., 6 and 24 hours); which is a common practice (Webb *et al.*, 2010). Final evaluation of dynamic interaction measures occurred at three sampling resolutions: 30 min, 6 hour, and 24 hours.

We delineated 95% volume contour home ranges using the kernel density estimate (Worton, 1989) and the *ad hoc* bandwidth, which assumes the resulting density surface is bivariate normal (Silverman, 1986). For each dyad, we calculated the area of overlap of the two home ranges (interpreted as the proportion of home range overlap – AOP; Millspaugh *et al.*, 2004) as a measure of static interaction that can be used as an a priori indicator of the potential for dynamic interaction in dyads. Similar to other studies, our hypothesis on the presence of dynamic interaction behaviour is based on a hierarchical approach where we first examine static interaction (home range overlap) between individuals (Figure 6.2), and subsequently look at finer resolution dynamic interactions. We predict little dynamic interaction in dyad 1 (AOP = 0.17), but greater dynamic interaction in dyads 2 and 3 (AOP = 0.67 and 0.57, respectively). We test these a priori predictions at all three sampling resolutions.

6.3.3 Calculating measures of dynamic interaction

All eight indices of dynamic interaction (Table 6.2) were computed for each of the dyads in the simulation study and using empirical GPS data from white-tailed deer. Several of the indices required the selection of parameter thresholds for identifying $T_{\alpha\beta}$ fixes and $S_{\alpha\beta}$ fixes. A t_c threshold of 1/2 the sampling resolution was used to determine simultaneous fixes, for example $t_c = 15$ min was used with the 30

[1]Dyads 1 and 3 were tracked using a 15 min sampling resolution, but here we resampled the temporal resolution to 30 min for consistency with dyad 2.

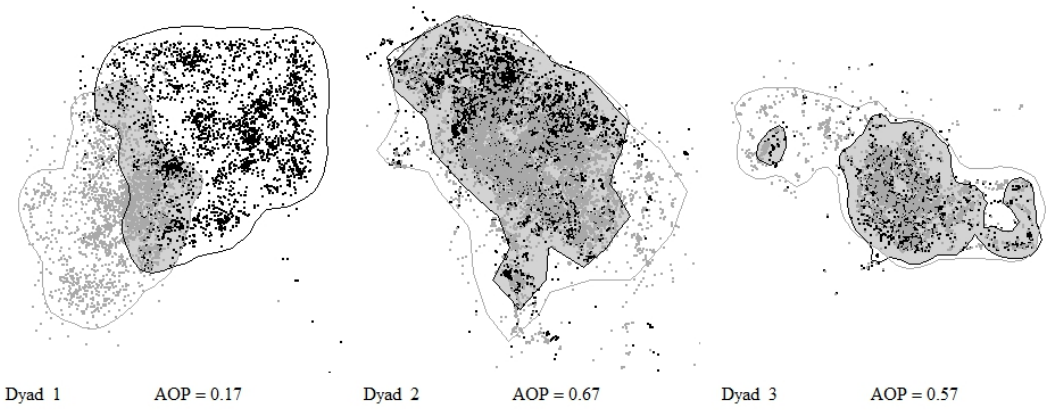

Figure 6.2: Empirical GPS telemetry data for three white-tailed deer dyads ($n = 6$ deer). Contours (grey and black polygons) represent 95% volume contour home ranges using kernel density estimates, along with static interaction measured as the area of overlap proportion (AOP) of the two individual home ranges. AOP is depicted as the grey shaded region. Deer in dyads 1 and 3 were tracked for approximately 3 months; while in dyad 2, deer were tracked approximately 6 months.

min white tailed deer telemetry data. In the simulation study, a distance threshold parameter of $d_c = 2 \times h$ was used, where h is the randomly selected movement length parameter in the CRW. Previous research on deer interactions used $d_c = 24$ m (Bertrand *et al.*, 1996). We selected a more conservative value of $d_c = 50$ m based on visual observations of deer and because there would be greater potential for identifying potential dynamic interactions at this spatial resolution for testing purposes. However, depending on specific hypotheses to test, the spatial threshold may be adjusted higher or lower, and comparison across spatial thresholds could help identify the scale at which processes are occurring across the landscape.

We evaluate results of the 1000 simulations, examining the mean and standard deviation of results from each index, for both the expected and unexpected scenarios, along with the number of significant values where appropriate. Histograms of index values for both the expected and unexpected dynamic interaction scenarios are presented to highlight the distribution of results. These results are used to explore the effect of sampling resolution on each dynamic interaction index and to compare among the various indices for measuring dynamic interaction in telemetry studies.

In the white-tailed deer examples, we examine each index in finer detail for each individual dyad. We explore the use of the local di statistic (Sec. 2.8) to examine temporal variations in dynamic interaction behaviour in these three dyads. At the

finest sampling interval (such as the 30 min sampling resolution here), the time-series plots of di can be noisy, making interpretation of patterns difficult. To circumvent this problem, we present the time-series plot of di for the 30 min data using a 24 hour moving average.

6.4 Results

6.4.1 Simulated Data

Using simulated data for expected and unexpected dynamic interactions provided a useful means for assessing the differences and similarities between the eight indices of dynamic interaction. Histograms for HAI, Prox, and Ca reveal that these three indices are each capable of identifying cases of expected and unexpected dynamic interaction (Figure 6.3). All three indices are comparable in terms of interpretation, but the Prox and Ca indices are essentially identical. The Prox, Ca, and HAI indices all appear to be robust to changes in sampling resolution because index values do not change markedly across sampling resolutions (Figure 6.3). The Don statistic also performed well, identifying interaction (i.e., attraction) in all 1000 of the expected scenarios (Figure 3). While seemingly positive, Don also is subject to Type I errors; dynamic interaction was identified in the unexpected simulations, which was more pronounced at the finer sampling resolutions, albeit minimal in occurrence (38 of 1000 simulations; Figure 6.3).

The Cs index also was successful at identifying correctly the presence of dynamic interaction in nearly all of the expected scenarios producing only minimal Type II error (Figure 6.3). However, Cs was susceptible to Type I error, which was more severe than Don in that Cs identified 344 of 1000 unexpected cases as dynamic interaction at the finest sampling resolution. The raw index values were consistent across sampling resolutions, and the histograms offer users a visual depiction of dynamic interaction, albeit qualitative in comparison to calculation of p-values.

The Lixn statistic performed poorest of all the eight indices in that it was highly susceptible to both Type I and Type II errors. Lixn failed to correctly identify dynamic interactions in many of the expected scenarios (Type II error), but tended to be reduced at finer sampling resolutions. More problematic is the fact that Lixn also produces a high level of Type I errors; Lixn identified $> 70\%$ (722/1000) of the unexpected scenarios as having significant dynamic interaction at the highest resolu-

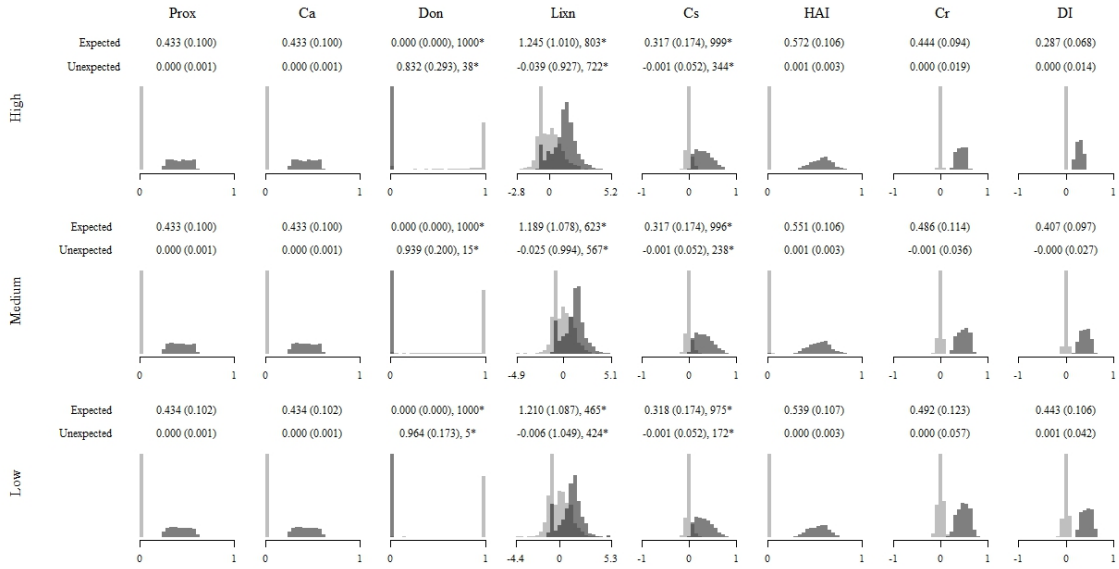

Figure 6.3: Results from simulations where 1000 scenarios were generated, each containing one dyad, under two circumstances: 1) dynamic interaction expected (medium grey histogram) and 2) dynamic interaction unexpected (light grey histogram). Dark grey histograms indicate where the index values overlap. Each scenario was examined at a high (100% of fixes), medium (25% of fixes), and low (10% of fixes) sampling resolution. Along with the histogram of index values, we present the mean, standard deviation (brackets), and number of significant results where appropriate ($p < 0.01$ - denoted by an asterisk).

tion. The rate of Type I error also appears to increase at finer sampling resolutions, making the use of Lixn problematic with modern high-resolution telemetry systems (e.g., GPS collars). The raw Lixn values appear to be robust to changes in sampling resolution, but given the overlap in histograms, interpretation would be misleading whether using qualitative or quantitative assessments (e.g., p-values).

As a path-based index of cohesive movement, Cr appears to be robust to changes in sampling resolution (Figure 6.3). DI, on the other hand, was more sensitive to changes in sampling resolution than Cr; lower DI values occurred at finer sampling resolutions (Figure 6.3). However, DI may be easier to interpret due to a relatively low variation in output values, and accurate assignment to dynamic interaction when one actually is present (conversely, no dynamic interaction for unexpected cases). Thus, at the global analysis level, it appears both Cr and DI provide relatively robust and similar results.

6.4.2 Empirical Data: White Tailed Deer

For dyad 1, 3 of the 8 indices indicated dynamic interaction in at least one sampling resolution (i.e., either 30 min or 6 hr), but no interaction at the coarsest sampling interval (24 hour) for all eight indices (Table 6.4). In general, indices did not show a strong and consistent indication that these two individual deer were exhibiting dynamic interaction. Also, for this particular dyad, indication of dynamic interaction appeared to disappear as sampling resolution became coarser when using Don, Cs, and Lixn; otherwise, indices did not reveal any interaction.

In dyad 2, the results showed much conflicting interpretation of whether dynamic interaction was present at the different sampling resolutions and with each index (Table 6.4). The Don and Cs indices identified significant attraction at all three sampling resolutions, Lixn indicated random use, and Prox, Ca, HAI, Cr and DI showed the absence of any dynamic interaction behaviour (Table 6.4). However, as Cr is similar to interpreting correlations, the values (from 0.095-0.16) might be interpreted as minimal correlation in movements. Thus, in dyad 2, it is particularly difficult to conclude whether or not dynamic interaction exists, which underscores the difficulty in selecting an appropriate index.

With dyad 3, results indicated substantial and consistent dynamic interaction among seven of the indices and at each sampling resolution (Table 6.4). Lixn suggested that there was random use of the shared area across the three sampling resolutions. Given the ubiquitous identification of dynamic interaction by 7 of the 8 indices, at all three sampling resolutions, we conclude that positive interactive behaviour (i.e., both attraction and cohesive movement) does exist in dyad 3.

Because of the varying levels of dynamic interaction among the 3 dyads of deer, we plotted time series of di (and Prox) to graphically represent the spatial and temporal patterns of: 1) no interaction (top plot; dyad 1), 2) infrequent or minimal interaction (middle plot; dyad 2), and 3) strong dynamic interaction (bottom plot; dyad 3) (Figure 6.4). In the plot of dyad 1, we first observe that distance between this pair never exceeds 800 m. However, the observed separation between the pair appears random because there is little variation in the plot of di (minimal variation around di = 0). In the plot for dyad 2, most simultaneous locations were within 1000 m of each other, however, in a few instances, locations of the two deer were $>$ 3000 m apart. Global statistics for the 8 indices revealed discrepancies in the dynamic interaction behavior in dyad 2. However, plotting di revealed that dyad 2 did in fact exhibit

Table 6.4: Results of dynamic interactions using empirical GPS data collected from white-tailed deer in Oklahoma, USA. Dynamic interactions were tested at three temporal resolutions (30 min, 6 hr, and 24 hr) for eight indices of dynamic interaction. Values with an asterisk(*) indicate significance at $p < 0.01$.

Dyad	**Sampling Resolution**			**interpretation**
1	30 min	6 hr	24 hr	
Prox	0.017	0.030	0.018	No proximity
Ca	0.014	0.025	0.016	No attraction
Don (p-val.)	0*	0.0042*	0.66	Attraction at 30 min and 6 hr
Cs	0.013*	0.02*	0.01	Attraction at 30 min and 6 hr
Lixn	-0.19	0.11	0.058	Simultaneous use of shared area at 30 min only
HAI	0.037	0.053	0.044	No attraction within shared area
Cr	-0.022	0.079	0.060	No correlation
DI	-0.004	0.021	-0.014	No dynamic interaction
2	30 min	6 hr	24 hr	
Prox	0.071	0.073	0.072	No proximity
Ca	0.069	0.07	0.07	No attraction
Don (p-val.)	0*	0*	0*	Attraction
Cs	0.10*	0.10*	0.10*	Attraction
Lixn	0.37	0.31	0.32	Random use of shared area
HAI	0.074	0.077	0.075	No attraction within shared area
Cr	0.095	0.16	0.11	Very low positive correlation
DI	0.029	0.11	0.028	No dynamic interaction
3	30 min	6 hr	24 hr	
Prox	0.57	0.57	0.58	Substantial proximity
Ca	0.53	0.52	0.54	Attraction
Don (p-val.)	0*	0*	0*	Attraction
Cs	0.40*	0.39*	0.36*	Attraction
Lixn	0.11	0.21	0.21	Random use of shared area
HAI	0.60	0.60	0.61	Attraction within shared area
Cr	0.59	0.66	0.72	Positive correlation
DI	0.28	0.57	0.61	Positive dynamic interaction (cohesive movement)

dynamic interaction on multiple occasions. Finally, the plot for dyad 3 revealed definitively the presence of strong dynamic interaction. Across much of the 3-month sampling period, the pair of male deer remained proximal for extended periods of time, with corresponding cohesive movement. Although dynamic interaction is occurring across most of the sampling period, this graph reveals periods of variable levels of dynamic interaction through time.

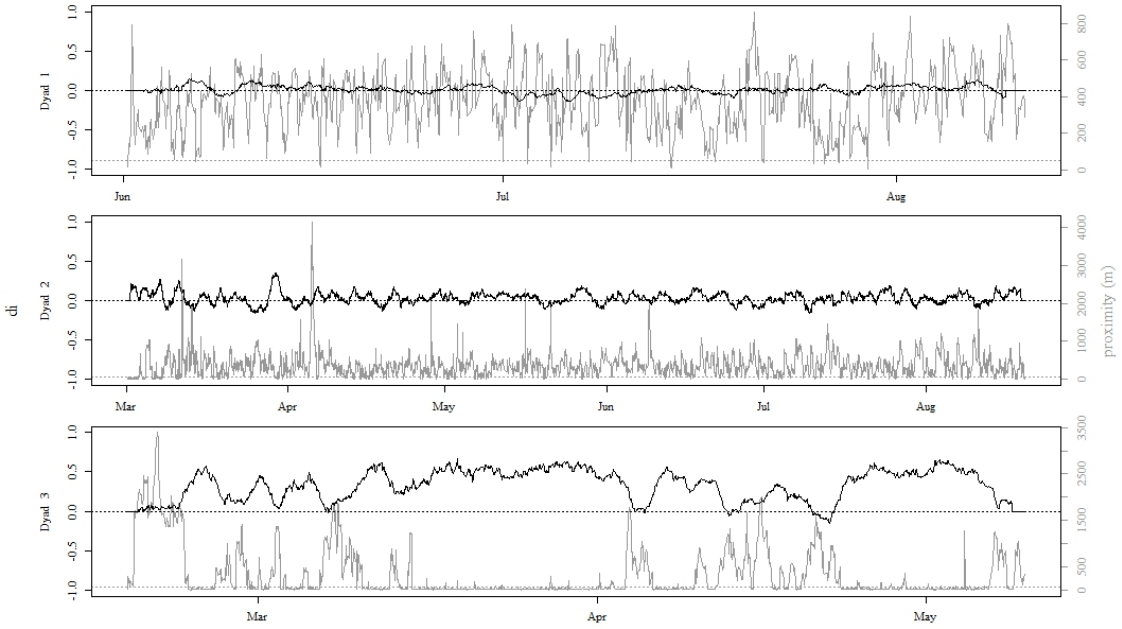

Figure 6.4: Time series plot of the local di statistic (in black; developed by Long and Nelson, 2013a) and proximity (meters; in grey) for the 30 min sampling interval for three white-tailed deer dyads ($n = 6$ deer). A 24 hr moving window average of di was used to minimize noise. The index di is presented on y-axis 1 and proximity (m) on y-axis 2. The black dotted line represents random interaction at di=0, whereas the grey dotted line represents the critical threshold for identifying proximal fixes ($d_c = 50$ m).

6.5 Discussion

6.5.1 Static vs. dynamic interaction

Results from this study can be used to guide selection of the most appropriate index for quantifying dynamic interaction in wildlife telemetry studies. Quantitative indices of dynamic interaction allow researchers to examine a wide range of questions relating to animal behaviour and general ecology, and go beyond typical static measures of interaction (e.g., home range overlap). There is a large body of literature employing dynamic interaction indices (see Table 6.1), but our study is the first we know of to characterize variation in inferences obtained by each method at different sampling resolutions.

Recent research has suggested that measures of static interaction (like AOP) can be used to estimate contact-rates and levels of dynamic interaction between wildlife,

important in modeling disease transmission (Robert *et al.*, 2012). Our analysis reveals the flaws in this assumption, as even moderate levels of static interaction may have highly variable contact rates and levels of dynamic interaction. We looked at three white-tailed deer dyads containing static interaction levels of AOP = 0.17, 0.67, and 0.57, respectively. However, only dyad 3 (AOP = 0.57) showed substantial dynamic interaction, which agrees with the seasonal biology of male deer during this time; male deer form bachelor groups during spring and summer (Hirth, 1977). Thus, we were able to identify a pair of deer belonging to the same bachelor group using dynamic interaction metrics, particularly di (see Figure 6.4). Further, while the converse may be true (i.e., no static interaction implies no dynamic interaction), our analysis suggests that the relationship between the level of static interaction and presence of dynamic interaction is complex and inferring dynamic interaction from static interaction can be misleading. Incorrect inferences may be most problematic for species inhabiting relatively large home range areas, where joint space use (i.e., home range overlap) may occur without individuals ever encountering one another from a temporal standpoint.

6.5.2 Comparison across indices

The Prox, Ca, and HAI indices produce nearly identical values in all cases, owing to the similarity in their derivation. As we portrayed alongside the temporal plot of di, a temporal plot of Prox is a meaningful local measure for identifying temporal dynamics in attraction behaviour (i.e., local changes in proximity). In our examples, HAI produced similar values to Prox and Ca and identical interpretation of attraction in both the simulation study, and in the white-tailed deer case study. Our results also suggest that three of the classical indices of dynamic interaction (Don, Cs, and Lixn) can be misleading, especially with high-resolution telemetry data. Misleading results are due to statistical testing procedures susceptible to Type I error (and in the case of Lixn, Type II error as well). Plots and contingency tables accompanying the Don statistic may be more useful for interpretation than computed p-values; for instance, when examining the effect of the dc parameter. Similarly, the Cs index may be useful as a non-statistical index as Cs was able to adequately separate between expected and unexpected dynamic interaction. Further, Cs measures a unique property of the data from the other indices by utilizing the raw distances between fixes rather than using a subjectively defined distance threshold (d_c). Unlike the other indices, Lixn index

tests for dynamic interaction based on simultaneous use of the shared area (Minta, 1992). However, despite examining co-occurrence of shared area, Lixn appears to be poorly suited for measuring dynamic interaction in wildlife telemetry datasets.

Cr and DI measure dynamic interaction as the cohesiveness in movement using a path-based approach. Cr values were consistent across sampling intervals, making it a suitable candidate for measuring dynamic interaction with modern telemetry datasets. The primary limitation of Cr is that it is dependent on measuring correlations relative to a mean path vector, which is typically not meaningful in the context of dynamic interaction analysis. DI was sensitive to changes in sampling resolution producing a lower index value at the highest resolution. The advantage of the DI approach is in examining spatial variation in dynamic interactions through the local level statistic – di.

6.5.3 Scale

Our simulation study outlined a key problem commonly encountered in wildlife movement analysis; inferences made at one scale of analysis do not necessarily hold at other scales (see Laube and Purves, 2011, for a more thorough discussion of this). Indices that do not formally test statistical significance appear to be less sensitive to varying scales (i.e., Prox, Ca, HAI, Cr, DI) when compared to indices that implement statistical tests (i.e., Don, Cs, Lixn). Beyond scale, inference among indices is impacted by the nature or level of dynamic interaction present. Most methods appear to successfully identify strong dynamic interaction when present (e.g., deer dyad 3) with greater inconsistencies occurring for infrequent or low levels of dynamic interaction (e.g., deer dyad 2). Our guidelines will help inform the selection of appropriate indices given varying scales and levels of dynamic interaction.

The implementation of the local-level statistic (i.e., di), along with time-series plots of proximity (e.g., in Figure 6.4), reveals information on infrequent (e.g., deer dyad 1), variable (e.g., deer dyad 2), and frequent (e.g., deer dyad 3) interactions. The prevalence of infrequent and variable interactions in wildlife is unknown, which typically is the motivating factor for assessing the level of dynamic interaction, and the di index represents a new metric for revealing both frequent and infrequent interactive behaviour. Animals are now routinely tracked for an extended duration (e.g., several months) with fine resolution telemetry fix rates (e.g., sub-hour sampling intervals), allowing the identification of rare and periodic interactive behaviour, which

has important implications on a wide range of studies such as disease spread (Böhm *et al.*, 2008), patterns of sociality (Gorman *et al.*, 2006), and predator-prey dynamics (Eriksen *et al.*, 2008).

We only superficially examine temporally local dynamics in di. More sophisticated analysis could involve other temporal variables to investigate more complex problems, such as circadian rhythms, seasonality, and weather factors. Variations in the level of dynamic interaction (measured through di) may be a result of different behaviour states that can relate to resting, foraging, or travelling behaviour (Dzialak *et al.*, Unpb). Including a quantitative characterization of movement behaviour (e.g., Jonsen *et al.*, 2005; Gurarie *et al.*, 2009; Patterson *et al.*, 2009) for comparison could enhance interpretation of changes in dynamic interaction when using the di measure. Similarly, by mapping the local statistic di, one can investigate the role of fine scale landscape and environmental variables on the observed patterns of dynamic interaction.

6.5.4 Statistical testing

In wildlife telemetry research, the effects of sampling resolution and autocorrelation are well documented, for example with home range delineation (Swihart and Slade, 1985; Seaman and Powell, 1996), habitat selection models (Otis and White, 1999; Nielsen *et al.*, 2002), and behaviour analyses (Boyce *et al.*, 2010). Yet, little is known about the effects of autocorrelated data and sampling resolution on indices of dynamic interaction. Serial autocorrelation in successive fixes of telemetry data (Swihart and Slade, 1985; Otis and White, 1999) increases with higher sampling frequency (Dray *et al.*, 2010) and invalidates use of methods where independence is assumed (Swihart and Slade, 1985). In the presence of highly autocorrelated data, statistical tests can be overly sensitive, producing false positives more frequently as n increases. While some have argued that the autocorrelation problem can be overcome by down-sampling telemetry data until it is functionally independent (Swihart and Slade, 1997), this procedure has been criticized due to loss of biologically relevant data (de Solla *et al.*, 1999). Alternately, it may be more effective to implement statistical methods that accommodate (de Solla *et al.*, 1999; Nielsen *et al.*, 2002) or are less sensitive (Fieberg, 2007) to autocorrelated structures. Computing dynamic interaction indices using frequent fixes (e.g., from GPS collars), autocorrelation in the dataset will be inherent, but also may be important for detecting 'true' contacts from false positives (i.e., Type I error).

Several of the indices we examined use formal statistical tests for the presence of dynamic interaction and therefore may suffer from meaningless statistical significance when biological significance is of greater importance (i.e., understanding behaviour above statistical significance when test data is influenced by power - e.g., with large sample sizes characteristic of GPS data - and autocorrelation). With Don, statistical expectations are computed based on measured distances, falling below d_c, of permutation's of the $n^2 - n$ non-simultaneous fixes, while Cs expectations are computed as the average raw distances of all n^2 fixes. The permutations required for statistical testing are not independent and lack the serially correlated spatial structure associated with movement trajectories (White and Harris, 1994). As such, these permutations represent an unrealistic expectation causing increased Type 1 error when data are correlated. Cs is more susceptible to Type I errors than Don, owing to the use of raw distance values in the permutation calculations.

With Lixn the statistical test is based on expectations derived from the frequency of total fixes found within each home range section (inhabited by α only, β only, and overlapping area). When the area of home range overlap is relatively large, Lixn is susceptible to Type I error; conversely, if the overlap is relatively small, Lixn may be susceptible to Type II error. Given the documented problems of computing reliable home range estimates (Hemson *et al.*, 2005; Börger *et al.*, 2006; Downs and Horner, 2008), the problem then lies with interpreting what Lixn is testing. In reality, Lixn is likely producing reliable results concerning the shared use of overlapping home range sections (Minta, 1992). The question remains, is the shared-use of a large home range area a useful measure of dynamic interaction in high resolution telemetry datasets where finer-scale movements are resolvable? The expectations for Lixn (and the spatial interaction metrics L_{AA} and L_{BB} proposed by Minta, 1992) can be computed from the relative area values of each home range section (although this approach was not explored here, as the frequency approach is recommended in Minta, 1992). Given that animals typically use their home range in a non-homogenous fashion (Samuel *et al.*, 1985), expectations derived in this alternate fashion may be misleading relative to actual space use intensity patterns - typically represented by a utilization distribution (Worton, 1989). An improved formulation for Lixn could derive the expected values as a function of the intersection (see Fieberg and Kochanny, 2005) of the two individual utilization distributions.

6.5.5 Guidelines

The role of classical hypothesis testing in ecological analysis continues to be questioned (e.g., Hobbs and Hilborn, 2006), and in the context of measuring dynamic interactions, our results suggest that non-statistical procedures provide greater insight into interactive behaviour patterns. The usefulness of Prox, Ca, and HAI owe to the simplicity of their (near identical) calculation and interpretation. Thus, it is wise to compute only one of Prox, Ca, or HAI. The Don and Cs indices suffer from the susceptibility of their statistical testing procedures to Type I errors, a problem magnified with high resolution GPS telemetry data. Cs, due its formulation based on raw distances, may still be useful to examine this separate property of the data; however, we suggest that a subjective interpretation be employed. The Don index measures attraction similar to Prox, Ca, and HAI, and we advocate the use of those methods over the Don index. The Lixn index suffers from the same statistical problems as Don and Cs, but also from the configurational problems related to the general task of home range delineation. Lixn performed poorly when viewed as a stand-alone index and provided contrasting results in the white-tailed deer case study, and as such we do not recommend its use.

Novel path-based measures of dynamic interaction (i.e., Cr, DI) examine the cohesiveness in movement segments, rather than proximity or arrangement of fixes represented as spatial points. Of these two metrics, Cr appeared to be least sensitive to sampling resolution, but did suffer from high variation in index values. DI provides novel insight into the spatial variation in dynamic interaction behavior, through the use of the local di statistic. Thus, di is most useful with high resolution GPS telemetry data due to its ability to examine spatial and temporal changes in dynamic interaction behaviour.

The following points can be used to guide decisions when studying dynamic interaction in wildlife telemetry datasets. First, it is necessary to identify the presence of temporal overlap in α and β to asses if a dyad offers potential for interactive behaviour. A measure of static interaction (e.g., AOP, as used here) can be used to assess joint space use. Second, either of Prox, Ca, or HAI can be used as an index of attractive/avoidance behaviour. Alternatively, Cs can be used as a measure of attraction/avoidance that is based on raw distances; however statistical inference should be avoided. Third, Cr or DI can be used as an index of the overall level of cohesion in movement segments. It is useful to compute proximity between all simultaneous fixes,

in order to examine local-level changes in proximity between α and β. Finally, di can be used to further investigate local-level variations in the cohesiveness of movement, especially within proximal episodes.

6.6 Conclusion

The simulated data (and subsequent analysis) allowed the identification of several indices that will be useful when applied to empirical wildlife telemetry data, including high resolution data such as those collected from GPS collars. The case study on white-tailed deer revealed that minimal AOP resulted in minimal dynamic interaction, but higher levels of AOP did not necessarily equate to dynamic interaction behavior because AOP relies on a 2-dimensional home range that does not account for the simultaneous use of these areas, which is inferred using both point- and path-based measures of interaction. The results from all eight indices corroborated what is known about the seasonal biology of white-tailed deer; dyad 3 was a pair of male deer of 3 years of age that were part of the same bachelor group. Although we used white-tailed deer as a case study to assess various indices of dynamic interaction and behaviour, these indices and guidelines can be applied much more broadly to animal behaviour studies, and even for the analysis of any two moving objects in space and time (e.g., dyads of vehicles, cell-phone users, athletes, etc.).

There still is much to be learned about animal behaviour. Despite recent technological advances, much research builds on early studies that relied on visual observations of animals (Hirth, 1977), which is limited by animal activity patterns, habitat use, and observation bias. For example, early accounts on the breeding behaviour of white-tailed deer were based on visual observations, but recent molecular genetic techniques have revealed greater insight into the breeding biology of deer (e.g., DeYoung *et al.*, 2009). Then, indices of dynamic interaction can be leveraged to further study rare behavioural events that are not readily observed visually while in the field. Other avenues of research that may benefit from recently developed techniques to study animal behaviour and contact rates (e.g., contact at d_c) include: studies of behavior and ecology, territory defense, determining mating and reproduction events, assessing disease spread through direct contact, interspecific competition for resources, and intra- and interspecific interactions. One of the most pervasive topics deals with potential for disease spread, particularly when wild animal species come into contact with domestic livestock (Wyckoff *et al.*, 2009). Combining behavioural indices with

spatially-explicit landscape data or genetic data will provide greater insight into unobservable phenomena that shape animal populations (demographics and dynamics) and long-term fitness measures (e.g., reproductive success). In future research, it will prove useful to combine local-level measurements of dynamic interaction (e.g., di, Long and Nelson, 2013a) with landscape data to determine where interactions are linked to (e.g., preferred resources on the landscape, dietary or niche overlap - within and between species, and the potential for disease spread).

In summary, the calculation of contact rates and measures of dynamic interaction offer promise in studying dynamically moving objects in a wide range of fields from pure scientific discovery and ethology to application and management. Herein, we have provided general points on the formulation, interpretation, and use of dynamic interaction indices that can be used to guide future research. To make these methods and indices as widely available as possible, we have implemented each of the eight indices in the R statistical computing environment, and made this code openly available from our website `http://www.geog.uvic.ca/spar/`.

Chapter 7

Conclusions

7.1 Conclusion

The primary goal of my research was to advance the science of movement analysis through the development and application of novel techniques and methods. As part of this process, I have contributed a systematic review of the quantitative movement analysis literature, and identified key challenges faced in movement analysis, along with areas for emerging research. My research has focused on two areas of movement research, namely the development and application of time geography theory, and the analysis of inter-object interactions in movement data, where I have made contributions through the development of methods, novel theory and movement models, and investigated the role of scale and statistical inference in analysis techniques. As well, I have designed analytical tools for implementing the methods I have developed (also for some existing techniques) to encourage the adoption of those techniques.

I conclude this research by outlining the various contributions I have made in the context of quantitative movement analysis. The contributions I have made to the movement analysis literature, the specific quantitative methods I have developed, and software tools I have generated are highlighted. I identify several key findings that have emerged from this research. Finally, throughout my research I have identified a number of areas for future research which I conclude with as areas to take my research moving forward.

7.1.1 Contributions

Review of Quantitative Movement Analysis

The movement literature is highly fragmented, owing to widespread interest in studying moving objects by a range of disciplines including geography, wildlife ecology, transportation, and computer science. Given this separation, it is difficult for those both within and entering the field of movement research to grasp the scope of currently available quantitative methods. With Chapter 2, I provide a unified review of the current state of quantitative methods for movement data, reviewing methods from a wide range of disciplines and applications. In this review I have identified seven classes of methods, relating to different problem contexts in movement research: time geography, path descriptors, similarity indices, pattern and cluster methods, individual-group dynamics, spatial field methods, spatial range methods. With Chapter 2, I present a number of key challenges facing movement research, notably geographys inability to move beyond static representations, and the importance of the temporal dimension in movement analysis (e.g., Neutens *et al.*, 2011). Further, I was able to identify areas where the current breadth of movement research is either inadequate, or primed for novel developments, which I use as the basis for my own research.

Chapter 2 has been successfully published (see Long and Nelson, 2013b) as a review paper in the International Journal of Geographical Information Science (IJGIS), the preeminent journal in the field of geographic information science (GIScience). The choice of IJGIS as a venue for this work was motivated not simply by its impact in GIScience, but also due to its broad, multi-disciplinary readership base. As such, this particular chapter, I believe, has allowed my work to expose the similarities associated with different areas of movement research. Hopefully, this leads to an increased recognition of the scope of movement research occurring across a range of disciplines and the development of universally accepted quantitative approaches.

Method Developments

In this research I have made significant contributions to the current state of quantitative movement analysis, tied predominantly to two underlying themes: advancing the use and application of time geography, and methods for quantifying interactive movement behaviour–termed dynamic interaction.

In Chapter 3, I present a time geographic approach for computing wildlife home

ranges, termed the PPA home range. The PPA home range is a logical approach for examining movement ranges in wildlife that uses existing movement theory from geography, namely time geography, applied to a common problem from wildlife movement ecology. The time geographic approach has long been utilized as a framework for quantitative movement analysis in human and transportation studies (e.g., Burns, 1979; Miller, 1991), but has not been widely adopted in wildlife ecology (but see Baer and Butler, 2000, for an exception). With Chapter 3, I aim to expose wildlife ecologists to the advantages of a time geographic approach when examining spatial ranges of movement (home ranges and other similar constructs, see for example Nelson *et al.*, Unpb). In a concerted effort to relay this idea to wildlife ecologists, I successfully targeted Chapter 3 toward a wildlife focused journal for publication (see Long and Nelson, 2012).

Apart from exposing wildlife ecologists to time geographic theory, this chapter presents two key methodological contributions. The first development is the PPA home range method itself. The PPA home range can be used as a stand-alone measure of home range, but more importantly can be used to examine omission and commission error in other more widely adopted home range methods. Consideration of omission and commission error is essential to spatial inferences in home range analysis (Sanderson, 1966); however this component of error is routinely overlooked in wildlife home range studies. The PPA home range is straightforward to compute and thus represents an easily implemented tool for examining omission and commission error. As it is customary to employ multiple home range methods in a given study (due to their ease of implementation in modern GIS software), the PPA home range effectively augments other approaches through the identification of omission and commission errors. A second contribution is the derivation of an objective, empirical approach to estimating the mobility parameter (v_{max}) from time geography using established statistical methods (e.g., Robson and Whitlock, 1964; van der Watt, 1980). Empirical methods for estimating v_{max} are potentially valuable with all movement applications employing methods from time geography. As demonstrated in Chapter 4, I use the empirical approach for computing v_{max} on simulated movement data along with real datasets pertaining to wildlife, cyclists, and athletes.

With Chapter 4, I present a novel development to the theoretical framework of time geography. The idea for kinetic-based probabilistic time geography was motivated by recent developments in time geographic theory, namely probabilistic time geography (Winter, 2009) and kinetic-based time geography (Kuijpers *et al.*, 2011). The

authors of kinetic-based time geography identify the synthesis of ideas from kinetic and probabilistic time geography as an area for future development. My motivation for exploring the skew-normal distribution as a model for movement probabilities in kinetic-based probabilistic time geography was spurred by a similar notion originally proposed by Prager and Yu (2005). Thus, my work represents a mainly theoretical advancement, contributing to the existing foundation of movement theory within the time geography framework (Hägerstrand, 1970; Miller, 2005; Winter and Yin, 2011), however I also provide several potential applications of the kinetic-based probabilistic time geography model.

With Chapter 5, I present a novel index (the DI index) for quantifying the level of inter-object interaction. Termed dynamic interactions, interactive behaviour relates directly to aspects of co-dependent movement, such as attraction, avoidance, and cohesion. The approach I develop is targeted specifically at examining cohesion, where many previous approaches have looked predominantly at attraction (e.g., Doncaster, 1990; Kenward *et al.*, 1993). I build upon the lone existing approach for examining cohesion in pairs of moving objects (termed correlation by Shirabe, 2006). The DI index improves upon Shirabes index by measuring cohesion in movement direction and velocity independently and by incorporating a spatially and temporally local level statistic (termed di) that facilitates a finer treatment of dynamic interaction behaviour. In fact, temporally local movement analysis, through time series plots, as I have proposed with the local di index, has been championed in numerous other facets of movement analysis (e.g., Dodge *et al.*, 2012). The DI and local di indices provide a novel contribution to the study of interactive movement behaviour, potentially useful in many application areas. Chapter 5 has also been successfully published (see Long and Nelson, 2013a).

In Chapter 6 I continue my exploration of dynamic interaction behaviour in movement datasets. Here, I specifically focus on the study of wildlife, where quantifying dynamic interactions has been of interest for some time (e.g., Macdonald *et al.*, 1980). This chapter provides a complete description of eight currently available methods for quantifying dynamic interaction behaviour, including the DI index from Chapter 5. Single measures of dynamic interaction are routinely implemented in applied research, but little research exists comparing across methods (but see Miller, 2012). I present a simulation experiment, along with empirical data, to examine the effects of sampling interval on the ability of each measure of dynamic interaction at correctly identifying expected or unexpected dynamic interaction patterns. The simulation study is the

first of its kind, as I was able to simulate both expected and unexpected dynamic interaction patterns. Future work could involve a refinement of the approach for simulating expected dynamic interaction. Using empirical data on white-tailed deer, I demonstrate the value of the local di index (from Chapter 5) for examining infrequent and periodic dynamic interaction behaviour in wildlife telemetry datasets. In summary, I find that three of the most popular methods (Don, Cs, and Lixn) are susceptible to Type I error, identifying dynamic interaction where there is none, an issue exacerbated by increasing sampling frequency.

Technology/Software

While the development of theory and methods is valuable in and of itself, it is useful to provide analytical tools to encourage adoption of methods by other researchers. Traditionally in GIS and spatial analysis, tools are developed as add-ons to existing GIS software (e.g., ArcGIS, ESRI, Redlands, CA), or as stand-alone modules (e.g., GEODA, Anselin *et al.*, 2006). The emergence of spatial methods within the R statistical computing environment (R Development Core Team, 2012) has added an additional toolset to the geographic analysis arsenal, and now many researchers are providing access to their methods through the development of R functionality and packages.

Through my research I have developed software tools for implementing the methods I have developed, specifically the PPA home range method from Chapter 3 and the DI index from Chapter 5. With the PPA home range method, I have developed a GIS add-on (or toolbar) that can be used with ArcGIS 9.x software. I have simultaneously developed functionality for computing the PPA home range in R. Similarly, I have also developed R functionality for the DI index from Chapter 5, along with the suite of other methods for exploring dynamic interaction patterns in wildlife telemetry studies (i.e., those included in the analysis of Chapter 6). I have provided access to all of these tools through the SPAR Lab website. Given that I have received numerous enquiries on both methods, it suggests that many other researchers are beginning to adopt these approaches as well. The documentation for these tools is provided in Appendix A, and includes the technical details of implementing each function, along with worked examples. In future work, I hope to incorporate the suite of tools for dynamic interaction analysis into a formal R package.

7.1.2 Key Findings

Based on the methods and research I have developed through my PhD I have identified several key findings related to movement analysis.

- The potential path area from time geography represents an effective tool for characterizing omission and commission error in wildlife home range analysis.
- Existing statistical methods for estimating the upper-bound of a distribution are useful for estimating the v_{max} parameter from time geography, and are dependent on the sampling-interval of the movement data.
- The skew-normal probability density function provides a more general model for probabilistic time geography, where the degree and direction of skewness relate to kinetic movement properties (i.e., speed and direction).
- Measuring dynamic interaction is enhanced through the use of a spatially and temporally local measure - such as di - that facilitates novel inference on infrequent and/or periodic interactive behaviour.
- Path-based metrics for dynamic interaction (such as Cr, DI and di) measure cohesiveness in movement, while existing point-based approaches measure attraction, typically related to spatial proximity.
- Three widely adopted metrics for dynamic interaction in wildlife studies (Don, Cs, Lixn) are all susceptible to Type I error, a problem exacerbated with finer sampling intervals (such as with GPS collars).

7.1.3 Future Research

The PPA home range approach from Chapter 3 has laid the foundation for more detailed time geographic analysis of wildlife movement. The PPA approach can be enhanced by considering dynamics in estimates of the v_{max} parameter, where changes in v_{max} may relate to changes in wildlife behaviour. Transportation geographers routinely implement variable v_{max} values in time geographic analysis where changes depend on, for example, mode of transportation (Neutens *et al.*, 2007). A dynamic v_{max} parameter would enhance home range analysis by directly relating the v_{max} parameter to changes in wildlife mobility. For example, female ungulates experience restricted mobility levels during parturition (Webb *et al.*, 2009); a dynamic v_{max} parameter

that accounts for these changes more appropriately characterizes space-use and home range with the PPA approach. One area for future work is to examine alternative means for computing a dynamic v_{max} for use with PPA home range analysis.

The SN model derived in Chapter 4 only scratches the surface in terms of potential for modeling object kinetics in probabilistic time geography. First, the development of an appropriate SN model for use with the space-time prism is required to effectively model movement where the start and end points of movement segments are known, as this is typically the case with movement data. Second, the SN model, as it stands, fails to consider the effect of object acceleration in the calculation of movement probabilities. In future research, my goal is to derive an appropriate formulation of the SN model for the space-time prism, building on what I have developed for use with the space-time cone. Similarly, I intend to modify the existing formulation to account for the role of acceleration in kinetic-based probabilistic time geography.

My future research on dynamic interactions will explore relationships between measured di values and ancillary spatial and temporal variables. Given our newfound ability to collect increasingly detailed movement information, improved lines of scientific questioning are warranted, capable of extracting valuable movement information. Inevitably, of interest is discovering the contextual relationships influencing movement, or more succinctly: the geography behind movement trajectories (Bogorny *et al.*, 2009). By studying the relationships between movement and contextual variables, researchers can begin to study the spatial and temporal processes driving observed movement patterns. This is especially of interest in wildlife movement analysis, where relationships between interactive behaviour and environmental variables such as weather, land cover, and topography can inform conservation decisions. Similarly, contextual variables such as barriers and other environmental objects influence movement, and the resulting potential for interaction. Thus, I hope to study how contextual objects influence interactive behaviour.

Bibliography

Aebischer, N., Robertson, P., and Kenward, R., 1993. Compositional analysis of habitat use from animal radio-tracking data. *Ecology*, 74 (5), 1313–1325.

Ahas, R., *et al.*, 2007. Seasonal tourism spaces in Estonia: Case study with mobile positioning data. *Tourism Management*, 28, 898–910.

Alt, H. and Godau, M., 1995. Computing the Frchet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, 5 (1-2), 75–91.

Alvares, L., *et al.*, 2007. A model for enriching trajectories with semantic geographical information. *In*: *GIS'07: Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems* Nov. 7-9, Seattle, WA: Association for Computing Machinery, New York, p. 8.

Andersson, M., *et al.*, 2008. Reporting leaders and followers among trajectories of moving point objects. *Geoinformatica*, 12, 497–528.

Andrews, H., 1973. Home range and urban knowledge of school-age children. *Environment and Behaviour*, 5, 73–86.

Andrienko, G., *et al.*, 2011. GeoVA(t) - Geospatial visual analytics: focus on time. Special issue of the International Cartographic Association Commission on GeoVisualization. *International Journal of Geographical Information Science*, 20 (10), 1453–1457.

Andrienko, G., Andrienko, N., and Wrobel, S., 2007. Visual analytics tools for analysis of movement data. *SIGKDD Explorations*, 9 (2), 38–46.

Andrienko, N. and Andrienko, G., 2007. Designing visual analytics methods for massive collections of movement data. *Cartographica*, 42 (2), 117–138.

Andrienko, N., Andrienko, G., and Gatalsky, P., 2005. Impact of data and task characteristics on design of spatio-temporal data visualization tools. *In*: J. Dykes, A. MacEachren and M. Kraak, eds. *Exploring Geovisualization*. New York: Elsevier, 201–222.

Anselin, L., 1995. Local indicators of spatial association - LISA. *Geographical Analysis*, 27 (2), 93–115.

Anselin, L., Syabri, I., and Kho, Y., 2006. GeoDa: An introduction to spatial data analysis. *Geographical Analysis*, 38 (1), 5–22.

Arellano-Valle, R. and Azzalini, A., 2008. The centred parameterization for the multivariate skew-normal distribution. *Journal of Multivariate Analysis*, 99, 1362–1382.

Atev, S., Masoud, O., and Papanikolopoulos, N., 2006. Learning traffic patterns at intersections by spectral clustering of motion trajectories. *In*: *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 9-15, Beijing, China, 4851–4856.

Atwood, T.C. and Weeks, H.P., 2003. Spatial home-range overlap and temporal interaction in eastern coyotes: the influence of pair types and fragmentation. *Canadian Journal of Zoology*, 81 (9), 1589–1597.

Azzalini, A., 1985. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12 (2), 171–178.

Azzalini, A. and Dalla Valle, A., 1996. The multivariate skew-normal distribution. *Biometrika*, 83 (4), 715–726.

Baer, L. and Butler, D., 2000. Space-time modeling of grizzly bears. *Geographical Review*, 90 (2), 206–221.

Baker, P.J. and Harris, S., 2000. Interaction rates between members of a group of Red Foxes (Vulpes vulpes). *Mammal Review*, 30 (3 & 4), 239–242.

Bandeira de Melo, L., *et al.*, 2007. Secret lives of maned wolves (Chrysocyon brachyurus Illiger 1815): as revealed by GPS tracking collars. *Journal of Zoology*, 271, 27–36.

Batty, M., Desyllas, J., and Duxbury, E., 2003. The discrete dynamics of small-scale events: agent-based models of mobility in carnivals and street parades. *International Journal of Geographical Information Science*, 17 (7), 673–697.

Bauman, P., 1998. The Wind Cave National Park elk herd: home ranges, seasonal movements, and alternative control methods. Thesis (PhD). South Dakota State University, Brookings, South Dakota, USA.

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

Benhamou, S., 2011. Dynamic approach to space and habitat use based on biased random bridges. *PloS one*, 6 (1), e14592.

Benkert, M., *et al.*, 2007. Finding popular places. *In*: T. Tokuyama, ed. *ISAAC 2007, LNCS 4835* Springer-Verlag, 776–787.

Benkert, M., *et al.*, 2008. Reporting flock patterns. *Computational Geometry*, 41, 111–125.

Bertrand, M.R., *et al.*, 1996. Effects of parturition on home ranges and social affiliations of female white-tailed deer. *Journal of Wildlife Management*, 60 (4), 899–909.

Bogorny, V., Kuijpers, B., and Alvares, L., 2009. ST-DMQL: A semantic trajectory data mining query language. *International Journal of Geographical Information Science*, 23 (10), 1245–1276.

Böhm, M., *et al.*, 2008. Dynamic interactions among badgers : implications for sociality and disease transmission. *Journal of Animal Ecology*, 77, 735–745.

Boots, B., 2002. Local measures of spatial association. *Ecoscience*, 9 (2), 168–176.

Börger, L., Dalziel, B., and Fryxell, J., 2008. Are there general mechanisms of animal home range behaviour? A review of prospects for future research. *Ecology Letters*, 11, 637–350.

Börger, L., *et al.*, 2006. Effects of sampling regime on the mean and variance of home range size estimates. *Journal of Animal Ecology*, 75 (6), 1393–1405.

Borruso, G., 2008. Network density estimation: a GIS approach for analysing point patterns in a network space. *Transactions in GIS*, 12 (3), 377–402.

Boyce, M.S., *et al.*, 2010. Temporal autocorrelation functions for movement rates from global positioning system radiotelemetry data.. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 365 (1550), 2213–2219.

Bradshaw, C., *et al.*, 1995. Winter peatland habitat selection by woodland caribou in northeastern Alberta. *Canadian Journal of Zoology*, 73, 1567–1574.

Brillinger, D., 2001. Time series: general. *In*: N. Smelser and P. Baltes, eds. *International Encyclopedia of the Social & Behavioral Sciences*. Berlin: Elsevier Science Ltd., 15724–15731.

Brillinger, D., 2007. A potential function approach to the flow of play in soccer. *Journal of Quantitative Analysis in Sports*, 3 (1), Article 3.

Brillinger, D., *et al.*, 2004. An exploratory data analysis (EDA) of the paths of moving animals. *Journal of Statistical Planning and Inference*, 122, 43–63.

Brillinger, D., *et al.*, 2002. Employing stochastic differential equations to model wildlife motion. *Bulletin Brazilian Mathematical Society*, 33 (3), 385–408.

Bromley, C. and Gese, E.M., 2001. Effects of sterilization on territory fidelity and maintenance, pair bonds, and survival rates of free-ranging coyotes. *Canadian Journal of Zoology*, 79 (3), 386–392.

Brotherton, P., *et al.*, 1997. Genetic and behavioural evidence of monogamy in a mammal, Kirk's dik dik (Madoqua kirkii). *Proceedings of the Royal Society of London: Series B Biological Sciences*, 264 (1382), 675–681.

Buchin, K., Buchin, M., and Gudmundsson, J., 2010. Constrained free space diagrams: a tool for trajectory analysis. *International Journal of Geographical Information Science*, 24 (7), 1101–1125.

Burns, L., 1979. *Transportation, Temporal, and Spatial Components of Accessibility*. Lexington, MA: Lexington Books.

Burt, W., 1943. Territoriality and home range concepts as applied to mammals. *Journal of Mammalogy*, 24 (3), 346–352.

Cagnacci, F., *et al.*, 2010. Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society B*, 365, 2157–2162.

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

Casaer, J., *et al.*, 1999. Analysing space use patterns by Thiessen polygon and triangulated irregular network interpolation: a non-parametric method for processing telemetric animal fixes. *International Journal of Geographical Information Science*, 13 (5), 499–511.

Chamberlain, M.J., Lovell, C.D., and Leopold, B.D., 2000. Spatial-use patterns, movements, and interactions among adult coyotes in central Mississippi. *Canadian journal of Zoology*, 78, 2087–2095.

Chrisman, N., 1998. Beyond the snapshot: Changing the approach to change, error, and process. *In*: M. Egenhofer and R. Golledge, eds. *Spatial and Temporal Reasoning in Geographic Information Systems*. New York, NY: Oxford University Press, 85–93.

Claussen, D., Finkler, M., and Smith, M., 1997. Thread trailing of turtles: methods for evaluating spatial movements and pathway structure. *Canadian Journal of Zoology*, 75, 2120–2128.

Cliff, A. and Ord, J., 1973. *Spatial Autocorrelation*. London: Pion Limited.

Cliff, A. and Ord, J., 1981. *Spatial Processes: Models and Applications*. London: Pion.

Cochran, W. and Lord, Jr, R., 1963. A radio-tracking system for wild animals. *Journal of Wildlife Management*, 27 (1), 9–24.

Codling, E., Plank, M., and Benhamou, S., 2008. Random walk models in biology. *Journal of the Royal Society Interface*, 5, 813–834.

Cole, L., 1949. The measurement of Interspecific associaton. *Ecology*, 30 (4), 411–424.

Cooke, P., 1979. Statistical inference for bounds of random variables. *Biometrika*, 66 (2), 367–374.

Coutts, A. and Duffield, R., 2010. Validity and reliability of GPS devices for measuring movement demands of team sports. *Journal of Science and Medicine in Sport*, 13, 133–135.

Craighead, F., 1982. *Track of the Grizzly*. New York, NY: Random House.

Cressie, N., 1993. *Statistics for Spatial Data*. New York: John Wiley & Sons Inc.

D'Auria, M., Nanni, M., and Pedreschi, D., 2005. Time-focused density-based clustering of trajectories of moving objects. *In*: *Proceeding of the Workshop on Mining Spatio-temporal Data (MSTD-2005)*, Porto, p. 14.

de Almeida Jácomo, A.T., *et al.*, 2009. Home Range and Spatial Organization of Maned Wolves in the Brazilian Grasslands. *Journal of Mammalogy*, 90 (1), 150–157.

de Lucca Siqueira, F. and Bogorny, V., 2011. Discovering chasing behavior in moving object trajectories. *Transactions in GIS*, 15 (5), 667–688.

de Solla, S.R., Bonduriansky, R., and Brooks, R.J., 1999. Eliminating autocorrelation reduces biological relevance of home range estimates. *Journal of Animal Ecology*, 68 (2), 221–234.

Delafontaine, M., *et al.*, 2011a. The impact of opening hours on the equity of individual space-time accessibility. *Computers, Environment and Urban Systems*, 35, 276–288.

Delafontaine, M., Neutens, T., and Van de Weghe, N., 2011b. Modelling potential movement in constrained travel environments using rough space-time prisms. *International Journal of Geographical Information Science*, 25 (9), 1389–1411.

Demšar, U. and Virrantaus, K., 2010. Space -time density of trajectories: exploring spatio-temporal patterns in movement data. *International Journal of Geographical Information Science*, 24 (10), 1527–1542.

Derekenaris, G., *et al.*, 2001. Integrating GIS, GPS and GSM technologies for the effective management of ambulances. *Computers, Environment and Urban Systems*, 25, 267–278.

Dettki, H., Ericsson, G., and Edenius, L., 2004. Real-time moose tracking: An internet based mapping application using GPS/GSM-collars in Sweden. *Alces*, 40, 13–21.

DeYoung, R., *et al.*, 2009. Molecular evaluation of the white-tailed deer (Odocoileus virginianus) mating system. *Journal of Mammalogy*, 90 (4), 946–953.

Dicke, M. and Burrough, P., 1988. Using fractal dimensions for characterizing tortuosity of animal trails. *Physiological Entomology*, 13, 393–398.

Diggle, P., 2003. *Statistical Analysis of Spatial Point Patterns*. 2nd New York, NY: Arnold.

Dixon, K. and Chapman, J., 1980. Harmonic mean measure of animal activity areas. *Ecology*, 61 (5), 1040–1044.

Dobson, J. and Fisher, P., 2003. Geoslavery. *IEEE Technology and Society Magazine*, 22 (1), 47–52.

Dodge, S., Laube, P., and Weibel, R., 2012. Movement similarity assessment using symbolic representation of trajectories. *International Journal of Geographical Information Science*, 26 (1), 1563–1588.

Dodge, S., Weibel, R., and Lautenschutz, A.K., 2008. Towards a taxonomy of movement patterns. *Information Visualization*, 7, 240–252.

Doncaster, C., 1990. Non-parametric estimates of interactions from radio-tracking data. *Journal of Theoretical Biology*, 143, 431–443.

Douglas, D.H., 1994. Least-cost path in GIS using an accumulated cost surface and slopelines. *Cartographica*, 31 (3), 37–51.

Downs, J. and Horner, M., 2012. Probabilistic potential path trees for visualizing and analyzing vehicle tracking data. *Journal of Transport Geography*, 23, 72–80.

Downs, J. and Horner, M., 2008. Effects of point pattern shape on home-range estimates. *Journal of Wildlife Management*, 72 (8), 1813–1818.

Downs, J. and Horner, M., 2009. A characteristic-hull based method for home range estimation. *Transactions in GIS*, 13 (5-6), 527–537.

Downs, J.A., 2010. Time-geographic density estimation for moving point objects. *LNCS*, 6292, 16–26.

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

Drewe, P., 2005. What about time in urban planning & design in the ICT age?. *In*: *CORP 2005 & Geomultimedia05*, Feb. 22-25, Vienna, Austria, 13–37.

Duda, R., Hart, P., and Stork, D., 2001. *Pattern Classification*. New York: John Wiley & Sons.

Dungan, J., *et al.*, 2002. A balanced view of scale in spatial statistical analysis. *Ecography*, 25 (5), 626–640.

Dykes, J., 2005. *Exploring Geovisualization*. Amsterdam: Elsevier.

Dykes, J. and Mountain, D., 2003. Seeking structure in records of spatio-temporal behaviour: visualization issues, efforts and applications. *Computational Statistics & Data Analysis*, 43, 581–603.

Dzialak, M., *et al.*, Unpb.. *Identifying functional habitat for brood-rearing greater sage-grouse. Unpublished Data.*. Technical report, The Samuel Roberts Noble Foundation.

Eagle, N. and Pentland, A., 2009. Eigenbehaviors: identifying structure in routine. *Behavioral Ecology and Sociobiology*, 63, 1057–1066.

Eiter, T. and Mannila, H., 1994. *Computing discrete Frchet distance.* Technical report, Technical University of Wien, CD-TR 94/64.

Elgethun, K., *et al.*, 2003. Time-location analysis for exposure assessment studies of children using a novel global positioning system instrument. *Environmental Health Perspectives*, 111 (1), 115–122.

Eriksen, A., *et al.*, 2008. Encounter frequencies between GPS-collared wolves (Canis lupus) and moose (Alces alces) in a Scandinavian wolf territory. *Ecological Research*, 24 (3), 547–557.

Erwig, M., *et al.*, 1999. Spatio-temporal data types: An approach to modeling and querying moving objects in databases. *Geoinformatica*, 3 (3), 269–296.

Fearnhead, P. and Taylor, B., 2011. On estimating the ability of NBA players. *Journal of Quantitative Analysis in Sports*, 7 (3), Article 11.

Fieberg, J. and Kochanny, C., 2005. Quantifying home-range overlap: The importance of the utilization distribution. *Journal of Wildlife Management*, 69 (4), 1346–1359.

Fieberg, J., *et al.*, 2010. Correlation and studies of habitat selection: problem, red herring or opportunity?. *Philosophical Transactions of the Royal Society B*, 365, 2233–2244.

Fieberg, J., 2007. Kernel density estimators of home range: smoothing and the autocorrelation red herring. *Ecology*, 88 (4), 1059–66.

Forer, P., 1998. Geometric approaches to the nexus of time, space, and microprocess: Implementing a practical model for mundane socio-spatial systems. *In*: M. Egenhofer and R. Golledge, eds. *Spatial and Temporal Reasoning in Geographic Information Systems*. New York: Oxford University Press, 171–190.

Frair, J., *et al.*, 2004. Removing GPS collar bias in habitat selection studies. *Journal of Applied Ecology*, 41, 201–212.

Frank, A., 1998. Different types of "times" in GIS. *In*: M. Egenhofer and R. Golledge, eds. *Spatial and Temporal Reasoning in Geographic Information Systems*. New York: Oxford University Press, 40–62.

Gannon, W., Sikes, R., and The Animal Care and Use Committee of the American Society of Mammalogists, 2007. Guidelines of the American Society of Mammalogists for the use of wild mammals in research. *Journal of Mammalogy*, 88 (3), 809–823.

Gao, Y., *et al.*, 2010. Algorithms for constrained k-nearest neighbor queries over moving object trajectories. *Geoinformatica*, 14, 241–276.

Gee, K.L., *et al.*, 1994. *White-tailed deer: their foods and management in the Cross Timbers. 2nd ed. Samual Roberts Noble Foundation Publication.* Technical report, Samuel Roberts Noble Foundation, Ardmore, Oklahoma.

Gee, K.L., Holman, J., and Demarais, S., 1999. A man-power efficient drop-net system for capturing white-tailed deer. *In*: *Abstracts of the 22nd Annual Southeast Deer Study Group Meeting*, p. 31.

Gehrt, S. and Fritzell, E., 1998. Duration of familial bonds and dispersal patterns for raccoons in south Texas. *Journal of Mammalogy*, 79 (3), 859–872.

Getz, W. and Wilmers, C., 2004. A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography*, 27, 489–505.

Giannotti, F., *et al.*, 2007. Trajectory pattern mining. *In*: *Knowledge Discovery and Data Mining*, San Jose, CA, USA, 330–339.

Girard, I., *et al.*, 2006. Balancing number of lcoations with number of individuals in telemetry studies. *The Journal of Wildlife Management*, 70 (5), 1249–1256.

Gonzalez, M., Hidalgo, C., and Barabasi, A.L., 2008. Understanding individual human mobility patterns. *Nature*, 453, 779–782.

Goodchild, M., 1992. Geographical information science. *International Journal of Geographical Information Systems*, 6 (1), 31–45.

Goodchild, M., 2008. Commentary: whither VGI. *GeoJournal*, 72, 239–244.

Gorman, T., *et al.*, 2006. Space use and sociality of river otters (Lontra canadensis) in Minnesota. *Journal of Mammalogy*, 87 (4), 740–747.

Gudmundsson, J. and van Kreveld, M., 2006. Computing longest duration flocks in trajectory data. *In*: *ACM-GIS'06, Nov. 10-11*, Arlington, VA, USA, p. 8.

Gudmundsson, J., Van Kreveld, M., and Speckmann, B., 2007. Efficient detection of patterns in 2D trajectories of moving points. *Geoinformatica*, 11, 195–215.

Gupta, R. and Gupta, R., 2004. Generalized skew normal model. *Test*, 13 (2), 501–524.

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.

Güting, R., Behr, T., and Xu, J., 2010. Efficient k-nearest neighbor search on moving object trajectories. *The VLDB Journal*, 19, 687–714.

Güting, R., *et al.*, 2000. A foundation for representing and querying moving objects. *ACM Transactions on Database Systems*, 25 (1), 1–42.

Güting, R. and Mamoulis, N., 2011. Special issue on data management for mobile services. *The VLDB Journal*, 20 (5), 641–642.

Güting, R. and Schneider, M., 2005. *Moving Objects Databases*. The Morgan Kaufmann Series in Data Management Systems New York, NY: Elsevier.

Hadjieleftheriou, M., *et al.*, 2003. On-line discovery of dense areas in spatio-temporal databases. *LNCS*, 2750, 306–325.

Hägerstrand, T., 1970. What about people in regional science?. *Papers of the Regional Science Association*, 24, 7–21.

Hallin, P., 1991. New paths for time-geography?. *Geografiska Annaler*, 73 (3), 199–207.

Hamer, D. and Herrero, S., 1990. Courtship and use of mating areas by grizzly bears in the front ranges of Banff National Park, Alberta. *Canadian Journal of Zoology*, 68, 2695–2697.

Hansen, M. and Riggs, R., 2008. Accuracy, precision, and observation rates of global positioning system telemetry collars. *Journal of Wildlife Management*, 72 (2), 518–526.

Hartenstein, W. and Iblher, H., 1967. Visitors' records: a method for analyzing the flow of people in the C.B.D.. *In*: U. of Amsterdam Sociological Department, ed. *Urban Core and Inner City*. Leiden: E.J. Brill, 497–502.

Hebblewhite, M., Merrill, E.H., and McDermid, G., 2008. A multi-scale test of the forage maturation hypohtesis in a partially migratory ungulate population. *Ecological Monographs*, 78 (2), 141–166.

Hemson, G., *et al.*, 2005. Are kernels the mustard? Data from global positioning system (GPS) collars suggests problems for kernel home-range analyses with least-squares cross-validation. *Journal of Animal Ecology*, 74 (3), 455–463.

Hengl, T., *et al.*, 2008. Geostatistical analysis of GPS trajectory data: space-time densities. *In*: *Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences* Shanghai, June 25-27: World Academic Union Press, 17–24.

Hirth, D., 1977. Social behavior of white-tailed deer in relation to habitat. *Wildlife Monographs*, 53, 3–55.

Hobbs, N.T. and Hilborn, R., 2006. Alternatives to statistical hypothesis testing in ecology: a guide to self teaching. *Ecological Applications*, 16 (1), 5–19.

Holly, B., 1978. The problem of scale in time-space research. *In*: T. Carlstein, D. Parkes and N. Thrift, eds. *Time and Regional Dynamics.*, Vol. 3 of *Timing Space and Spacing Time* New York: John Wiley & sons, 5–18.

Horne, J.S., *et al.*, 2007. Analyzing animal movements using Brownian bridges. *Ecology*, 88 (9), 2354–2363.

Hornsby, K. and Egenhofer, M., 2002. Modeling moving objects over multiple granularities. *Annals of Mathematics and Artificial Intelligence*, 36, 177–194.

Hunter, A., El-Sheimy, N., and Stenhouse, G., 2005. Up close and grizzly: GPS/camera collar captures bear doings. *GPS World*, Feb. 2005, 24–31.

Huttenlocher, D., Klanderman, G., and Rucklidge, W., 1993. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15 (9), 850–863.

Hwang, J.R., Kang, H.Y., and Li, K.J., 2005. Spatio-temporal similarity analysis between trajectories on road networks. *In*: J. Akoka, S. Liddle, I.Y. Song, M. Bertolotto, I. Comyn-Wattiau, W.J. van den Heuvel, M. Kolp, J. Trujillo, C. Kop and H. Mayr, eds. *Perspectives in Conceptual Modeling, ER Workshops, LNCS 3770* Springer-Verlag, 280–289.

Jacobs, J., 1974. Quantitative Measurement of Food Selection : A Modification of the Forage Ratio and Ivlev's Electivity Index. *Oecologia*, 14 (4), 413–417.

Jennrich, R. and Turner, F., 1969. Measurement of non-circular home range. *Journal of Theoretical Biology*, 22, 227–237.

Jensen, C., Lu, H., and Yang, B., 2010. Indoor - a new data management frontier. *IEEE Data Engineering Bulletin*, 33 (2), 12–17.

Jeung, H., *et al.*, 2008. Discovery of convoys in trajectory databases. *In*: *Proceedings VLDB*, Auckland, New Zealand, Aug. 23-28, 1068–1080.

Johnston, R., 2003. Geography: a different sort of discipline?. *Transaction of the Institute of British Geographers*, 28, 133–141.

Johnston, R., 1997. W(h)ither spatial science and spatial analysis. *Futures*, 29 (4/5), 323–336.

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., Myers, R., and Mills Flemming, J., 2003. Meta-analysis of animal movement using state-space models. *Ecology*, 84 (11), 3055–3063.

Kalbfleisch, J.G., 1985. *Probability and Statistical Inference*. Springer Texts in Statistics New York, NY: Springer New York.

Kalnis, P., Mamoulis, N., and Bakiras, S., 2005. On discovering moving clusters in spatio-temporal data. *In*: C. Medeiros, M. Egenhofer and E. Bertino, eds. *9th International Symposium on Advances in Spatial and Temporal Databases, LNCS 3633* Springer-Verlag, 364–381.

Kareiva, P. and Shigesada, N., 1983. Analyzing insect movement as a correlated random walk. *Oecologia*, 56, 234–238.

Karlin, M. and Chadwick, J., 2011. Measures of Space use and Association of Two Unrelated Male Red Wolves in a Shared Area Measures of space use and association of two unrelated male red wolves in a shared area. *Mammal Study*, 36 (3), 147–153.

Kenward, R., Marcstrom, V., and Karlbom, M., 1993. Post-nestling behaviour in goshawks, Accipiter gentilis: II. Sex differences in sociality and nest-switching. *Animal Behaviour*, 46, 371–378.

Kenward, R.E. and Hodder, K., 1998. Red squirrels (Sciurus vulgaris) released in conifer woodland: the effects of source habitat, predation and interactions with grey squirrels (Sciurus carolinensis). *Journal of Zoology*, 244 (1), 23–32.

Kernohan, B., Gitzen, R., and Millspaugh, J., 2001. Analysis of animal space use and movements. *In*: J. Millspaugh and J. Marzluff, eds. *Radio Tracking and Animal Populations*. New York: Academic Press, 125–166.

Kie, J.G., *et al.*, 2010. The home-range concept: are traditional estimators still relevant with modern telemetry technology?. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 365 (1550), 2221–2231.

Knighton, R. and Claramunt, C., 2001. An Aeronautical Temporal GIS for Post-Flight Assessment of Navigation Performance. *Transactions in GIS*, 5 (1), 53–66.

Knox, E., 1964. The detection of space-time interactions. *Journal of the Royal Statistical Society (Series C): Applied Statistics*, 13 (1), 25–30.

Kondo, K. and Kitamura, R., 1987. Time-space constraints and the formation of trip chains. *Regional Science and Urban Economics*, 17, 49–65.

Kraak, M., 2003. The space-time cube revisited from a geovisualization perspective. *In*: *Proceedings 21st International Cartographic Conference*, August 10-16, Durban, South Africa, 1988–1995.

Krumm, J., Letchner, J., and Horvitz, E., 2007. Map matching with travel time constraints. *In*: *SAE World Congress*, Detroit, MI, April 16-19, p. 11.

Kuijpers, B., Miller, H., and Othman, W., 2011. Kinetic space-time prisms. *In*: *19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* Nov. 1-4, Chicago, IL: Association for Computing Machinery, 162–170.

Kumar, S. and Stokkeland, J., 2003. Evolution of GPS technology and its subsequent use in commercial markets. *International Journal of Mobile Communications*, 1 (1/2), 180–193.

Kwan, M., 1998. Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis*, 30 (3), 191–216.

Kwan, M., 1999. Gender, the home-work link, and space-time patterns of nonemployment activities. *Economic Geography*, 75 (4), 370–394.

Kwan, M., 2004. GIS methods in time-geographic research: Geocomputation and geovisulization of human activity patterns. *Geografiska Annaler*, 86 B (4), 267–280.

Kwan, M. and Lee, J., 2004. Geovisualization of human activity patterns using 3D GIS: A time-geographic approach. *In*: M. Goodchild and D. Janelle, eds. *Spatially Integrated Social Science: Examples in Best Practice*. Oxford: Oxford University Press, 48–66.

Langran, G. and Chrisman, N., 1988. A framework for temporal geographic information. *Cartographica*, 25 (3), 1–14.

Laube, P., *et al.*, 2007. Movement beyond the snapshot - Dynamic analysis of geospatial lifelines. *Computers, Environment and Urban Systems*, 31, 481–501.

Laube, P., Imfeld, S., and Weibel, R., 2005. Discovering relative motion patterns in groups of moving point objects. *International Journal of Geographical Information Science*, 19 (6), 639–668.

Laube, P. and Purves, R., 2011. How fast is a cow? Cross-scale analysis of movement data. *Transactions in GIS*, 15 (3), 401–418.

Laube, P., van Kreveld, M., and Imfeld, S., 2004. Finding REMO - Detecting relative motion patterns in geospatial lifelines. *In*: P. Fisher, ed. *Developments in Spatial Data Handling, Proceedings of the 11th International Symposium on Spatial Data Handling*. Berlin: Springer, 201–215.

Laver, P. and Kelly, M., 2008. A critical review of home range studies. *Journal of Wildlife Management*, 72 (1), 290–298.

Lenntorp, B., 1976. *Paths in Space-Time Environments: a Time Geographic Study of Movement Possibilities of Individuals*. Lund Studies in Geography: Series B, Human Geography, no. 44 Stockholm: Gleerup.

Lenntorp, B., 1999. Time-geography - at the end of its beginning. *GeoJournal*, 48, 155–158.

Liu, J., *et al.*, 2009. Automatic player detection, labeling and tracking in broadcast soccer video. *Pattern Recognition Letters*, 30, 103–113.

Long, J.A. and Nelson, T.A., 2012. Time geography and wildlife home range delineation. *The Journal of Wildlife Management*, 76 (2), 407–413.

Long, J.A. and Nelson, T.A., 2013a. Measuring dynamic interaction in movement data. *Transactions in GIS*, 17 (1), 62–77.

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

Lu, W.L., Okuma, K., and Little, J.J., 2009. Tracking and recognizing actions of multiple hockey players using the boosted particle filter. *Image and Vision Computing*, 27 (1-2), 189–205.

Lu, W.L., *et al.*, 2011. Identifying players in broadcast sports videos using conditional random fields. *In*: *IEEE Conference on Computer Vision, CVPR'11*, 2, p. 8p.

Macdonald, D., Ball, F., and Hough, N., 1980. The evaluation of home range size and configuration using radio tracking data. *In*: C. Amlaner and D. Macdonald, eds. *A Handbook on Biotelemetry and Radio Tracking*. Oxford: Pergamon Press, 405–424.

Mace, R. and Waller, J., 1997. Spatial and temporal interaction of male and female grizzly bears in northwestern Montana. *The Journal of Wildlife Management*, 61 (1), 39–52.

Mark, D., 1998. Geospatial lifelines. *In*: *Integrating Spatial and Temporal Databases*, 98471, Dagstuhl Seminars, p. 12.

Mennis, J., Viger, R., and Tomlin, C., 2005. Cubic map algebra functions for spatiotemporal analysis. *Cartography and Geographic Information Science*, 32 (1), 17–32.

Merrill, E.H., *et al.*, 2010. Building a mechanistic understanding of predation with GPS-based movement data. *Philosophical Transactions of the Royal Society B*, 365, 2279–2288.

Merrill, S. and Mech, L., 2003. The usefulness of GPS telemetry to study wolf circadian and social activity. *Wildlife Society Bulletin*, 31 (4), 947–960.

Miller, H., 1991. Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems*, 5 (3), 287–301.

Miller, H., 2003. What about people in geographic information science?. *Computers, Environment and Urban Systems*, 17, 447–453.

Miller, H., 2005. A measurement theory for time geography. *Geographical Analysis*, 37, 17–45.

Miller, H., 2010. The data avalanche is here. Shouldn't we be digging?. *Journal of Regional Science*, 50 (1), 181–201.

Miller, H. and Bridwell, S., 2009. A field-based theory for time geography. *Annals of the Association of American Geographers*, 99 (1), 49–75.

Miller, H. and Wu, Y., 2000. GIS Software for measuring space-time accessibility in transportation planning and analysis. *Geoinformatica*, 4 (2), 141–159.

Miller, J., 2012. Using spatially explicit simulated data to analyze animal interactions: A case study with brown hyenas in Northern Botswana. *Transactions in GIS*, 16 (3), 271–291.

Millspaugh, J.S.J., *et al.*, 2004. Comparability of three analytical techniques to assess joint space use. *Wildlife Society Bulletin*, 32 (1), 148–157.

Millspaugh, J., *et al.*, 1998. Some comments on spatial independence in studies of resource selection. *Wildlife Society Bulletin*, 26 (2), 232–236.

Minta, S., 1992. Tests of spatial and temporal interaction amoung animals. *Ecological Applications*, 2 (2), 178–188.

Minta, S., 1993. Sexual differences in spatio-temporal interaction among badgers. *Oecologia*, 96 (3), 402–409.

Mintsis, G., *et al.*, 2004. Applications of GPS technology in the land transportation system. *European Journal of Operational Research*, 152 (2), 399–409.

Miskelly, F., 2005. Electronic tracking of patients with dementia and wandering using mobile phone technology. *Age and Ageing*, 34, 497–499.

Mitchell, M. and Powell, R., 2008. Estimated home ranges can misrepresent habitat relationships on patchy landscapes. *Ecological Modelling*, 216, 409–414.

Mosnier, A., *et al.*, 2003. Habitat selection and home-range dynamics of the Gaspe caribou: a hierarchical analysis. *Canadian Journal of Zoology*, 81, 1174–1184.

Nams, V.O., 2005. Using animal movement paths to measure response to spatial scale. *Oecologia*, 143 (2), 179–88.

Nanni, M. and Pedreschi, D., 2006. Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems*, 27, 267–289.

Nathan, R., *et al.*, 2008. A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Science*, 105 (49), 19052–19059.

Nelson, T., *et al.*, Unpb. *Mapping wildilfe movement patterns: Dominant utilization zones. Unpublished Data.*. Technical report, Spatial Pattern Analysis & Research Lab, Department of Geography, University of Victoria.

Neutens, T., Schwanen, T., and Witlox, F., 2011. The prism of everyday life: Towards a new research agenda for time geography. *Transport Reviews*, 31 (1), 25–47.

Neutens, T., *et al.*, 2010. Evaluating the temporal organization of public service provision using space-time accessibility analysis. *Urban Geography*, 31 (8), 1039–1064.

Neutens, T., *et al.*, 2007. Human interaction spaces under uncertainty. *Transportation Research Record*, 2021, 28–35.

Nicholson, K.L., *et al.*, 2011. Spatial and temporal interactions of sympatric mountain lions in Arizona. *European Journal of Wildlife Research*, 57 (6), 1151–1163.

Nielsen, S., *et al.*, 2002. Modeling grizzly bear habitats in the Yellowhead ecosystem of Alberta: Taking autocorrelation seriously. *Ursus*, 13, 45–56.

Nielsen, S., *et al.*, 2003. Development and testing of phenologically driven grizzly bear habitat models. *Ecoscience*, 10 (1), 1–10.

Noyon, V., Claramunt, C., and Devogele, T., 2007. A relative representation of trajectories in geographical spaces. *Geoinformatica*, 11, 479–496.

Okabe, A., Satoh, T., and Sugihara, K., 2009. A kernel density estimation method for networks, its computational method and a GIS-based tool. *International Journal of Geographical Information Science*, 23 (1), 7–32.

Openshaw, S., 1984. *The Modifiable Areal Unit Problem*. CATMOG-38 Norwich, England: GeoBooks.

Orians, G. and Pearson, N., 1979. On the theory of central place foraging. *In*: D. Horn, G. Stairs and R. Mitchell, eds. *Analysis of Ecological Systems*. Columbus, OH: Ohio State University Press, 155–178.

O'Sullivan, D., Morrison, A., and Shearer, J., 2000. Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. *International Journal of Geographical Information Science*, 14 (1), 85–104.

O'Sullivan, D. and Unwin, D., 2010. *Geographic Information Analysis*. 2nd Hoboken, NJ: John Wiley & Sons.

Othmer, H., Dunbar, S., and Alt, W., 1988. Models of dispersal in biological systems. *Journal of Mathematical Biology*, 26, 263–298.

Otis, D. and White, G., 1999. Autocorrelation of location estimates and the analysis of radiotracking data. *The Journal of Wildlife Management*, 63 (3), 1039–1044.

Palma, A., *et al.*, 2008. A clustering-based approach for discovering interesting places in trajectories. *In*: *Proceedings of the 2008 ACM Symposium on Applied Computing* Fortaleza, Ceara, Brazil: ACM, p. 6.

Parkes, D. and Thrift, N., 1975. Timing space and spacing time. *Environment and Planning A*, 7, 651–670.

Patterson, T.A., *et al.*, 2009. Classifying movement behaviour in relation to environmental conditions using hidden Markov models. *The Journal of Animal Ecology*, 78 (6), 1113–1123.

Pearson, K., 1905. The problem of the random walk. *Nature*, 72 (1865), 294.

Peuquet, D., 1994. It's about time: a conceptual framework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of American Geographers*, 84 (3), 441–461.

Peuquet, D., 2002. *Representations of Space and Time*. New York, NY: The Guilford Press.

Pfoser, D., Jensen, C., and Theodoridis, Y., 2000. Novel approaches to the indexing of moving object trajectories. *In*: *26th Conference on Very Large Databases*, Cairo, Egypt, Sep. 10-14, p. 12p.

Poole, K., 1995. Spatial organization of a lynx population. *Canadian Journal of Zoology*, 73, 632–341.

Powell, R., 2000. Animal home ranges and territories and home range estimators. *In*: L. Boitani and T. Fuller, eds. *Research Techniques in Animal Ecology: Controversies and Consequences*. New York: Columbia University Press, 65–110.

Prager, S., 2007. Environmental contextualization of uncertainty for moving objects. *Computers, Environment and Urban Systems*, 31, 303–316.

Prager, S. and Yu, B., 2005. Contextualized probability for approximation of spatiotemporal data distributions. *In*: *International Conference on Cybernetics and Information Technologies, Systems and Applications (ISAS CITSA)*, Orlando, Fl., 318–322.

Pred, A., 1981. *Space and Time in Geography*. Lund Studies in Geography, ser. B, Human Geography, no. 48 Gleerup: The Royal University of Lund.

Pultar, E., *et al.*, 2010. EDGIS: a dynamic GIS based on space time points. *International Journal of Geographical Information Science*, 24 (3), 329–346.

R Development Core Team, 2012. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raper, J., 2002. The dimensions of GIScience. *In*: *Second International Conference of Geographic Information Science, GIScience 2002.*, Boulder, CO, Last retrieved: May 2011, ¡http://www.soi.city.ac.uk/ raper/research/GIScience2002-OHs-pub.ppt¿.

Raper, J., *et al.*, 2007. A critical evaluation of location based services and their potential. *Journal of Location Based Services*, 1 (1), 5–45.

Rempel, R., Rodgers, A., and Abraham, K., 1995. Performance of a GPS animal location system under boreal forest canopy. *Journal of Wildlife Management*, 59 (3), 543–551.

Righton, D. and Mills, C., 2006. Application of GIS to investigate the use of space in coral reef fish: a comparison of territorial behaviour in two Red Sea butterflyfishes. *International Journal of Geographical Information Science*, 20 (2), 215–232.

Robert, K., Garant, D., and Pelletier, F., 2012. Keep in touch: Does spatial overlap correlate with contact rate frequency?. *Journal of Wildlife Management*, p. 6.

Robson, D. and Whitlock, J., 1964. Estimation of a truncation point. *Biometrika*, 51 (1-2), 33–39.

Rodriguez, D., Brown, A., and Troped, P., 2005. Portable global positioning units to complement accelerometry-based physical activity monitors. *Medicine & Science in Sports & Exercise*, 37 (11), (Suppl.) S572–S581.

Rose, C., 1977. Reflections on the notion of time incorporated in Hägerstrand's time-geographic model of society. *Tijdschrift voor Economische en Sociale Geografie*, 68 (1), 43–50.

Samuel, M., Pierce, D., and Garton, E., 1985. Identifying areas of concentrated use within the home range. *Journal of Animal Ecology*, 54 (3), 711–719.

Sanderson, G., 1966. The study of mammal movements: A review. *Journal of Wildlife Management*, 30 (1), 215–235.

Sawyer, H., Lindzey, F., and McWhirter, D., 2005. Mule deer and pronghorn migration in Western Wyoming. *Wildlife Society Bulletin*, 33 (4), 1266–1273.

Sawyer, H., *et al.*, 2007. Habitat selection of Rocky Mountain Elk in a nonforested environment. *Journal of Wildlife Management*, 71 (3), 868–874.

Schwartz, C., *et al.*, 2010. Contrasting activity patterns of sympatric and allopatric black and grizzly bears. *Journal of Wildlife Management*, 74 (8), 1628–1638.

Seaman, D. and Powell, R., 1996. An evaluation of the accuracy of kernel density estimators for home range analysis. *Ecology*, 77 (7), 2075–2085.

Shao, F., Cai, S., and Gu, J., 2010. A modified Hausdorff distance based algorithm for 2-dimensional spatial trajectory matching. *In*: *The 5th International Conference on Computer Science & Education*, Aug. 24-27, Hefei, China, 166–172.

Sharp, N., 1997. Timed running speed of a cheetah (Acinonyx jubatus). *Journal of Zoology*, 241 (3), 493–494.

Shaw, S.L., Yu, H., and Bombom, L., 2008. A space-time GIS approach to exploring large individual-based spatiotemporal datasets. *Transactions in GIS*, 12 (4), 425–441.

Sheppard, E., 2001. Quantitative geography: representations, practices, and possibilities. *Environment and Planning D: Society and Space*, 19, 535–554.

Shirabe, T., 2006. Correlation analysis of discrete motions. *In*: M. Raubal, H. Miller, A. Frank and M. Goodchild, eds. *GIScience 2006. LNCS, vol. 4197.*, Vol. 4197 Berlin: Springer-Verlag, 370–382.

Shoshany, M., Even-Paz, A., and Bekhor, S., 2007. Evolution of clusters in dynamic point patterns: with a case study of Ants' simulation. *International Journal of Geographical Information Science*, 21 (7), 777–797.

Shoval, N., *et al.*, 2008. The use of advanced tracking technologies for the analysis of mobility in alzheimer's disease and related cognitive diseases. *BMC Geriatrics*, 8, 7.

Shoval, N. and Isaacson, M., 2007. Sequence alignment as a method for human activity analysis in space and time. *Annals of the Association of American Geographers*, 97 (2), 282–297.

Silverman, B., 1986. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.

Sinha, G. and Mark, D., 2005. Measuring similarity between geospatial lifelines in studies of environmental health. *Journal of Geographical Systems*, 7, 115–136.

Skellam, J., 1951. Random dispersal in theoretical populations. *Biometrika*, 38, 196–218.

Smulders, M., Spatial-temporal analysis of grizzly bear habitat use. Master's thesis, University of Victoria, 2009. .

Spaccapietra, S., *et al.*, 2008. A conceptual view on trajectories. *Data & Knowledge Engineering*, 65, 126–146.

Stenhouse, G., *et al.*, 2005. Grizzly bear associations along the eastern slopes of Alberta. *Ursus*, 16 (1), 31–40.

Stewart Hornsby, K. and Cole, S., 2007. Modeling moving geospatial objects from an event-based perspective. *Transactions in GIS*, 11 (4), 555–573.

Swihart, R. and Slade, N., 1985. Influence of sampling interval on estimates of home-range size. *Journal of Wildlife Management*, 49 (4), 1019–1025.

Swihart, R. and Slade, N., 1997. On testing for independence of animal movements. *Journal of Agricultural, Biological, and Environmental Statistics*, 2 (1), 48–63.

Tester, J. and Heezen, K., 1965. Deer response to a drive census determined by radio tracking. *BioScience*, 15 (2), 100–104.

Tew, T. and Macdonald, D., 1994. Dynamics of space use and male vigour amongst wood mice, Apodemus sylvaticus, in the cereal ecosystem. *Behavioral Ecology and Sociobiology*, 34 (5), 337–345.

Thomas, J.J. and Cook, K.A., eds. , 2005. *Illuminating the path: The research and development agenda for visual analytics*. Vol. 54. National Visualization and Analytics Center.

Tomkiewicz, S., *et al.*, 2010. Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society B*, 365, 2163–2176.

Tremblay, Y., *et al.*, 2006. Interpolation of animal tracking data in a fluid environment. *The Journal of Experimental Biology*, 209, 128–140.

Turchin, P., 1998. *Quantitative Analysis of Movement*. Sinauer Associates.

Urbano, F., *et al.*, 2010. Wildlife tracking data management: a new vision. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365 (1550), 2177–85.

van der Watt, P., 1980. A note on estimation bounds of random variables. *Biometrika*, 67 (3), 712–714.

Van Winkle, W., 1975. Comparison of several probabilistic home-range models. *Journal of Wildlife Management*, 39 (1), 118–123.

Verhein, F. and Chawla, S., 2006. Mining spatio-temporal association rules, sources, sinks, stationary regions and thoroughfares in object mobility databases. *In*: *Proceedings of the 11th International Conference on Database Systems for Advanced Applications (DASFAA)*, Vol. Vol. 3882, Lecture Notes in Computer Science Springer, 187–201.

Verhein, F. and Chawla, S., 2008. Mining spatio-temporal patterns in object mobility databases. *Data Mining & Knowledge Discovery*, 16, 5–38.

Viswanathan, G., *et al.*, 1996. Lévy flight search patterns of wandering albatrosses. *Nature*, 381, 413–415.

Vlachos, M., Gunopulos, D., and Kollios, G., 2002. Robust similarity measures for mobile object trajectories. *In*: *5th International Workshop on Mobility in Databases and Distributed Systems (MDDS)*, Aix-en-Provence, France, 721–726.

Webb, S., *et al.*, 2010. Measuring fine-scale white-tailed deer movements and environmental influences using GPS collars. *International Journal of Ecology*, 2010 (Article ID 459610), 12p.

Webb, S., *et al.*, 2009. Using fractal analyses to characterize movement paths of white-tailed deer and response to spatial scale. *Journal of Mammalogy*, 90 (5), 1210–1217.

Weiss, R. and Boutourline, S., 1963. *Fairs, pavilions, exhibits, and their audiences.* Technical report, IBM Corporation.

Wells, L., 1981. Real-time missile tracking with GPS. *In*: *National Aerospace Meeting*, Apr. 8-10, Washington, DC, 56–61.

Wentz, E., Campbell, A., and Houston, R., 2003. A comparison of two methods to create tracks of moving objects: linear weighted distance and constrained random walk. *International Journal of Geographical Information Science*, 17 (7), 623–645.

White, P.C.L. and Harris, S., 1994. Encounters between red foxes (Vulpes vudpes): implications for territory maintenance , social cohesion and dispersal. *Journal of Animal Ecology*, 63 (2), 315–327.

Wiens, J., Crist, T., and Milne, B., 1993. On quantifying insect movements. *Environmental Entomology*, 22 (4), 709–715.

Winter, S., 2009. Towards a probabilistic time geography. *In*: M. Mokbel, P. Scheuermann and W. Aref, eds. *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* Seattle, WA: ACM, New York, NY, 528–531.

Winter, S. and Yin, Z., 2010. Directed movements in probabilistic time geography. *International Journal of Geographical Information Science*, 24 (9), 1349–1365.

Winter, S. and Yin, Z., 2011. The elements of probabilistic time geography. *Geoinformatica*, 15 (3), 417–434.

Wolf, J., Guensler, R., and Bachman, W., 2001. Elimination of the travel diary: Experiment to derive trip purpose from global positioning system travel data. *Transportation Research Record*, 1768, 125–134.

Wolfer, D., *et al.*, 2001. Extended analysis of path data from mutant mice using the public domain software Wintrack. *Physiology & Behavior*, 73, 745–753.

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

Worton, B., 1989. Kernel methods for estimating the utilization distribution in home-range studies. *Ecology*, 70 (1), 164–168.

Wyckoff, A.C., *et al.*, 2009. Feral swine contact with domestic swine: a serologic survey and assessment of potential for disease transmission.. *Journal of Wildlife Diseases*, 45 (2), 422–9.

Yanagisawa, Y., Akahani, J., and Satoh, T., 2003. Shape-based similarity query for trajectory of mobile objects. *In*: M.S. Chen, P. Chysanthis, M. Sloman and A. Zaslavsky, eds. *Mobile Data Management, LNCS 2574*. Berlin: Springer-Verlag, 63–77.

Yu, B. and Kim, S., 2006. Interpolating and using most likely trajectories in moving-objects databases. *In*: S. Bressan, J. Kung and R. Wagner, eds. *DEXA 2006, LNCS 4080* Springer-Verlag, 718–727.

Zhang, Z., Huang, K., and Tan, T., 2006. Comparison of similarity measures for trajectory clustering in outdoor surveillance scenes. *In*: *IEEE International Conference on Pattern Recognition*, 1135–1138.

Appendix A

Documentation of R Tools

PPA Home Range

May 16, 2013

`calcVmax` *Calculate estimates of a mobility parameter for object movement data*

Description

This function computes various estimates for the v_{max} parameter for object movement data, following Long and Nelson (2012).

Usage

```
calcVmax(traj, k=5, alpha=0.05)
```

Arguments

`traj`	an object of the class `ltraj` which contains the time stamped animal telemetry data. Note this object must be a `type II ltraj` object. For more information on objects of this type see `help(ltraj)`.
`k`	value for the *k* parameter in the van der watt (1980) method.
`alpha`	value for the α parameter for the upper and lower C.I. methods.

Details

This function is used to compute a set of estimates of the v_{max} parameter for comparison across methods.

Value

This function returns a data frame with the values from the Robson and Whitlock (1964) and van der Watt (1980) methods for estimating the upper bounds of a distribution, along with the lower and upper $(1 - \alpha) * 100\%$ confidence interval limits.

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Long, J.A., Nelson, T.A. 2012. Time geography and wildlife home range delineation. *Journal of Wildlife Management*. 76(2):407-413.

Robson, D.S., Whitlock, J.H. 1964. Estimation of a truncation point. *Biometrika* 51:33-39.

van der Watt, P. 1980. A note on estimation bounds of random variables. *Biometrika* 67(3):712-714.

See Also

PPAHomeRange

Examples

```
## simulate an animal trjectory:
traj <- simm.crw(1:100)
calcVmax(traj,k=6,alpha=0.1)
```

`PPAHomeRange` *Calculation of the PPA Home Range*

Description

This function computes the PPA measure of animal home range from time stamped telemetry data. It can be broadly defined as that area accessible to an animal along it's path, given telemetry fixes and a known level of mobility.

Usage

```
PPAHomeRange(traj, method = "Robson", k = 5, alpha = 0.05,
             manualVmax, tol, ePoints)
```

Arguments

`traj`	an object of the class `ltraj` which contains the time stamped animal telemetry data. Note this object must be a `type II ltraj` object. For more information on objects of this type see `help(ltraj)`.
`method`	method for computing the v_{max} parameter determining the object mobility; can be one of several options: `"Robson"` for the Robson & Whitlock (1964) method, `"RobsonLL"` for the R & W (1964) lower $(1-\alpha)*100\%$ C.I. limit, `"RobsonUL"` for the R & W (1964) upper $(1-\alpha)*100\%$ C.I. limit, `"vanderWatt"` for the van der Watt (1980) method, `"vanderWattLL"` for the van der Watt (1980) lower $(1-\alpha)*100\%$ C.I. limit, `"vanderWattUL"` for the van der Watt (1980) upper $(1-\alpha)*100\%$ C.I. limit, `"manual"` to select your own value.
`k`	value for the *k* parameter in the van der watt (1980) method.
`alpha`	value for the α parameter if using upper or lower C.I. methods.
`manualVmax`	value for the v_{max} parameter if using the `manual` method.

tol	Often missing fixes will result in long time intervals between fixes; the tol option can be used to ignore those segments longer than the specified value. Default is simply the maximum observed time interval (in seconds), so all segments are included.
ePoints	number of vertices used to construct each ellipse. More points will necessarily provide a more detailed ellipse but slow computation. Default is 360.

Details

This function can be used for calculating the PPA home range as described by Long and Nelson (2012). The PPA home range method is only appropriate for wildlife telemetry data that is characterized by a reasonably short time interval between fixes.

Value

An object of class SpatialPolygons is returned that can be used to map the animals PPA home range.

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Long, J.A., Nelson, T.A. 2012. Time geography and wildlife home range delineation. *Journal of Wildlife Management*. 76(2):407-413.

Robson, D.S., Whitlock, J.H. 1964. Estimation of a truncation point. *Biometrika* 51:33-39.

van der Watt, P. 1980. A note on estimation bounds of random variables. *Biometrika* 67(3):712-714.

See Also

calcVmax

Examples

```
## simulate an animal trjectory:
traj <- simm.crw(1:100)

hr <- PPAHomeRange(traj, method = "vanderWatt", k = 7)
plot(hr)
```

DI Tools

February 13, 2013

`Cole` *Compute Cole statistic for dynamic interaction*

Description

This function measures the dynamic interaction between two moving objects following the methods first described by Cole (1949), and more recently employed by Bauman (1998).

Usage

```
Cole(traj1, traj2, tc=0, dc=100)
```

Arguments

`traj1`	an object of the class `ltraj` which contains the time-stamped movement fixes of the first object. Note this object must be a `type II ltraj` object. For more information on objects of this type see `help(ltraj)`.
`traj2`	same as `traj1`.
`tc`	temporal tolerance limit (in seconds) for defining when two fixes are simultaneous or together. Parameter passed to function `GetSimultaneous`.
`dc`	distance tolerance limit (in appropriate units) for defining when two fixes are spatially together.

Details

This function can be used to calculate the Cole (1949) measure of dynamic interaction between two animals. Termed a coefficient of association, the Cole statistic tests the number of fixes the animals are observed together against the total number of fixes following:

$$Ca = \frac{2AB}{A+B}$$

where A (respectively B) is the number of times animal 1 (resp. 2) are observed, and AB is the number of times the two animals are observed together. Several works, including Bauman (1998) have suggested that Ca > 0.5 indicates affiliation or fidelity, while Ca < 0.5 indicates no association between the two animals. Note that this function calls `GetSimultaneous` to identify the temporal component of identifying when fixes together.

Value

This function returns a numeric result of the Ca statistic.

`Ca`	The coefficient of association, see **Details**

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Bauman, P.J. (1998) The Wind Cave National Park elk herd: home ranges, seasonal movements, and alternative control methods. M.S. Thesis. South Dakota State University, Brookings, South Dakota, USA.

Cole, L.C. (1949) The measurement of interspecific association. *Ecology*. **30**, 411–424.

See Also

Doncaster, Kenward, Minta, GetSimultaneous

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:100)
traj2 <- simm.crw(1:100)

Cole(traj1,traj2,dc=2)
```

---

di *Measuring dynamic interaction in movement data*

Description

This function measures the dynamic interaction between two moving objects. It calculates the local level di statistic for movement displacement, direction, and overall. Output is appended to the first `ltraj` object input into the function. It also can compute time- and/or distance-based weighting schemes following Long and Nelson (2012).

Usage

```
di(traj1, traj2, alpha=1, TimeWeight=FALSE, DistWeight=False)
```

Arguments

`traj1`	an object of the class `ltraj` which contains the time-stamped movement fixes of the first object. Note this object must be a `type II ltraj` object. For more information on objects of this type see `help(ltraj)`.
`traj2`	same as `traj1`.
`alpha`	value for the α parameter in the formula for di_d.
`TimeWeight`	whether or not to compute the weights for time-based weighting.
`DistWeight`	whether or not to compute the weights for distance-based weighting.

Details

This function can be used for calculating the dynamic interaction (di) statistic as described in Long and Nelson (2012). The di statistic can be used to measure the local level of dynamic interaction between two moving objects. Specifically, it measures dynamic interaction in movement direction and displacement. This method requires two movement datasets with simultaneous fixes.

Value

This function returns an object of the class `ltraj` that corresponds to the first `ltraj` object input into the function (i.e., `traj1`). The resulting values are appended to the object `infolocs` metadata as `di`, `di.theta`, and `di.d` representing dynamic interaction overall, in direction (azimuth), and in displacement respectively. The `infolocs` information can be accessed using the function `ld()` converting `ltraj` objects into a regular dataframe. If time- and/or distance-based weighting is selected, the corresponding weights are included as additional columns `t.weight` and `d.weight`, respectively. Please see Long and Nelson (2012) for a more detailed description of the the di method.

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Long, J.A., Nelson, T.A. 2012. Measuring dynamic interaction in movement data. *Transactions in GIS*. DOI: 10.1111/j.1467-9671.2012.01353.x

See Also

Shirabe

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:100)
traj2 <- simm.crw(1:100)

di1 <- di(traj1, traj2)
di2 <- di(traj1, traj2, TimeWeight=TRUE)
plot(ld(di1)$date,ld(di1)$di.theta)

#Compute the global DI statistic
DI <- mean(ld(di1)$di, na.rm=TRUE)
DI.TimeWeighted <- sum(ld(di2)$di * ld(di2)$t.weight, na.rm=TRUE)
```

Doncaster *Compute Doncaster measure of dynamic interaction*

Description

This function measures the dynamic interaction between two moving objects following the methods outlined by Doncaster (1990).

Usage

```
Doncaster(traj1, traj2, dc=100)
```

Arguments

traj1	an object of the class ltraj which contains the time-stamped movement fixes of the first object. Note this object must be a type II ltraj object. For more information on objects of this type see help(ltraj).
traj2	same as traj1, but should contain simultaneous fixes to traj1 – see the function GetSimultaneous.
dc	critical distance, in appropriate units, for examining the interaction of the two animals using a Chi-squared test.

Details

This function can be used to compute the Doncaster (1990) methods for measuring dynamic interaction between two animals. The Doncaster method tests the proportion of simultaneous fixes that are below dc against that which would be expected based on the distribution of distances between all fixes.

Value

This function first returns a plot, for distance values ranging from 0 to the maximum distance separating two fixes, of the observed proportion of simultaneous fixes below each distance value. The expected values based on all fixes are also included. Second, a list is returned that contains the contingency table of simultaneous fixes (paired) and non-paired fixes below and above dc, along with the associated *p*-value from the Chi-squared test.

conTable	contingency table showing frequency of paired and non-paired fixes above and below dc.
p.value	*p*-value from the Chi-squared test of conTable.

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Doncaster, C.P. (1992) Non-parametric estimates of interaction from radio-tracking data. *Journal of Theoretical Biology*, **143**: 431-443.

See Also

Cole, Kenward, Minta, GetSimultaneous

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:100)
traj2 <- simm.crw(1:100)

Doncaster(traj1,traj2,dc=3)
```

GetSimultaneous *Derive the simultaneous fixes for movement analysis*

Description

This function identifies and extracts simultaneous fixes, within a given tolerance limit, between two movement datasets

Usage

```
GetSimultaneous(traj1, traj2, tc=0)
```

Arguments

traj1	an object of the class ltraj which contains the time-stamped movement fixes of the first object. Note this object must be a type II ltraj object. For more information on objects of this type see help(ltraj).
traj2	same as traj1.
tc	tolerance limit (in seconds) allowing discrepency between fixes that the algorithm will determine as simultaneous.

Details

This function is used to determine the simultaneous fixes between two movement datasets facilitating further analysis. The time-stamps of the simultaneous fixes from traj1 are given to traj2 in the output.

Value

A list of two ltraj objects, containing simultaneous fixes.

Author(s)

Jed A. Long <jlong@uvic.ca>

See Also

di, shirabe

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:100)
traj2 <- simm.crw(seq(1:100,by=2))

simul <- GetSimultaneous(traj1,traj2)
tr1 <- simul[1]
tr2 <- simul[2]
```

HAI *Calculate Half-weight Association Index*

Description

This function computes the Half-weight Association Index for examining the presence of dynamic interaction in wildlife telemetry studies. This implementation follows that outlined in the paper Atwood and Weeks (2003).

Usage

```
HAI(traj1, traj2, hr1, hr2, dc=50)
```

Arguments

traj1	an object of the class ltraj which contains the time-stamped movement fixes of the first object. Note this object must be a type II ltraj object. For more information on objects of this type see help(ltraj).
traj2	same as traj1, but should contain simultaneous fixes to traj1 – see the function: GetSimultaneous.
hr1	home range polygon associated with traj1. Required to be an object that coerces to class SpatialPolygons.
hr2	same as hr1, but for traj2.
dc	critical distance, in appropriate units, for determining whether the two animals are together.

Details

This function can be used to test for the presence of dynamic interaction within the shared area (often termed the overlap zone) of the two animals home ranges. It can be interpreted as an extension of the Minta (1992) statistics, focusing spatially on interactions within the shared home range area. It's calculation is identical to that for Ca, but considers only those fixes in the shared area.

Value

This function returns the numeric value of the HAI statistic. Values near 1 indicate attraction within the shared home range area, while values near 0 indicate avoidance within this shared area.

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Atwood, T.C. and Weeks Jr., H.P. (2003) Spatial home-range overlap and temporal interaction in eastern coyotes: The influence of pair types and fragmentation. *Canadian Journal of Zoology*, **81**: 1589-1597.

Minta, S.C. (1992) Tests of spatial and temporal interaction among animals. *Ecological Applications*, **2**: 178-188.

See Also

Minta, GetSimultaneous

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:1000)
traj2 <- simm.crw(1:1000)

#compute home ranges using MCP
pts1 <- SpatialPoints(cbind(traj1[[1]]$x,traj1[[1]]$y))
pts2 <- SpatialPoints(cbind(traj2[[1]]$x,traj2[[1]]$y))
hr1 <- gBuffer(gConvexHull(pts1),width=0.001)
hr2 <- gBuffer(gConvexHull(pts2),width=0.001)

HAI(traj1,traj2,hr1,hr2,dc=2)
```

Kenward *Compute Kenward statistic for dynamic interaction*

Description

This function measures the dynamic interaction between two moving objects following the methods outlined by Kenward et al. (1993).

Usage

```
Kenward(traj1, traj2)
```

Arguments

`traj1`	an object of the class `ltraj` which contains the time-stamped movement fixes of the first object. Note this object must be a `type II ltraj` object. For more information on objects of this type see `help(ltraj)`.
`traj2`	same as `traj1`, but should contain simultaneous fixes to `traj1` – see the function: `GetSimultaneous`.

Details

This function can be used to calculate the Kenward et al. (1993) measure of dynamic interaction between two animals. The Kenward statistic tests the observed mean distance between simultaneous fixes against that expected by the overall distribution of distances between all fixes. Kenward et al. (1993) propose a sociality coefficient as a useful metric for exploring dynamic interaction, given by:

$$Cs = \frac{D_E - D_O}{D_O + D_E}$$

Where D_O is the mean observed distance between simultaneous fixes, and D_E is the mean expected distance between all fixes. Values for Cs closer to 1 indicate attraction, while values for C_s closer to -1 indicate avoidance. Values of Cs near 0 indicate that the two animals' movements have no influence on one another.

Further, the difference between the observed and expected distances are compared using a paired signed-rank test (both one-sided tests, indicative of attraction or avoidance). It should be noted that the function `GetSimultaneous` can be used to obtain simultaneous fixes from two trajectories.

Value

This function returns a list of objects representing the calculated values from the Kenward statistic and associated *p*-values from the signed rank test.

`Do`	The mean distance of simultaneous fixes
`De`	The mean expected distance, from all fixes
`Cs`	The coefficient of sociality, see **Details**
`p.Attract`	One sided *p*-value from signed rank test, testing for attraction
`p.Avoid`	One sided *p*-value from signed rank test, testing for avoidance

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Kenward, R.E., Marcstrom, V. and Karlbom, M. (2012) Post-nestling behaviour in goshawks, *Accipiter gentilis: II*. Sex differences in sociality and nest-switching. *Animal Behaviour*. **46**, 371–378.

See Also

Doncaster, Minta, Cole, GetSimultaneous

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:100)
traj2 <- simm.crw(1:100)

Kenward(traj1,traj2)
```

`Minta` *Compute Minta statistics for dynamic interaction*

Description

This function measures the dynamic interaction between two moving objects following the methods outlined by Minta (1992).

Usage

```
Minta(traj1, traj2, hr1, hr2, method="frequency", tc=0)
```

Arguments

`traj1`	an object of the class `ltraj` which contains the time-stamped movement fixes of the first object. Note this object must be a `type II ltraj` object. For more information on objects of this type see `help(ltraj)`.
`traj2`	same as `traj1`.
`hr1`	home range polygon associated with `traj1`. Required to be an object that coerces to class `SpatialPolygons`.
`hr2`	same as `hr1`, but for `traj2`.
`method`	method for computing the marginal distribution from which expected values are computed. If `method = "frequency"`, the marginal values are calculated based on the number of all fixes within the shared and unshared portions of the home ranges. If `method = "spatial"`, the marginal values are calculated based on areas of the shared and unshared portions of the home ranges. Note that this selection can greatly influence the results.
`tc`	time tolerance limit for determining simultaneous fixes (in seconds), to be passed to the function `GetSimultaneous`.

Details

This function can be used to calculate the Minta (1992) measures of dynamic interaction between two animals. The Minta statistic tests how the two animals utilize a shared area of their home ranges (defined as the intersection). Three coefficients are produced L_{AA}, L_{BB}, and L_{ixn}. Each of these statistics are based on a contingency table that compares the observed frequency of those fixes that are simultaneous to the expectations derived from all fixes (if `method="frequency"`) or from the area overlap proportions (if `method="spatial"`) – see Minta (1992) for more details. A Chi-squared statistic can then be used to examine the significance between the observed and expected use of the shared portion.

Minta (1992) suggests the following interpretations of the coefficients. When L_{AA} is near 0, the first animal's use of the shared area is random (or as expected). When $L_{AA} > 0$ it signifies spatial attraction to the shared area, or greater than expected use. When $L_{AA} < 0$ it signifies spatial avoidance of the shared area, or less than expected use. Interpretation of L_{BB} is the same as for L_{AA} with respect to the second animal. L_{ixn} tells us far more about the nature of the interaction between the two individuals. As L_{ixn} nears 0, both animals use the shared area randomly, with regards to the other animal. If $L_{ixn} > 0$ the animals use the shared area more *simultaneously*, whereas if $L_{ixn} < 0$ it is an indication of *solitary* use, or avoidance. This is why L_{ixn} is termed the temporal interaction coefficient. A Chi-squared test can be used to identify the significance of the L_{AA}, L_{BB}, and L_{ixn} values.

Value

This function returns a list of objects representing the calculated values from the Minta statistic and associated *p*-values from the Chi-squared test.

`pTable`	contingency table showing marginal probabilities of expected use based on the selecton of the `method` parameter.
`nTable`	contingency table showing observed frequency of use of the shared area based on simultaneous fixes.
`oTable`	the odds for each cell in the contingency table.
`Laa`	the calculated value of the L_{AA} statistic

p.AA	the associated *p*-value
Lbb	the calculated value of the L_{BB} statistic
p.BB	the associated *p*-value
Lixn	the calculated value of the L_{ixn} statistic
p.IXN	the associated *p*-value

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Minta, S.C. (1992) Tests of spatial and temporal interaction among animals. *Ecological Applications*, **2**: 178-188.

See Also

Cole, Doncaster, Kenward, GetSimultaneous

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:100)
traj2 <- simm.crw(1:100)

#compute home ranges using MCP
pts1 <- SpatialPoints(cbind(traj1[[1]]$x,traj1[[1]]$y))
pts2 <- SpatialPoints(cbind(traj2[[1]]$x,traj2[[1]]$y))
hr1 <- gBuffer(gConvexHull(pts1),width=0.001)
hr2 <- gBuffer(gConvexHull(pts2),width=0.001)

Minta(traj1,traj2,hr1,hr2,method="frequency")
Minta(traj1,traj2,hr1,hr2,method="spatial")
```

Prox *Perform Proximity Analysis*

Description

This function computes the proportion of fixes that are proximal, based on some spatial threshold – dc (Bertrand et al. 1996). It also facilitates local-level proximity analysis.

Usage

```
Prox(traj1, traj2, dc=50, dataframe=FALSE)
```

Arguments

traj1	an object of the class ltraj which contains the time-stamped movement fixes of the first object. Note this object must be a type II ltraj object. For more information on objects of this type see help(ltraj).
traj2	same as traj1, but should contain simultaneous fixes to traj1 – see the function: GetSimultaneous.
dc	critical distance, in appropriate units, for determining whether the two animals are together/proximal.
dataframe	logical value indicating whether or not a dataframe, containing the distance between each simultaneous fix, should be returned.

Details

This function can be used to test for the presence of attraction (via proximity) in wildlife telemetry data. Prox is simply the proportion of simultaneous fixes below the threshold distance – dc. The output dataframe can be useful for examining variation in proximity through time.

Value

If dataframe=FALSE (the default) Prox returns the numeric value of the Prox index. If dataframe=TRUE Prox returns a list, the first argument being the numeric value of the Prox index, the second argument being the dataframe (containing the date/times of simultaneous fixes, and the distance between fixes at each time.

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Bertrand, M.R., DeNicola, A.J., Beissinger, S.R, Swihart, R.K. (1996) Effects of parturition on home ranges and social affiliations of female white-tailed deer. *Journal of Wildlife Management*, **60**: 899-909.

See Also

Ca, HAI, GetSimultaneous

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:1000)
traj2 <- simm.crw(1:1000)
#compute Prox
Prox(traj1,traj2,dc=2)
```

Shirabe *Calculate the Shirabe correlation coefficient for movement data*

Description

This function computes the correlation statistic for movement data as presented in the paper by Shirabe (2006). The statistic is essentially a Pearson product-moment correlation statistic formulated for use with movement data. To be valid this method requires two movement datasets with simultaneous fixes.

Usage

```
Shirabe(traj1, traj2)
```

Arguments

traj1	an object of the class ltraj. For more information on objects of this type see help(ltraj).
traj2	same as traj1.

Details

This function can be used to measure the level of dynamic interaction (termed correlation) between a pair of simultaneously moving objects. The statistic is sensitive to interaction in both movement direction (azimuth) and displacement, but is unable to disentangle the effects of these components.

Value

This function returns the Shirabe (2006) correlation statistic for two moving objects.

Author(s)

Jed A. Long <jlong@uvic.ca>

References

Shirabe, T. 2006. Correlation analysis of discrete motions. In: Raubal, M., Miller, HJ, Frank, AU, and Goodchild, M. eds. GIScience 2006, LNCS 4197. Berlin: Springer-Verlag; 370-382.

See Also

di

Examples

```
## simulate two movement trajectories ##
traj1 <- simm.crw(1:100)
traj2 <- simm.crw(1:100)
#simulate a third trajectory
traj3 <- simm.crw(1:100)
```

```
Shirabe(traj1,traj2)
Shirabe(traj1,traj3)
```