UCGE Reports
Number 20258

Department of Geomatics Engineering

Sensor-based Animal Tracking

(URL: http://www.geomatics.ucalgary.ca/research/publications/GradTheses.html)

by

Andrew Hunter

September 2007

UNIVERSITY OF CALGARY

Sensor-based Animal Tracking

by

Andrew Hunter

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOMATICS ENGINEERING
CALGARY, ALBERTA
SEPTEMBER, 2007

Abstract

The advent of Global Positioning System (GPS) technologies has provided wildlife researchers with new insights into the movement and habitat utilization patterns of wildlife species by being able to provide vast quantities of detailed location data. However, current wildlife tracking techniques have numerous limitations, as GPS locations can be biased to an unknown extent because animals move through habitats that are often denied GPS signals. This can result in some habitat types being under sampled or not sampled at all. Additionally, researchers using GPS tracking systems cannot understand what behaviour an animal is exhibiting at each GPS position without either relying on extensive field data or statistical techniques that may infer behaviour. Overall these issues, and others, limit the knowledge that can be derived from the data currently being collected by GPS collars alone. To address these limitations, a dead reckoning solution (called the NavAid) has been developed to augment GPS tracking collars, which enables both the acquisition of continuous movement trajectories for animals under study, and the collection of digital images on a user-defined schedule along travel routes. Analysis of an animal's velocity allows one to identify different types of movement behaviours that can be associated with foraging, searching for food, and locomotion between patches. In addition, the ability to capture continuous paths allows researchers to identify habitat that is important to a species, and habitat that is not — something that is not possible when relying solely on GPS. This new system weighs approximately 220 g and can be deployed on most conventional collar systems for a wide range of species. This thesis presents the research and development of this new system over the past four years, along with preliminary findings from field work carried out on grizzly bears (*Ursus arctos*) in the foothills of the Canadian Rocky Mountains. Analysis of tracking data suggests that animals select different types of habitat for different purposes, that foraging occurs at movement rates of less than 52m/minute, searching for food between movement rates of 52 m/minute and 223 m/minute and locomotion, or active walking between foraging sites at movement rates greater than 223 m/minute.

Acknowledgements

To Kerry, my wife and partner in life; to my three girls: Lauren, Ellissa and Sarah. Thank you for your constant love and support; for putting up with the grumpy days and the time together that has been missed. Without your encouragement and continual support this journey that we are on would not be possible.

To Dr. Naser El-Sheimy, thank you for your guidance, never-ending support, and trust. It has been immensely appreciated.

To Gordon Stenhouse, the Foothills Model Forest and its research partners. Your willingness to continue providing financial, field and data support has allowed me to pursue this work. Without your support it would not have been possible to move forward, or past some of the disappointments that we have encountered.

To my defence committee, thank you for your comments, suggestions and support. Your contribution has made this work better.

To Elk Valley Coal, NSERC, Focus Corporation, the Department of Geomatics Engineering and the University of Calgary, your financial support has helped to smooth over many obstacles that would have been insurmountable had your assistance not been available.

To Greg McDermid, Jerome Cranston, Julie Duval, Karen Graham, Bernie Goski, Robin Munro, and Patricia Jaroslawski, thank you for your time and effort to provide assistance and feedback throughout this research.

To Rob Scorey, Clint Stern and everybody else from the Engineering Machine shop who was dragged into manufacturing pieces for me; to Bruce Wright, Robert Thomson and Sean Murphy for electronic and software assistance. Thank you very much.

Lynn Raaflaub, Rebeca Quiñonez-Piñón and everybody else who has passed through EN F 124, thank you for the many discussions. It may not have always seemed like it, but getting to see another perspective did help.

To Mum and Dad, thank you for teaching me the value of family, the value of working hard, and the value of taking time to appreciate both. To Austin and Molly-Jo, thank you for always being available when we needed your support.

Table of Contents

List of Tables

List of Figures and Illustrations

List of Symbols and Abbreviations

Symbol	Definition
b_{f}	Accelerometer bias
I_f	Accelerometer measurement
e_{f}	Accelerometer noise
X_i, Y_i, Z_i	Axes of reference frame defined by subscript: G - global; A -animal; S - sensor; M - magentometer, h - local level plane
H	Corrected magnetometer output
df	Degrees of freedom
E, N	Easting and northing coordinates of local mapping frame
SSE	Error sum of squares
e_{φ}, e_{ω}	Errors in roll and pitch
$\hat{s}$	Estimated distance or stride length
$\hat{P}(j \mid \mathbf{x}_u)$	Estimated probability that unit u belongs to population j - a posterior probability
$\hat{n}$	Estimated step count
$\hat{l}$	Estimated stride length
y	Expeceted outcome
f	Foraging bouts
ψ	Heading (bearing)
HR	Hit rate
S_1, S_2	Linear and non-linear accelerometer scale factors
δg	Local gravity anomaly
θ_{Mag}	Magnetic Declination
θ_{Dip}	Magnetic Dip

Symbol	Definition
ψ_m	Magnetic heading
H_{i_0}	Magnetometer axes bias
$A \rightarrow B$	Mapping of coordinate sytem A to coordinate system B
N	Matrix representing non-orthogonality of accelerometer axes
β	Model coefficents
r	Movement rate
k	Number of events to be observed
$P(r \mid \lambda)$	Poisson process — probability of observing r given an average movement rate of λ
$\pi(x)$	Resource estimation function
$w(x)$	Resource selection function
φ, ω, κ	Roll, pitch and yaw
$\theta_Z^{A \rightarrow S}$	Rotation from coordinate frame A to coordinate frame S about the Z axis of S
D_{uj}^{*2}	Sample distance between unit u and group j centroid of the error covariance matrix
S_i	Scale factor for each magnetometer axis
s	Searching bouts
N_r	Total number of movements at rate r
SST	Total sum of squares
f	True specific force
v	Velocity
$\psi_{u,s}$	Wavelet function with scale and translation parameters of s and u

Abbreviations	Definition
ADC	Analog to digital converter
AIC	Akaike's Information Criteron
aLAI	August Leaf Area Index
BL	Barren land
BLF	Broad leaf forest
BMU	Bear Management Unit
CC	Crown closure
CCF	Closed coniferous forest
CI	Confidence interval
CMOS	Complementary Metal-Oxide Semiconductor
D2E	Distance to Edge
D2R	Distance to roads
D2W	Distance to water
DEM	Digital Elevation Model
DOP	Dilution of Precision
DR	Dead reckoning
EPE	Estimated Position Error
EV	Elevation variation
FFT	Fast Fourier Transform
FMFGBRP	Foothills Model Forest Grizzly Bear Research Program
FR	Forest regeneration
G	Global reference frame
G	Greenness — Tassel cap
GIS	Geographic Information System
GISc	Geographic Information Science
GLS	Generalized Least Squares

Abbreviations	Definition
GPS	Global Positioning System
I/O	Input output port
I2C	Inter-Integrated circuit
IAGA	International Association of Geomagnetism and Aeronomy
IC	Integrated Circuit
IDT	Integrated Decision Tree
IGRF	International Geomagnetic Reference Field
IR	Infra-red
LAI	Leaf Area Index
LB	Lower bound of confidence interval
LC	Land cover
Li-SOCI2	Lithium-thionyl chloride
LOO	Leave one out rule used for classification
MAUP	Modifiable area unit problem
MEMS	Micro-electro-mechanical Systems
MF	Mixed forest
MT	Mean temperature
NavAid	Navigation Aid
NLR	Nonlinear Regression
NR	Net solar radiation
OCF	Open coniferous forest
OR	Outgoing solar radiation
OW	Open wetland
PDA	Predictive discriminant analysis
PS	Power spectrum
R	Risk

Abbreviations	Definition
R, P, S	Variables representing Role, pitch and slope of accelerometer
RAM	Random Access Memory
RF	Radio frequency
ROC	Receiver Operating Curve
S	Shrubs
SAI	Slope Aspect Index
SDW	Shadow
SPI	Serial peripheral interface
TW	Treed wetland
UART	Universal Asynchronous Receiver/Transmitter
UB	Upper bound of confidence interval
UERE	User Equivalent Range Error
UH	Upland herbaceous
UPGMA	Unweighted Pair-Group Methods using Arithmetic Averages
USB	Universal serial bus
W	Water
WGS84	World Geodetic System 1984
YV	Variable representing Y accelerometer variance
ZF	Variable representing Z accelerometer frequency
ZR	Variable representing Z accelerometer range
ZV	Variable representing Z accelerometer variance

Chapter 1

Introduction

Modelling of space and time within the domain of geographic information science (GISc) is replete with studies from a broad range of applications that associate time and space, ranging from land use or land ownership studies, transportation, epidemiological research, to environmental modelling[1].

Since Langran and Chrisman (1988), whose work explicitly enabled spatio-temporal data to be incorporated into a geographic information systems[2] (GIS),

---

1 M. F. Worboys, "A Unified Model of Spatial and Temporal Information," *The Computer Journal* 37, no. 1 (1994) C. Claramunt et al., "Database Modelling for Environmental and Land Use Change," in *Geographical Information and Planning: European Perspectives*, ed. S. Geertman, S. Openshaw, and J. Stillwell (New York: Springer-Verlag Ltd., 1999); S. Openshaw, I. Turton, and J. MacGill, "Using the Geographical Analysis Machine - Analyze Limiting Long-term Illness Census Data," *Geographical and Environmental Modelling* 3, no. 1 (1999); A. S. Fotheringham and M. Wegener, *Spatial Models and GIS: New Potential and New Models* (London: Taylor and Francis, 2000); J. L Mennis, D. J. Peuquet, and L. Qian, "A Conceptual Framework for Incorporating Cognitive Principles into Geographical Database Representation," *International Journal of Geographical information Science* 14, no. 6 (2000); P. Laube, "A Classification of Analysis Methods for Dynamic Point Objects in Environmental GIS" (paper presented at the 4th AGILE Conference - GI in Europe: Integrative, Interoperable, Interactive, Brno, Czech Republic, April 19 – 21 2001)

2 G. Langran and N. R. Chrisman, "A Framework for Temporal Geographic Information," *Cartographica* 25, no. 3 (1988)

many spatial models have been extended to include temporal information[3]. These systems have largely viewed the world in an historical context, in that they store facts regarding events that have taken place in the past[4].

From these space-time constructs flows information about spatial processes[5]. Location can be used to further scientific understanding by providing variability in explanatory variables, for example, habitat data combined with species population data can be used for the analysis of a species preference for different types of habitat. To this end, one of the primary objectives of this work is to provide wildlife researchers with a tool, called the NavAid throughout this work, that can provide detailed data about grizzly bear trajectories.

To provide some context for this research at a higher level, we are interested, ultimately, in land use planning with a particular emphasis on wildlife management. In a broad sense, planning deals with elements of the physical environment, both built and natural; the purpose of which is to achieve

3 D. J. Peuquet, "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, no. 3 (1994); D. J. Peuquet and N. Duan, "An Event-based Spatiotemporal Data Model (ESTDM) for Temporal Analysis of Geographical Data," *Journal of Geographical information Science* 9, no. 7 - 24 (1995); Worboys, "A Unified Model of Spatial and Temporal Information," ; Y. Bédard et al., "Adapting Data Models for the Design of Spatio-Temporal Databases," *Computer, Environment and Urban Systems* 20, no. 1 (1996); K. Hornsby and M. J. Egenhofer, "Qualitative Representation of Change" (paper presented at the Conference on Spatial Information Theory, COSIT '97, Laurel Highlands, PA, 1997); M. Erwig et al., "Abstract and Discrete Modeling of Spatio-Temporal Data Types" (paper presented at the 6th ACM Symposium on Advances in Geographic Information Systems, Washington D.C., November, 1998 1998); M. Erwig et al., "Spatio-Temporal Data Types: An Approach - Modeling and Querying Moving Objects in Databases," *GeoInformatica* 3, no. 3 (1999); and R. H. Güting et al., "A Foundation for Representing and Querying Moving Objects," *ACM Transactions on Database Systems* 25, no. 1 (2000) to name a few.

4 O. Wolfson et al., "Updating and Querying Databases that Track Mobile Units," *Distributed and Parallel Databases* 7, no. 3 (1999)

5 R. Haining, *Spatial Data Analysis: Theory and Practice* (Cambridge: Cambridge University Press, 2003) pg. 16

environments that meet the desires of its citizens[6]. Most problems addressed by planning arise from the external effects that one instance of development has upon others. Hence, a common theme of planning is dealing with the problems associated with quantitative and qualitative changes resulting from growth and development, i.e., changing land use. Our interest is in understanding these effects with respect to wildlife, grizzly bears in particular, in the Rocky Mountains of western Canada.

Currently in Alberta, grizzly bears are listed as being "under review as a threatened species" by Alberta's Endangered Species Conservation Committee[7]. At the national level, grizzly bear status is of "special concern"[8] as grizzly bear habitat is at risk from expanding industrial, residential and recreational developments in western Canada. Current research puts the number of grizzly bears in Alberta at less than 500[9], and despite the recognition of population declines and the importance of secure habitats, current management is largely based on a 1988 assessment of land cover and human disturbance[10] rather than animal occurrence and risk.

---

6 G. Hodge, *Planning Canadian Communities: An Introduction to the Principles, Practice, and Participants*, 4 ed. (Scarborough: Thompson Nelson, 2003)

7 Alberta Sustainable Resource Development, "Species Currently Listed under the Wildlife Act and New Species Assessed by the ESCC," (Edmonton: Alberta Government, 2007)

8 COSEWIC Committee on the Status of Endangered Wildlife in Canada, "COSEWIC Species Database: Bear, Grizzly," (Government of Canada, 2002)

9 G.B. Stenhouse, M.S. Boyce, and J. Boulanger, "Report on Alberta Grizzly Bear Assessment of Allocation," (Hinton: Alberta Sustainable Resource Development, Fish and Wildlife Division, Alberta, 2003)

10 Ibid.

Land use planning is part of our culture; it is based on long standing principles about how a community should develop, and how it should protect the public interest during the process of development. Important public activities are seldom adopted overnight. They grow out of experience, tradition, and ideals that date back several centuries[11]. To ensure that these principles are met, in addition to physical constraints and governmental control, planning must consider the array of social purposes that land use must address.

While planning legislation may prescribe a planning process, it is people, individually and in groups, who make it reality. To borrow from Aldo Leopold[12], we as a culture have a long history acting as "conqueror of the land", when perhaps we would be better served if we acted merely as members of the land community[13]. This is not to say that we wish to prevent alteration, management, and use of natural resources, "but it does affirm their right to continued existence."

It is within this framework that this research is set. In order to develop planning policy that can address wildlife management issues, we need to understand the relationship between the life history traits of a species and its habitat use[14]. As discussed by Nielson et al (2006), without understanding such

---

11 Hodge, *Planning Canadian Communities: An Introduction to the Principles, Practice, and Participants*

12 Aldo Leopold, *A Sand Country Almanac and Sketches Here and There* (New York: Oxford University Press, 1949) pg. 204

13 Aldo Leopold's concept of a land ethic extends community to include, in addition to humans, the soil, water, plants and animals that collectively make up the land.

14 Alan B. Franklin et al., "Climate, Habitat Quality, and Fitness in Northern Spotted Owl Populations in Northwestern California," *Ecological Monographs* 70, no. 4 (2000)

functions, one risks assuming that animal occurrence relates directly to habitat quality, something that is not always the case[15]. For instance, some sites considered high in habitat quality from an occupancy standpoint may be low from a survival and/or recruitment standpoint. Such phenomena have been described as ecological traps[16], and, for many species we lack the data required to understand these interactions[17].

Why the concern with grizzly bears (*Ursus arctos*)? Grizzly bears are an important umbrella species[18] that have declined substantially throughout much of North America in the past century[19], largely due to vulnerability from late maturation, low density, and low reproductive rates[20]. In addition, habitat loss is often more challenging for large animals[21], and animals that cover vast expanses of territory generally encounter a broader range of habitats, which exposes the animal to greater risks[22].

---

15 Scott Eric Nielsen, Gordon B. Stenhouse, and Mark S. Boyce, "A Habitat-based Framework for Grizzly Bear Conservation in Alberta," *Biological Conservation* 130, no. 2 (2006)

16 Therese M. Donovan and Frank R. Thompson Iii, "Modeling The Ecological Trap Hypothesis: A Habitat and Demographic Analysis for Migrant Songbirds," *Ecological Applications* 11, no. 3 (2001)

17 Nielsen, Stenhouse, and Boyce, "A Habitat-based Framework for Grizzly Bear Conservation in Alberta,"

18 R. F. Noss, "Indicators for Monitoring Biodiversity: A Hierarchical Approach," *Conservation Biology* 4, no. 4 (1990); P. H. Williams and K. J. Gaston, "Measuring more of Biodiversity: Can Higher-Taxon Richness Predict Wholesale Species Richness?" *Biological Conservation* 67, no. 3 (1994)

19 Bruce N. McLellan et al., "Rates and Causes of Grizzly Bear Mortality in the Interior Mountains of British Columbia, Alberta, Montana, Washington, and Idaho," *The Journal of Wildlife Management* 63, no. 3 (1999)

20 Andy Purvis et al., "Predicting Extinction Risk in Declining Species," *Proceedings of the Royal Society of London - B* 267, no. 1456 (2000)

21 S Herrero, "Introduction," in *Bears: Status Survey and Conservation Action Plan*, ed. C. Servheed, S. Herrero, and B. Peyton (Gland, Switzerland: IUCN, 1999)

22 Marcel Cardillo, "Biological Determinants of Extinction Risk: Why are Smaller Species Less Vulnerable?" *Animal Conservation* 6, no. 01 (2003)

The long-term objective of this work is to gain insights into the behaviour of grizzly bears, as behaviour is the primary means by which an animal copes with environmental challenges[23], whether they are brought on naturally, or are induced as a result of anthropogenic development. An animal can produce intricate movement patterns[24], which, if we can understand them, will have broad implications in the field of wildlife and ecology conservation[25].

Wildlife Modelling

A wide-range of techniques has been used to understand animal movement. A significant body of work has utilized patch based[26], or landscape connectivity[27] approaches derived from island biogeography theory[28]. Recent studies have also shown that patch size and landscape spatial structure influence how an animal population will disperse[29]. While these methods help to identify the elements

---

23 Michael L. Morrison, "A Proposed Research Emphasis to Overcome the Limits of Wildlife-Habitat Relationship Studies," *The Journal of Wildlife Management* 65, no. 4 (2001)

24 I. D. Jonsen, R. A. Myers, and J. Mills Flemming, "Meta-Analysis of Animal Movement using State-Space Models," *Ecology* 84, no. 11 (2003)

25 M. Bélisle and C. C. St. Clair, "Cumulative Effects of Barriers on the Movements of Forest Birds," *Conservation Ecology* 5, no. 2 (2001)

26 See D. J. Bender, L. Tischendorf, and L. Fahrig, "Using Patch Isolation Metrics to Predict Animal Movement in Binary Landscapes," *Landscape Ecology* 18, no. 1 (2003) R. E. Russell, R. K. Swihart, and Z. Feng, "Population Consequences of Movement Decisions in a Patchy Landscape," *Oikos* 103, no. 1 (2003) R. L. Schooley and J. A. Wiens, "Finding Habitat Patches and Directional Connectivity," *Oikos* 102, no. 3 (2003)

27 See L. Tischendorf and L. Fahrig, "How Should we Measure Landscape Connectivity?" *Landscape Ecology* 15, no. 7 (2000a) L. Tischendorf and L. Fahrig, "On the Usage and Measurement of Landscape Connectivity," *Oikos* 90, no. 1 (2000b) or B. J. Goodwin and L. Fahrig, "How does Landscape Structure influence Landscape Connectivity?" *Oikos* 99, no. 3 (2002)

28 R. H. MacArthur and E. O. Wilson, *The Theory of Island Biogeography* (Princeton: Princeton University Press, 1967)

29 J. A. Wiens, R. L. Schooley, and Jr. R. D. Weeks, "Patchy Landscapes and Animal Movements: Do Beetles Percolate?" *Oikos* 78, no. 2 (1997) I. D. Jonsen and P. D. Taylor, "Fine-scale Movement Behaviours of Calopterygid Damselfies are Influenced by Landscape Structure: An Experimental Manipulation," *Oikos* 88, no. 3 (2000)

that need to be present in a landscape, they are unable to define the appropriate quantity and distribution of those elements[30].

Alternative methods have taken a narrower view and have focused on the requirements of a particular species, typically an umbrella species[31] whose requirements for persistence encapsulate those of an array of additional species[32]. Early work in ecological literature focused on point pattern analysis, largely as an exploratory technique, to describe the distribution of the phenomena being studied, but the techniques provide few insights into the processes that drive an animal's behaviour.

Spatial patterns are studied because they can reveal information about the dynamics of a process[33]. The fundamental question to be answered is whether an observed pattern is random or not. Departure from randomness implies that there is some process that is causing the distribution observed[34]. Quadrat analysis[35], distance measures such as Nearest Neighbour[36] and the *K*-function[37], and hierarchical cluster analysis techniques such as the Unweighted Pair-Group

---

30 R. J. Lambeck, "Focal Species: A Multi-species Umbrella for Nature Conservation," *Conservation Biology* 11 (1997)

31 Noss, "Indicators for Monitoring Biodiversity: A Hierarchical Approach," ; Williams and Gaston, "Measuring more of Biodiversity: Can Higher-Taxon Richness Predict Wholesale Species Richness?"

32 Lambeck, "Focal Species: A Multi-species Umbrella for Nature Conservation,"

33 P. Hasse, "Spatial Pattern Analysis in Ecology based on Ripley's K-Function: Introduction and Methods of Edge Correction," *Journal of Vegetation Science* 6, no. 4 (1995)

34 P. Legendre and M. - J. Fortin, "Spatial Pattern and Ecological Analysis," *Vegetatio* 80 (1989)

35 P. Greig-Smith, "The use of Random and Contiguous Quadrats in the Study of the Structure of Plant Communities," *Annals of Botany* 16, no. 62 (1952)

36 P. J. Clark and F. C. Evans, "Distance - Nearest Neighbour as a Measure of Spatial Relations in Populations," *Ecology* 35 (1954)

37 B. D. Ripley, "The Second Order Analysis of Stationary Point Processes," *Journal of Applied Probability* 13 (1976)

Methods using Arithmetic Averages (UPGMA)[38] have all been applied to behavioural studies in an attempt to tease out more information. However, many of these techniques are limited by the scale at which they are studied.

Point pattern and simple distance-based techniques have approached the study of species behaviour from a population driven perspective, i.e., top-down. Alternatively many studies of animal behaviour have taken a bottom-up approach, and focused on the movement of individual animals. Techniques have ranged from random walks[39], to stochastic differential equations[40], to mechanistic models for the delineation of home ranges[41].

Random walk techniques break movement of an animal down into two components: move length and turning angle between successive moves, along with their associated distributions[42]. In a random walk, turning angles are typically drawn from a uniform distribution. By constraining movement between known locations, and concentrating the direction of movement by assuming a

---

38 See H. C. Romesburg, *Cluster Analysis for Researchers* (Belmont: Life Long Learning Publications, 1984) R. Bethke et al., "Population Delineation of Polar Bears Using Satellite Collar Data," *Ecological Applications* 6, no. 1 (1996) and M. K. Taylor et al., "Delineating Canadian and Greeenland Polar Bear (Ursus maritimus) Populations by Cluster Analysis of Movements," *Canadian Journal of Zoology* 79 (2001)

39 P. M. Kareiva and N. Shigesada, "Analyzing Insect Movement as a Correlated Random Walk," Oecologia 56, no. 234 - 238 (1983) E. A. Wentz, A. F. Campbell, and R. Houston, "A Comparison of Two Methods - Create Tracks of Moving Objects: Linear Weighted Distance and Constrained Random Walk," International Journal of Geographical Information Science 17, no. 7 (2003)et al., 2003)

40 H. K. Preisler et al., "Stochastic Differential Equations: A Tool for Studying Animal Movement" (paper presented at the IUFRO 4.11 Conference on Forest Biometry, Modelling and Information Science, Greenwich, UK, June 25-29 2001)

41 P. Turchin, "Translating Foraging Movements in Heterogeneous Environments into Spatial Distribution of Foragers," Ecology 72, no. 4 (1991) M. A. Lewis and J. D. Murray, "Modelling Territoriality and Wolf-Deer Interactions," Nature 366, no. 23 (1993) P. R. Moorcroft, M. A. Lewis, and R. L. Crabtree, "Home Range Analysis Using a Mechanistic Home Range Model," Ecology 80, no. 5 (1999)

42 Kareiva and Shigesada, "Analyzing Insect Movement as a Correlated Random Walk,"

non-uniform distribution of turning angles, a correlated random walk can be generated. Both the random walk and the correlated random walk models assume that there is no autocorrelation in sequential step lengths or turning angles, i.e., the processes are first order markov chains; as such, the length and turning angle of a step are independent from those of the previous step[43]. The advantage of this technique is that behavioural changes, such as foraging or directed walks, which influence movement length and turning angles[44], can be analyzed with respect to expected net displacement. But for animals that are difficult to observe, what distribution should be used when applying such models?

Both stochastic differential equations and mechanistic models draw from the mathematics of random walk analysis. Mechanistic models attempt to incorporate additional behavioural attributes of the animal under investigation (territoriality for example), in which one partial differential equation is used to model the animal's spread through a territory, a second models its tendency to move towards the centre of its home range in response to other animals in its vicinity, and a third model describes the spatial pattern of scent markings left by other animals[45]. The mechanistic models reviewed have been used to describe the home range of the particular animal being studied, whereas stochastic differential

43 P. Turchin, *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants* (Sunderland, Massachutsetts: Sinauer Associates Inc., 1998)

44 J. N. M. Smith, "The Searching Behaviour of Two European Thrushes: I. Description and Analysis of Search Paths," *Behaviour* 48 (1974)

45 Lewis and Murray, "Modelling Territoriality and Wolf-Deer Interactions," Moorcroft, Lewis, and Crabtree, "Home Range Analysis Using a Mechanistic Home Range Model,"

equations have been directed towards the analysis of an animal's movement. Both methods have been reduced to first order models but can be extended if needed. The primary advantage of these techniques is that home range patterns can be described using individual level patterns of movement and behaviour. Few limitations were mentioned in the literature reviewed, although Moorcroft et al (1999) noted that boundary conditions for the models need to be selected with care in order to achieve realistic results[46]. Van Horne (2002) also observed that models that are very specific may not be useful across time and space unless the models are very detailed[47]. They may, however, allow the identification of processes that are important in driving the behaviour under observation. These processes can then be incorporated back into more general models that have greater predictive power.

During the 1990s there was a shift within the realm of spatial ecology towards the use of logistic regression techniques[48] for regional scale analysis. With this method, presence sites are compared against absence sites, sampled from the background landscape, in order to predict the probability of use of a resource. In effect, logistic regression divides space into two portions, habitable and uninhabitable, at a particular probability threshold. However, the technique

46 Moorcroft, Lewis, and Crabtree, "Home Range Analysis Using a Mechanistic Home Range Model,"

47 B. van Horne, "Approaches to Habitat Modelling: The Tension Between Pattern and Process and Between Specificity and Generality," in *Predicting Species Occurrences: Issues of Accuracy and Scale*, ed. J. M. Scott, et al. (Washington, USA: Island Press, 2002)

48 D. E. Stauffer, "Linking Populations and Habitats: Where have we been? Where are we going?" in *Predicting Species Occurrences: Issues of Accuracy and Scale*, ed. J. M. Scott, et al. (Washington, USA: Island Press, 2002)

is not well suited to categorical data often used in logistic models[49], such as vegetation, and models developed at one scale or location, and then applied at another, may give misleading results[50]. In addition, Keating and Cherry (2004) have recently raised concerns about the misapplication of logistic regression models in habitat studies. They suggested that most use-availability studies violate the assumption that presence and absence data are drawn randomly from the available habitat, and that in such a situation logistic regression can result in unreliable species-environment models[51].

Grizzly Bear Research

The idea that there is some "causal" relationship between the location of an animal and the habitat that surrounds each location has been investigated for some time[52]. The foraging behaviour of grizzly bear in the Yellowhead ecosystem of Alberta has been studied extensively over the last eight years by the Foothills Model Forest Grizzly Bear Research Program (FMFGBRP). The results of Resource Selection Function modelling show that there is some variation among individual grizzly bears with respect to significance of parameter estimates and

49 R. Dettmers, D. A. Buehler, and J. B. Bartlett, "A Test and Comparison of Wild-life Habitat Modeling Techniques for Predicting Bird Occurrence at a Regional Scale," in *Predicting Species Occurrences: Issues of Accuracy and Scale*, ed. J. M. Scott, et al. (Washington, USA: Island Press, 2002)

50 P. J. Heglund, "Foundations of Species-Environment Relations," in *Predicting Species Occurrences: Issues of Accuracy and Scale*, ed. J. M. Scott, et al. (Washington, USA: Island Press, 2002)

51 K. A. Keating and S. Cherry, "Use and Interpretation of Logistic Regression in Habitat-Selection Studies," *Journal of Wildlife Management* 64, no. 4 (2004)

52 See Stauffer, "Linking Populations and Habitats: Where have we been? Where are we going?" for a summary.

goodness-of-fit tests[53]. At the population level, during spring, grizzly bears prefer high greenness areas (Tasselled-cap transformation), medium sized streams, and alpine habitats, while non-vegetated areas and young regenerating burns are avoided. During the summer, alpine regions, recent burn stands, cut blocks, open forest, herbaceous areas and shrub-wetlands are preferred, while young regenerating forest burns are avoided[54]. The underlying cause of the differing spring and summer models is believed to be resource switching between roots and grasses such as sweet vetch roots of *hedysarum spp*, horsetail (*Equisetum spp.*) and monocots (grasses and sedges) during the spring, to fruits such as soopolallie (*Shepherdia canadensis*) and mountain huckleberry (*Vaccinium membranaceum*) during the summer, and then back to roots and grasses in the fall when the berries are depleted[55]. Grizzly bears are also opportunistic in that they will feed on carrion should the opportunity exist. Most often this occurs during late spring[56]. Hamer (1985) has also highlighted a number of other factors such as reproductive activity, familiarizing offspring with available feeding sites, social status, and exploration for new sources of food as possible influences on grizzly bear movement[57].

53 Scott E. Nielsen et al., "Modeling Grizzly Bear Habitats in the Yellowhead Ecosystem of Alberta: Taking Autocorrelation Seriously," *Ursus* 13 (2002)

54 Ibid.

55 R. H. M. Munro et al., "Seasonal and Diel Patterns of Grizzly Bear Diet and Activity in West-Central Alberta," *Journal of Mammalogy* 87, no. 6 (2006), J. A. Nagy et al., *Population Characteristics of Grizzly and Black Bears in West-Central Alberta*, ed. Alberta Environment Centre, vol. AECV88-R1 (Vegrevill, Alberta, Canada: Government of Alberta, 1989)

56 Munro et al., "Seasonal and Diel Patterns of Grizzly Bear Diet and Activity in West-Central Alberta,"

57 J. D. W. Hamer, "Feeding Ecology of Grizzly Bear in the Cascade and Panther Valleys of Banff National Park, Alberta" (University of Calgary, 1985)

Smith (1974) has identified three general feeding phases: searching, feeding, and ingesting[58] and searching can be broken into three additional components: locomotion, scanning and specialized searching[59] (e.g., turning over logs). Zollner and Lima (1999) and MacIntyre and Wiens (1999) have found that the best non-systematic way for an animal to encounter a new patch is a nearly straight-line search path, and that optimal behaviour should be a combination of slow, sinuous searches within high-density resources, and faster walks between high-density resources[60].

A review of GPS data undertaken during the early stages of this research suggest similar movement patterns can be observed in the telemetry data of grizzly bear G20, a seven-year-old female. G20's home range fell within Alberta's Bear Management Units (BMU) Gregg, McLeod, Maskuta and McPherson in the north-western corner of the FMFGBRP study area (see Figure 1). G020 was fitted with a Televilt[61] GPS-Simplex radio-collar that was programmed to obtain positions at four-hour intervals. Data was analyzed from den emergence (April 19, 2002) until July 31, 2002, by comparing the change in position overtime.

---

58 Smith, "The Searching Behaviour of Two European Thrushes: I. Description and Analysis of Search Paths," B. A. Nolet and W. M. Mooj, "Search Paths of Swans Foraging on Spatially Autocorrelated Tubers," *Journal of Animal Ecology* 71 (2002)

59 Smith, "The Searching Behaviour of Two European Thrushes: I. Description and Analysis of Search Paths,"

60 P. A. Zollner and S. L. Lima, "Search Strategies for Landscape-level Interpatch Movements," *Ecology* 80, no. 3 (1999); N. E. MacIntyre and J. A. Wiens, "Interactions between Landscape Structure and Animal Behaviour: the Roles of Heterogeneously Distributed Resources and Food Deprivation on Movement Patterns," *Landscape Ecology* 14 (1999)

61 Televilt TVP Positioning AB, *GPS-Simplex™, Televilt TVP Positioning AB* (Lindesburg: Televilt TVP Positioning AB, 2001)

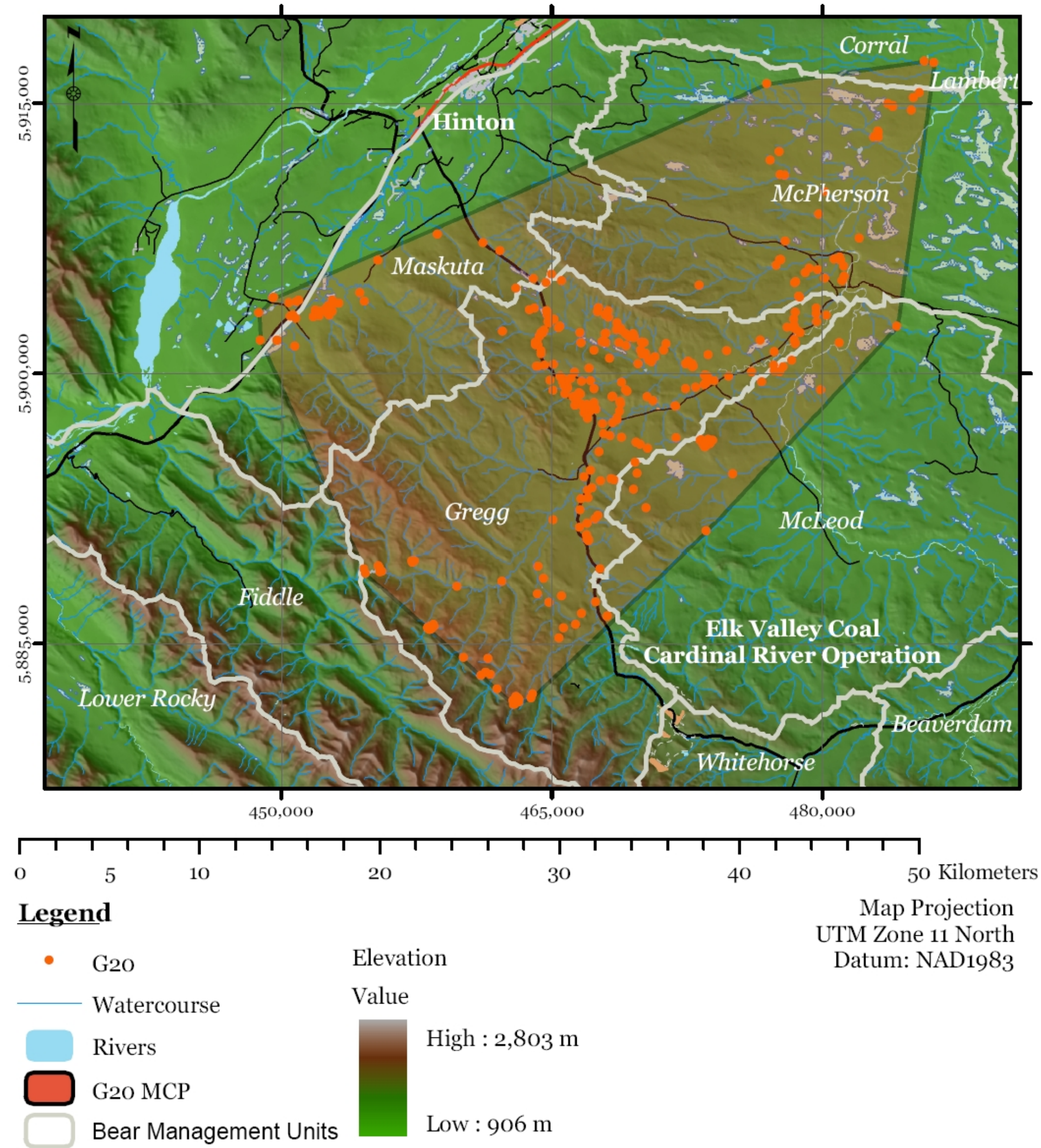


Figure 1: G020 tracking data from 2002

Following Laube (2001) and Laube et al (2005), the analysis transformed the three dimensional data (northing, easting, time) into a two dimensional representation by reducing the spatial component to an inter-event Euclidean distance matrix where the *x* and *y* axes of the matrix represents linear time[62]. By this process the transformation resulted in a matrix of spatial distances for every time-point to all other time-points. Low values in the matrix represented locations that were close together in space at the particular time interval. High values represent large differences in position over the temporal interval. Both visually and through the use of cross correlation analysis it was possible to identify clustering at multiple temporal scales (square patterns in Figure 2(a)), multiple use sites (repetition of square in triangular pattern in Figure 2(a)), use of a similar path in opposite directions (cross pattern in Figure 2(a)), and periods relating to nearly straight line path movements between clusters (larger squares of similar colour in Figure 2(b)). In addition there were often periods just prior

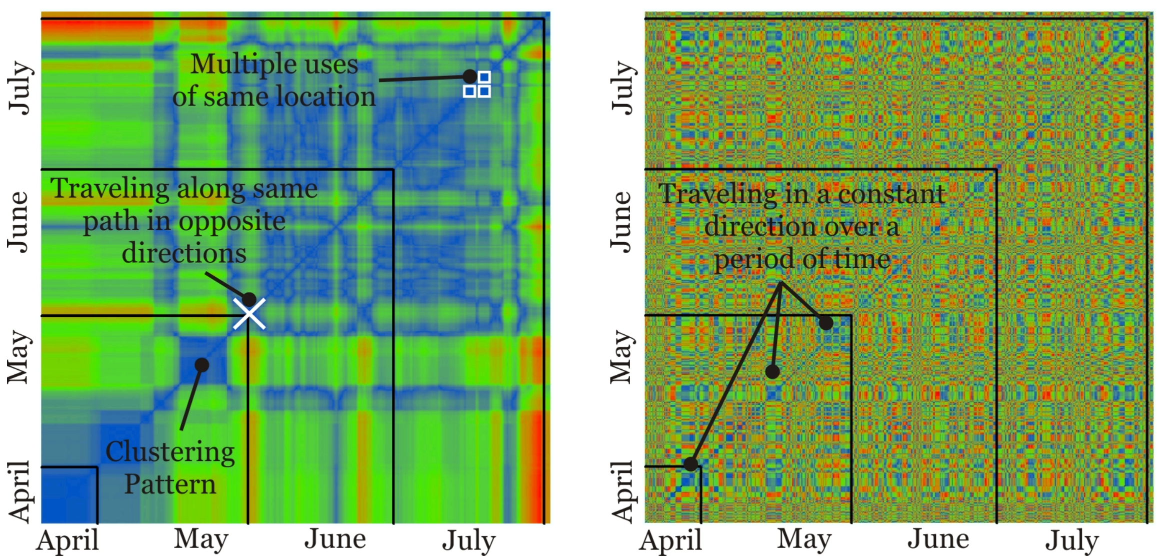


Figure 2: Transformed movement data for G020. (a) Time - Distance - Distance Plot (b) Time – Heading plot

---

62 Laube, "A Classification of Analysis Methods for Dynamic Point Objects in Environmental GIS" and P. Laube, S. Imfeld, and R. Weibel, "Discovering Relative Motion Patterns in Groups of Moving Point Objects," *International Journal of Geographical Information Science* 19, no. 6 (2005)

to arriving at a new patch, or just prior to leaving a patch, where movement was more random than directed (Figure 3). But as with many of the techniques discussed, interpretation was hampered by the scale, in this case the temporal interval, at which the data was acquired. Many sampling designs have been promoted for the analysis of wildlife processes. Designs such as simple random sampling are suitable if the sample size is sufficiently large to ensure that all classes of habitat use are adequately represented[63]. Often, however, it is impractical to follow such sampling procedures[64]. For example, given site

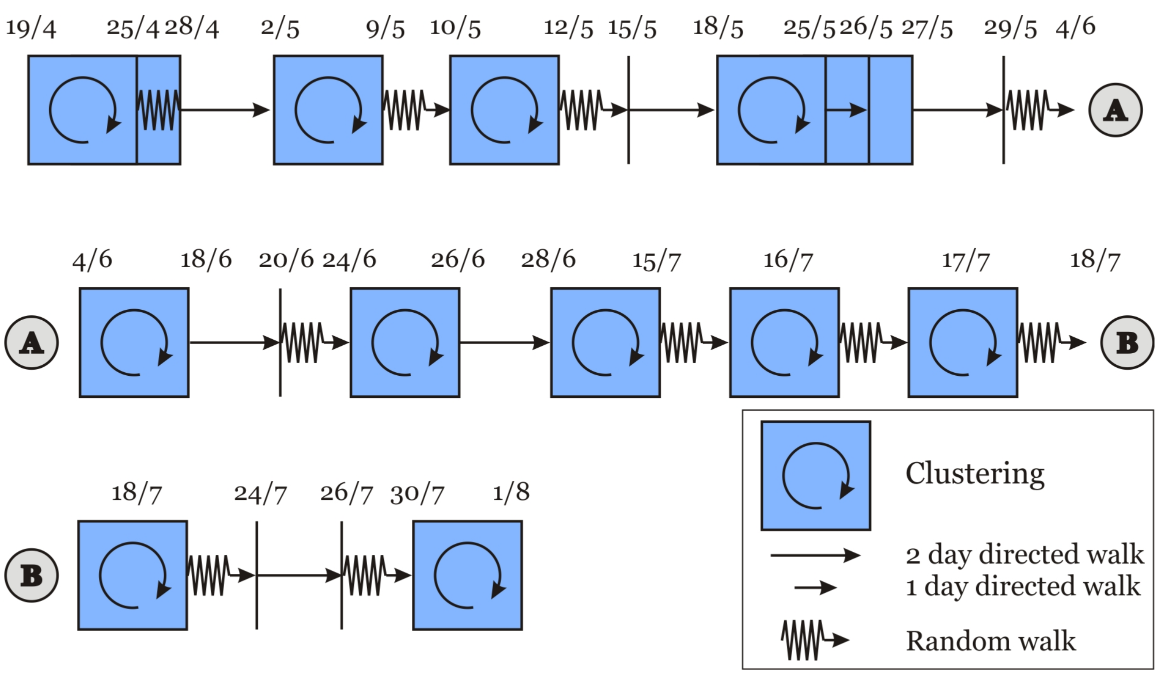


Figure 3: Summary of G020 movements

63 G. M. Foody, "Status of Land Cover Classification Accuracy Assessment," *Remote Sensing of Environment* 80 (2002)

64 T. C. Edwards, G. G. Moisen, and D. R. Cutler, "Assessing Map Accuracy in a Remotely Sensed Ecoregion-scale Cover Map," *Remote Sensing of Environment* 63 (1998)

conditions (particularly in mountainous areas), it may be difficult to use randomly located sites, which results in ground data collection being restricted to locations that provide easiest access.

Researchers influenced by financial and/or practical constraints often required alternative sample designs. Methodologies range from 'windshield' surveys to techniques based on double sampling[65] and cluster sampling[66]. While there is an obvious desire to balance statistical requirements with practicalities[67], the choice of sampling design influences the accuracy of an analysis technique employed[68], particularly if the sampling scheme fails to sample over the full range of the process under investigation.

Animal Tracking

As has been noted, tracking of grizzly bears has been undertaken with the aid of Global Positioning System (GPS) telemetry. It has been widely reported[69] that

---

65 M. A. Kalkhan, R. M. Reich, and T. J. Stohlgren, "Assessing the Accuracy of Landsat Thematic Mapper Classification using Double Sampling," *International Journal of Remote Sensing* 19, no. 2049 - 2060 (1998)

66 S. V. Stehman, "Basic Probability Sampling Designs for Thematic Map Accuracy Assessment," *International Journal of Remote Sensing* 20 (1999)

67 Edwards, Moisen, and Cutler, "Assessing Map Accuracy in a Remotely Sensed Ecoregion-scale Cover Map,"

68 Stehman, "Basic Probability Sampling Designs for Thematic Map Accuracy Assessment,"

69 See W. J. Rettie and P. D. McLoughlin, "Overcoming Radiotelemetry Bias in Habitat Selection Studies," *Canadian Journal of Zoology* 77, no. 8 (1999); C. Dussault et al., "Evaluation of GPS Telemetry Collar Performance for Habitat Studies in the Boreal Forest," *Wildlife Society Bulletin* 27, no. 4 (1999); I. A. R. Hulbert, "GPS and its Use in Animal Telemetry: The Next Five Years" (paper presented at the International Conference on Tracking Animals with GPS, Aberdeen, Scotland, 12 - 13 March 2001); R. J. Gau et al., "Uncontrolled Field Performance of Televilt GPS-Simplex (TM) Collars on Grizzly Bears in Western and Northern Canada," *Wildlife Society Bulletin* 32, no. 3 (2004a); Robert G. D'Eon et al., "GPS Radiotelemetry Error and Bias in Mountainous Terrain," *Wildlife Society Bulletin* 30, no. 2 (2002); Christopher L. Jerde and Darcy R. Visscher, "GPS Measurement Error Influences on Movement Model Parameterization," *Ecological Applications* 15, no. 3 (2005); L. D. Mech and S. M. Barber, "Critique of Wildlife Radio-Tracking and its Use in National Parks: A Report to the U.S. National Park Service," ed. Northern Prairie Wildlife Research Center U.S. Geological Survey (Jamestown, N.D.: Northern Prairie Wildlife Research Center Online, 2002)

while GPS can provide more accurate, and more frequent, animal locations under all weather conditions it remains prone to non-random errors that are prevalent in other radio tracking techniques. These include telemetry bias, habitat bias, and biases attributed to particular animals. Telemetry bias may result from the animal going undetected in some habitat types. Telemetry error may also be greater in some habitats that may result in errors that exceed patch sizes of the specific habitat. This becomes a particular problem when working in areas of high latitude, especially when the animal moves through north facing slopes[70]. In these instances the number of satellites visible to the animal may be substantially reduced, or non-existent, and/or the geometry of visible satellites may be poor, thereby reducing the quality of the telemetry data. With GPS data, the data points are serially correlated as GPS are programmed to acquire positions at specific times, whereas with standard radio tracking they often are not. Standard radio tracking positions are generally acquired when an animal exhibits a behaviour that is of interest to the researcher, whereas GPS data is not. The result is that the use of GPS data has an adverse effect on the quality and statistical legitimacy of models that are developed from the data because the model implicitly assumes that the species under observation is reacting to the local environment surrounding each GPS position equally.

70 With the deployment of new satellite positioning systems such as Russia's GLASNOSS system, and the European Union's Galileo system, the effect of poor satellite geometry should be reduced. This should result in high position fix rates and improved accuracy simply because there are more satellites visible to the positioning receiver.

When an animal is sleeping, location fixes may not be possible because no satellites are visible to the GPS receiver, but researchers have no way of knowing that this is the reason for the loss of data. Equally, we do not know what the animal is doing at the time that a location is obtained; we just know that it was at a certain location at a certain time.

Given the techniques used to study grizzly bear behaviour and resource selection, the inability to relate grizzly bear activity to habitat use severely affects the reliability of parameter estimates derived from GPS data. Data quality and accuracy become important as many of these models are spatially explicit, including individual-based movement models, that describe animal movements for long periods of time will be hampered by GPS-induced bias and error, which can only weaken the conclusions that are drawn from them[71].

Spatial Modelling

It is well known that spatial modelling techniques are confounded by the effect of spatial dependence resulting in a loss of efficiency and increased model variance, which in turn results in less reliable parameter estimates[72]. Finding the degree of spatial association, or spatial autocorrelation, allows the researcher to determine the level of spatial dependence in their data, and therefore avoid the many pitfalls that arise from auto-correlated data. Statistically, spatial dependence suggests

71 Wolf M. Mooij and Donald L. DeAngelis, "Error Propagation in Spatially Explicit Population Models: a Reassessment," *Conservation Biology* 13, no. 4 (1999)

72 S. Bertazzon, "Spatial and Temporal Autocorrelation in Innovation Diffusion Analysis" (paper presented at the International Conference on Computational Science and its Applications, Montreal, Canada, May 18 - May 21 2003)

that many statistical tools and inferences are inappropriate[73]. For example, the use of Ordinary Least Squares regression (OLS) to predict a consequence (e.g., preferred habitat for grizzly bears) assumes that the observations have been selected randomly from a normal distribution. However, if the observations are spatially clustered in some way, estimates obtained from the OLS estimator will be biased and overly precise. They will be biased because the areas with higher concentration of events will have a greater impact on model estimates and they will overestimate precision because, since events tend to be concentrated, there are actually fewer numbers of independent observations than are being assumed[74]. The common approach to resolving spatial dependence is the use of Generalized Least Squares (GLS) that incorporates a model of spatial dependency via a weighting matrix[75]. Just how spatial dependence is defined within the weight matrix remains an issue of debate, and for animal tracking applications, it is a debate that is hindered further because of the difficulty to discriminate between different animal behaviours when basing a study on GPS positions alone.

General Objectives

If we were to organize wildlife management planning within a pyramid, given increasing public demand for access to ecosystem resources, both recreational

73 A. Getis, "Spatial Statistics," in *Geographical Information Systems: Principles, Techniques, Management and Applications*, ed. P. A. Longley, et al. (New York: John Wiley & Sons, Ltd., 1999)

74 D. A. Griffith and L. J. Layne, *A Casebook for Spatial Statistical Data Analysis: A Compilation of Analyses of Different Thematic Data Sets* (New York: Oxford University Press, 1999)

75 L. Anselin, *Spatial Econonmetrics: Methods and Models* (New York: Kluwer Academic, 1988)

and economic, one could place land use planning at the top. Effective planning policy is implemented by government, but it is put in place by the community that the government serves. In order for either group to arrive at a point where consensus allows the implementation of plans with well-defined goals, they require sound, science-based support[76], such that they may justify implementation of a particular plan. Hence, under land use planning sits government and the public, under them, strong science to provide support for strong plans. Good science is based on good data[77]. Environmental scientists have worked to improve theoretical concepts, and to develop tools that provide the best information possible, but many challenges continue to exist[78], many of which can be traced to the data acquisition techniques that are currently available for wildlife applications.

As pointed out by Backus (2006), environmental scientists are becoming increasingly aware that our environment is not spatially uniform, and that animals do not act/react in the same manner to different stimuli[79]. But many of the models that are currently used assume that space is homogeneous and that different animals of the same species react over time with invariance to this environment. This, in no way is meant to denigrate scientific advances that have

---

76 Vickie Backus, "Comprehensive Conservation Modeling: A Spatially Explicit Individual-Based Approach using Grizzly Bears as a Case Study" (University of Utah, 2006)

77 Haining, *Spatial Data Analysis: Theory and Practice*

78 Michael A. Huston, "Critical Issues for Improving Predictions," in *Predicting Species Occurrences: Issues of Accuracy and Scale*, ed. J. M. Scott, et al. (Washington, USA: Island Press, 2002)

79 Backus, "Comprehensive Conservation Modeling: A Spatially Explicit Individual-Based Approach using Grizzly Bears as a Case Study"

been, and continue to be made within the wildlife sciences, as in reality there is often no other practical choice.

But given the advances in technology, new and better data acquisition techniques are inevitable. These advances should be made available to wildlife researchers so that they are able to better address these issues. So for the wildlife management planning pyramid, this research forms part of its foundation.

As noted by Stauffer (2002), despite having powerful analytical tools available, there are limits to the precision of models developed in the wildlife sciences because of the noise inherent in the processes under investigation[80]. Huston (2002) adds that if data is not matched to the spatial and temporal dimension of the processes being measured then strong models are unlikely to be realized[81]. In effect, these are measurement issues, and it is the opinion of the author that the use of current wildlife tracking technology limits researchers' ability to reap the benefits of current advances in measurement technology. Until researchers are able to make use of improved positioning technology these issues are likely to continue. As such, the following problem statement can be derived from this argument and provides the main focus of the research presented by this thesis:

Due to an inability to classify animal location according to animal activity, or behaviour, current modelling endeavours suffer from

80 Stauffer, "Linking Populations and Habitats: Where have we been? Where are we going?"
81 Huston, "Critical Issues for Improving Predictions,"

assumptions relating animal occurrence and habitat quality which may not be correct, and that in order to address this issue we must develop tools that can provide more comprehensive spatial and temporal information about animal behaviour.

Based on this problem statement a secondary problem statement can be formulated as follows:

Given the limitations of current data acquisition techniques for tracking wildlife, understanding of animal behaviour lags behind what could be possible if full advantage is taken of recent advances in positioning technology.

The following sections proceed from these statements by deriving the research goals, objectives and scope that are addressed in this thesis, followed lastly by an outline of the thesis' structure.

Research Goals

Based on the problem statements above, this thesis predominantly focuses on the development of a tool and techniques that can aid our understanding of wildlife behaviour. The primary and most important research goal can be stated as follows:

This thesis primarily aims to develop a technology solution for the identification of animal trajectories and to use this tool to provide researchers with the continuous routes of an animal.

Furthermore, this thesis will attempt to identify different categories of animal behaviour. Hence, a second research goal can be stated as follows:

> *This thesis aims to identify animal behaviours in order that realistic management plans may be formulated based on rules that consider relationships between habitat use and the purpose of the use.*

Following from these research goals the subsequent research question can be derived with respect to the primary research goal:

> *Can technology available for pedestrian navigation be applied in a wildlife environment in order to implement a dead reckoning navigation solution for animals using step detection?*

Following the second research goal, a second research question can be derived as follows:

> *Assuming animal routes can be estimated, can animal behaviour be identified from the data collected by the NavAid?*

A third research question can also be derived from research goal two as follows:

> *Assuming the identification of behavioural information is possible; do grizzly bears exhibit different selection policies depending upon their current behaviour?*

Research Scope

Clearly, this chapter has covered a potentially enormous research field. In order to make the research effort focused and feasible, the scope of this work has been reduced to three dimensions: 1) development of the NavAid that includes the necessary technology and software for the estimation of grizzly bear trajectories; 2) identification and analysis of rules for characterization of their behaviour; and, 3), analysis of habitat selection given estimates of grizzly bear behaviour.

Thesis Content

In chapter two an extensive literature review is presented on the subjects of animal locomotor activity and sensor-based dead reckoning technologies and algorithms. The objective of the chapter is to provide a set of guidelines within which animal tracking technology must fit, starting with a review of animal limb action, as this is the mechanism used to identify animal behaviour. This is followed by the basic mathematical rules of dead reckoning and signal processing necessary to identify movement using inertial and magnetic sensing technology.

Chapter three concentrates on the physical implementation of the NavAid developed for animal tracking. In particular, the chapter focuses on calibration of the NavAid to enable error sources inherent in the technology to be accounted for during the processing phase following data collection.

In chapter four an analysis using data currently available to grizzly bear researchers is undertaken to test the hypothesis that habitat selection is, at least in part, dependent upon the locomotion behaviour of the animal. A two-process movement model is applied to GPS data for G098, a ten-year-old male grizzly bear. Logistic regression is then used to develop habitat use models that identify habitat characteristics most strongly correlated with each locomotion behaviour.

The hypothesis is that if animals select habitat for use based on their current locomotion behaviour, then different behaviours will be correlated with different environmental characteristics.

In chapter five, predictive discriminant analysis is used to develop a set of rules for the characterisation of sensor data acquired by the NavAid. Once characterized, trajectories of G040, a female grizzly bear tracked during the spring of 2006 are computed. The estimated velocity of G040 as she moves throughout her home range is then applied to a three-process movement model to determine threshold that can be used to identify different locomotion behaviours.

The body of the thesis concludes with chapter six. A review of the significant contributions, potential uses and future research direction are given.

Chapter 2

Locomotion, Signal Processing and Dead Reckoning Models

Introduction

In this chapter we will describe how animal locomotor activity and Micro-electro-mechanical Systems (MEMS) sensors coupled with GPS can be combined to derive the trajectory of an animal by dead reckoning — the estimation of an animals current location based upon knowledge of it position at an earlier time, and the distance and direction that it has travelled since that time. The chapter commences with a review of animal gaits and forelimb motion followed by a description of dead reckoning mechanization using step detection. The various reference frames that are encountered during the mechanization process are described followed by a number of techniques that may be used for the identification of steps, stride length and heading. The chapter concludes with a brief review of error sources and their effect on dead reckoning. Animal locomotion draws from animal biomechanics literature, whereas the step detection review draws from the personal navigation literature. While the personal navigation literature is not directed towards animals, because the pattern of forces that are generated by an animal's forelimb is similar to that of a person, it is suggested that the techniques reviewed are sufficiently general that they can be applied to animal locomotion as well. Particularly, when you consider that humans and grizzly bears have a very similar foot action – we are plantigrade – we both walk on the soles of our feet, with a heel-to-toe action.

Animal Behaviour

The ability to move, at some stage in the life cycle, is fundamental to success in life[1]. Hence, the simplest definition of behaviour is movement[2], whether it is the movement of legs when walking, the head when eating, or the throat when threatened (vocalization). Consequently, animal behaviour consists of a series of patterns that can be recognized if they are performed often enough, and in similar enough form.

We can assume that animals move in order to find better environments than the one that they are in. In order for animals to determine when to move they need mechanisms to tell them that their current environment is not suitable and equally they require mechanisms that tell them where to locate a superior environment[3]. There is anecdotal evidence that some animals sense

---

1 Andrew Sugden and Elizabeth Pennisi, "When to Go, Where to Stop," *Science* 313, no. 5788 (2006)

2 Mark Ridley, *Animal Behaviour: A Concise Introduction* (Oxford: Blackwell Scientific Publications, 1986)

3 See John R. Krebs et al., "Optimal Prey Selection in the Great Tit (*Parus major*)," *Animal Behaviour* 25, no. 1 (1977) for an optimal diet model.

environmental change[4] that may, or may not, precede environmental improvement. Animals sense changing seasons through changes in the length of daylight hours[5]. Animals will emigrate when overcrowding occurs to regulate their population[6]. Some animals such as bee-killing digger wasps have been shown to develop mental maps of landmarks[7] from which they navigate. Salmon return to their birthplace to spawn. Pigeons can find their way home, even when taken to a release site under conditions that preclude learning. It is presumed that they use the sun and the earth's magnetic field (they can navigate on cloudy days)[8]. Hence, observation of animals can tell us many things about them.

Because a captive animal is usually too constrained by its artificial environment to perform even a small fraction of the activities of which it is capable, the study of free moving animals has been shown to be more influential[9] in terms of identifying behaviours. Furthermore, according to Martin and Bateson (1986) factors that have been shown to influence behaviour in an experimental environment may not be factors that influence the behaviour of free living individuals[10].

---

4 Horses will become agitated and sheep will locate shelter as a storm approaches. Dogs appear to be able to sense the onset of earthquakes.

5 The sheep breeding season is determined by the shortening number of daylight hours and reducing temperature. See Ian R. Gordon, "The Ewe's Oestrous Cycle and Seasonal Breeding Activity," in *Controlled Breeding in Farm Animals* (Oxford, UK: Pergamon, 1983)

6 MacArthur and Wilson, *The Theory of Island Biogeography*

7 James L. Gould, "The Locale Map of Honey Bees: Do Insects Have Cognitive Maps?" *Science* 232, no. 3 (1986)

8 S. T. Emlen, in *Avian Biology*, ed. D. S. Farner and J. R. King (New York: Academic Press, 1975)

9 Paul Martin and Patrick Bateson, *Measuring Behaviour: An Introductory Guide* (Cambridge: Cambridge University Press, 1986)

10 Ibid.

From an operational perspective, the knowledge gained from understanding behaviour patterns of a species is valuable for scientifically based conservation and wild-life management, because it enables researchers to identify the major variables that control behaviour patterns[11].

However, field studies can be problematic. Historical techniques have often resulted in animals under observation going undetected for long periods of time, which can wreak havoc on even the best-laid plans for systematic recording over a period of time[12]. New techniques making use of GPS technology can provide an abundance of data, but have also been found to be fraught with new limitations[13] and performance issues in the field[14].

For example, systematic location of an animal is often corrupted through the loss of GPS signals under certain habitats; the degradation of accuracy due to poor satellite geometry; or the failure of the GPS receiver to see sufficient satellites in order to obtain a location. In addition, while more animal locations are typically obtained during a field season when using GPS, because GPS

---

11 Gordon Stenhouse, March 28, 2005, personal communication

12 Martin and Bateson, *Measuring Behaviour: An Introductory Guide*

13 Ron Moen et al., "Effects of Moose Movement and Habitat Use on GPS Collar Performance," *The Journal of Wildlife Management* 60, no. 3 (1996); Dussault et al., "Evaluation of GPS Telemetry Collar Performance for Habitat Studies in the Boreal Forest," ; D'Eon et al., "GPS Radiotelemetry Error and Bias in Mountainous Terrain," ; Mech and Barber, "Critique of Wildlife Radio-Tracking and its Use in National Parks: A Report to the U.S. National Park Service," ; Rettie and McLoughlin, "Overcoming Radiotelemetry Bias in Habitat Selection Studies,"

14 Aaron P. Di Orio, Richard Callas, and Robert J. Schaefer, "Performance of Two GPS Telemetry Collars under Different Habitat Conditions," *Wildlife Society Bulletin* 31, no. 2 (2003); Robert J. Gau et al., "Uncontrolled Field Performance of Televilt GPS-Simplex™ Collars on Grizzly Bears in Western and Northern Canada," *Wildlife Society Bulletin* 32, no. 3 (2004b)

positions are not continuous[15], and because the animals are not physically observed, it is not possible to associate the location of an animal with a particular behaviour. This requires researchers to make assumptions that may not be correct.

As indicated by this work's main problem statement, we are interested in an animal's use of its environment. What habitat does the animal require? Are there places that an animal avoids? What does an animal choose in terms of security? But how can these types of questions be addressed when current field techniques can only provide a small portion of the data that is necessary? As outlined in Chapter 1, the solution promoted by this work is the development of hardware that can acquire the continuous trajectory of an animal. In order to understand how the technology is to be used, some background as to how animals move is required.

Animal Gaits

Given the initial research question identified in chapter 1, "can pedestrian navigation systems be applied in a wildlife environment, whereby step detection is used to estimate an animal's trajectory", it is necessary to first discuss animal gaits. As noted above, one of the defining characteristics of animals is their movement[16]. Active foraging for food sources, movement to avoid a stressful

---

15 To conserve GPS power, animal tracking collars attempt to first obtain a 2D fix within a certain time frame, 180 seconds for example. If a 2D fix is acquired within the allotted time, the units than attempt to obtain a 3D fix within a certain time of acquiring the 2D fix, typically 20 seconds. If no fix was obtained, the system shuts down until the next scheduled attempt.

16 Andrew A. Biewener, *Animal Locomotion* (Oxford: Oxford University Press, 2003) pg. 1

environment, active pursuit of prey or avoidance of predators, and finding a mate are all behaviours that animals engage in by means of locomotion.

A study of animal locomotion depends on understanding the physical principles that influence how animals move. In terms of physical properties that a terrestrial animal must overcome, mass-related gravitational forces are the most problematic[17].

As with human locomotion, animals utilize their muscles to generate forces that are transmitted to the ground by means of a skeleton. When an animal moves the forces exerted by the limbs on the ground rise and fall during limb support, and are zero whenever there are no limbs on the ground. As a result the maximum force exerted on the ground by a single limb is always higher than those sustained when an animal is standing at rest. If the limbs are kept on the ground for a longer period of time, smaller forces are required, but this limits speed of movement. To move faster, animals must move their limbs more rapidly, reducing the time that a limb is in contact with the ground and thereby increasing the magnitude of force that must be generated against the ground[18]. Peak ground forces acting on an individual limb may be less than body weight when an animal moves slowly, but as an animal moves more quickly they can become much greater than the animal's body weight[19].

---

17 Ibid. pg. 3

18 P. P. Gambaryan, *How Mammals Run: Anatomical Adaptations*, trans. Hilary Hardin (New York: John Wiley & Sons, Ltd., 1974)

19 Biewener, *Animal Locomotion* pg.47 - 49

Locomotion of quadrupeds has a long history with the first formal studies attributed to Goiffon and Vincent (1779) in which they studied the pounding of horses' hooves by attaching a bell with a specific tone to each hoof[20]. Marei[21] later developed a pneumatic automatic recorder that enabled him to determine the sequence of a horse's footfalls and estimate the duration of support for each limb. From this work he characterised the main forms of locomotion of animals[22].

Following Biewener (2003) and Gambaryan (1974), locomotion gaits are defined by the relative timing of support among the limbs of an animal during a stride[23]. Changes in gait are associated with movement at different speeds and typically involve a discontinuous change in limb movement and/or mechanics of the support. While an animal's gait might be better described as a continuum that includes a number of classes of movement, three general classes of gait commonly referred to are walking, trotting, and galloping.

Walking gaits usually involve overlapping periods of support among the limbs. Limb duty cycle[24] is typically greater than 0.5. For quadrupeds this means that walking incorporates periods during which three limbs are in contact with the ground, providing a stable base of support. During this gait an animal

---

20 Source: Gambaryan, *How Mammals Run: Anatomical Adaptations* , pg. 14

21 E. J. Marei, "Mechanics of the Animal Organisms *(Russian)*," *St.Petersburg State University* (1875)

22 Source: Gambaryan, *How Mammals Run: Anatomical Adaptations* , pg. 15

23 Biewener, *Animal Locomotion* and Gambaryan, *How Mammals Run: Anatomical Adaptations*

24 Limb duty cycle is the ratio of time that a limb is in contact with the ground and the time that it takes to complete a stride, see Biewener, *Animal Locomotion*

maintains its balance because its centre of gravity falls within the triangular area of support represented by the limbs. When walking, the animal is considered to be in a state of static equilibrium. However, as an animal's speed increases, it must rely on dynamic balancing mechanisms[25], as generally the animal will only have two feet on the ground at one time.

Animals can change speed within a gait, but to move over a greater range of speeds they must change gait. When an animal changes gait from a walk to a trot, its stride period decreases, as will limb duty cycle.

Trotting gaits are typically characterized by duty cycles of less than 0.5, hence there are rarely overlapping support periods between alternating support limbs. A trot gait consists of the diagonal fore and hind limbs moving in unison, contacting the ground at the same time, and leaving the ground before the opposing diagonal limb pair begin their support phases. In order to increase speed, stride frequency rather than stride length is typically increased, resulting in shorter duty cycles, and an increase in downward force on the ground[26].

A gallop, sometimes referred to as a canter, allows animals to move at greater speeds than can be achieved at a trot. The transition from a trot to a gallop involves a relative shift in the support phases of the fore and hind legs, such that the two fore legs are more or less in phase, followed by the two hind legs. By shifting the phase of limb support to allow fore and hind legs to act

---

25 Ibid. pg. 54

26 Ibid. pp. 56-57

together as pairs, galloping animals are able to increase their stride length to a greater extent than is possible by rotation of their limbs alone. This is achieved by flexion and extension of the spinal column, and rotation of the shoulder and pelvis. However, larger animals tend to have a more rigid spine than smaller, lighter animals, as larger animals must support the increased weight of their trunk. Typically, increases in speed at a gallop mainly involve increases in stride length with little increase in stride frequency.

At a slow gallop, one fore leg lands slightly ahead of the other, followed by a similar pattern of support by the hind legs. The phase difference is more often greater between the fore legs than the hind legs. At faster gallops the two fore legs and two hind legs progressively land more in phase and the limb duty cycle decreases.

Because of the reduced duty cycle, galloping involves aerial phases, which may intervene between one or both sets of limb support. The aerial phases are a necessary consequence of the increased stride length that animals use to increase their speed[27]. Consequently, the force exerted upon the ground by the limbs increases as the animal increases the speed of its gallop.

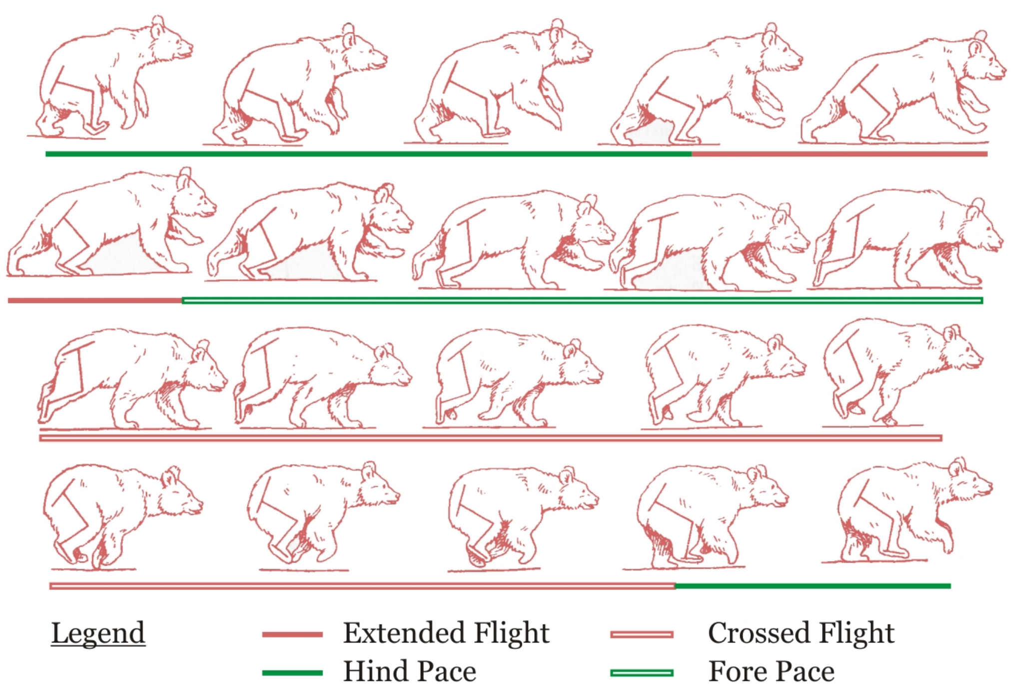
When contrasting the general gaits described above with that of the grizzly bear (*ursus arctos*) the major difference occurs at the gallop[28]. When a grizzly bear gallops the extended flight phase is comparatively short (approximately 15%

---

27 Ibid. pp. 57-58

28 Gambaryan, *How Mammals Run: Anatomical Adaptations* pg. 211

of the distance they cover in one cycle), and the crossed phase is lengthened (approximately 40% of the distance they cover in one cycle)[29]. In comparison, a dog's extended flight is 55%, and their crossed flight is 18%; and a cat 80% and 0% respectively – hind step and crossed flight are virtually absent in a cat. It is presumed that these differences are due to the relatively high force that a grizzly's forelimbs can exert during a gallop as opposed to other quadrupeds[30]. Visually, this makes their backs appear more arched during a gallop (see Figure 4) when compared to other animals with similar skeletal structure, such as dogs and wolves.



Adapted from Gambaryan (1985)

Figure 4: Movement phases of grizzly bear (*ursus arctos*) while galloping

---

29 Ibid. pg. 209

30 Ibid. pg. 209

Step Detection and Dead Reckoning

The purpose of this section is to analyze and summarize grizzly bear limb motion and the technologies and algorithms that were used in the design of a grizzly bear navigation aid (NavAid) for GPS assisted dead reckoning navigation. In order for such a system to function satisfactorily under normal conditions it requires the integration of a number of technologies, including GPS, accelerometers and magnetometers, so that continuous positioning and orientation of the system is possible (see Chapter 4 for details).

In the following sections, we will review limb motion of an animal, and the various techniques, which allow for the identification of steps using various sensors, under the assumption that the NavAid is attached to a collar that is placed around the animal's neck (see Figure 5). It is assumed that the collar is oriented in the local level frame — the axes are aligned with north and east and

Figure 5: Collar installation on a grizzly bear

Forelimb Locomotor Activity

A large component of this study is the use of MEMS accelerometers to identify the steps that an animal takes. In order to understand how these sensors are able to identify stepping motion it is necessary to have an understanding of the phases and periods of forelimb activity during locomotion. The ideal location to measure limb activity is either on the foot, or directly on top of the shoulder of a quadruped. However, for practical reasons, the sensors are mounted on a collar placed around the neck of an animal.

Following Gambaryan (1974), forelimb movement can be split into two phases, support and transit[31]. Each phase consists of two periods. The support phase has a preparatory period and a starting period, while the transit phase has a drawing up period and an adjustment period. During the preparatory period movement takes place in the joints of the limb to prepare the leg muscles to move the body forward (stage 1 on Figure 6), which occurs during the starting period (stage 2 on Figure 6). During the drawing up period of the transit phase, the limb is lifted upwards and forward (stage 3 on Figure 6), and lastly, during the adjustment period the limb sinks down again until it is placed on the ground (stage 4 on Figure 6). Accordingly, in the phase of support the limbs act to propel the body forward, while in the phase of transit the limb gets ready for the next locomotor cycle.

[31] Gambaryan, *How Mammals Run: Anatomical Adaptations* , pp. 63 - 65

As depicted in Figure 6, as the animal's foot is placed on the ground it exerts a downward force that peaks at the end of period 1, the downward force then reduces to zero by the time the limb is raised off the ground during period 3. This action results in a vertical acceleration at the shoulder that can be observed by a set of MEMS accelerometers. Figure 6 also indicates that periods 1 and 4 will tend to exhibit a downward acceleration, whereas periods 2 and 3 will exhibit an upward acceleration.

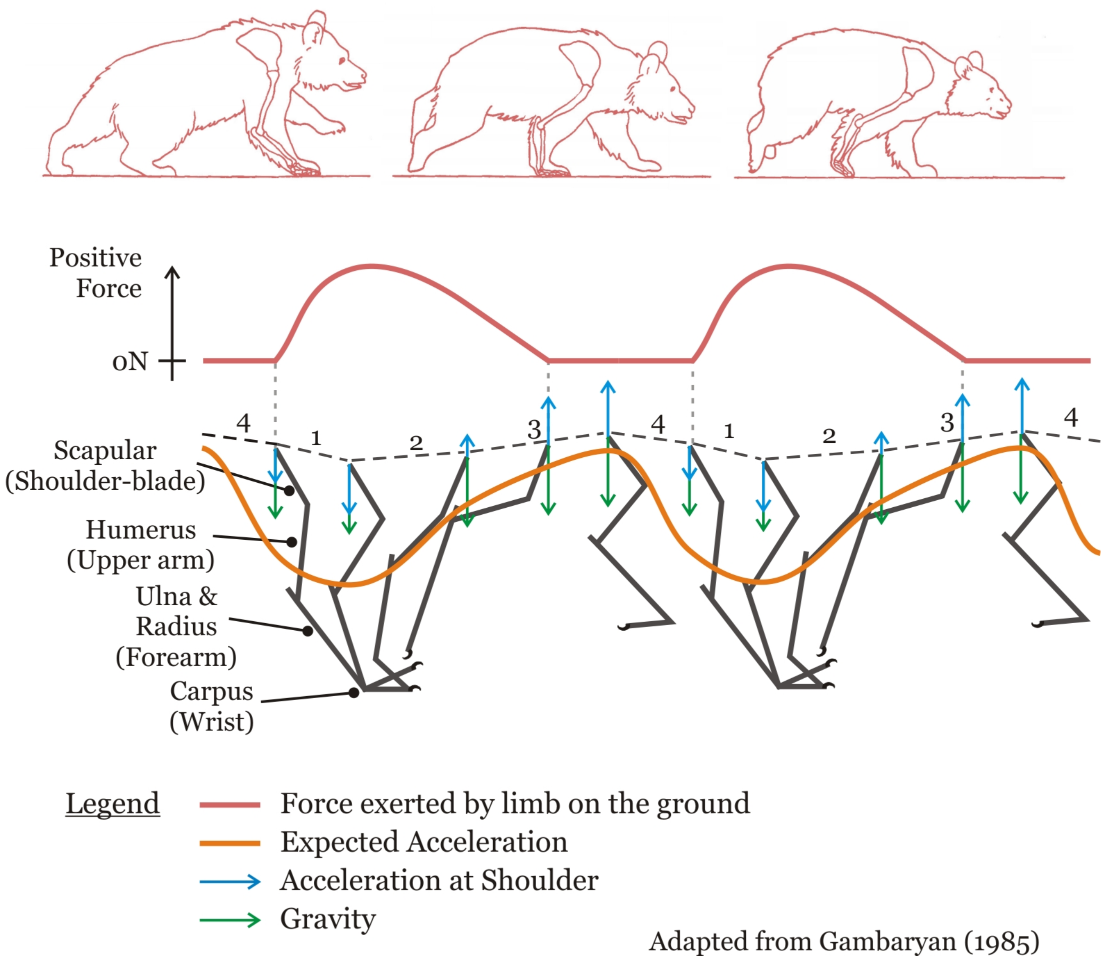


Figure 6: Phases and periods of forelimb activity

When the acceleration at the shoulder is combined with acceleration due to gravity, we obtain an expected acceleration signal from which to identify steps.

Dead Reckoning Navigation

The primary objective of the NavAid is to determine the animal's location and displacements within a reference frame that relates the animal to objects in the surrounding environment. This global reference frame, G, is a fixed right handed coordinate system aligned for convenience with the familiar geographic directions East, North, and Up as shown in Figure 7. The origin of the global reference frame is at an arbitrary but known position.

The goal of the NavAid is to track the animal's position as it moves through the global reference frame. The animal's reference frame, A, is defined by the animal's right hand side (X_A), the direction they are facing (Y_A), and the animal's zenith (Z_A) as shown in Figure 7. Through sensor measurements, i.e., the accelerometer and magnetometer, we can determine the orientation of the animal frame, A, with respect to the global reference frame, G.

The outputs of the sensors that measure accelerations and magnetic field are given with respect to a third reference frame, the sensor frame, S. This is also a right-handed coordinate system defined by the directions of the positive output from the sensors. Lastly, a fourth sensor frame exits at the sensor level. As there are multiple sensor units it is important that we know the relative orientations of each of the sensors in order to minimize the effect of biases that exist if the sensor components are not aligned correctly during manufacture.

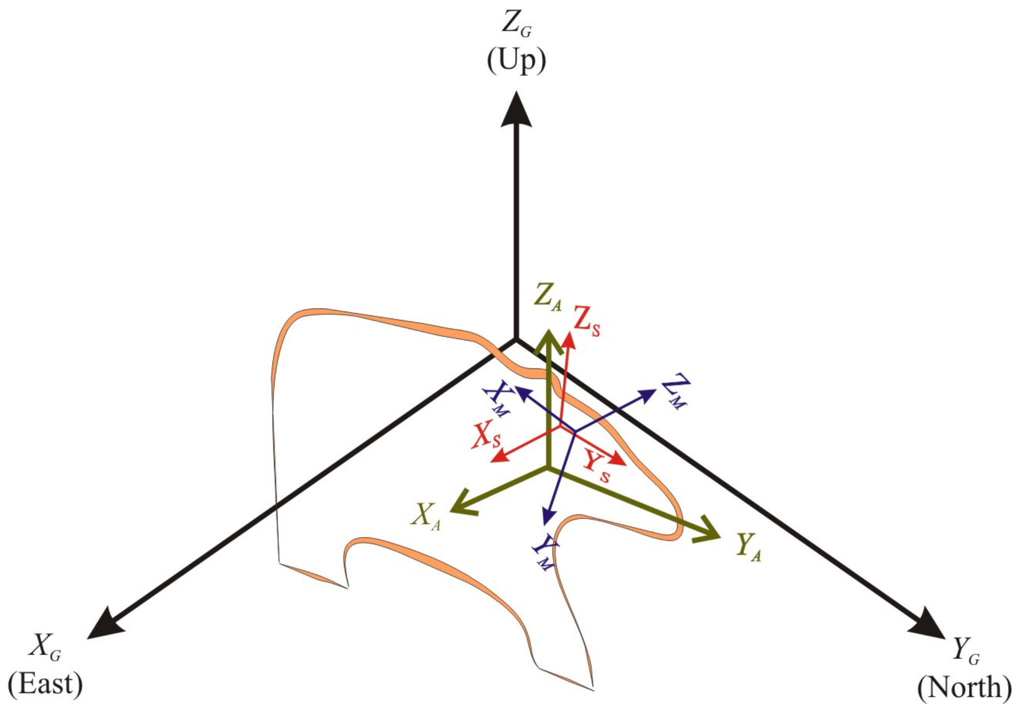


Figure 7: Coordinate frames for dead reckoning navigation

For this work it has been assumed that the accelerometer sensor axes define the S frame, and that the magnetometer axes, the M frame, must first be mapped, or rotated, to the S frame. Figure 7 shows the relationship among these four reference frames. In the case shown, the sensors are mounted on the animal's neck.

In sensor-based navigation, we attempt to solve the problem of relating the position and orientation of the animal frame to the global frame $(A \to G)$, by solving three intermediate orientations relating each frame to the sensor frame $(M \to S)$, $(A \to S)$ and then $(S \to G)$. The animal's movements cause S to translate and rotate with respect to G.

In practice, when the sensors are fixed within the NavAid case, e.g., on microprocessor boards inside a case, this means that M and S are coupled but not necessarily aligned. Hence, a calibration is required to establish their relative orientation. Equally, because the case is fixed to the animal, A and S are also coupled. In an ideal world any misalignment between A and S should also be determined. However, this is not practical for this application. We can assume that the animal is vertical when it is moving, hence rotation about Y and X can be estimated dynamically from the accelerometer data. Rotation about the Z axis is more problematic, but can essentially be dealt with in the same manner as magnetic declination, by assuming $\theta_Z^{A \to S}$, the rotation of A to S about the Z axis of S, is constant.

Dead Reckoning Mechanization

For dead reckoning (DR) navigation applications using step detection, sensor data are mechanized via a step counting and DR solution so that errors are proportional to the distance traveled, rather than to time[32]. In addition, step detection procedures and stride length estimation tend to give better results than mathematical integration of accelerometer outputs[33].

There is a range of sensors that may be used for the implementation of a step detection system. Sensors may include a GPS receiver for absolute[34]

---

32 R. Sterling, "Development of a Pedestrian Navigation System using Shoe Mounted Sensors" (University of Edmonton, 2004)

33 Q. Ladetto, V. Gabaglio, and B. Merminod, "Two Different Approaches for Augmented GPS Pedestrian Navigation" (paper presented at the International Symposium on Location Based Services for Cellular Users: Locellus, Munich, Germany, 2001)

34 Absolute, given the constraints of GPS and the mapping frame.

positioning, accelerometers for the detection of steps and estimation of stride length; magnetometers for determination of heading; gyroscopes for assisting with heading determination and assessment of local magnetic anomalies; and barometers for determination of changes in the height. However, given that we do not expect there to be substantial magnetic anomalies of the type found within urbanized environments[35], nor are we concerned with elevation changes at this time, the NavAid solution for this work has been restricted to accelerometers and magnetometers only.

The process of navigation by DR can be divided into four essential components (see Figure 8): step detection; stride length estimation; heading determination; and calibration. Step detection techniques can be categorized

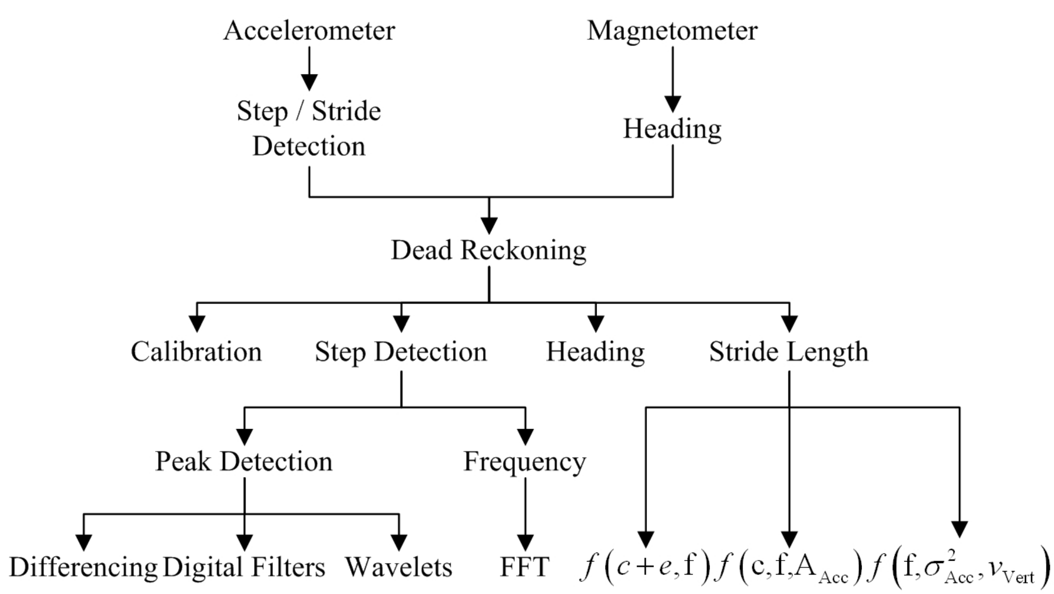


Figure 8: Data sources and analysis components for dead reckoning

35 Common anomalies include things such as electrical transformers, vehicle engine blocks, etc.

into two methodologies: (a) pre-processing of the accelerometer signal followed by peak detection techniques, such as differencing; (b) or alternatively, frequency analysis of the accelerometer data for the identification of step rate.

When implementing step detection navigation aids for people, during pre-processing, accelerometer data are typically band limited to 0.5–3.5 Hz[36], which are the frequencies of interest for human locomotion. As little is known about grizzly bear step frequencies it was proposed that this work would use a slightly wider range of frequencies, 0.33–4.5 Hz, when pre-processing the data[37]. Simple DR systems use a constant stride length to determine the distance traveled. However, analysis of the human gait shows that as people walk faster, their stride length increases, as such, a variable stride length model has been shown to improve the accuracy of distance traveled[38]. It is presumed that this will also be the case for animals given that during certain gaits they also increase their stride length to move more rapidly[39]. Many models exist for estimating stride length[40],

---

36 C. V. C. Bouten et al., "A Triaxial Accelerometer and Portable Data Processing Unit for the Assessment of Daily Physical Activity," *IEEE Transactions on Biomedical Engineering* 44, no. 3 (1997)

37 A review of grizzly bear video would suggest that this range of step frequencies will cover most grizzly bear movements. While there have been many studies of animal locomotion it would appear that those studies were more interested in the mechanics of animal gaits, as opposed to the timing.

38 S.-W. Lee and K. Mase, "Recognition of Walking Behaviors for Pedestrian Navigation" (paper presented at the Proceedings of 2001 IEEE Conference on Control Applications (CCA01), Mexico City, Mexico, 2001b); and R. C. Wagenaar and W. J. Beek, "Hemiplegic Gait: A Kinematic Analysis using Walking Speed as a Basis," *Journal of Biomechanics* 25, no. 9 (1992)

39 Gambaryan, *How Mammals Run: Anatomical Adaptations*

40 S. H. Shin and H. S. Hong, "MEMS Based Personal Navigator Equipped on User's Body" (paper presented at the ION GNSS 18th International Technical Meeting of the Satellite Division, Long Beach, CA, 2005); R. Jirawimut et al., "A Method for Dead Reckoning Parameter Correction in Pedestrian Navigation Systems" (paper presented at the 18th IEEE Instrumentation and Measurement Conference, 2001); and R. W. Levi and T. Judd, "Dead Reckoning Navigational System using Accelerometers to Measure Foot Impacts," (United States: 1996)

and all models used some combination of constant stride length, step frequency, accelerometer amplitude, vertical velocity, power spectrum (PS), etc., to estimate stride length.

As we have seen from the section on animal gaits, as the animal adopts a gait that allows it to move more rapidly, its duty cycle reduces, resulting in greater force being applied by the animal to the ground, hence when we know little about an animal's stride, we would expect that surrogates such as stride frequency, acceleration range, acceleration variance, or the PS would provide valuable information to assist with the determination of stride length.

In the following sections both hardware and DR algorithms that may be implemented in a NavAid are described. General limitations of the techniques are provided along with the mathematical models for DR used for the generation of locomotion vectors for the grizzly bear navigation system.

Dead Reckoning

Dead reckoning (DR) is the determination of a new position from the knowledge of a previous known position utilizing current distance and heading information. As such, DR consists of three important components: the prior absolute position of the user at time $t-1$, $\left(E_{t-1}, N_{t-1}\right)$, the estimated distance traveled by the animal since time $t-1$, $\left(\hat{s}_{[t-1,t]}\right)$, and the animal's heading (ψ) clockwise from north during the period $[t-1, t]$ (see Figure 9). The coordinates $\left(E_t, N_t\right)$ of a new position with respect to a previously known position $\left(E_{t-1}, N_{t-1}\right)$ can be computed as follows:

$$\begin{aligned} E_t &= E_{t-1} + \hat{s}_{[t-1,t]} \sin\psi_{t-1} \\ N_t &= N_{t-1} + \hat{s}_{[t-1,t]} \cos\psi_{t-1} \end{aligned} . \tag{2.1}$$

Consequently, the DR solution implemented for the NavAid captures the trajectory of a moving object in three-dimensional space (two-dimensional geometry with time in this case). As described in Trajcevski et al. (2004), a trajectory can be represented by a sequence of points[41] $(E_1, N_1, t_1), (E_2, N_2, t_2), \ldots, (E_n, N_n, t_n), (t_1 < t_2 < \ldots < t_n)$.

For a given trajectory, T , its projection into the XY plane is the route of T . A trajectory defines the position of an object as an implicit function of time,

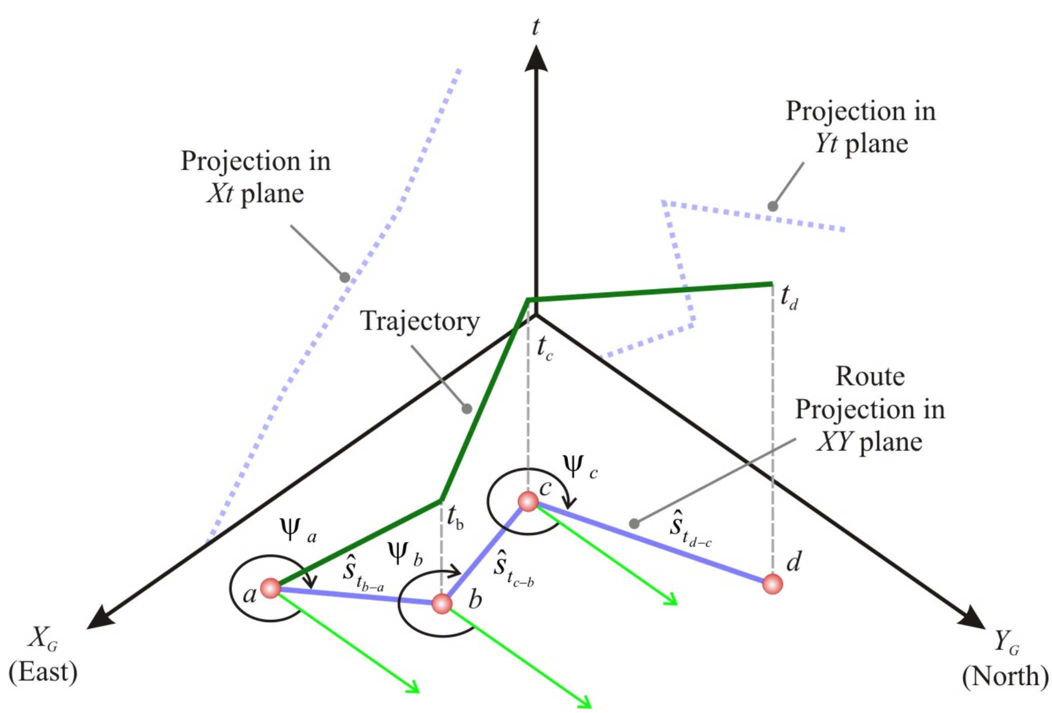


Figure 9: Principle of dead reckoning

---

41 G. Trajcevski et al., "Managing Uncertainty in Moving Objects Databases," *ACM Transactions on Database Systems (TODS)* 29, no. 3 (2004)

which can be calculated using the velocity of the object between (E_t, N_t) and (E_{t+1}, N_{t+1}) as shown in (2.2)

$$v_t = \frac{\sqrt{(E_{t+1} - E_t)^2 + (N_{t+1} - N_t)^2}}{t_{t+1} - t_t}. \tag{2.2}$$

Hardware

As indicated in Dead Reckoning above, the main hardware components that are required for the implementation of this animal NavAid include accelerometers and magnetometers. In addition, a GPS receiver is necessary to provide periodic locations within the general coordinate frame so that errors within the NavAid system can be bounded within some limit. The following sections describe briefly each of these components.

Accelerometers

An accelerometer measures acceleration forces via the application of Newton's second law of motion. That is, an accelerometer measures the force required to keep a proof, or reference mass, in its original position when the proof mass experiences a displacement due to an external force[42]. Micro-electro-mechanical systems (MEMS) based accelerometers generally utilize piezoresistive or capacitive technology. A piezoresistive system measures the strain due to an external force on the cantilever element that attaches the proof mass to the sensor housing. Piezoresistive accelerometers are sensitive to temperature variations and drift due to errors introduced by long-term integration, and

[42] J. P. Lynch et al., "Design of Piezoresistive MEMS-Based Accelerometer for Integration with Wireless Sensing Unit for Structural Monitoring," *Journal of Aerospace Engineering* 16, no. 3 (2003)

therefore, require the addition of compensation circuitry. Nevertheless, piezoresistive accelerometers exhibit low-noise at high frequencies. Capacitive accelerometers detect motion via a differential capacitor whose balance is disrupted by the movement of the proof mass. But while historically less expensive to manufacture, capacitive accelerometers exhibit high noise levels[43], which can result in a signal that is more difficult to interpret. Regardless of the type of accelerometer, its signal can be modeled as follows[44]:

$$I_f = \mathrm{f} + b_\mathrm{f} + \left(S_1 + S_2\mathrm{f}\right)\mathrm{f} + N\mathrm{f} + \delta \mathrm{g} + e_\mathrm{f} \tag{2.3}$$

where, I_f is the accelerometer measurement, f is the true specific force, b_f is accelerometer bias, S_1 and S_2 are linear and non-linear scale factor errors, N is a matrix representing the non-orthogonality of the accelerometer axes, $\delta\mathrm{g}$ is the local gravity anomaly (derived from the theoretical gravity value) and e_f is the accelerometer noise, which is assumed to follow a first order Gauss-Markov process.

Because the NavAid only requires the identification of steps, the primary concern is to locate the peaks (due to steps) within the accelerometer signal. As such, errors associated with accelerometer scale factors, gravity anomalies, and non-orthogonality do not need to be considered. So, a simplified version of (2.3) can be implemented as follows

$$I_f = \mathrm{f} + b_\mathrm{f} + e_\mathrm{f} \tag{2.4}$$

43 Ibid.

44 E-H. Shin and N. El-Sheimy, "A New Calibration Method for Strapdown Inertial Navigation Systems," *Zeitschrift fur Vermessungswesen* 127, no. 1 (2002)

The accelerometer information in the NavAid can also be used to estimate stride length through double integration of the accelerations observed by the sensor, but this works best if the sensor is attached to the foot[45]. Accelerometer data is also useful for estimating inclination of the sensor frame with respect to the global frame.

Magnetometers

Magneto-resistive or Magneto-inductive sensors may be used to measure the Earth's magnetic field at a location on the earth. As a sensor is rotated through the magnetic field, its resistance, or inductance, will vary as the magnetic field changes parallel to the sensor[46]. If we treat the magnetic field as a unit vector, h, then x and y components of h in the local level plane will point to magnetic north. Magnetic north varies from astronomic/geodetic[47] north by an angle θ_{Dec}, the magnetic declination. In addition, the magnetic field is tilted with respect to the local level plane by an angle θ_{Dip}, the magnetic dip angle[48].

45 When using accelerometers in this manner, error in position is compounded through the integration process; it increases $\propto t^2$, and will become problematic if not controlled.

46 M.J. Caruso and L. S. Withanawasan, "Vehicle Detection and Compass Applications using AMR Magnetic Solutions" (paper presented at the Sensors Expo, Baltimore, MD, 1999)

47 For the purposes of this work we shall assume that Geodetic and Astronomic north are the same. From a positioning perspective, prior to GPS, Astronomic North was the most accurate, convenient, and repeatable North reference. Astronomic North is based on the direction of gravity and the rotation axis of the earth, whereas Geodetic North is based on a Geodetic Datum, or mathematical approximation of the earth. The difference between the two is defined by the LaPlace Correction, $n \tan\phi$, where n is the deflection of the vertical in the prime vertical, and ϕ is the latitude at the point of interest.

48 Caruso and Withanawasan, "Vehicle Detection and Compass Applications using AMR Magnetic Solutions"

Magnetometers measure the Earth's magnetic field using dual or tri-axial sensors[49]. Assuming the magnetometer is levelled with respect to the local level frame, magnetic direction $\left(\theta_{Mag}\right)$ can be calculated using the horizontal components (x and y axes) of a magnetic sensor as follows

$$\theta_{Mag} = \tan^{-1}\left(\frac{Y_h}{X_h}\right), \tag{2.5}$$

where, Y_h and X_h are the horizontal components of the observed magnetic field. In order to account for roll (φ) and pitch (ω) between the sensor frame and the local level frame, the sensor observations should be rotated by the respective angles (roll and pitch) using (2.6) and (2.7), prior to the application of (2.5).

$$X_h = X\cos(\omega) + Y\sin(\varphi)\sin(\omega) - Z\cos(\varphi)\sin(\omega), \tag{2.6}$$

$$Y_h = Y\cos(\varphi) + Z\sin(\varphi). \tag{2.7}$$

X , Y and Z are the magnetic readings from the magnetometers sensor axes. To determine the heading (ψ_m), clockwise from magnetic north, we may adapt (2.5) as follows:

$$\begin{aligned}
&\text{if } X_h = 0 \wedge Y_h > 0 \rightarrow \psi_m = 0 \\
&\text{elseif } X_h = 0 \wedge Y_h < 0 \rightarrow \psi_m = 180 \\
&\text{else } \psi_m = 180 - \tan^{-1}\left(\frac{Y_h}{X_h}\right) + 90\left(\frac{-X_h}{\left|X_h\right|}\right)
\end{aligned} \tag{2.8}$$

where ψ_m is the magnetic azimuth of the NavAid. ψ_m may still contain errors due to the residual errors in roll $\left(e_\varphi\right)$, pitch $\left(e_\omega\right)$, and magnetic dip $\left(\theta_{Dip}\right)$. The magnitude of the error can be determined as follows

49 M.J. Caruso, "Applications of Magnetic Sensors for Low Cost Compass Systems" (paper presented at the Positioning, Location and Navigation Symposium (PLANS) 2000, San Diego, CA, 2000)

$$e_m = -e_\varphi \tan\theta_{Dip} \cos\psi_m - e_\omega \tan\theta_{Dip} \sin\psi_m \tag{2.9}$$

This heading error has directional dependence and increases with pitch angle. In addition, the larger the dip angle the greater the heading error; this limits the usefulness of a magnetometer for heading estimation as a user approaches the magnetic poles[50]. Astronomic, or grid, azimuth can then be determined using

$$\psi = \psi_m + \theta_{Dec} \tag{2.10}$$

Declination and dip can be obtained from magnetic field models developed by organizations such as the International Association of Geomagnetism and Aeronomy (http://www.iugg.org/IAGA) (IAGA) that provide various models for estimating magnetic declination at a particular location. In 2005, the IAGA released the 10th generation of their current model, the International Geomagnetic Reference Field (IGRF)[51].

Magneto-resistive sensors tend to be sensitive to temperature, whereas magneto-inductive sensors swap sensor polarity during an observation to remove the effect of temperature on the output signal[52]. Magnetometers are sensitive to hard and soft iron effects, which distort the local magnetic field, and therefore the heading estimated by a magnetometer in such a field[53]. Hard iron effects result

50 Caruso and Withanawasan, "Vehicle Detection and Compass Applications using AMR Magnetic Solutions" ; W. H. Campbell, ""Magnetic" Pole Locations on Global Charts are Incorrect," *EOS Transactions, AGU* 77, no. 36 (1996)

51 F.J. Lowes, *IAGA Division V-MOD Geomagnetic Field Modeling: IGRF Proper Use* (IAGA Working Group VMOD, 2005 [cited May 3 2006]); available from http://www.ngdc.noaa.gov/IAGA/vmod/igrfhw.html

52 3-Axis Magneto-Inductive Sensor Driver and Controller with SPI Serial Interface, PNI Corporation, Santa Rosa, CA

53 Q. Ladetto and B. Merminod, "Digital Magnetic Compass and Gyroscope Integration for Pedestrian Navigation" (paper presented at the 9th Saint Petersburg International Conference on Integrated Navigation Systems, Saint Petersburg, Russia, May 27-29 2002)

from magnetic materials in the device itself or in the environment where the device will be mounted. These are generally constant in magnitude and their effect is to add field components along the axes of the magnetometer. Soft iron effects are due to the induced magnetic fields that distort the Earth's magnetic field measurements. These effects will vary depending on the orientation and location of the NavAid[54].

Because hard iron disturbances are constant for a particular NavAid, they can be estimated via calibration, and hence their effects can be removed. The calibration process estimates the biases in both X and Y axes, and the scale factor required to ensure the magnitude of both axes are equal. This will ensure that the vector product of the axes will produce a circle centered at [0, 0] when the magnetometer is rotated through 360° (see Figure 10 below). The calibration model, therefore, is

$$\begin{aligned} H_X &= S_X\left(H_X^m + H_{X_0}\right) \\ H_Y &= S_Y\left(H_Y^m + H_{Y_0}\right) \end{aligned} \tag{2.11}$$

where H is the corrected magnetometer outputs, S is the scale factor for each axes, H^m is the observed output from the axes, and H_{X_0} and H_{Y_0} are the biases for each axes.

The primary advantages of using magnetometers for the estimation of heading are that they provide absolute direction with long-term accuracy and

54 Caruso and Withanawasan, "Vehicle Detection and Compass Applications using AMR Magnetic Solutions"

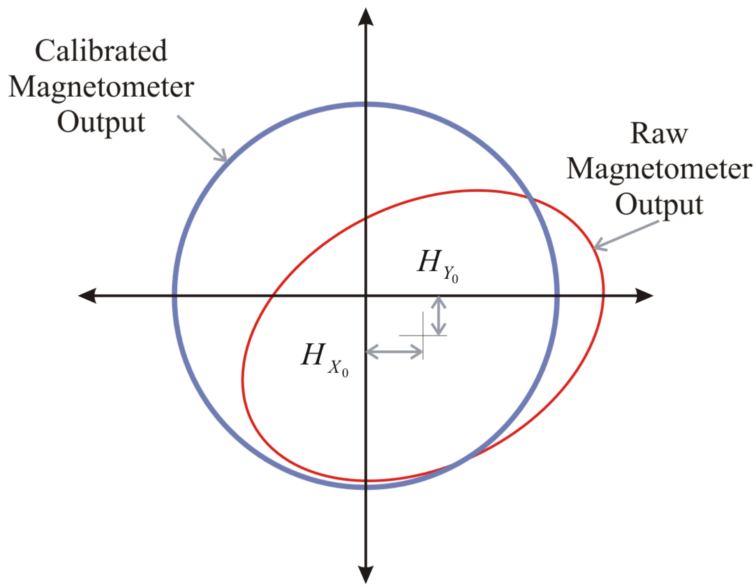


Figure 10: The effect of hard iron anomalies on magnetometer output

repeatable results. However, external magnetic disturbances can result in unpredictable outcomes[55].

To summarize, the grizzly bear NavAid solution consists of the following hardware:

1. One tri-axial accelerometer;
2. One tri-axial magnetometer;
3. A microprocessor for managing the sensors; and
4. On-board memory for data storage.

Step Detection Techniques

Step detection techniques are a group of procedures that estimate the time of each stride. As outlined in Figure 8, these processes can be separated into step

55 Ladetto, Gabaglio, and Merminod, "Two Different Approaches for Augmented GPS Pedestrian Navigation"

detection, stride length estimation, heading determination, and calibration. In addition, a number of step detection techniques require signal pre-processing and smoothing to facilitate interpretation. Each of these stages is summarized in the following sub-sections.

Signal Pre-Processing

Signal pre-processing is necessary to aid in the interpretation of sensor outputs. The most common techniques available for smoothing data include moving average filters, digital filters, and wavelets.

Moving Average Filter

A common technique used in many fields to smooth high frequency noise in a data stream is to take the average of a moving window. A window of size n (n is odd) is chosen and the window is centered on each data point. A weighted average of the data points that fall within the window replaces the data point of interest. The window then moves forward repeating the process one point at a time.

The level of smoothing depends on the size of the window and the weighting function employed. A larger window will produce a smoother signal, at the expense of the loss of high frequency information. Many different weighting functions may be used, these could include uniform, triangular, exponential, median, etc., and each will have a different effect. For example, an exponentially weighted moving average filter gives greater weight to more recent measurements[56].

Wavelets

Wavelets are an extension of Fourier analysis. They are mathematical functions that cut up data into different frequency components, and then analyse each component with a resolution matched to its scale[57]. This ability to analyse a non-stationary signal in both frequency and time is what makes wavelets attractive. A wavelet is a function $\upsilon \in \mathrm{L}^2(\mathbb{R})$ [58] with a zero average

$$\int_{-\infty}^{+\infty} \upsilon(t)\,dt = 0 \tag{2.12}$$

that is normalized, $\|\upsilon\| = 1$, and is centred in the neighbourhood of $t = 0$ [59]. Equation (2.12) implies that $\upsilon(t)$ is an oscillating function that tends to zero in some finite time. A family of wavelets can be generated from a mother wavelet by scaling and translating the mother wavelet using

$$\upsilon_{u,s}(t) = \frac{1}{\sqrt{s}}\upsilon\left(\frac{t-u}{s}\right) \tag{2.13}$$

---

56 F. R. Johnston and P. J. Harrison, "Discount Weighted Moving Averages," *The Journal of the Operational Research Society* 35, no. 7 (1984)

57 Barbara B. Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making*, 2nd ed. (Natick, Massachusetts: A K Peters, 1998)

58 $\mathrm{L}^2(\mathbb{R})$ is a finite energy function such that $\int \left| f(t)^2\,dt < +\infty \right|$, and returns a real number.

59 Stéphane Mallet, *A Wavelet Tour of Signal Processing*, 2nd ed. (New York: Academic Press, 1999) pg. 59

Where s is a scale parameter, and u a translation parameter. The wavelet transform at scale s and position u can be computed by correlating a signal with a wavelet atom

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \upsilon^{*}\left(\frac{t-u}{s}\right) dt \,. \tag{2.14}$$

It is the wavelet atoms that measure the variation of a signal in the neighbourhood of u whose size is proportional to s [60]. The mother wavelet is initially dilated and convolved with the signal to identify high frequency components in the signal. The wavelet is then rescaled by powers of two to progressively identify lower frequency components[61].

When using wavelets to denoise a data stream you decompose the data to produce a set of coefficients; some of the wavelet coefficients correspond to fine details in the data set. If the details are small enough, they may be omitted without substantially affecting the main features of the data set. Hence the process of denoising a signal requires setting to zero all coefficients that are less than a particular threshold. The wavelet coefficients are then used in an inverse wavelet transformation to reconstruct the smoothed data set[62].

---

60 Ibid. pg. 80

61 Amara Graps, "An Introduction to Wavelets," *IEEE Computational Science and Engineering* 2, no. 2 (1995)

62 Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making* pg.63, and David L. Donoho et al., "Wavelet Shrinkage: Asymptopia?" *Journal of the Royal Statistical Society. Series B (Methodological)* 57, no. 2 (1995)

Using wavelet analysis, Aminian et al. (2002) found that the swing phase of a person's stride could be reliably detected by gyroscope signals[63]. Ladetto, (2000) utilized a Myer basis function with wavelets to smooth accelerometer data during the step detection process[64]. The raw data was decomposed to level 4.

An issue with wavelets is the myriad of basis functions from which to choose. Do you reach for something familiar, or do you search for a new, better basis function? Wavelets with many vanishing points[65] tend to concentrate signal information into fewer coefficients[66], which is useful for denoising and compression. But they require larger support, which requires more computation[67]. Daubechies orthogonal wavelets require the smallest support for a given number of vanishing points. Many authors[68] suggest the use of a Best Basis algorithm to select the most appropriate basis function for a signal.

Signal Processing Filters

A low-pass filter is a filter that passes low frequency signals but attenuates signals with frequencies higher than the cut-off frequency. The actual amount of attenuation for each frequency varies from filter to filter. An ideal low-pass filter

---

63 K. Aminian et al., "Spatio-temporal Parameters of Gait Measured by an Ambulatory System using Miniature Gyroscopes," *Journal of Biomechanics* 35, no. 5 (2002)

64 Q. Ladetto, "On Foot Navigation: Continuous Step Calibration Using Both Complementary Recursive Prediction and Adaptive Kalman Filtering" (paper presented at the ION GPS 2000, Salt Lake City, Utah, September 19 - 22. 2000)

65 More vanishing points tends to equate to more oscillations Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making* pg. 244

66 Ibid.

67 p vanishing points require $(2p-1)$ points of support, see Ingrid Daubechies, *Ten Lectures of Wavelets* (Philadelphia: SIAM, 1992) pg. 256

68 Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making* ; Daubechies, *Ten Lectures of Wavelets* ; Mallet, *A Wavelet Tour of Signal Processing*

will be square with a vertical pass-band stop-band transition at the selected threshold frequency. In practice, a vertical pass-band stop-band transition is not possible; hence the transition will be curved to some degree resulting in leakage at the threshold frequency[69]. The Butterworth filter is the preferred low-pass filter as it most closely resembles an ideal filter.

Band-pass filters can also be implemented to remove both low frequency and high frequency signal noise at the pre-processing stage. Previous studies have shown that the frequencies of interest for human locomotion range from 0.5 to 3.5 Hz[70], and, as mentioned earlier, a similar pass-band criterion will be applied to a digital filter for use with large animals. For example, Lee and Mase (2001a) implemented a second order digital band pass elliptical filter with a 0.5 Hz to 5.0 Hz cut-off[71]. This filter provides sharp cut-off thresholds within the pass-band while limiting the amplitude of ripple in both the pass and stop bands.

A Butterworth filter can also be used if a flat pass-band (no ripple) is required[72]. However, the Butterworth filter does not provide as sharp a cut-off frequency for the pass/stop-bands as does an Elliptical filter. A Butterworth filter produces no ripple in the magnitude response of the filter and has a monotonically decreasing magnitude function with respect to the frequency. In

69 B. P. Lathi, *Linear Systems and Signals* (Berkeley: Cambridge University Press, 1992) pg.326

70 Bouten et al., "A Triaxial Accelerometer and Portable Data Processing Unit for the Assessment of Daily Physical Activity,"

71 S-W. Lee and K. Mase, "Recognition of Walking Behaviours for Pedestrian Navigation" (paper presented at the IEEE Conference on Control Applications (CCA01), Mexico City, Mexico, September 2001a)

72 Lathi, *Linear Systems and Signals* pg. 326

addition, the Butterworth is the only filter that maintains the same shape for higher order versions, but it has a slower roll off and therefore requires a higher order filter in order to obtain a comparable threshold effect to that of the Elliptical filter[73].

Peak Detection

There are several methods available to detect steps using peaks in the accelerometer signal. Techniques may make use of just the Z accelerometer data (vertical motion), or they may combine the Z and Y accelerometer data to observe vertical and forward motion. It is generally assumed that the wearer of the NavAid does not move side to side. Four common techniques used to determine when a step has been taken are described.

Peak Identification via Differencing and Thresholds

In this method, either just vertical, or vertical and forward signals can be used for step detection. When both signals are processed together the two peaks are shifted in time but follow each other for every real step, i.e., during each step, a vertical (Z accelerometer) peak will be followed by a forward (Y accelerometer) peak. After signal pre-processing, the two accelerometer signals can be added together, and then the peaks can be determined via differencing consecutive data points and utilizing a fixed absolute threshold. In Ladetto (2000) data were collected at 40 Hz, but no information was provided regarding the window size

73 Ashfaq Khan, *Digital Signal Processing Fundamentals* (Da Vinci Engineering Press, 2005 [cited 23 June 2007]); available from http://www.ebrary.com.ezproxy.lib.ucalgary.ca/corp/index.jsp pg. 261

that was used to identify the location of peaks[74]. Assuming a step frequency of 0.5 – 3.5 Hz, one could expect a maximum window size of 80 data points.

Absolute Value, Peak to Peak Time and Correlation

Lee and Mase (2001a) collected data at 50 Hz and used a sliding window of 25 samples for peak detection[75]. Major peaks and pits were identified in both the vertical and the forward signals. A step was then identified if the following three criteria were met within the sliding window.

1. The absolute difference between consecutive peaks and pits for both the vertical and forward signals was greater than some threshold;
2. The time between two consecutive steps was greater than some minimum threshold; and
3. Vertical autocorrelation reached some pre-defined threshold. This was to ensure that only steps were detected and not some other sudden body movement.

Differentiation

In this method, orientation of the accelerometer does not matter. The main issue is to determine an appropriate window size that will cover one complete locomotor cycle. Once determined, it is then a matter of identifying when the signal has a slope equal to zero. In effect, the algorithm identifies the "starting period" of a locomotor cycle (stage 2 in Figure 6). Once the signal has been

74 Ladetto, "On Foot Navigation: Continuous Step Calibration Using Both Complementary Recursive Prediction and Adaptive Kalman Filtering"

75 Lee and Mase, "Recognition of Walking Behaviours for Pedestrian Navigation"

smoothed with a low pass filter it is then differentiated. The differentiated signal during the "starting period" will have a slope approximately equal to zero (0)[76].

Zero Crossings

In Käppi et al. (2001), steps were identified by determining when the smoothed vertical accelerometer data crosses zero after the mean of the vertical signal has been removed, i.e., when $Z^* = Z_{smoothed} - \overline{Z} = 0$, and the first derivative of Z^* is positive[77]. In this instance a low-pass filter was used to smooth the data.

Issues

All the papers reviewed for these techniques were somewhat vague in their descriptions relating to window sizes and thresholds, in addition sampling rates ranged from 16 Hz to 50 Hz. It is anticipated that some calibration would be required to determine appropriate values as walking styles will likely differ from animal to animal, however, this is not generally practical. One suspects that these parameters should ideally be dynamic as there are many factors that influence stride patterns – steep terrain, hot weather, fatigue, rain or snow, to name a few. Given the limited knowledge of grizzly bear locomotion, this would seem impractical at this stage; hence it is expected that step count accuracies in the order of 2% to 3% as reported by Ladetto (2000) and Lee and Mase (2001a) is unlikely at this stage.

---

76 S. Y. Cho et al., "A Personal Navigation System using Low-Cost MEMS/GPS/Fluxgate" (paper presented at the ION 59th Annual Meeting/ CIGTF 22nd Guidance Test Symposium, Albuquerque, NM, 23 - 25 June 2003 2003)

77 J. Käppi, J. Syrjärinne, and J. Saarinen, "MEMS-IMU Based Pedestrian Navigator for Handheld Devices" (paper presented at the ION GPS 2001, Salt Lake City, UT, 11-14 September 2001 2001)

Fast Fourier Transform (FFT)

FFT is an algorithm used to determine the frequency components of a time varying signal. The FFT separates the waveform into a sum of sinusoids of different frequencies[78]. This transform is useful for stationary signals, as an FFT cannot provide both frequency information and timing information relating to changes in frequency. In order to obtain frequency information from a non-stationary signal it is necessary to combine the FFT with a moving window algorithm[79]. The FFT algorithm is somewhat computationally intensive, hence when employed in a NavAid the accelerometer data should first be assessed to determine if the NavAid is moving, typically via some (variance) threshold. The FFT algorithm produces a Discrete Fourier Transform of the signal, which consists of a series of discrete frequencies and their relative strength[80]. The discrete frequency with the largest magnitude is the dominant frequency in the signal. In terms of step detection, if this frequency falls within an acceptable range, i.e., 0.33 Hz to 4.5 Hz, then it is taken as the step frequency[81].

Stride Determination

Studies of human locomotion have concluded that stride length is a function of step frequency, velocity, and slope of the ground[82]. As such, if optimal results are

---

78 B. Oran Bigham, *Fast Fourier Transform and Its Applications*, ed. Alan V. Oppenheim (Upper Saddle River, NJ: Prentice Hall, 1988)

79 Mallet, *A Wavelet Tour of Signal Processing* pp. 60 - 70

80 Lathi, *Linear Systems and Signals* pp. 480 - 482

81 K. Macheiner, K. Legat, and B. Hofmann-Wellenhof, "Testing a Pedestrian Navigator" (paper presented at the European Navigation Conference, GNSS 2004, Rotterdam, The Netherlands, 2004)

82 Lee and Mase, "Recognition of Walking Behaviours for Pedestrian Navigation" ; Wagenaar and Beek, "Hemiplegic Gait: A Kinematic Analysis using Walking Speed as a Basis," ; and Macheiner, Legat, and Hofmann-Wellenhof, "Testing a Pedestrian Navigator"

to be achieved, estimation of stride length is essential. Most pedestrian based applications however have tended to disregard slope, and concentrate on stride length as a function of step frequency as it is straightforward to determine using a windowed FFT. The following sections briefly review some algorithms that have been reported for estimating stride length.

Constant Stride Length plus some Variation

The simplest model is to adopt a constant stride length. However, as with the human gait, it is expected that stride length will differ depending on the type of terrain that is being traversed. In addition, a number of studies by Lee and Masw (2001a), Wagenaar and Beek (1992), Cho et al. (2003) and Macheiner et al. (2004) have shown that under the same walking conditions stride length will increase as velocity increases, and equally, different walking conditions also affect stride length[83]. Therefore, in order to account for this variation, the stride length can be augmented with a variable component that represents the variation in a user's stride, $\hat{l} = l + e_l$. The error component could be Gaussian noise, or a Gauss Markov process. The difficulty is to determine the standard deviation (σ) for the Gaussian distribution, or σ and ρ (correlation coefficient) for the Gauss Markov process, so that stride variation is representative of the general user population, as it is expected that factors such as age and sex of a user will affect

83 Cho et al., "A Personal Navigation System using Low-Cost MEMS/GPS/Fluxgate" , Lee and Mase, "Recognition of Walking Behaviours for Pedestrian Navigation" , Macheiner, Legat, and Hofmann-Wellenhof, "Testing a Pedestrian Navigator" , Wagenaar and Beek, "Hemiplegic Gait: A Kinematic Analysis using Walking Speed as a Basis,"

the users stride length. Hence, population estimates will require a substantial sample size that includes both sexes and a wide range of ages.

Step Frequency and Length

A linear relationship was been used to relate step frequency to stride length[84], $\hat{l} = a + bf + e_l$ where a and b are constants and f is the step frequency. Macheiner et al. (2004) developed a mathematical relationship between the step frequency and the stride length by curve fitting for various test subjects under different ground conditions[85].

Neural Network (NN)

Cho et al. (2003) implemented a NN to estimate stride length. The model assumed a non-linear relationship between stride frequency, accelerometer signal variance and ground slope from which a set of weights were determined and then combined to estimate stride length. The NN also included a feedback loop which provided a means of updating the weights to reflect changes in the users' gait. However, the system required calibration in order to estimate the mean acceleration during a stationary phase, and the mean and variance of accelerations for a stride. Under good GPS signal conditions, Cho et al. (2003) determined that the system was accurate to ±3% of the distance travelled.

84 Macheiner, Legat, and Hofmann-Wellenhof, "Testing a Pedestrian Navigator"
85 Ibid.

Calibration

The DR navigation solutions require calibration to determine step rate and stride length accurately. Generally, the calibration process requires that the user of the DR system traverses a path of a known distance and the number of steps are physically observed[86]. In many of the pedestrian applications reviewed, GPS was often available. As such, Kalman filtering was used to merge the distance and heading information obtained from the sensors with the GPS data, and therefore they were able to update their calibration parameters in real-time[87]. Clearly, stride calibration and integration of data streams via Kalman filtering is a beneficial method for producing reliable results.

Ideally, step count and stride length calibration would be possible under a range of environmental conditions, uphill, downhill, wet ground, treed ground, etc. However, with respect to grizzly bear tracking, these methods pose some problems. While it is recognized that calibration would be beneficial in terms of quality of results, due to logistical problems, expense to physically observe an animal and personal safety, it is difficult to undertake this type of calibration for a reasonable time in the animal's natural environment.

---

86 Lee and Mase, "Recognition of Walking Behaviours for Pedestrian Navigation" ; Ladetto and Merminod, "Digital Magnetic Compass and Gyroscope Integration for Pedestrian Navigation" ; Aminian et al., "Spatio-temporal Parameters of Gait Measured by an Ambulatory System using Miniature Gyroscopes," ; Cho et al., "A Personal Navigation System using Low-Cost MEMS/GPS/Fluxgate"

87 Ladetto, Gabaglio, and Merminod, "Two Different Approaches for Augmented GPS Pedestrian Navigation" ; Macheiner, Legat, and Hofmann-Wellenhof, "Testing a Pedestrian Navigator" ; J. W. Kim et al., "A Step, Stride and Heading Determination for the Pedestrian Navigation System," *Journal of Global Positioning Systems* 3, no. 1 - 2 (2004)

Given these issues a simplistic approach to calibration has been adopted for this work. The earlier discussion on forelimb locomotor activity shows that as an animal moves more rapidly, either within a gait, or between gaits, the force that it must exert on the ground has to increase. This must translate into greater accelerations. Hence, for this work it is proposed that we investigate a number of attributes derived from the accelerometer data to determine which of them can act as a surrogate for stride length variation.

Once steps have been identified and stride length estimated, each DR segment is calculated from a $[0,0]$ origin using (2.1). Calibration of the DR segment is then performed using a four parameter affine transformation

$$\begin{bmatrix} E \\ N \end{bmatrix} = \begin{bmatrix} E_0 \\ N_0 \end{bmatrix} + \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x_{DR} \\ y_{DR} \end{bmatrix} \tag{2.15}$$

where E and N are the transformed coordinates in the general mapping frame (see Figure 11), E_0 and N_0 are the GPS coordinates at the start of a DR route segment (in the general mapping frame), $s = \dfrac{\sqrt{\left(E_{GPS2} - E_{GPS1}\right)^2 + \left(N_{GPS2} - N_{GPS1}\right)^2}}{\sqrt{\left(x_{DR2} - x_{DR1}\right)^2 + \left(y_{DR2} - y_{DR1}\right)^2}}$, is a scale factor that maps the distance travelled by a DR segment to the distance between the GPS points that fix the DR segment in the general mapping frame, and $\beta = \alpha - \gamma$ is the rotation of the DR segment to fit its GPS anchor points; $\alpha = \tan^{-1}\left(\dfrac{x_{DR2} - x_{DR1}}{y_{DR2} - y_{DR1}}\right) + C$, $\gamma = \tan^{-1}\left(\dfrac{E_{GPS2} - E_{GPS1}}{N_{GPS2} - N_{GPS1}}\right) + C$, and C places the azimuth in the correct quadrant (also see equation(2.8)). In Figure 11 the green path is the path estimated directly from the sensors, and the blue path is the transformed path fitted to the GPS data. In a perfect system, $\hat{s}/\hat{n}$ is the stride length (see equation (2.17)), and β will be composed of magnetic declination, θ_{Dec}, and misalignment between the sensor frame and the animal frame, $\theta_Z^{A \to S}$, about the Z axis of the sensor frame (see Figure 7).

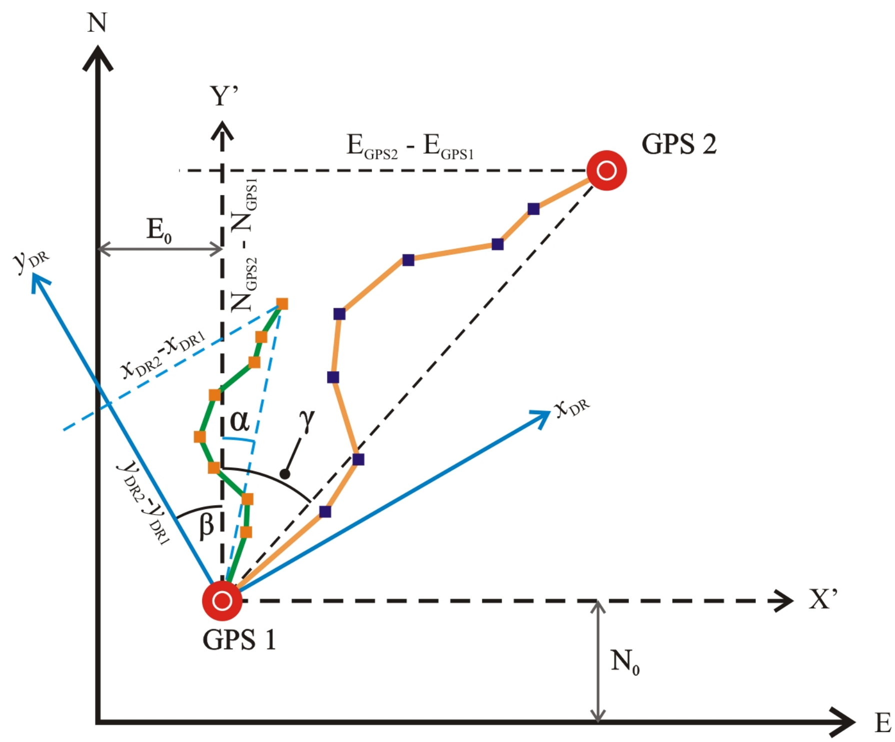


Figure 11: Geometry of a four parameter affine transformation

A NavAid Error Model

As described above, DR is the determination of a new position from the knowledge of a previous known position utilizing current distance and heading information, as such, (2.8) and (2.1) are the mathematical models used for DR.

Regardless of the peak detection method used for estimating the occurrence of a step, there will always be a certain chance that too few or too many steps are estimated. Any step count error will result in an erroneous estimated position of the NavAid. The simplest assumption is that missed or additional steps will be random, and follow a normal distribution. Hence, the step count model can be described by

$$\hat{n} = n + e_n , \tag{2.16}$$

where $\hat{n}$ is the estimated number of steps, n is the true number, and e_n, the step count error, is $N\left(0, \sigma_n^2\right)$.

Even with the use of a specific foot mounted inertial sensor[88], it is difficult to estimate stride length exactly. As such, the effect of stride length error must also be assessed. As with step counting a simple model for stride length can be defined as

$$\hat{l} = l + e_s , \tag{2.17}$$

where $\hat{l}$ is the estimated stride length, l is the true stride length, and e_l, stride length error, is $N\left(0, \sigma_l^2\right)$.

It then follows that the distanced travelled equals

$$\hat{s} = \hat{n}\hat{l} , \tag{2.18}$$

and the along-track error due to the step length and step count errors is

---

88 R. Sterling, K. Fyfe, and G. Lachapelle, "Evaluation of a New Method of Heading Estimation for Pedestrian Dead Reckoning using Shoe Mounted Sensors," *The Journal of Navigation* 58, no. 1 (2005)

$$e_s = S\sqrt{\left(\frac{e_n}{\hat{n}}\right)^2 + \left(\frac{e_l}{\hat{l}}\right)^2} \;. \tag{2.19}$$

Following the heading model defined in (2.10), heading determination using magnetometers consists of two components, magnetic heading, and declination. In addition, we know from (2.9) that the magnitude of the error in magnetic heading is determined by the actual heading, the error in estimating roll and pitch, and the local magnetic dip angle. Assuming that the error in pitch and roll are equal, constant, and uncorrelated, and that the magnetic dip angle is constant for the region within which the NavAid is to be used, then we can assume that the distribution of magnetic heading error is $N\left(0,\sigma_{\psi_m}^2\right)$, likewise, the distribution of the magnetic declination error is $N\left(0,\sigma_{\theta_{Dec}}^2\right)$. As such, the error model for heading can be described by

$$\hat{\psi} = \psi + e_\psi \,, \tag{2.20}$$

where $\hat{\psi}$ is the estimated heading, ψ is the true heading, and e_ψ, the heading error, is $N\left(0,\sigma_\psi^2\right)$, which is obtained by pooling the error distributions for magnetic heading and magnetic declination.

By applying the law of error propagation for uncorrelated observations, given the math models described by (2.1) and the Taylor expansion, the error propagation model for the eastings of a segment of a route is

$$\begin{aligned}\sigma_{E_{i+1}} &= \sqrt{\left(\frac{\partial E_{i+1}}{\partial \hat{s}_{i+1}}\right)^2 \sigma_{s_{i+1}}^2 + \left(\frac{\partial E_{i+1}}{\partial \hat{\psi}_{i+1}}\right)^2 \sigma_{\psi_{i+1}}^2} \\ &= \sqrt{\left(\sin\hat{\psi}\right)^2 e_{s_{i+1}}^2 + \left(\hat{s}_{i+1}\cos\hat{\psi}\right)^2 e_{\psi_{i+1}}^2}\end{aligned} \;. \tag{2.21}$$

Similarly

$$\begin{aligned}\sigma_{N_{i+1}} &= \sqrt{\left(\frac{\partial N_{i+1}}{\partial \hat{s}_{i+1}}\right)^2 \sigma_{s_{i+1}}^2 + \left(\frac{\partial N_{i+1}}{\partial \hat{\psi}_{i+1}}\right)^2 \sigma_{\psi_{i+1}}^2} \\ &= \sqrt{\left(\cos\hat{\psi}\right)^2 e_{s_{i+1}}^2 + \left(-\hat{s}_{i+1}\sin\hat{\psi}\right)^2 e_{\psi_{i+1}}^2}\end{aligned} \tag{2.22}$$

To ensure homogeneous dimensions, heading must be introduced to (2.21) and (2.22) in radians, either directly, or by dividing e_ψ by $\rho = 206,264.8$, the number of seconds in one radian. As route segments accumulate the error in eastings and northings will propagate according to

$$\begin{aligned}\sigma_E &= \sqrt{\sum_{i=1}^{p} \sigma_{E_i}^2} \\ \sigma_N &= \sqrt{\sum_{i=1}^{p} \sigma_{N_i}^2}\end{aligned} \tag{2.23}$$

where p is the total number of segments.

The error model will form an ellipse centred on the estimated position, where σ_E and σ_N are the dimensions of the ellipse axes. In order to achieve a 95% confidence interval, the semi major and semi minor axes should be multiplied by 1.96. This model has been depicted graphically in Figure 12.

Error Budget

The literature suggests that step counts can be estimated with an accuracy of 95% on level ground[89]. Ladetto (2000) reported a step count/stride length accuracy of 2%, while Lee and Mase (2001a) obtained an accuracy of 3% via the

89 Lee and Mase, "Recognition of Walking Behaviours for Pedestrian Navigation"

implementation of a neural network solution. Stride length accuracy on its own has not been reported in the literature reviewed. One suspects that accuracies of distance walked in the range of 2% to 3% is perhaps optimistic given the controlled environments under which the tests were likely performed, and the difficulty of calibrating the system when on a grizzly bear. Consequently, a simulation to investigate error propagation has been undertaken using an optimistic, one sigma error rate of 5% for step counts and stride length. If we assume a GPS fix rate of 30 minutes[90], an average stride length of 1.0 m, and an

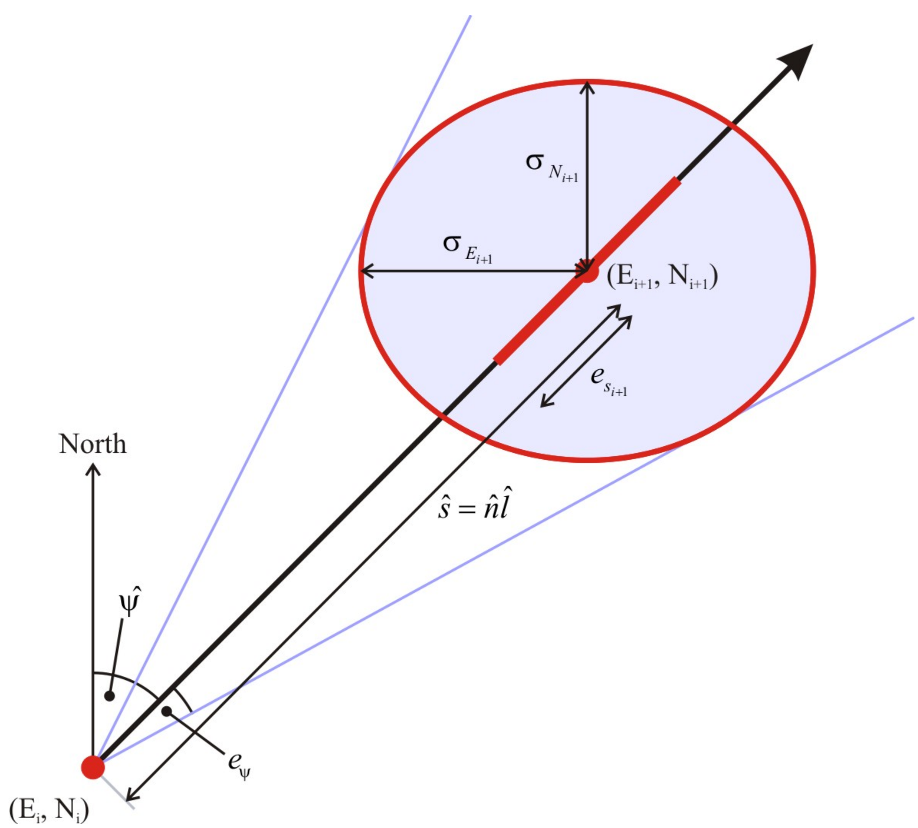


Figure 12: Geometry of NavAid error due to stride and heading error

90 30 minutes during mornings and evenings is a typical protocol used by the Foothills Model Forest for tracking grizzly bears.

average velocity of 5.5 m per minute[91], it is possible to develop a distribution for the expected distance travelled during a 30-minute period, $\left(\hat{s} = 165.0\text{m}, \sigma_s = 11.7\text{m}\right)$. Magnetic disturbances can cause significant errors in the heading. However, if properly calibrated a Root Mean Square (RMS) error of approximately 6° is possible[92] . The simulation that follows consists of a ten-segment route. It was assumed that a GPS fix was available at the start of the first segment and then again at the start of the sixth segment. In addition, the GPS positions have also been assumed to contain no error in order to highlight the errors attributed to dead reckoning. For the purpose of this simulation, headings were selected randomly from between 0° and 135°. Figure 13 depicts the route and the expected errors as they propagate along the route. It is evident from Figure 13 and Table 1 that given the specifications used, the primary source of error is due to errors in the distance travelled (i.e., errors attributed to stride length and step counts).

91 This rate is derived from existing GPS data, See Chapter 3.

92 J-H. Wang and Y. Gao, "A New Magnetic Compass Calibration Algorithm using Neural Networks," *Measurement of Science and Technology* 17, no. 1 (2006)

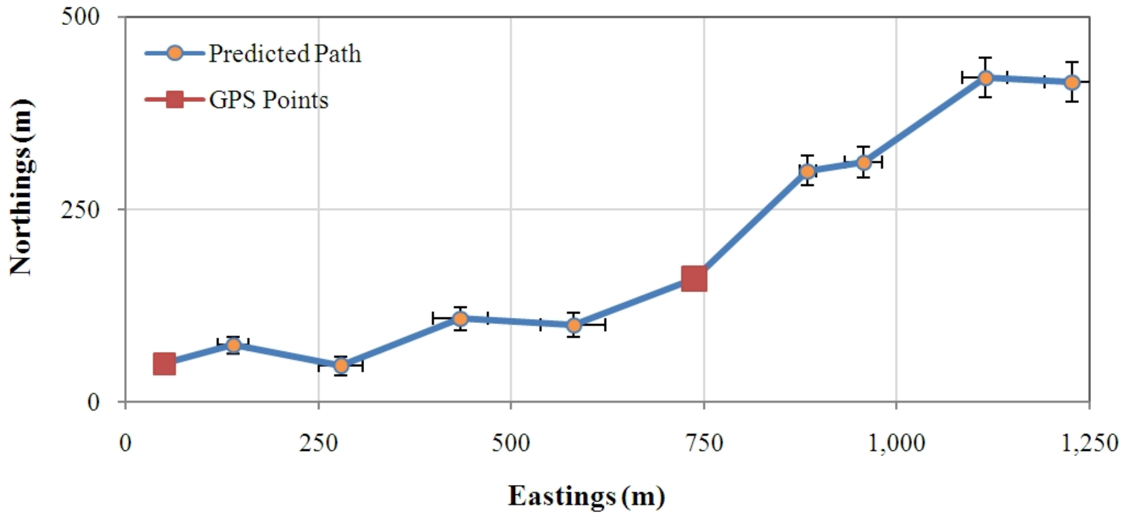


Figure 13: Propagation of error (95% CI) due to step count, stride length, and heading errors

Table 1: Tabulation of error propagation results

Point	$\hat{s}$	$\hat{\psi}$	East	North	σ_E	σ_N	$\sum\sigma_E$	$\sum\sigma_N$	$2\sigma_E$	$2\sigma_N$
1			50.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0
2	158.1	62°	139.6	74.2	9.9	5.5	9.9	5.5	19.5	10.7
3	160.6	100°	279.2	46.8	11.0	2.6	14.9	6.0	29.1	11.8
4	175.7	67°	435.0	108.6	10.3	4.7	18.1	7.6	35.5	15.0
5	165.0	93°	580.5	100.3	11.2	1.8	21.3	7.9	41.7	15.4
6	168.5	68°	738.4	159.5	0.0	0.0	0.0	0.0	0.0	0.0
7	157.3	27°	885.0	300.4	5.3	10.0	5.3	10.0	10.4	19.6
8	164.4	86°	956.8	311.4	11.2	1.9	12.4	10.2	24.3	20.0
9	176.7	46°	1,114.6	421.3	8.1	7.9	14.8	12.9	29.0	25.3
10	160.0	92°	1,228.3	415.8	11.2	1.7	18.6	13.0	36.4	25.5

Concluding Remarks

It is evident from this review that if we could place the accelerometers on the animal's feet, and the magnetometers on the animal's spine, then obtaining good results would be feasible. However, this is not the case. The sensor unit must be attached to a collar that is placed around the animal's neck. It is expected that this will make observation of steps more problematic as the sensors are not attached directly to the limb. As was observed from a review of a grizzly bear video, in addition to grazing as they walk (up/down accelerations of their neck and rotations of their head as they feed on grasses and herbs), grizzly bears exhibit a side-to-side motion as they walk. Given the fact that the collar is not fixed to the animal it seems reasonable to expect that the accelerometer outputs will be confounded with other signals unrelated to locomotion.

Like humans, grizzly bears spend time looking at things as they walk along. It is expected that this will compound the variation that is observed in the magnetometer data. What the effect of a collar that has rotated substantially around the animal's neck is unknown (X and Z axes of the magnetometer will be reversed if the rotation approaches 180°).

As suggested during the review of the various techniques, many of the procedures rely on certain window sizes from within which to identify steps, etc. What is the correct window size? Should it be constant over a range of locomotion gaits? Animal researchers using accelerometers to identify behaviours, or posture (but not gait) would suggest a variable window[93].

Thresholds for the identification of certain gaits also appear to be problematic. While three general gaits have been described, in reality each gait consists of a broad range of rhythms that will affect the accelerations that are observed. Selection of an inappropriate threshold may obscure the signal needed for the identification of steps. However, unless an animal can be observed for a period of time in its natural habitat, identification of an appropriate window size and signal threshold level for different gaits is likely to be driven by the researchers "intuition" for this stage of the research. This will likely limit the ability of the techniques discussed to identify steps and estimate stride length accurately.

93 See Shinichi Watanabe et al., "A new technique for monitoring the detailed behaviour of terrestrial animals: A case study with the domestic cat," *Applied Animal Behaviour Science* 94, no. 1-2 (2005) for example.

Chapter 3

Grizzly Bear NavAid

Introduction

In this chapter we review the objectives of the grizzly bear NavAid; the durability testing that was undertaken to determine appropriateness of the design for its proposed field environment; the hardware configuration of the NavAid; and the calibration of the system to address systematic errors attributed to misalignment of sensors during fabrication.

The NavAid

We have developed a simple NavAid solution to address a number of issues that limit the utility of current animal tracking methods that use GPS only. The objective of this phase of the research was to develop a sensor unit that was small and robust enough to sustain the rigours of life on a grizzly bear. Aside from compactness and durability, a primary objective was to develop a NavAid that was capable of lasting two seasons in the field. Guideline 28 of the "Guidelines On: The Care and Use of Wildlife"[1] states that

> *"… devices placed on wildlife should weigh no more than five percent of the animal's bodyweight, and where feasible, the device should be made as light as practical."*

The primary concern with weight is that it may have a detrimental affect on animal behaviour, and the heavier the unit, the greater the chance that the animal will persist in trying to remove it.

Case Development

Throughout this research a number of prototype cases have been developed from a range of materials. The initial unit was fabricated from fibreglass, but the fibreglass proved to be relatively heavy and difficult to seal adequately. The second version was constructed from polyurethane, which required that the electronics inside the case be encapsulated in an epoxy resin, as polyurethane does not provide sufficient structural strength on its own. While lighter, the case was quite flexible, which eventually resulted in the seal between the lid and the case failing due to a slight gap between the encapsulated electronics and the inside of the lid. In addition, the manufacture and encapsulation process was time consuming, as encapsulation required that each unit be placed in a vacuum in order to remove trapped air from the potting compound. After discussions with the technical staff at the University of Calgary Engineering Workshop, it was

---

1 *Guidelines On: The Care and Use of Wildlife*, Canadian Council on Animal Care (2003)

determined that the most appropriate material to manufacture the cases from would be aluminium. Figure 14 shows the design that has been deployed in 2007. The case, screws and glass window weigh 103 g.

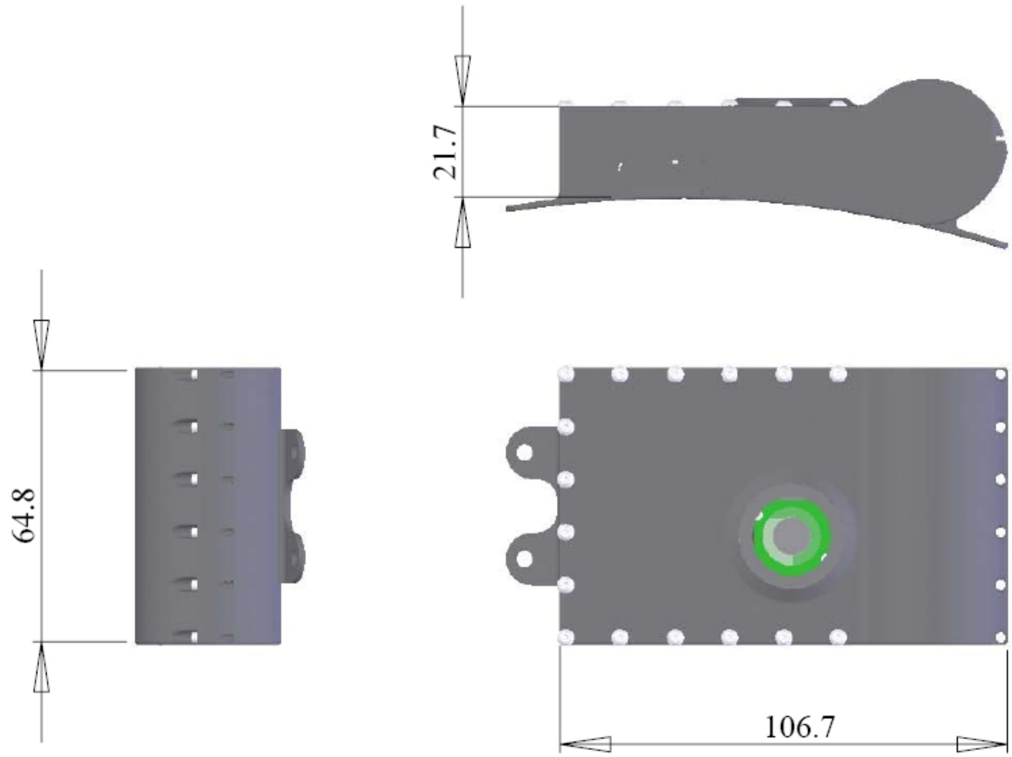


Figure 14: Current case design, v. 2007 (units = mm)

During development of the case and hardware, destruction tests were undertaken to determine the unit's ruggedness. A battery of tests included a vibration test[2]; a drop test[3], a temperature test[4], and a submersion test[5]. The vibration test lasted

2 This test was performed in the Civil Engineering Laboratory EN E 130, using the Syntron Vibrating Table (Model: VP86C; SN: 572523 2 81); the Syntron Electric Controller (Model: C2B; SN G45260 2 81); a National Instruments data logger card (Model: DAQCard 6062E; SN: 1056885) and National Instruments VI Logger software.

3 The drop test was performed in the Mobile Multi-Sensor Systems Research Laboratory, EN E 30.

4 The temperature test was performed in the Mobile Multi-Sensor Systems Research Laboratory, EN E 30 using the ESPEC Criterion Chamber ECT-3 (SN: 056708) with a Watlow F4 Controller.

5 Typically, a bucket with a minimum of 30 cm of water. Testing was always during the winter; hence by morning the water normally had at least 1 cm of ice on top.

for 2.5 hours at 120 Hz, with a typical force of 1.7 g[6] (5 on controller), and bouts of 3.6 g (10 on controller) for one minute at 15-minute intervals for the first two hours, and then 7.1 g (15 on controller) for the last 30 minutes. The only issue encountered during vibration testing was that the storage media for sensor data and images, SD cards, had a tendency to disengage from their locking mechanism. This was resolved by placing foam padding on top of the cards prior to installation of the lid.

During the drop test the unit was dropped by hand from a height of 0.76 m onto a 10 mm piece of plywood (5-ply), giving an impact velocity of 3.8 m/s (~14 km/h). A case was dropped 26 times on 6 faces, 8 corners and 12 edges. When dropped on a corner the case made indentations that were ~0.5 mm deep[7], which resulted in an estimated de-acceleration of ~152 g. However, it was noted that the case generally rotated upon impact, in particular when the case was dropped on an edge or corner, hence it is expected that most impacts were less than 152 g. No issues were encountered during a drop test.

During the submersion test the case was dropped in the water tank in the Civil Engineering Hydrology Lab for 30 minutes (~75 cm deep), or overnight in a bucket of water (at least 30 cm deep). The purpose of this test was to check for leaks in the seals. While submerged the air in the case cools creating a vacuum. If there are leaks in the seals or o-rings, the vacuum will draw water into the case.

6 According to accelerometer data recorded during the initial test.

7 As best I could measure.

The o-rings proved most problematic, as they tended to catch and distort when the plugs that provided access to the electronics were tightened when the lid was sealed.

The unit was also tested under a range of temperature conditions prior to deployment. Temperature normals[8] for Lake Louise suggest that average daily temperature could range from -21°C to +20°C in the Foothills region (extremes: -53°C to +35°C). An ESPEC Criterion Chamber ECT-3 (SN: 056708) with a Watlow F4 Controller was used to test the system over temperatures ranging from -30°C to +40°C. Temperature was extended towards warmer temperatures because we know from HOBO temperature sensors on the grizzly bear collars that temperatures could reach as high as 38°C during the summer, and appear to only go below zero degrees on older, male grizzly bears.

In order to ensure that we obtained reasonable temperature readings on the MSP430 microprocessor, it was necessary to maintain temperature at a particular level for at least 1.25 hours when the temperature was increasing, and one hour when decreasing. This ensured that temperature was stable for at least 15 minutes during each temperature setting. The main issue encountered was the identification of cold solder joints[9]. Four of the five cameras tested would stop functioning between -16°C and -20°C, the other camera stopped at -5°C. All

---

8 Environment Canada, *Canadian Climate Normals 1971-2000* (Environment Canada, 17 October 2005 2004 [cited 17 October 2005]); available from http://www.climate.weatheroffice.ec.gc.ca/climate_normals/index_e.html

9 A cold joint is a joint in which the solder does not make good contact with the component lead or printed circuit board pad. Essentially a bad connection that will fail, or increase noise within the hardware.

cameras restarted when the temperature was increased. It is presumed that the camera that stopped working at the higher temperature did so because of manufacturing issues particular to that camera. It should be noted that the operating temperature range required to ensure a stable image for the cameras was 0°C to 50°C with a maximum range of -10°C to 70°C[10]. All other electronics functioned throughout the temperature range tested.

Early prototype cases used plain glass for the window through which images are acquired. The window was recessed into the lid and held in place by an insert nut. This configuration proved to be problematic as it tended to become filled with debris. With the current prototype, the case is fabricated so that the lens protrudes slightly from the front face of the case in an attempt to minimize the accumulation of debris. We have also used Pilkington Activ™ Glass, which is "self cleaning". The glass is coated with a titanium oxide[11] compound that reacts to UV light to decompose organic matter that is attached to the glass. When the glass gets wet, the oxidized compound is washed away.

Hardware Development

An[12] an ST Microelectronics 3-axis Linear Accelerometer (±2 g/±6 g; Model: LIS3L02AS4[13]) for the detection of steps with a PNI ASIC 3-axis

---

10 OmniVision, "OV7640 Color CMOS VGA (640 x 480) Camera Chip," (Sunnyvale, CA: OmniVision, 2005)

11 Pilkington, *Glass that cleans itself!* (Pilkington Activ, 2006 [cited 24 October 2006]); available from http://www.pilkington.com/about+pilkington/education/glass+that+cleans+itself.htm

12 The physical integration of hardware components and software development to control the hardware was undertaken by Bruce Wright, Research Associate to Dr. Naser El-Sheimy.

13 LIS2L02AS4 - MEMS Inertial Sensor:2-Axis - ±2g/±6g Linear Accelerometer, STMicroelectronics, Geneva, Switzerland

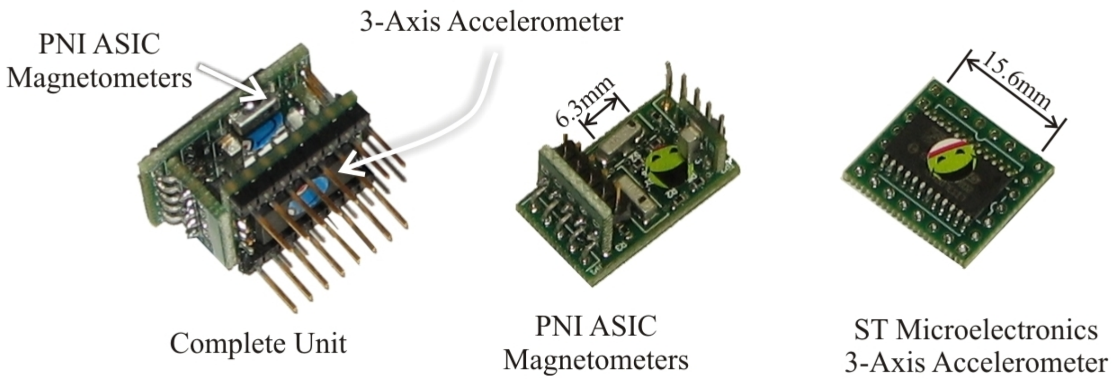


Figure 15: NavAid developed for grizzly bear dead reckoning

magneto-inductive sensor/IC (Model: 11096 plus SEN-S65 sensors[14]) for heading determination (see Figure 15 for the physical sensor layout, Figure 16 for a view of the completed hardware and case used in 2006, and Figure 17 for a block diagram of the components). These sensors augment the Televilt[15] Simplex and Tellus GPS collars used by the Foothills Model Forest Grizzly Bear Research Program (FMFGBRP) to create a dead reckoning (DR) system, the NavAid, for the acquisition of the continuous paths of an animal. The purpose of these additions to the GPS collar is two-fold: first they provide a means of addressing GPS bias due to canopy-induced data loss[16], as location information is now continuous; and secondly, they provide information that allows a researcher to assess animal behaviour through the analysis of trajectories – low movement

14 3-Axis Magneto-Inductive Sensor Driver and Controller with SPI Serial Interface

15 TVP Positioning AB, Bandygatan 2, SE-71134 Lindesberg, Sweden, http://www.positioning.televilt.se

16 See Mech and Barber, "Critique of Wildlife Radio-Tracking and its Use in National Parks: A Report to the U.S. National Park Service," ; D'Eon et al., "GPS Radiotelemetry Error and Bias in Mountainous Terrain," ; Dussault et al., "Evaluation of GPS Telemetry Collar Performance for Habitat Studies in the Boreal Forest," ; Gau et al., "Uncontrolled Field Performance of Televilt GPS-Simplex™ Collars on Grizzly Bears in Western and Northern Canada," and others.

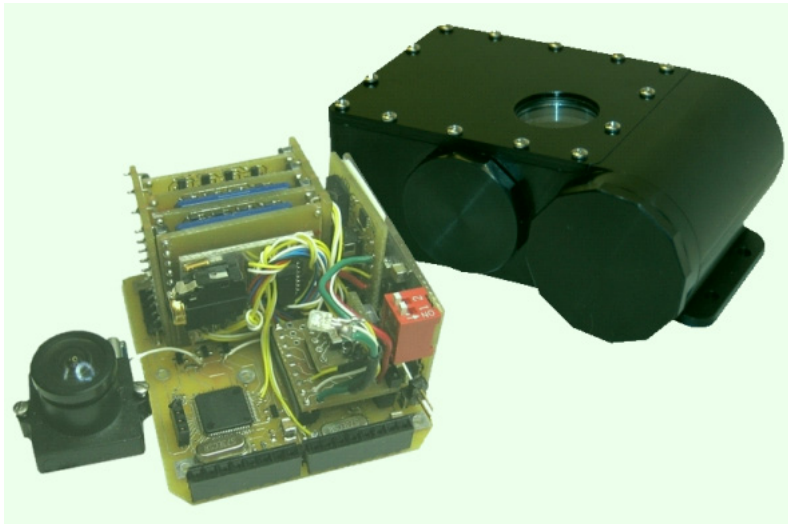


Figure 16: Complete sensor unit (power, memory, camera, sensors) and case, v. 2006

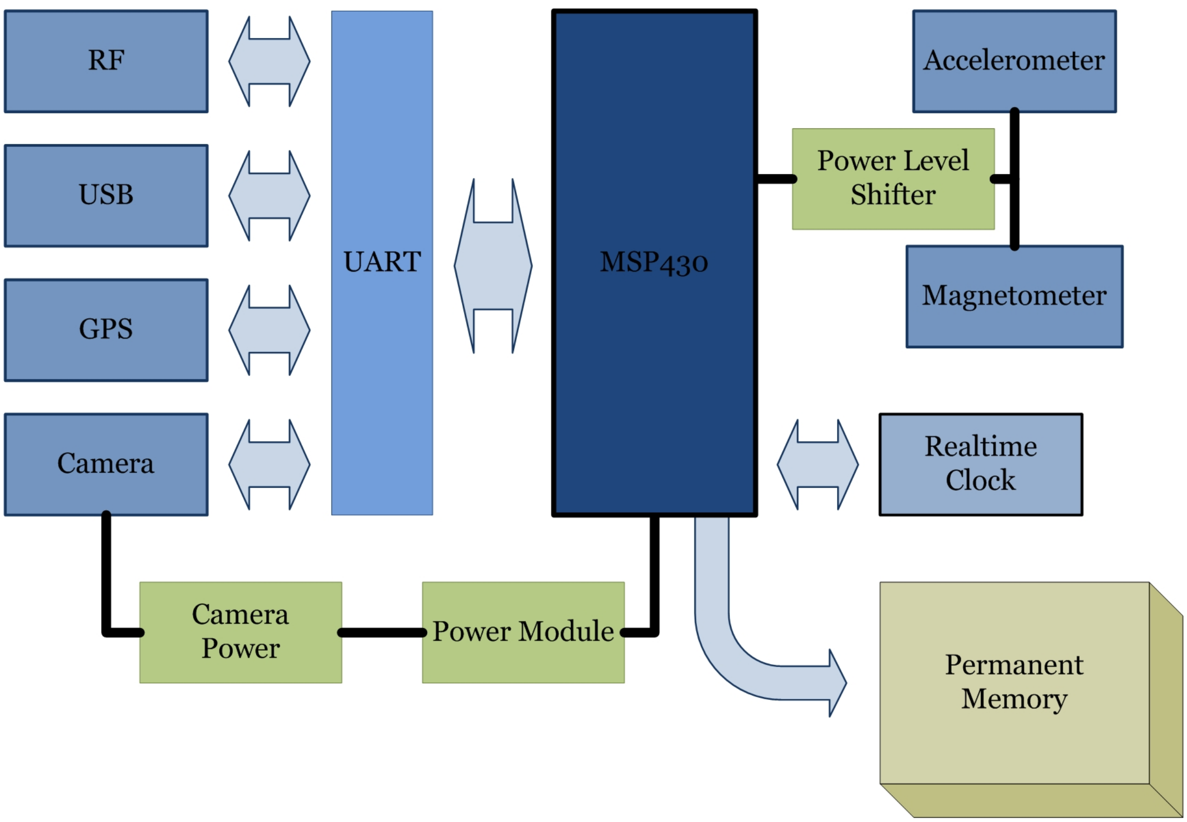


Figure 17: Block diagram of NavAid hardware

velocities combined with highly convoluted paths indicate searching and feeding behaviour, whereas higher movement rates and more directed trajectories indicate movement between feed sites[17].

It is proposed that this will allow a researcher to associate habitat preferences to activities such as feeding and locomotion, and to gain behavioural insights into issues related to animal security and interaction.

In addition to the DR sensors, the NavAid also includes a low-end CMOS digital camera manufactured by Commedia[18]. The C328-7640 is a camera module based on OmniVision's[19] OV7640 VGA camera chip (640 by 480). The OV7640 chip is a low voltage sensor that outputs 8-bit images at rates of up to 30 frames per minute. COMedia have mated the sensor with OmniVision's OV528 chip to provide JPEG compressed images. The current lens that we are using with the camera unit is the CL4022IR (focal length = 4.0 mm; F-Stop = 2.2) with an IR Cut filter. Throughout development various lenses have been trialed. This particular lens is a trade off between the field of view — capture as much information as possible — and maximum aperture — the animals spend a reasonable amount of time under the forest canopy, hence low light conditions are common. The purpose of the camera is to provide contextual information to locations along the route of an animal.

---

17 Smith, "The Searching Behaviour of Two European Thrushes: I. Description and Analysis of Search Paths,"

18 COMedia Ltd., Rm 802, Nan Fung Ctr, Castle Peak Rd, Tsuen Wan, NT., Hong Kong. Tel: (852) 2498 6248 Fax: (852) 2414 3050, Email: Sales@comedia.com.hk, Web: http://www.comedia.com.hk

19 OmniVision, 1341 Orleans Drive, Sunnyvale, California 94089, USA http://www.ovt.com/

The accelerometer sensors, magnetometer sensors, and camera are controlled by an MSP430F169 microprocessor from Texas Instruments[20]. The MSP430 is an 8 MHz, 16-bit ultra-low power (Off Mode (RAM retention): 0.2 µA; Standby Mode: 1.1 µA; Active Mode: 330 µA) programmable microprocessor. It includes an internal real-time clock, one synchronous I/O port for either SPI, UART or I²C communication[21] and one synchronous I/O port for either SPI or UART communication, a watchdog for the implementation of automatic restarts; two timers for controlling software and peripheral devices, two digital to analogue 12 Bit converters, 60 KB of Flash memory, 2 KB RAM, and an internal temperature sensor.

In addition to the primary data acquisition features, the NavAid also includes a power board to control power supply to the various components. Many of the components function at different voltages; therefore, in order to minimize power use a specialized board has been developed to perform power level shifting. The objective here has been to maximize the life expectancy of the power supply. To enable this, each component was designed to operate at the lower end of its voltage range. The accelerometer was set at 2.5 V (min. 2.4); the magnetometer is at 2.3 V (min. 2.2); the memory is at 2.8 V (min. 2.7); the MSP430 is at 2.7-3 V (min. 1.8, but frequency/processor speed is limited by

---

20 MSP430x15x, MSP430x16x, MSP430x161x Mixed Signal Microcontroller Ver. E, Texas Instruments Inc., Dallas, Texas , www.ti.com

21 SPI - Serial Peripheral Interface, a synchronous "4 wire" port ; UART – A universal asynchronous "2 wire" port; I2C – Inter-Integrated Circuit often used for communication between low speed peripherals and a microprocessor.

voltage); the GPS is at 2.8 V (min. 2.7); and the camera is at 3.0 V (min 3.0). Power is provided by one Saft[22] LSH 33600C Lithium-thionyl chloride (Li-SOCI2) D cell rated at 3.6 V and 18.5 Ah.

The final major component incorporated into the NavAid is memory. At present the NavAid is designed to hold four micro SD cards for storage of raw data. The prototype in Figure 16 used standard SD Cards for data storage, however the 2007 version has since been shrunk so as to provide a more compact package.

Throughout development, a considerable amount of time has been spent field-testing the system on animals. Initial field tests were on Llamas, then horses, followed by sheep. The main objective here was to test software reliability over longer periods of time. Field-testing identified many software issues that have resulted in the implementation of a more reliable software solution. Field-testing also gave an opportunity to look at the data streams that were being generated by the NavAid. The issue of timing was often problematic, and the long-term effects of the multiple interrupts used to activate the different components could not have been adequately addressed had we not had the opportunity to test the system on animals in an outdoor environment.

As an aside, Llamas are quite gentle animals that are easy to work with, but they are also difficult to keep a collar on, as their head is smaller than their neck. When they put their heads down the collar falls off, so many hours were spent

---

22 SAFT, 12, rue Sadi Carnot, 93170 BAGNOLET, France http://www.saftbatteries.com

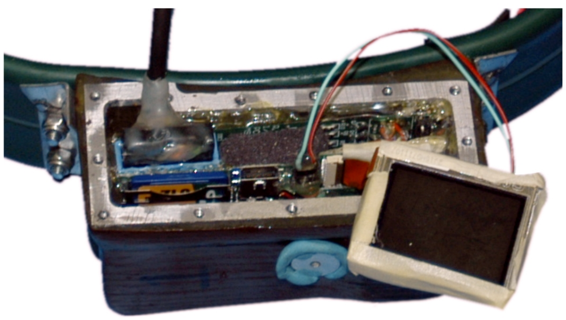
trying to find a collar in the grass. Horses too are easy to work with, but mounting the collar was somewhat problematic, and the horses we had access to tended to wait at the gate for people or food. Sheep are more difficult to catch, especially without a trained sheep dog, or yards. However, the sheep did provide insights into the type of movements to expect from the collar when placed on an animal with a tendency to graze.

The first prototype placed on a grizzly bear contained only a camera. Many of the photos that were acquired were blurry, even in what appeared to be good light conditions. After observing the sheep it was noted that as the animal tears the grass, the NavAid performs a distinct rotation, approximately around the X axis, and the photos exhibited a similar blurring pattern. This type of information has helped immensely to interpret the data streams that have been obtained from grizzly bears.

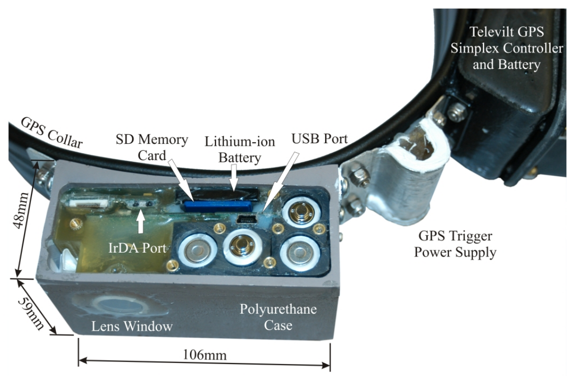
Physical implementation of the NavAid has been challenging. Testing various prototypes on grizzly bears over the four years that this research has been in progress has resulted in a number of physical challenges, notably waterproofing the NavAid case. Figure 18 depicts the various versions that have been developed for this work. The first prototype was fabricated using fibreglass. An aluminium plate was imbedded in the rim of the case so that the lid could be screwed on securely. Two cases were then deployed in the summer of 2003. However, both cases failed within two weeks because they filled with water.

Upon retrieval it was found that cracks had formed between the aluminium and fibreglass.

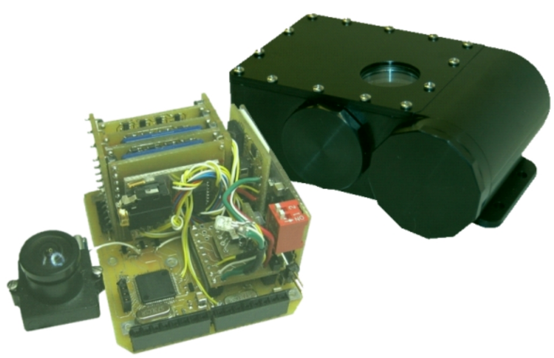
Version two was constructed from polyurethane, but as mentioned earlier, a meniscus formed along the outer edge of the electronics block during encapsulation. The meniscus resulted in the lid of the case flexing when under pressure. Two units were deployed in 2004 on female grizzly bears with cubs.



Version 1, 2003



Version 2, 2004



Version 3, 2006

Version 4, 2007

Figure 18: Various versions of the NavAid prototype

Inspection of the cases when they were released from the animals suggested that they had been chewed by the cubs, which resulted in the silicon seal failing after a few days.

Over the winter of 2005 the system was redesigned to include accelerometers and magnetometers. The redevelopment took considerably longer than expected and the five units planned for 2005 were not able to be deployed. Five units were eventually constructed for deployment in 2006; however, only two animals, G040 and G008, were captured for tracking. The collar from G040 was retrieved mid August of 2006 as G040 was relocated by Parks Canada. G008 denned before its collar could be released. This was later retrieved in early April of 2007.

Upon receipt of G040's collar it was found that the NavAid failed after 15 days of operation on the animal due to battery failure. Analysis of the units power demand at 10 kHz identified a spike approaching 650 mA when the camera turned on, however, the fuse in the battery was only rated for 250 mA, hence the failure. Earlier power testing had been carried out at a lower sampling rate that was too slow to catch the spike in power.

During the 2007 field capture season, the five units deployed in 2006 were redeployed along with two of the latest version of the NavAid. Ten new units were manufactured for 2007.

Sensor Calibration

As described in the previous chapter, the primary objective of a DR system is to determine the location and displacements of a moving object within a reference frame that relates the moving object to objects in the surrounding environment. In order to be able to do this reliably the DR sensors require calibration.

The ST Microelectronics LIS3L02AS4 accelerometer provides an analog signal representing accelerations in the directions of its axes. The LIS3L02AS4 allows the user to set full-scale at either ±2 g or ±6 g over a bandwidth of 1.5 KHz[23]. The PNI ASIC 3-axis magneto-inductive sensor is capable of sensing ±1100 µT (±11 Gauss) at a resolution of up to 0.015 µT, depending on configuration[24]. In terms of signal reliability, the performance indicators of concern for both the accelerometer and the magnetometer are sensor bias and scale factor. Typically, deviations from theoretical values are due to sensor imperfections and misalignment of axes during manufacture. Because of these imperfections sensor calibration is required in order to determine if the sensors are performing as expected, and to remove the effect of systematic errors that result from sensor imperfections such as misalignment.

Accelerometer Calibration

When implementing a DR system based on step detection, the purpose of accelerometer calibration is to determine if the sensors are functioning as

23 LIS2L02AS4 - MEMS Inertial Sensor:2-Axis - ±2g/±6g Linear Accelerometer

24 3-Axis Magneto-Inductive Sensor Driver and Controller with SPI Serial Interface

expected. Misalignment of the axes does not have a detrimental affect on performance of the sensors when using peak detection methods. Specification of ST Microelectronics accelerometer bias and scale factor is provided in terms of Zero-g Level and Sensitivity (see Table 2). Zero-g Level is considered to be the actual sensor output when no acceleration is present (bias), whereas sensitivity refers to the gain of the sensor (scale factor).

A six-position static calibration test was performed on five units during March 2006 in the Mobile Multi-Sensor Systems Research Laboratory at the University of Calgary (EN E 30). During the test each axis of the accelerometer was observed in an up and down direction with respect to gravity. For each of the six positions, averaged values where obtained from five minutes of observation at 32 Hz. Vdd[25] was set at 3 V; hence the design specification for Sensitivity was 0.60 ± 0.06 V/g, and Zero-g Level was 1.50 ± 0.15 V. With regards to Figures 19 to 21, the red line represents the expected output from the sensors for the

Table 2: ST Microelectronics accelerometer specifications

Parameter	**Test Condition**	**Min.**	**Typ.**	**Max.**	**Unit**
Acc. Range	Full Scale = 2g	± 1.8	± 2.0		g
Sensitivity	Full Scale = 2g	(Vdd/5)-10%	Vdd/5	(Vdd/5)+10%	V/g
Zero-g Level	T = 25°C	(Vdd/2)-10%	Vdd/2	(Vdd/2)+10%	V
Acc. Noise Den.	Full Scale = 2g		50		μg/√Hz

25 Vdd refers to the voltage of the power supply, in this case it is a "drain" voltage.

respective graph, and the error bars represent the expected range (one standard deviation).

Table 3 lists the results of the Zero-g Level test. This data is also depicted in Figure 19. Observed Zero-g values are all well within the design expectation, with the blue unit exhibiting the least bias over the three axes, the yellow unit having the least range in bias and the orange unit the greatest range.

Table 4 and Figure 20 illustrate the sensitivity results. As with Zero-g, all sensors were well within the design specification, although on average their sensitivity is slightly lower than expected, i.e., it would appear to be biased downwards. This is probably due to the sensor being operated near its minimum design voltage.

Table 5 and Figure 21 show the noise density results. In terms of meeting sensor design specifications, all units tested had lower observed noise densities than expected. In fact the noise density was less than half of that expected. This is good, because it means that the sensors should be able to observe finer details in the acceleration signal, but as with the sensitivity test it may also be tied to the

Table 3: Observed zero-g level (V)

Unit	**X**	**Y**	**Z**
Blue	1.542	1.489	1.495
Green	1.507	1.493	1.533
Yellow	1.512	1.536	1.506
Orange	1.560	1.483	1.530
Red	1.470	1.458	1.476

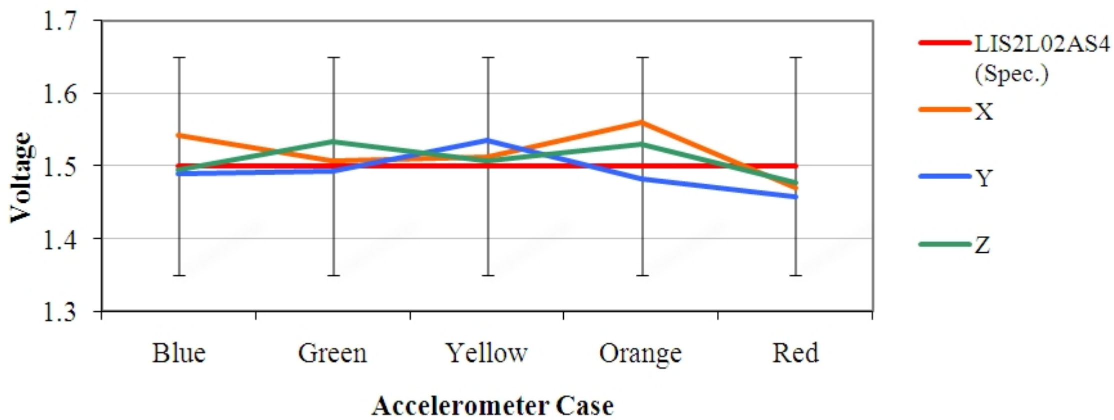


Figure 19: Observed zero-g levels

Table 4: Observed sensitivity (V/g)

Unit	X	Y	Z
Blue	0.583	0.582	0.576
Green	0.579	0.582	0.566
Yellow	0.569	0.573	0.573
Orange	0.575	0.584	0.580
Red	0.579	0.580	0.593

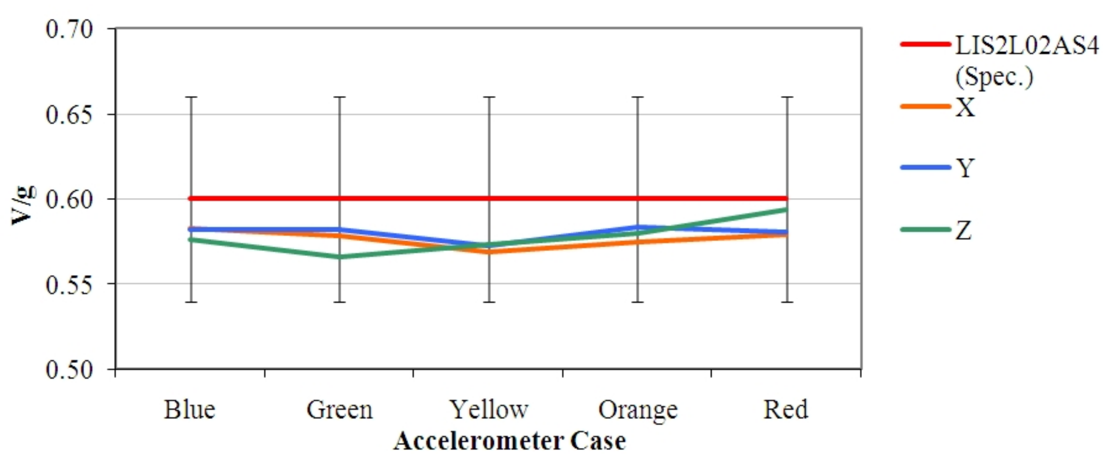


Figure 20: Observed sensitivity

Table 5: Observed noise density (μg/√Hz)

Unit	X	Y	Z
Blue	17.948	23.273	25.797
Green	17.644	18.876	22.649
Yellow	23.512	25.114	24.889
Orange	24.618	14.854	27.389
Red	14.693	15.635	13.595

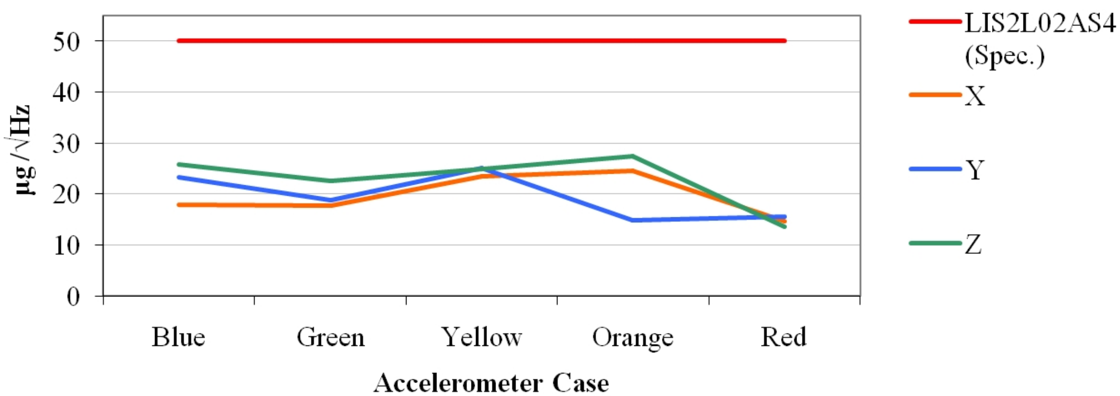


Figure 21: Accelerometer noise density

Table 6: Bias and scale factor for ST Microelectronics LIS3L02AS4 accelerometer

	Bias			Scale Factor		
Unit	X_b	Y_b	Z_b	X_{SF}	Y_{SF}	Z_{SF}
Red	41.58	57.59	32.71	1.0359068	1.0337266	1.0110295
Blue	-57.26	15.48	7.32	1.0295458	1.0305331	1.0415046
Green	-9.73	10.17	-44.47	1.0370517	1.0309736	1.0600225
Orange	-81.57	23.74	-40.46	1.0438955	1.0279658	1.0341622
Yellow	-15.80	-48.40	-8.29	1.0539890	1.0473627	1.0468808

fact that the sensor is operating at the bottom of its voltage range.

In terms of raw accelerometer output, a 12-bit data stream, these values equate to bias and scale factor estimates for the sensors as tabulated in Table 6 below. These biases and scale factors have been applied to the raw accelerometer data acquired from the sensor units prior to analysis.

Magnetometer Calibration

In order to obtain a correct heading when using a magnetic compass, calibration of the magnetometer is essential. There are two relatively straightforward methods of calibration that can be used. The first requires a reference heading against which you compare the output from the magnetometer. When a reference heading is unavailable, an alternative is to level and rotate the magnetometer through 360° to find the minimum and maximum values of the X and Y axes observations. These four values can then be used to compute the magnetometer scale factors and biases based on the knowledge that the locus of error-free measurement on the X and Y axes is a circle[26].

The design specification for the PNI ASIC 3-axis magneto-inductive sensor provides a range of graphics that give a general indication of the expected results[27]. However, no specific specifications (math model) were provided in the documentation provided with the sensors, aside from a resolution of 0.03 µT (1/Gain) when the Period Select is set to 2048 and a Gain of 27– 38 counts/µT at

[26] Caruso, "Applications of Magnetic Sensors for Low Cost Compass Systems"

[27] 3-Axis Magneto-Inductive Sensor Driver and Controller with SPI Serial Interface

Vdd = 3.0 V with a resistance of 33 Ω. Period Select is the number of cycles that a sensor is differenced to determine a magnetometer reading. Gain varies depending on the voltage and resistance used. For example, when Vdd = 5.0 V and resistance = 53 Ω, Gain should fall in the range of 18 – 26 counts/μT. For the configuration tested (Vdd = 3.3 V, resistance equals 43 Ω), inspection of the 3 V and 5 V graphs for a 43Ω resistor indicates that bias should, at worst, be ~275 counts.

A calibration test was performed in March 2006 in the back yard of 616 Woodbine Boulevard SW Calgary (Latitude (WGS84): 50° 56' 40"; Longitude (WGS84): 114° 08' 04", Elevation: 1,119 m). The expected horizontal intensity of the magnetic field at this location, based on the 2005 International Geomagnetic Reference Field model, is 16.13839 μT[28]. A wooden jig was constructed that allowed five units to be rotated at the same time through 360 degrees in 10 degree steps. The holes in the jig that determined the 10 degree steps were drilled using a turntable with a smallest micrometer reading of 15'. The jig was attached to a wooden fence post in the rear of the property (see Figure 22), in a location as far as possible (> 25 m) from visible magnetic anomalies such as electrical transformers, and levelled using a spirit level. For each position, average values where obtained from one minute of observation at 1 Hz. Prior to calibration, raw magnetometer observations were first corrected for non-

28 Magnetic field calculated using the U.S. Geological Surveys on-line geomagnetic field calculator, GeoMag. The calculator can be accessed from http://geomag.usgs.gov/models/models/.

orthogonality[29] of the axes, and then mathematically rotated so that the axes of the magnetometers were aligned with the axes of the accelerometer. Finally roll and tilt estimates from the accelerometers were used to rotate the magnetometer observations so that they were in terms of the local level frame.

Both Gain (33.1 ± 0.4 counts/μT) and Resolution (0.0302 ± 0.0003 μT) were within their expected ranges at the 95% confidence level. While the X Bias is somewhat higher than expected, given the limited information in the sensor specification documentation, we are unable to draw any conclusions regarding

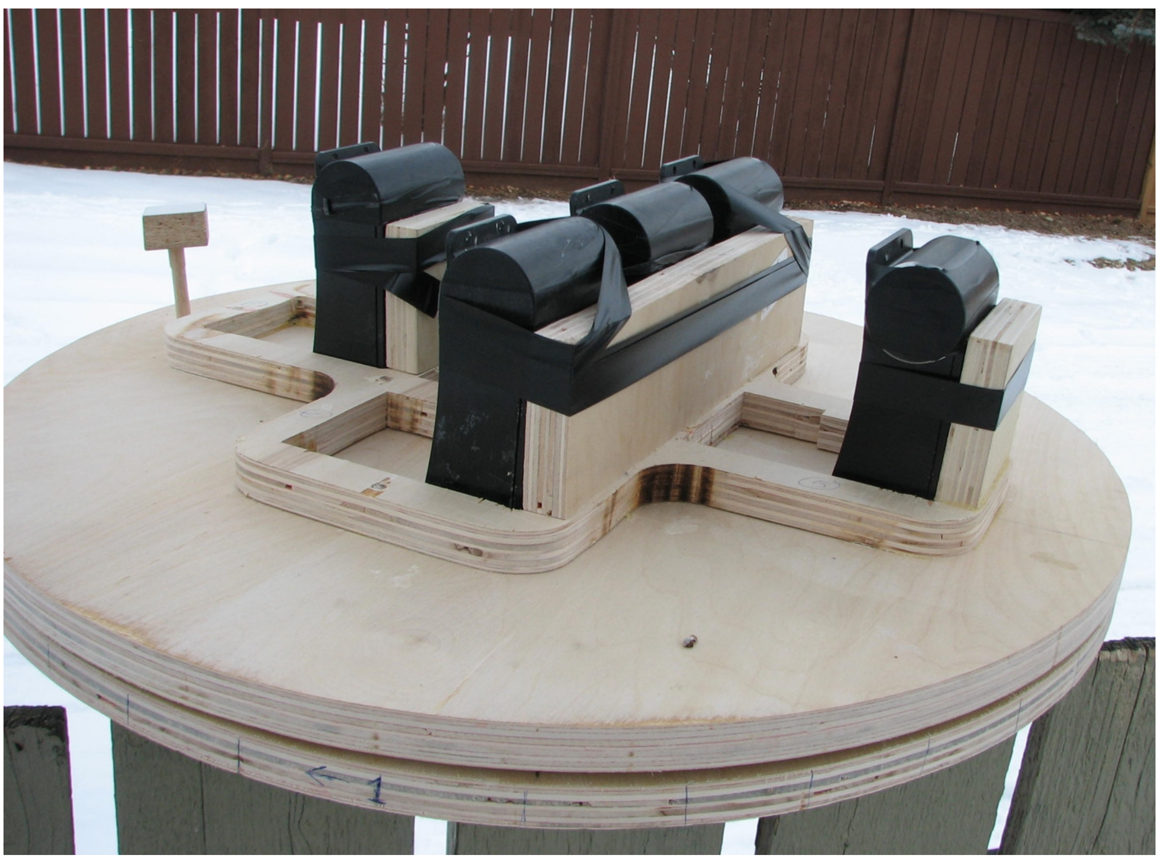


Figure 22: Calibration of magnetometers

29 See the next section describing calibration of the sensors for non-orthogonality.

the significance of the observed difference except to say that it is close to the maximum expected.

In terms of raw magnetometer output, 16 bit data stream, bias and scale factor estimates for the Green unit are set out in Table 7 below. Figure 24 maps out the raw data and adjusted data used to derive the information in Table 7.

Table 7: Bias and scale factor for PNI ASIC 3-axis magneto-inductive sensor

	Bias		**Scale Factor**	
Unit	X_b	Y_b	X_{SF}	Y_{SF}
Green	325.2	-176.0	1.000000	1.186947

Given the limited accuracy of determining magnetic north at the calibration site (good quality orienteering compass), the calibration data suggests that the orientation of the NavAid was biased by 2° 00' West, with a dispersion[30] (1σ) of 1° 33'. Assuming a normal error distribution, we can state that, at a confidence level of 95%, there is no significant difference between the magnetic direction observed by the orienteering compass used to align the jig with magnetic north, and the magnetic heading measured by the NavAid.

Orientation of Sensor Axes

Axis misalignment results from the imperfect manufacture and mounting of sensors. The effect of these imperfections is non-orthogonality of the axes defining the sensor frame. As a result, each axis is affected by the other two axes

30 See R.J. Yamartino, "A Comparison of Several "Single-Pass" Estimators of the Standard Deviation of Wind Direction," *Journal of Climate and Applied Meteorology* 23, no. 1362 - 1366 (1984) for estimation procedure.

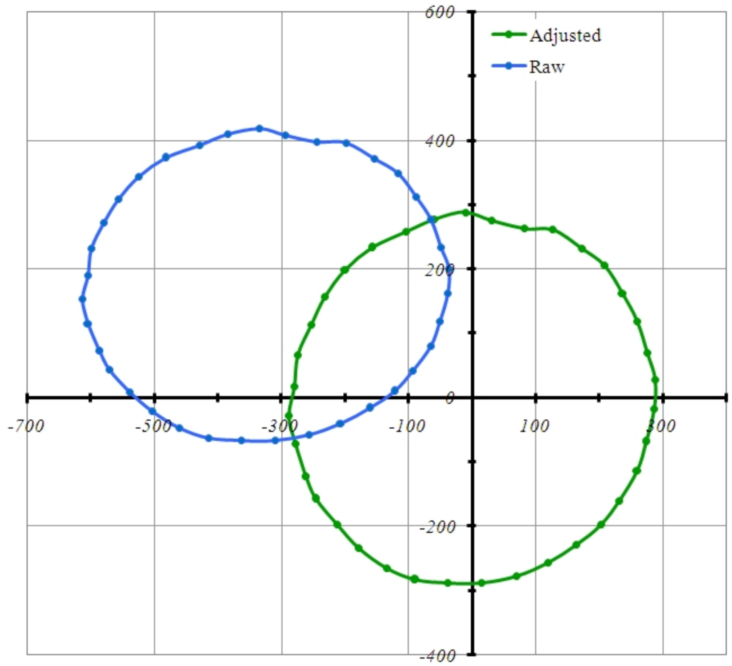


Figure 23: Raw and adjusted magnetometer data

in the sensor frame[31]. In general, axes misalignment can be estimated through calibration and then accounted for prior to making use of a sensor's output.

For this application, we are not concerned with any misalignment of the accelerometer axes, as correcting for such error has minimal benefit when implementing peak detection for distance estimation. However, we do need to know the relative orientation of the magnetometer axes with respect to the accelerometer axes. It has been assumed that the accelerometer axes are parallel to the exterior surface of the accelerometer chip packaging (see Figure 24). You will also note from Figure 24 that the accelerometer axes form a left hand coordinate system. As the general mapping frame is a right hand coordinate

31 Naser El-Sheimy, "Inertial Techniques and INS/GPS Integration," (Calgary: The Department of Geomatics Engineering, The University of Calgary, 2003)

system, the Z axis of the accelerometer must first be inverted,

$Z_{RH}^{g} = g + \left(g - Z_{LH}^{g}\right)$, g is gravity.

Rotations for non-orthogonality, and misalignment of the magnetometer axes with the accelerometer chip were determined using the Mitutoyo Bright STRATO-7106 Coordinate Measuring Machine[32] in the Department of Mechanical Engineering, University of Calgary. While the accuracy of the equipment is quite high, the components that we are measuring are very small, and quite awkward to measure. The planes of two surfaces (defined by four points), at right angles to each other, of each sensor were measured to determine the long axis of the sensors, i.e., the intersection of each pair of planes. The deviation[33] of the observed points to their respective planed range from 0.0014 mm to 0.0935 mm $\left(\bar{x} = 0.0226 \text{ mm}, s = 0.03568 \text{ mm}, \text{n=6}\right)$. While the smaller deviances are satisfactory, the larger deviances will likely have an affect on the accuracy of the magnetometer alignment, given that a magnetometer sensor is 6.25 mm x 2.25 mm x 2.25 mm. This suggests that the orientation of each plane could be within ±01°00' at 1σ, of the estimated plane. It was also found that if the magnetometer was assumed to be aligned with the accelerometer, the bias in the Y axis remained at a similar level, but the bias in X axis was reduced by 40 counts to -295, and the X and Y scale factors reversed!

---

32 The Bright SRATO has a specified accuracy of 1µm under at 18°C to 22°C. See product documentation at Mitutoyo America Corporation, http://www.mitutoyo.com/home.aspx

33 For the Bright SRATO deviation is defined as the maximum distance from the plane from all points used to define the plane.

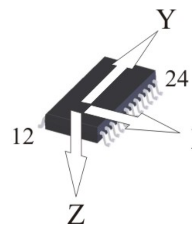


Figure 24: Direction of accelerometer axes

Following Shin and El-Sheimy (2002), and as depicted in Figure 25, the X axis was held fixed and the Y axis rotated in the XY plane so that the axes were at 90°, and then rotate the Z axis in both the XZ and YZ planes so that it was orthogonal to both. Hence, non-orthogonality of the magnetometer axes can be corrected using[34]

$$\begin{bmatrix} 1 & 0 & 0 \\ -\sin\theta_{YZ} & \cos\theta_{YZ} & 0 \\ -\sin\theta_{ZY} & \cos\theta_{ZY}\sin\theta_{ZX} & \cos\theta_{ZY}\cos\theta_{ZX} \end{bmatrix}. \tag{3.1}$$

Following Mikhail et al. (2001), a standard photogrammetric approach to rotating the magnetometer sensor axes to match the accelerometer axes was adopted using three sequential rotations[35]: ω about the X axis; ϕ about the once rotated Y axis; and κ about the twice rotated Z axis. This results in a single rotation matrix equal to

$$\begin{bmatrix} \cos\phi\cos\kappa & \cos\omega\sin\kappa+\sin\omega\sin\phi\cos\kappa & \sin\omega\sin\kappa-\cos\omega\sin\phi\cos\kappa \\ -\cos\phi\sin\kappa & \cos\omega\sin\kappa-\sin\omega\sin\phi\cos\kappa & \sin\omega\cos\kappa-\cos\omega\sin\phi\sin\kappa \\ \sin\phi & -\sin\omega\cos\phi & \cos\omega\cos\phi \end{bmatrix} \tag{3.2}$$

Table 9 summarizes the rotation parameters for one case.

---

34 Shin and El-Sheimy, "A New Calibration Method for Strapdown Inertial Navigation Systems,"

35 Edward M. Mikhail, James S. Bethel, and J. Chris McGlone, *Introduction to Modern Photogrammetry* (New York: John Wiley and Sons Inc., 2001)

Table 8: Non-orthogonality parameters for magnetometer

Unit	θ_{YZ}	θ_{ZY}	θ_{ZX}
Green	$-0°19'$	$2°26'$	$0°32'$

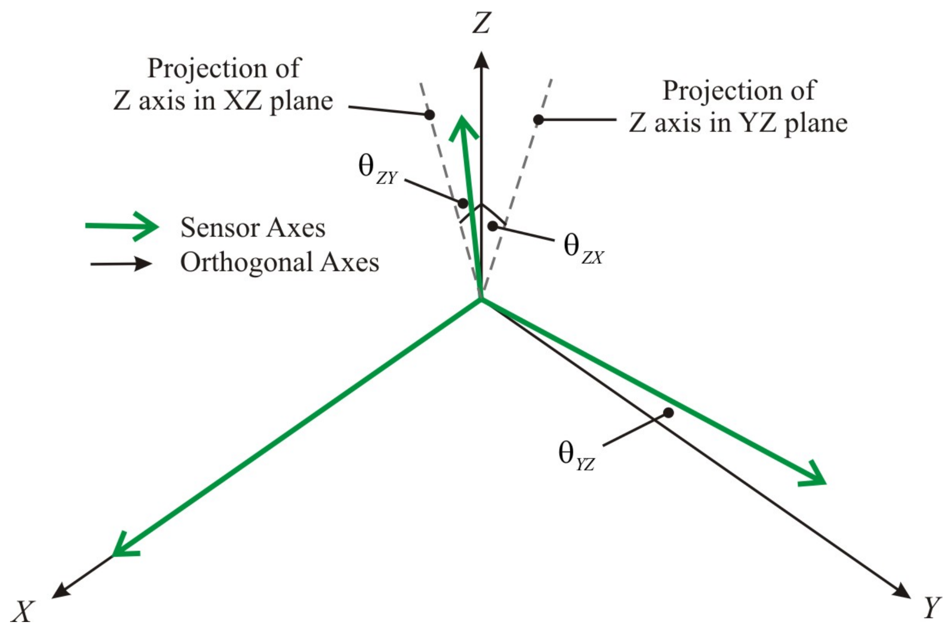


Figure 25: Non-orthogonality of sensor axes

Table 9: Rotation parameters for magnetometer

Unit	ω	ϕ	κ
Green	$-0°29'$	$-0°25'$	$-1°46'$

Concluding Remarks

As alluded to, this phase of the research has been a long, frustrating, and a sometimes disheartening process. However, the tools that we have to work with at the moment paint a broad picture that produces a broad answer. To address this, we have developed a tool that can provide the type of data that preliminary investigations undertaken in the next chapter suggest can answer animal behaviour questions more clearly. It provides the data that is expected of it, and it has survived all the abuse that we can muster. Hence this component of the research has directly addressed the primary objective of this thesis, which is the development of a technology solution for the tracking of animals. Admittedly, whether it can survive a grizzly bear for an extended period is still open to debate, but it does look promising.

Chapter 4

Preliminary Analysis of Animal Movement and Selection Preference

Introduction

This chapter summarizes preliminary findings derived from tracking data for G098, a ten-year-old male grizzly bear, over a portion of the summer of 2005 (66 days). Analysis of the animal's movement rate indicates that a velocity of 6.5 m/minute [LB: 5.5 m/min.; UB: 7.7 m/min.] [1] is the threshold that can be used to separate foraging from locomotion. Characterization of forage and locomotion sites points towards G098 having a preference for different environmental and land cover characteristics depending on its current movement state.

Study Area

G098 was captured southwest of the Ya-Ha-Tinder Ranch in the Municipality of Big Horn, Alberta. Following release, the first GPS position was acquired on June

[1] 95% Confidence Interval

16 at 6 p.m. with the final position being acquired on August 21 at 5 a.m. approximately 7 km north and 22 km east of the capture site. Following installation of the collar G098 headed in a SSE direction for 10 days (~31 km) and then in a northeasterly direction for approximately 4 weeks (~37 km) until he reached a cut block area, where he stayed until the collar failed.

G098 occupied an area that includes lower and upper boreal cordilleran and Subalpine ecoregions, which are typically characterized by closed forests of lodgepole pine (*Pinus contorta*), engelmann spruce (*Picea engelmannii*), trembling aspen (*Populus tremuloides*), and balsam poplar (*Populus balsamifera*)[2]. A summary of the land cover classification[3] based on a 95% home range polygon indicates that G098's home range consisted primarily of closed coniferous forest (43%), followed by mixed forest (23%) and shrubs (9%). The remaining land cover types were classified as open coniferous forest, broad leaf forest, forest regeneration, upland herbaceous, open wetland, treed wetland and barren land. The eastern limit of the animal's home range tends to coincide with the transition from lower boreal forest to semi developed agricultural land. Land use within G098's home range also includes forestry, oil and gas exploration and development, hunting, trapping, and all-terrain vehicle use, i.e., use covers the full range of human activities and disturbances.

---

2 W. L. Strong, *Ecoregions and Ecodistricts of Alberta* (Edmonton: Alberta Forestry, Lands and Wildlife, 1992)

3 Derived from the Foothills Model Forest IDT Land Cover data set, see S. E. Franklin et al., "An Integrated Decision Tree Approach (IDTA) - Classification of Landcover in support of Grizzly Bear Habitat Analysis in the Alberta Yellowhead Ecosystem," *Canadian Journal of Remote Sensing* 27, no. 6 (2001).

The home range was estimated using a Biweight Kernel[4] on a 300 m grid using the volume criteria. Least Squares Cross Validation was used to determine the optimum bandwidth (h = 6.8 km). Following Silverman (1986), the data was also standardized prior to analysis using Unit Standardization in an attempt to account for variation in distribution of northings and eastings that is attributed to physical terrain barriers[5].

G098 Tracking Data

G098 was tracked using a Televilt GPS-Satlink[6] collar set to acquire hourly positions, giving a total of 1,596 expected positions. However, only 74.1% (1,182) of these positions were obtained[7]. A comparison of the GPS Dilution of Precision (DOP) data for G098 with that obtained over a similar period, and in a similar location, for G096 — a five year old female — (see Table 10) indicates that the quality of positions for G098 is significantly worse than that obtained for G096[8] (t Statistic = 5.22, $p < 0.000$, $v = 1,442$, $\alpha = 0.05$). However, it should be noted that the fix rate for G098 was substantially higher than G096 (~32.6%). This may be partially explained by the considerably higher maximum DOP observed on G098. This indicates that because the DOP mask for G098 was higher than the DOP mask

4 B. W. Silverman, *Density Estimation for Statistics and Data Analysis* (London: Chapman & Hall, 1986)

5 Ibid.

6 TVP Positioning AB, Bandygatan 2, SE-71134 Lindesberg, Sweden, http://www.positioning.televilt.se

7 This fix rate is comparable to a study of GPS fix rates in BC, Alberta, and the Northwest Territories by Gau et al., "Uncontrolled Field Performance of Televilt GPS-Simplex™ Collars on Grizzly Bears in Western and Northern Canada,"

8 G096 was tracked using a Televilt Simplex collar. This is the standard GPS/VHF collar that has been used by the Foothills Model Forest Grizzly Bear Program. G098 is the only animal to be tracked using a Televilt GPS Satlink collar to date; hence comparison of similar collars is not possible.

for G096, under similar conditions, G098 would be expected to obtain more fixes than G096.

Table 10: GPS DOP statistics

	G098	G096
Average DOP	4.3	3.2
Standard Deviation DOP	3.3	1.8
Maximum DOP	21	11
Minimum DOP	1	1

Geometrically, DOP is related to the volume formed by the intersection points of the user-satellite vectors, with the unit sphere centered on the user[9]. Larger volumes give smaller DOPs. Lower DOP values generally represent better position accuracy. But, a lower DOP value does not automatically mean a low position error. The quality of a GPS-derived position estimate depends upon both the measurement geometry as represented by DOP values, and errors caused by signal strength, ionosphere effects, multipath, etc. However this requires that you know the magnitude of a number of error sources, which is unlikely with the grizzly bear data.

In short, the Estimated Position Error (EPE) of a GPS fix is[10]:

$$\mathrm{EPE}_{\sigma} = \mathrm{DOP} \times \mathrm{UERE}_{\sigma} \tag{4.1}$$

9 Jr. J. J. Spiker, "GPS Signal Structure and Performance Characteristics," in *Global Positioning System*, ed. P. M. Janiczek (Alexandria, VA: The Institute of Navigation, 1980)

10 E. H. Martin, "GPS User Equipment Error Models," in *Global Positioning System*, ed. P. M. Janiczek (Alexandria, VA: The Institute of Navigation, 1980)

User Equivalent Range Errors (UERE) can be grouped into general classes as follows[11]:

1. Ephemeris data – uncertainty in the location of the satellite;
2. Satellite clock – bias or uncertainty in the transmitted clock;
3. Ionosphere - uncertainty due to ionospheric signal delay effects;
4. Troposphere - uncertainty due to tropospheric delay effects;
5. Multipath - errors caused by multipath disturbances at the antenna; and
6. Receiver – clock uncertainties, reference station uncertainties, etc.

In situations where there is limited information regarding the extent of these errors for a particular observation session, then a "rule of thumb" is to multiply the observed DOP by the standard deviation of the GPS reciever[12].

Tests performed on five Televilt GPS Simplex[13] collars used by the Foothills Model Forest indicate that the standard deviation ranges between 6.8 m and 12.5 m[14] giving a pooled standard deviation of 8.0 m. Therefore, given the DOP values listed in Table 10, if we assume a 95% confidence level, the estimated horizontal error for G098 ranges from ±17.7 m to ±329.3 m, with an average estimated error of ±55.0 m.

---

11 B. Hofmann-Wellenhof, H. Lichteneggar, and J. Collins, *Global Positioning System: Theory and Practice* (New York: Springer-Verlag Wien, 1997), Martin, "GPS User Equipment Error Models," , and J. J. Spiker, "GPS Signal Structure and Performance Characteristics,"

12 C. Johnson and P.Ward, "GPS Application to Seismic Oil Exploration," in *Global Positioning System*, ed. P. M. Janiczek (Alexandria, VA: The Institute of Navigation, 1980)

13 TVP Positioning AB, Bandygatan 2, SE-71134 Lindesberg, Sweden, http://www.positioning.televilt.se. The Simplex collars are now obsolete, and are no longer being produced.

14 Two sets of tests were carried out between July 21 and August 15 of 2005 (2 units), and February 23 and March 10, 2006 (3 units) using the southernmost pillars on the roof of Engineering's F Block at the University of Calgary. GPS fixes were obtained at five-minute intervals 24 hours a day for the duration of the tests. The individual number of observations varied due to different start and end times and battery failure. Individual results were as follows: collar 151.965, tested in 2005 – σ_{xy}=8.4 m (n=5,811); collar 151.375, tested in 2005 – σ_{xy}=12.7 m (n=1,025); collar 150.621, tested in 2006 – σ_{xy}=7.0 m (n=3,611); collar 150.235, tested in 2006 – σ_{xy}=7.4 m (n=3,611); collar 151.593, tested in 2006 – σ_{xy}=6.8 m (n=3,607).

Movement Analysis from GPS Data

Following Foreman and Gordon (1986), Taylor et al. (1993) and Tischendorf and Fahrig (2000b), we can expect that forage patches are connected by movement corridors[15]. By being able to partition grizzly bear movement into locomotion (corridor movement) and specialized search movement (patch movement) it is anticipated that we can account for more of the variation in the measured model currently used for grizzly bear resource selection functions. Silby et al. (1990) and Johnson et al. (2002) have proposed non-linear curve fitting models to differentiate between feeding bouts[16], or locomotion movement and specialized searching movement[17]. This technique fits a curve to the log transformed frequency distribution of an animal's movement velocity or feeding frequency. Major inflections along the curve provide a means of differentiating between the different types of processes, or movements.

If we assume a single process model, then behaviour is not split into bouts, and all movement is generated by a single Poisson process. In this instance, the average movement rate is equal to λ. For a Poisson process, the chance of observing a particular movement rate, r, is

$$P(r \mid \lambda) = \frac{e^{-\lambda}\lambda^{r}}{r!}. \tag{4.2}$$

---

15 R. T. T. Forman and M. Gordon, *Landscape Ecology* (New York: John Wiley & Sons, Ltd., 1986); P. D. Taylor et al., "Connectivity is a Vital Element of Landscape Structure," *Oikos* 68, no. 3 (1993); Tischendorf and Fahrig, "How Should we Measure Landscape Connectivity?"

16 R. M. Silby, H. M. R. Nott, and D. J. Fletcher, "Splitting Behaviour into Bouts," *Animal Behaviour* 39 (1990)

17 C. J. Johnson et al., "Movement Parameters of ungulates and Scale-specific Responses to the Environment," *Journal of Animal Ecology* 71 (2002)

If N_r is the total number of occurrences that occur before rate r is observed then

$$P\left(N_r = k\right) = \frac{e^{-\lambda r}\left(\lambda r\right)^k}{k!}, \tag{4.3}$$

hence the chance of obtaining no events at rate r is $e^{-\lambda r}$, and the probability of a movement rate being between r and $r+1$ is $\lambda e^{-\lambda r}$. If there are N events in total, then the expectation is that the number of movement rates between r and $r+1$ is $N\lambda e^{-\lambda r}$ [18].

In order to fit a two-process model to the data, it required the use of nonlinear curve fitting techniques. In this instance the Nonlinear Regression (NLR) procedure in SPSS 15.0 was used. As the data to be fitted are log frequencies, the appropriate form of the model used to describe motion behaviour is[19]:

$$y = \log_e(N_f \lambda_f e^{-\lambda_f r} + N_s \lambda_s e^{-\lambda_s r}) \tag{4.4}$$

where N is the total number of counts for each type of movement, λ is now the probability that an event occurs in the next movement rate interval, and r is the movement rate. f represents foraging bouts, s represents searching, or locomotion between foraging sites, and y is the expected number of movements that occur during each discrete interval of the movement rates.

As outlined above, this model of locomotion is generated by two Poisson processes, or behaviours: a less frequently occurring process representing a movement bout (large scale movement), and a more frequently occurring process

---

18 R. A. Johnson, *Miller & Freund's Probability and Statistics for Engineer*, 6th ed. (New Jersey: Prentice-Hall, 2000)

19 Silby, Nott, and Fletcher, "Splitting Behaviour into Bouts,"

representing searching or feeding bouts (small scale movement). Movement rate, velocity, serves as a measure by which frequent and less frequent behaviour can be identified. In this instance data describing the frequency of rates of movement has been generated from successive GPS collar position fixes of the animal.

For G098, 1,181 movements were observed by GPS between June 16 and August 21, 2005. Movement rates ranged from 0 m per minute to 59 m per minute, with an average movement rate of 5.2±0.5 m[20] per minute. Prior to the data collection period severe flooding occurred in this area, however, during the data collection period it was extremely dry[21].

The raw data is a frequency distribution (histogram) of movement rates. Typically, slow movements are more frequent, and faster movements are progressively less frequent (see Figure 26). Because of the sampling regime used in the field some higher velocities will probably not occur at all as averaging velocity over an hour will tend to hide high movement rates of an animal. However, zeros could not be used in the analysis because it was carried out on log-transformed data. In order to address this issue a constant value of two (2) was added to all counts. An alternative way to circumvent this issue is to use wider movement rate intervals. For this analysis the first option was used. The logarithm of the frequencies was then taken, to equalize the variances at different movement rates.

---

20 95% Confidence Interval

21 Dr. Caterina Valeo, August 28, 2007, *personal communication*

Log~ (frequency) (hereafter called y) is plotted against velocity (metres/minute) in Figure 26.

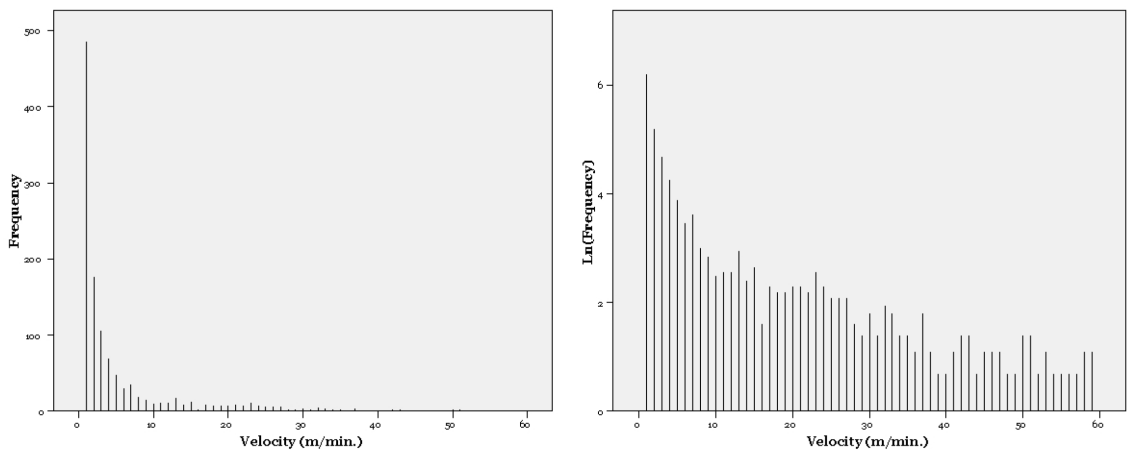


Figure 26: (a) Frequency, (b) and $\log_e$ frequency versus velocity (m/min.) for grizzly bear G098 collected over 6 week period

When a two-process dataset is plotted in this way they roughly fit a broken-stick model, i.e., two joined straight lines. Each line has a negative slope. As explained above, the higher velocities belong mostly to the less frequently occurring process, searching, and the slow velocities belong mostly to the more frequently occurring process, foraging, provided $N_f\lambda_f >> N_s\lambda_s$.

Hence the low velocities generated by $z = N_f\lambda_f e^{-\lambda_f r}$, can be fit by $y = \log_e(z) = \log_e(N_f\lambda_f - \lambda_f r)$, and similarly the higher velocities fit to $y = \log_e(N_s\lambda_s - \lambda_s r)$. The slope of the steep line is $-\lambda_f$, and of the shallow line $-\lambda_s$. The vertical axis intercept of the steep line is approximately $\log_e(N_f\lambda_f)$ and the shallow line $\log_e(N_s\lambda_s)$, so N_s and N_f can be estimated from a knowledge of the y-intercept and the slope of each line. In each case, $N = (1/\lambda)\exp(\text{y intercept})$.

It is clear from Figure 26 that the $\log_e$ frequency distribution is not a single straight line, as would be expected of a single-process model. To see whether the data might fit a two-process model, the model was fit as follows. The left-hand 10 points of Figure 26b were replotted in Figure 27a. The points appear to fall on a straight line with equation $y = 5.979 - 0.367r$ (from linear regression, $F_{1,8} = 128.06,\ p = 0.000$). Hence, $\lambda_f = 0.367$, and $N_f = (1/0.367)e^{5.978} = 1,075$. The right-hand 49 points of Figure 26b are replotted in Figure 27b. These points appear to follow a straight line with equation $y = 2.982 - 0.041r$ ($F_{1,47} = 161.91,\ p = 0.000$). Hence, $\lambda_s = 0.041$, and $N_s = (1/0.041)e^{2.982} = 481$.

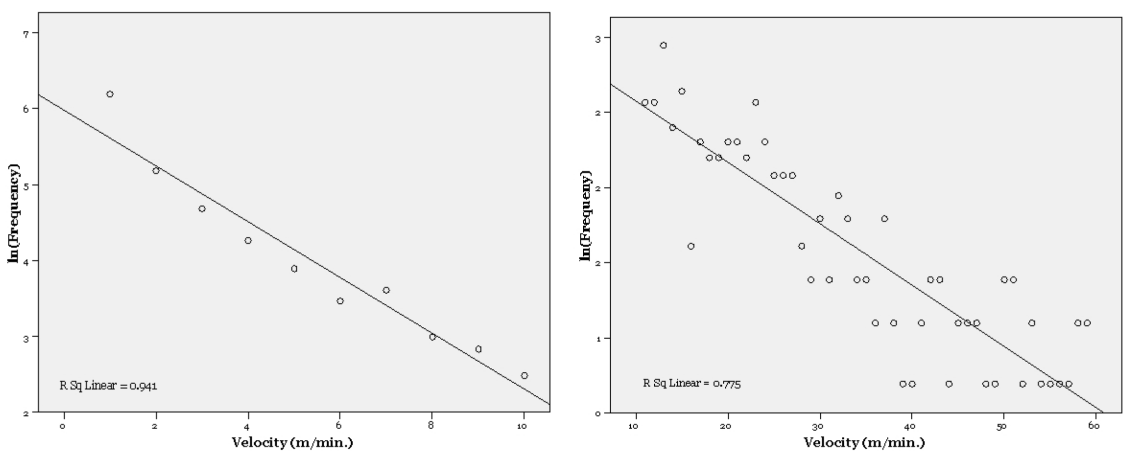


Figure 27: The data of Figure 27 is split into two for further analysis. (a) The left-hand points plotted are for velocities less than 10 m/min. (b) The right-hand points are for all velocities greater than 10 m/min.

It should be noted that if movement behaviour is generated by two processes as described above, it is inevitable that some events will be assigned to the wrong process. Our objective is therefore to minimize misassignment.

Following Slater and Lester (1982), we can minimize the total number of misassigned movements by rearranging (4.4) to obtain[22]

$$r_c = \frac{1}{\lambda_f - \lambda_s} \ln \frac{N_f \lambda_f}{N_s \lambda_s}. \qquad (4.5)$$

Slater and Lester (1982) also showed that the expected number of movement bouts misassigned when the bout criterion is r_c is

$$N_f e^{-\lambda_f r_c} + N_s \left(1 - e^{-\lambda_s r_c}\right) \qquad (4.6)$$

The output from the NLR procedure (in SPSS Release 15.0[23]) gives the parameter estimates, the parameter correlation matrix, and the total sum of squares associated with the four parameter model of equation (4.4) (see Table 11 and Figure 28).

The model accounts for 98% $\left(R^2 = 1 - \frac{SSE}{SST_c}\right)$ of the variation in the data. The F-statistic for the two-process model is 1,097.8, with $\mathrm{df}_1 = 3$[24], $\mathrm{df}_2 = 55$, and $P = 0.000$. The critical value for this test, given a confidence interval of 95%, is 2.773, clearly the model is worth fitting to the data of Figure 26.

All model coefficients are significant, with $N_f = 1{,}071.319$ $[\mathrm{LB}: 666.033; \mathrm{UB}: 1{,}476.605]$, $\lambda_f = 0.579$ $[\mathrm{LB}: 0.420; \mathrm{UB}: 0.738]$, $N_s = 478.598$

---

22 P. J. B. Slater and N. P. Lester, "Minimising Errors in Splitting Behaviour into Bouts," *Behaviour* 79 (1982)

23 SPSS Inc., 233 S. Wacker Drive, 11th floor, Chicago, IL 60606-6307, USA, http://www.spss.com

24 SPSS reports the total sum of squares for the four parameter model without subtracting the sum of squares accounted for by the mean of the data points. It is more conventional to subtract the sum of squares accounted for by the mean, and attribute one less degree of freedom for the sum of squares attributed to the model.

Table 11: Output from the non-linear curve fitting procedure NLR

Parameter	Estimate	Std. Error	95% Conf. Interval	
			Lower Bound	*Upper Bound*
N_f	1,071.319	202.234	666.033	1,476.605
λ_f	0.579	0.079	0.420	0.738
N_s	478.598	26.032	426.429	530.767
λ_s	0.040	0.003	0.034	0.046

Asymptotic Correlation Matrix

	λ_f	N_s	λ_s
N_f	0.593	0.053	0.044
λ_f		0.367	0.355
N_s			0.599

ANOVA

Source	Sum of Squares	df	Mean Squares	F	P
Regression	302.995	3	100.998	1097.808	0.000
Residual	5.040	55	0.092		
Uncorrected Total	308.035	59			
Corrected Total	80.898	58			

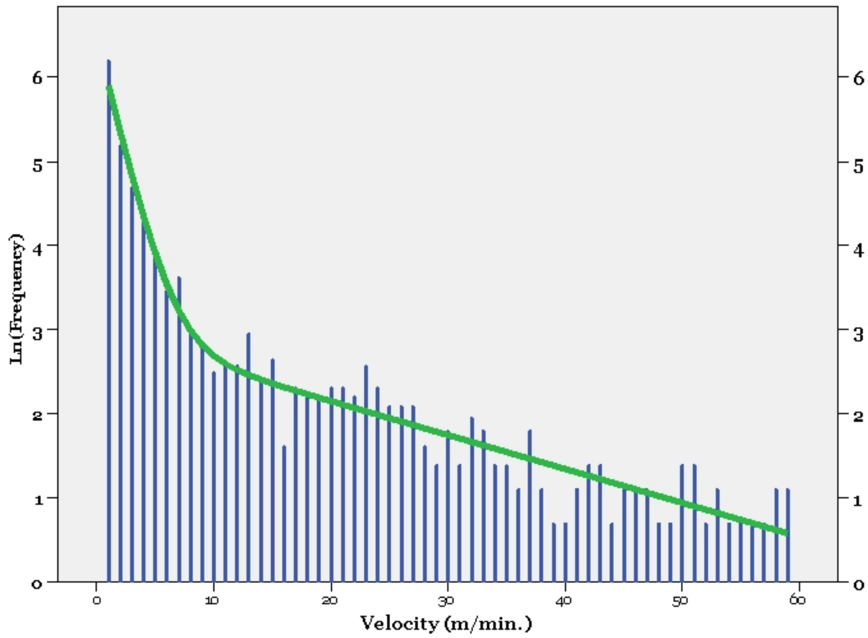


Figure 28: Log_e of frequency of G098's velocity including model

$[\text{LB}:426.429;\text{UB}:530.767]$, $\lambda_s = 0.040$ $[\text{LB}:0.034;\text{UB}:0.046]$ confirming that they all contribute significantly to this model. The correlation matrix (Table 11) shows moderate correlation between N_f and λ_f, 0.593, and N_s and λ_s, 0.599. While these correlations are not excessive, the correlations are approaching a limit that makes it difficult to attribute effects to one or the other, as a consequence their standard errors may be higher than they otherwise might have been.

The movement behaviour criterion, r_c, using equation (4.5) was 6.5 m/minute [95% CI: LB: 5.5 m/min.; UB: 7.7 m/min.] and the expected number of movement bouts misassigned (see equation (4.6)) was 134, or 11.4% of the total movement bouts.

An attempt to identify a three-process model to account for foraging, searching and locomotion behaviours was also undertaken by modifying equation (4.4) to obtain

$$y = \ln(N_f \lambda_f e^{-\lambda_f r} + N_s \lambda_s e^{-\lambda_s r} + N_l \lambda_l e^{-\lambda_l r}) \tag{4.7}$$

where f, s and l now represent foraging, searching and locomotion respectively. However, the model was unable to converge. It was assumed that this was because there were very few data points at the higher movement rates, often only one count. In addition, there were large gaps between the higher movement rates that essentially acted as missing data and therefore weakened the model's ability to be identified. This result indicates that the current sampling regime employed by the FMFGRP is inadequate if we wish to identify the range of movement behaviours suggested by Silby et al. (1990) and Johnson et al. (2002).

Yet the two behaviour model has allowed G098's GPS positions to be partitioned into two groups representing search, or forage, sites (movement rates < 6.5 m/minute; green points and white outlines in Figure 29) and locomotion sites (movement rates > 6.5 m/minute; orange points and red outlines in Figure 29). Figure 29 depicts this partitioning of GPS positions along with the home range associated with each type of movement. As described earlier, the 95% home range polygon was calculated using a biweight kernel[25].

---

25 Silverman, *Density Estimation for Statistics and Data Analysis*

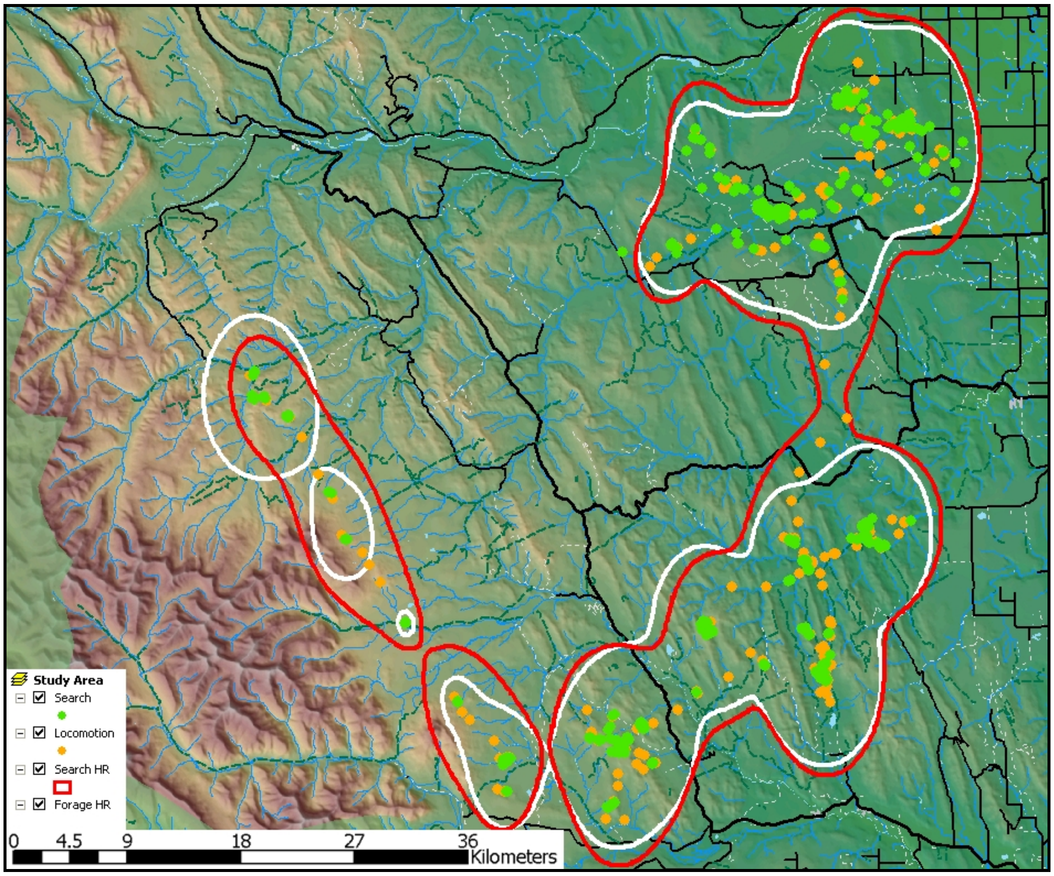


Figure 29: Searching/forage v. locomotion sites - G098, 2005

Species Presence Absence Data

Topographic and vegetation metrics from multi-spectral satellite imagery have been employed in previous species–environment models[26]. These metrics provided accurate surrogates for vegetation condition, structure, productivity and biomass[27]. Thus, in addition to IDT[28] Land Cover data, a number of additional

---

26 See Richard D. Mace et al., "Landscape Evaluation of Grizzly Bear Habitat in Western Montana," *Conservation Biology* 13, no. 2 (1999), and Carlos Carroll, Reed F. Noss, and Paul C. Paquet, "Carnivores as Focal Species for Conservation Planning in the Rocky Mountain Region," *Ecological Applications* 11, no. 4 (2001)

27 Eric. P. Crist and Richard C. Cicone, "A Physically Based Transformation of Thematic Mapper Data - The TM Tasseled Cap," *IEEE Transactions on Geoscience and Remote Sensing* GE-22, no. 3 (1984); John R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective* (Upper Saddle River, NJ.: Prentice Hall, 2005)

spatial data sets have been used in the development of these preliminary models. Leaf area index (LAI), crown closure and (Tassel Cap) greenness[29] were obtained from the Foothills Model Forest. An animal risk surface was developed following Nielson et al. (2004), which considers human access, water, land cover edge features, terrain ruggedness and greenness[30]. Various terrain indices such as the Terrain Roughness Index[31], Slope Aspect Index[32], Shannon Index[33] (provides a measure of slope and aspect diversity), Elevation Standard Deviation[34], etc. were developed along with distance to water, distance to roads and distance to edge of land cover type. The Slope Aspect Index gives maximum preference to southwest facing slopes and minimum preference to northeast facing slopes. In addition the index is scaled to take into account slope, with flatter slopes scaled towards one (1) and steeper slopes towards zero (0). Hence, low gradient slopes with a southwest aspect were assigned values near to one (1) and north east steep slopes were assigned values near to zero (0). Finally, numerous solar radiation indices were

---

28 Franklin et al., "An Integrated Decision Tree Approach (IDTA) - Classification of Landcover in support of Grizzly Bear Habitat Analysis in the Alberta Yellowhead Ecosystem,"

29 Greg J. McDermid, "Remote Sensing for Large-Area, Multi-Jurisdictional Habitat Mapping" (University of Waterloo, 2005)

30 Scott E. Nielsen, M. S. Boyce, and G. B. Stenhouse, "Grizzly Bears and Forestry I. Selection of Clearcuts by Grizzly Bears in West-central Alberta, Canada," *Forest Ecology and Management* 199, no. 1 (2004a)

31 Christian Nellemann and Raymond D. Cameron, "Effects of Petroleum Development on Terrain Preferences of Calving Caribou," *Arctic* 49, no. 1 (1996)

32 Scott E. Nielsen and Alan Haney, "Gradient Responses for Understory Species in a Bracken Grassland and Northern Dry Forest Ecosystem of Northeast Wisconsin," in *Transactions of the Wisconsin Academy of Sciences, Arts and Letters*, ed. William J. Urbrock (1998)

33 C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal,* 27 (1948)

34 John P. Wilson and John C. Gallant, *Terrain Analysis: Principles and Applications* (New York: John Wiley & Sons, Ltd., 2000) pg. 74

developed using SRAD[35]. As described in Wilson and Gallant (2000), the software estimates potential solar radiation as a function of latitude, slope, aspect, topographic shading, and time of year, and then modifies the estimate using information about monthly average cloudiness and sunshine hours[36] (local data was obtained from weather stations at Kananaskis, Banff, Lake Louise, Madden, Olds and Rocky Mountain House). Temperature is extrapolated across the surface and corrected for elevation using lapse rates, slope/aspect effects via shortwave radiation ratio, and vegetation effects via leaf area index. Daily outgoing long wave irradiance was calculated from surface temperature and daily incoming long wave irradiance was calculated from air temperature and the fraction of visible sky at each grid point. These short and long wave radiation fluxes can be used to estimate the surface energy budget at each grid point for a specified period, in this case from June 16 to August 21, 2005.

These radiation estimates were developed because the surface energy budget has been shown to have a large effect on vegetation diversity and biomass production[37]. Leaf area index[38], crown closure[39] and greenness[40] were included

---

35 See Ian D. Moore, "Terrain Analysis Programs for the Environmental Sciences (TAPES)," *Agricultural Systems and Information Technology* 4 (1992), and Wilson and Gallant, *Terrain Analysis: Principles and Applications*

36 Wilson and Gallant, *Terrain Analysis: Principles and Applications* pp. 91 -106

37 R. B. Hutchins et al., "The Influence of Soils and Microclimate on Vegetation of Forested Slopes in Eastern Kentucky," *Soil Science* 121, no. 4 (1976); Ian D. Moore, T. W. Norton, and Jann E. Williams, "Modelling Environmental Heterogeneity in Forested Landscapes," *Journal of Hydrology* 150, no. 2 - 4 (1993); S. E. Franklin et al., "Change Detection and Landscape Structure Mapping Using Remote Sensing," *The Forestry Chronicle* 78, no. 5 (2002)

38 Joseph D. White et al., "Measurement and Remote Sensing of LAI in Rocky Mountain Montane Ecosystems," *Canadian Journal of Forest Research* 27, no. 11 (1997)

39 Crist and Cicone, "A Physically Based Transformation of Thematic Mapper Data - The TM Tasseled Cap,"

for similar reason, i.e., variation in each of these leads to variation in vegetation diversity. Following various authors, the terrain indices[41], risk surface[42], and distances to water[43], roads[44], or forest edges[45] were assumed to be indicators of animal security and movement corridors[46].

Resource Selection Model Development

A logistic regression model of the form

$$w(x) = \exp\left(\beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p\right) \tag{4.8}$$

has been used to characterize the environment variables for both forage and locomotion sites. $w(x)$ represents the resource selection functions and β_i the

40 Michael L. Gibeau et al., "Grizzly Bear Response to Human Development and Activities in the Bow River Watershed, Alberta, Canada," *Biological Conservation* 103, no. 2 (2002), Richard D. Mace et al., "Relationships Among Grizzly Bears, Roads and Habitat in the Swan Mountains, Montana," *Journal of Applied Ecology* 33, no. 6 (1996); Mace et al., "Landscape Evaluation of Grizzly Bear Habitat in Western Montana," Nielsen et al., "Modeling Grizzly Bear Habitats in the Yellowhead Ecosystem of Alberta: Taking Autocorrelation Seriously,"

41 Javier Naves et al., "Endangered Species Constrained by Natural and Human Factors: the Case of Brown Bears in Northern Spain," *Conservation Biology* 17, no. 5 (2003)

42 Scott E. Nielsen et al., "Modelling the Spatial Distribution of Human Caused Grizzly Bear Mortalities in the Central Rockies Ecosystem of Canada," *Biological Conservation* 120, no. 1 (2004b)

43 Nielsen et al., "Modeling Grizzly Bear Habitats in the Yellowhead Ecosystem of Alberta: Taking Autocorrelation Seriously," Jeanette. C. Theberge, "Scale-dependent Selection of Resource Characteristics and Landscape Pattern for Female Grizzly Bear in the Eastern Slopes of the Canadian Rocky Mountains" (University of Calgary, 2002)

44 B. N. McLellan and D. M. Shackleton, "Grizzly Bears and Resource-Extraction Industries: Effects of Roads on Behaviour, Habitat Use and Demography," *Journal of Applied Ecology* 25, no. 2 (1988); Carrie L. Roever, "Grizzly Bear (Ursus Arctos L.) Selection of Roaded Habitats in a Multi-Use Landscape " (University of Alberta, 2007)

45 Nielsen et al., "Modeling Grizzly Bear Habitats in the Yellowhead Ecosystem of Alberta: Taking Autocorrelation Seriously," Theberge, "Scale-dependent Selection of Resource Characteristics and Landscape Pattern for Female Grizzly Bear in the Eastern Slopes of the Canadian Rocky Mountains"

46 Zollner and Lima, "Search Strategies for Landscape-level Interpatch Movements," ; Sophie Charrier, Sandrine Petit, and Francoise Burel, "Movements of Abax parallelepipedus (Coleoptera, Carabidae) in woody habitats of a hedgerow network landscape: a radio-tracing study," *Agriculture, Ecosystems & Environment* 61, no. 2-3 (1997)

selection coefficients estimated from the environmental predictors x_i[47]. Equation (4.8) enables an estimating function to be defined as,

$$\pi(x) = \frac{\exp\left(\beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p\right)}{1 + \exp\left(\beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p\right)} \tag{4.9}$$

This structure was implemented following the resource selection function literature[48], as pseudo-absences were sampled from GIS data and not directly measured.

The extent of the 95% home range polygon (described above) constrained the analytical frame to reflect the areas utilized by G098, and was used to select random pseudo-absence sites. The non-use sites (0: n_{Forage} = 2,394; $n_{\text{Locomotion}}$ = 1,152) characterized the environmental variables within the forage and locomotion study sites to provide a background against which to test the importance of the observed use-sites to G098[49]. The non-use sample sizes should, in theory, reduce the chance of spurious model results by ensuring an adequate representation of available environmental variables[50].

Given the large number of variables, a set of covariates for an initial candidate model was selected based on large t statistics $(P < 0.25)$ obtained from univariate

47 B. F. J. Manly et al., *Resource Selection by Animals: Statistical Design and Analysis for Field Studies*, 2nd ed. (Dordrecht, the Netherlands: Kluwer Academic Press, 2002)

48 Ibid. and Mark S. Boyce et al., "Evaluating resource selection functions," *Ecological Modelling* 157, no. 2-3 (2002)

49 Boyce et al., "Evaluating resource selection functions,"

50 D. R. Stockwell and A. T. Peterson, "Bias in Biodiversity Data," in *Predicting Species Occurrences: Issues of Accuracy and Scale*, ed. J. M. Scott, et al. (Washington, USA: Island Press, 2002); Simon Ferrier et al., "Extended Statistical Approaches to Modelling Spatial Pattern in Biodiversity in Northeast New South Wales. I. Species-Level Modelling," *Biodiversity and Conservation* 11, no. 12 (2002)

logistic models[51]. The initial candidate model for forage sites comprised of a number of continuous covariates including maximum temperature (MT), August LAI (aLAI), slope aspect index (SAI), net radiation (NR), crown closure (CC), distance to roads (D2R), distance to water (D2W), digital elevation model (DEM), distance to edge (D2E), greenness (G) and risk (R). In addition, a categorical classification of the study area derived from object-oriented analysis of Landsat orthomosaic imagery representing land cover (LC) [52] was also included. The classification legend for this data was: closed coniferous forest [LC(CCF); open coniferous forest [LC(OCF)]; broadleaf forest [LC(BLF)]; mixed forest [LC(MF)]; forest regeneration [LC(FR)]; upland herbaceous [LC(UH)]; shrubs [LC(S)], open wetland [LC(OW)]; treed wetland [LC(TW)]; water [LC(W)]; barren land [LC(BL)]; and shadow [LC(SDW)].

The initial candidate model for locomotion sites included maximum temperature, August LAI, elevation variation (EV), outgoing radiation (OR), crown closure, distance to roads, distance to water, DEM, Distance to edge, greenness, risk and land cover.

Logistic regression, using S-Plus 7.1[53], was undertaken for both forage and locomotion models. The candidate models were tested for multi-collinearity using

---

51 David W. Hosmer and Stanley Lemeshow, *Applied Logistic Regression*, 2 ed. (John Wiley & Sons, Ltd., 2000) (p. 95)

52 Franklin et al., "An Integrated Decision Tree Approach (IDTA) - Classification of Landcover in support of Grizzly Bear Habitat Analysis in the Alberta Yellowhead Ecosystem,"

53 Insightful Corporation, Global Headquarters, 1700 Westlake Avenue North, Suite 500, Seattle, WA 98109-3044, USA

Pearson's r [54]. When two variables were highly correlated, $|r| \geq 0.7$, the variable with the lower univariate model P-value was maintained for further analysis. Forward stepwise logistic regression was then used to identify and remove non-significant variables.

Results

Multivariate analysis of the Foraging model indicates that overall the foraging model was significant $\left(\chi^2 = 447.68, \text{df=16}, P = 0.000\right)$, and produces a c index of 0.736 (area under receiver operating curve (ROC)). The area under the ROC curve provides a measure of discrimination, which is the likelihood that a presence site will have a higher probability $\left(P\left(y=1\right)\right)$ than a pseudo absence site[55]. According to Hosmer and Lemeshow (2000), a c index between 0.7 and 0.8 is considered acceptable discrimination, while 0.8 to 0.9 is considered excellent[56].

The model indicated that foraging was positively associated with water (D2W), edge features (D2E), August leaf area index (aLAI), and crown closure (CC) (i.e., negative coefficients for distance to feature), while negatively associated with the slope aspect (SAI) and net radiation (NR) indices (Table 12: bold covariates are significant). That is, G098's preference was for cooler northeast facing slopes that are near to water and have a more complete canopy. Of the Land Cover types, mixed forest (MF), regenerating forest (RF), upland herbaceous (UH) and barren land (BL) played a positive role in foraging selection.

---

54 Hosmer and Lemeshow, *Applied Logistic Regression*

55 Ibid. (p. 162)

56 Ibid.

Analysis of the locomotion model indicates that overall the model was also significant $\left(\chi^2 = 236.17, \text{df=14}, P = 0.000\right)$, and produced a c index of 0.751, which is also acceptable discrimination. Analysis of the model for Locomotion sites indicates that G098's resource selection/preference differed depending on whether

Table 12: Table of coefficients for forage covariates

Multivariate Model - Forage: Presence = 798, Pseudo-Presence = 2,394

Covriate	β_i	*S.E.*	*Wald Z*	*OR*	*95% CI*	*G*	*p*
Intercept	3.62	0.82	4.40	37.17	(7.43, 185.93)		0.000
NR	**-0.05**	**0.01**	**-9.27**	**0.95**	**(0.94, 0.96)**	**170.18**	**0.000**
aLAI	**0.384**	**0.04**	**9.20**	**1.47**	**(1.35, 1.59)**	**120.20**	**0.000**
LC(OCF)	-0.28	0.30	-0.96	0.77	(0.43, 1.33)		0.335
LC(BLF)	0.36	0.20	1.84	1.43	(0.98, 2.11)		0.065
LC(MF)	**0.72**	**0.11**	**6.38**	**2.06**	**(1.65, 2.57)**		**0.000**
LC(FR)	**0.53**	**0.27**	**1.96**	**1.69**	**(1.00, 2.87)**		**0.050**
LC(UH)	**1.24**	**0.29**	**4.28**	**3.46**	**(1.96, 6.09)**		**0.000**
LC(S)	0.29	0.19	1.51	1.34	(0.92, 1.97)		0.130
LC(OW)	-0.06	0.57	-0.11	0.94	(0.31, 2.85)		0.910
LC(TW)	0.09	0.32	0.27	1.09	(0.59, 2.02)		0.785
LC(W)	0.89	0.60	1.48	2.44	(0.752, 7.91)		0.139
LC(BL)	**0.93**	**0.30**	**3.09**	**2.54**	**(1.41, 4.59)**	**74.44**	**0.002**
SAI	**-0.23**	**0.09**	**-2.55**	**0.80**	**(0.67, 0.95)**	**5.91**	**0.011**
D2W	**-0.00**	**0.00**	**-7.43**	**1.00**	**(1.00, 1.00)**	**61.45**	**0.000**
CC	**0.01**	**0.00**	**2.69**	**1.01**	**(1.00, 1.02)**	**6.23**	**0.007**
D2E	**-0.00**	**0.00**	**-2.94**	**1.00**	**(1.00, 1.00)**	**9.25**	**0.003**
Model L.R.		**d.f.**	**p**	**C**			
447.68		16	0.000	0.736			

the animal was searching or foraging. The Locomotion model indicated that locomotion was positively associated with distance to water (D2W) and August LAI (aLAI), both of which are common to the forage model. In addition, locomotion is negatively associated with maximum temperature (MT). Significant land cover variables have also changed, indicating that treed wetland (TW) and barren land (BL) play a role in where G098 moves between forage sites. When exhibiting locomotion behaviour, the model indicates that G098 has a preference for barren land, but will tend to avoid treed wetland. One could presume that barren land is preferred for locomotion as it is easier to traverse, whereas treed wetland is not. The locomotion model also "weakly" identifies open wetland (OW) as an avoidance area; it is very close to being significant at the $\alpha = 0.05$ level. A tabulation of these results is listed in Table 13 below (bold covariates are significant).

The locomotion model suggests that cooler areas near water that are open (few trees), and are either barren, or have reasonable green vegetation such as horsetails (*Equisetum*), graminoids and forbs[57], which may include plants such as cow parsnip (*Heracleum lanatum*), clover (*Trifolium*), and alfalfa (*Medicago sativa*)[58], are preferred.

57 Graminoids are all grasses and grass-like plants, including sedges and rushes. Forbs are flowering plants with a non-woody stem that is not a grass. Most wild and garden flowers, herbs and vegetables are forbs.

58 See Munro et al., "Seasonal and Diel Patterns of Grizzly Bear Diet and Activity in West-Central Alberta,"

Probability surfaces for both models using the significant covariates identified in Table 12 and Table 13 were computed in ArcGIS v. 9.0[59]. The probability surfaces have been rescaled to range from 1 (low preference) to 5 (high preference).

Table 13: Table of coefficients for locomotion covariates

Multivariate Model - Locomotion: Presence = 384, Pseudo-Presence = 1,152

Covriate	β_i	*S.E.*	*Wald Z*	*OR*	*95% CI*	*G*	*p*
Intercept	32.26	4.33	7.44	$1.02*10^{14}$	$(2.15*10^{10}, 4.82*10^{17})$		0.000
MT	**-1.68**	**0.21**	**-7.95**	**0.19**	**(0.126, 0.28)**	**82.63**	**0.000**
D2W	**-0.00**	**0.00**	**-7.30**	**1.00**	**(1.00, 1.00)**	**54.13**	**0.000**
LC(OCF)	-0.47	0.44	-1.08	0.62	(0.27, 1.47)		0.282
LC(BLF)	-0.05	0.28	-0.17	0.95	(0.56, 1.64)		0.865
LC(MF)	0.06	0.17	0.37	1.06	(0.76, 1.48)		0.712
LC(FR)	-0.46	0.36	-1.29	0.63	(0.312, 0.13)		0.198
LC(UH)	0.40	0.31	1.29	1.49	(0.81, 2.72)		0.197
LC(S)	0.11	0.21	0.54	1.12	(0.74, 1.68)		0.591
LC(OW)	*-2.02*	*1.03*	*-1.95*	*0.13*	*(0.02, 0.93)*		*0.051*
LC(TW)	**-1.63**	**0.63**	**-2.58**	**0.20**	**(0.06, 0.66)**		**0.010**
LC(W)	0.79	0.82	0.97	2.20	(0.45, 10.78)		0.331
LC(BL)	**0.85**	**0.37**	**2.32**	**2.34**	**(1.14, 4.78)**		**0.020**
LC(SH	-0.54	26.67	-0.02	2.30	$(0.00, 2.26*10^{7})$	36.74	0.984
aLAI	**0.45**	**0.06**	**7.61**	**1.56**	**(1.39, 1.75)**	**62.66**	**0.000**

Model L.R.	**d.f.**	**p**	**C**
236.17	14	0.000	0.751

59 ESRI, 380 New York Street, Redlands, CA 92373-8100, http://www.esri.com/

The search/foraging surface (Figure 30) shows strong patterns of high selection along the north eastern slopes of the Rocky Mountain and foothills zones of the study area, whereas the locomotion surface (Figure 31) highlights a clear preference for the river and stream network. Hence, not only is there a difference in the covariates that describe each model, there is also a distinct visual difference in the spatial distributions of high-probability sites for foraging and locomotion.

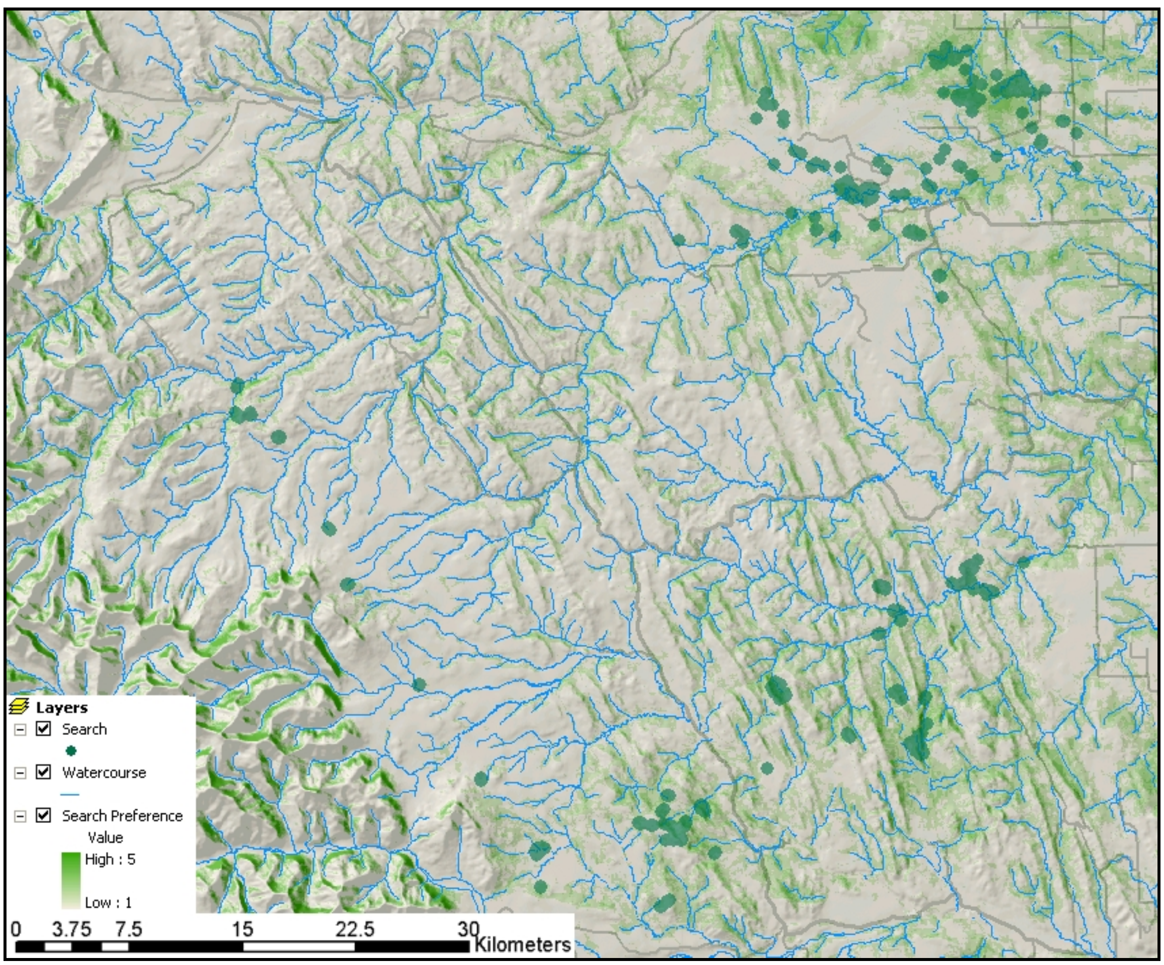


Figure 30: Spatial representation of forage model

Discussion

Given that these models are limited to data from just one animal, that the location data does not reflect the full range of expected movement rates[60], and that some 26% of the expected data is missing, these models must be treated with caution.

However, it is apparent that these models show promise in that they support the hypothesis that animals adjust their selection policy depending on what they

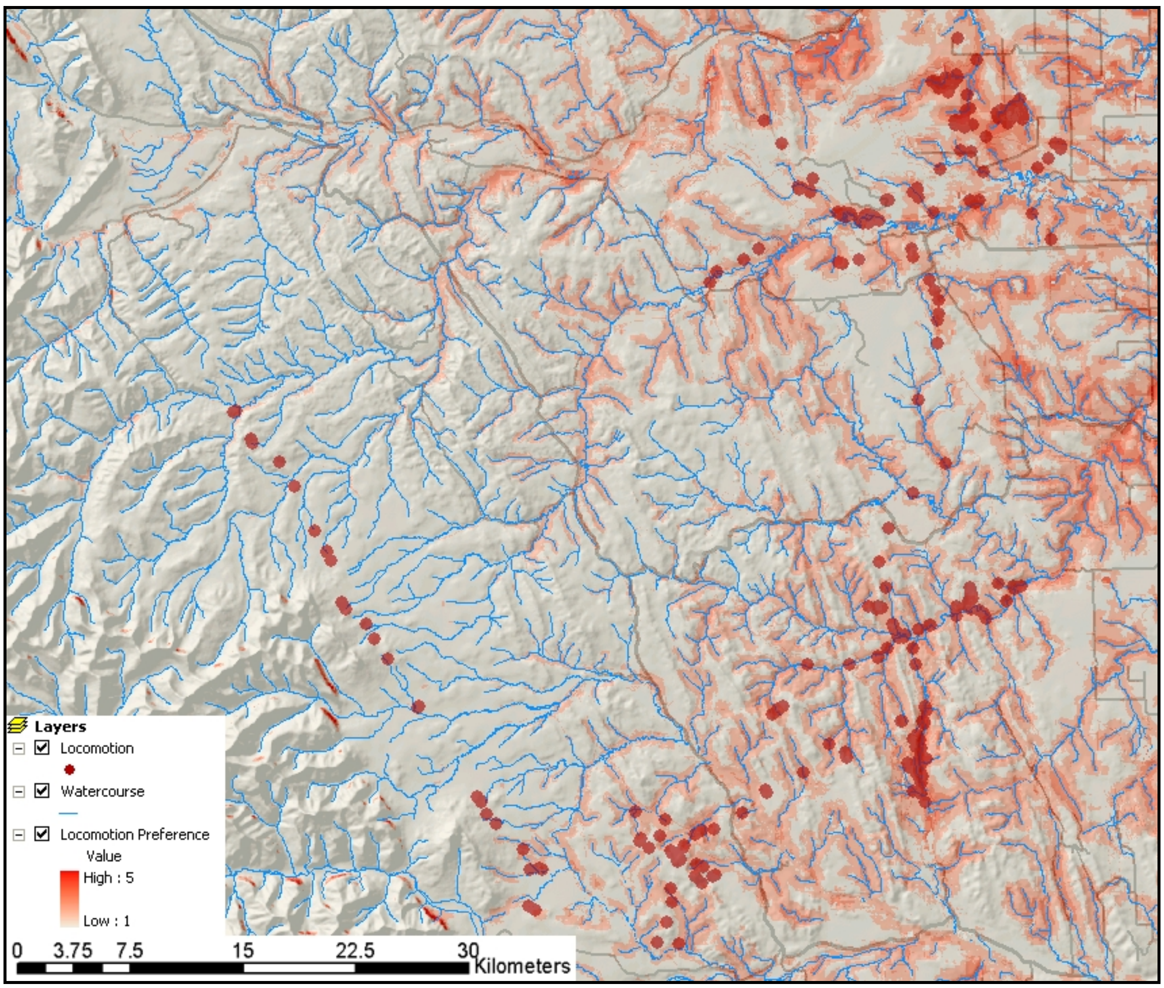


Figure 31: Spatial representation of locomotion model

60 It is believed that a grizzly bears can gallop at speeds close to 55 km/h (~916 m/min.) for 100 m or so, but I have been unable to substantiate this.

are doing (i.e., foraging, looking for new food sources, etc.), or perhaps, because their local environmental stimuli have changed, they change their behaviour.

From a modelling perspective this may allow the researcher to address issues of scale and frame bias more succinctly. If spatial scale and extent can be more clearly defined, then this will enable the analyst to test hypotheses and see more clearly to what extent the data supports the set of hypotheses. As described by Dobson (1999) a statistical model in information theory terms represents measurements that consist of signal and noise[61]. The signal is the mathematical description of the important features of the data, while the noise is that part of the data that is unexplained by the signal component. Whilst it is expected that the analyst has good knowledge of the generating processes behind a set of hypothesis, it is not always practical, or possible, to observe all the processes at work. This will affect the amount of unexplained information that remains upon completion of an analysis. Equally, if the areal framework chosen to investigate a process is not meaningful, it is probable that the level of unexplained information will be greater than necessary[62]. Within the spatial sciences, this problem is referred to as the modifiable areal unit problem (MAUP)[63].

MAUP consists of two distinct effects on the properties of estimators: those associated with the scale of the analysis, and those associated with the particular

---

61 Annette. J. Dobson, *An Introduction to Generalized Linear Models* (Boca Raton: Chapman and Hall, 1999)

62 Haining, *Spatial Data Analysis: Theory and Practice* p. 150.

63 S. Openshaw and P.J. Taylor, "The Modifiable Areal Unit Problem.," in *Quantitative Geography: A British View*, ed. Neil Wrigley and R. Bennett (London: Routledge and Kegan Paul, 1981)

partition, given the scale of the analysis. Scale may include the areal extent of an analysis, the number of sub-areas that a study area is partitioned into, or, particularly when using logistic regression techniques common to wildlife analysis, the ratio between observed and pseudo-absence/available sites. It has been well documented that levels of association will vary if scale-dependent processes are influencing outcomes[64]. In effect, measures of association at any given scale confound different scales of association.

As with many statistical techniques based on regression or correlation, the best way to minimize the amount of unexplained information is to ensure that the analysis model is as complete as possible. Failure to do so will result in biased estimates of variables that are included in the model. This in turn will tend to reduce the efficiency and explanatory power of the resulting model. This suggests that individual models for each type of movement behaviour will provide the researcher, and by extension the wildlife manager, with more powerful models and better information.

Concluding Remarks

Given that within the broad categories of covariates: solar radiation, terrain, security and land cover, there are reasonably high correlations, it is possible to select numerous other candidate models, which would be better assessed using a

64 D. E. Jelinski and J. Wu, "The Modifiable Areal Unit Problem and Implications for Landscape Ecology," *Landscape Ecology* 11, no. 3 (1996); Openshaw and Taylor, "The Modifiable Areal Unit Problem.," ; S. Openshaw, *The Modifiable Areal Unit Problem. CATMOG 38* (Norwich, England.: GeoBooks, 1984); D. Holt, D. G. Steel, and M. Tranmer, "Area Homogeneity and the Modifiable Area Unit Problem," *Geographical Systems* 3 (1996)

"best subsets" logistic regression procedure and Akaike's Information Criterion (AIC). As such, this analysis can only be considered indicative of a procedure that could be undertaken to improve the quality of models produced from animal tracking data, as we anticipate that there will likely be differences in movement rates of different animals depending on their age, sex, and reproductive state. In addition changes in environmental factors may produce interactions within the model that have not been considered.

It is not the intention of this work to conclude definitively that the models presented above are the best models. But the models do suggest that G098 makes use of different environmental and topographical stimuli depending on its current movement behaviour. As such, this chapter addresses the third question of this research, do grizzly bears exhibit different selection policies depending on their current movement behaviour. Given the limitations identified, the results of this section provide support for this hypothesis. From a research perspective, this is a key point, as previously it has been difficult to discriminate between animal behaviours when using GPS point data alone. To be able to do so will be extremely important in terms of understanding an animal's use of the landscape, and response to changing environmental conditions and human activity.

By partitioning an animal's position data and trajectories into different types of behaviours it is anticipated that the resulting models will produce superior results. But until an animal's continuous trajectory can be captured these models will continue to be limited by incomplete data, and the bias inherent in current animal tracking methods. As stated above, the approach taken in this chapter is a step forward, but what is needed is continuous path data that provides information regarding animal behaviour along the path, hence the development of the animal NavAid.

Chapter 5

Classification of Step Data

Introduction

The purpose of this analysis is to assess the possibility of predicting an animal's locomotion state using an initial candidate set of seven signal characteristics derived from the accelerometer stream of a NavAid while on a grizzly bear. Once locomotion behaviour has been determined, a series of algorithms will be used to identify the animal's trajectory.

Analysis of Accelerometer Data for Animal Tracking

The data acquired for this test was from grizzly bear G040. G040 is an eight-year-old female grizzly bear that had previously been tracked during 2001 to 2003. She was captured on April 15, 2006, with the first GPS position being acquired at 1:15 p.m. The collar was removed on July 27, 2006, with the last GPS position obtained at 11:16 a.m. The NavAid data used for this analysis was collected from April 15, 2006 until May 1, 2006. The geometric center of her home range during this time was 117°26' W, 53°05' N, and covered an area approximately 16 km (east-west) by 10 km (north-south). Throughout the time that she was tracked, her home range was centred on the same general location, which is the Elk Valley Coal Mine on the Cardinal River, approximately 42 km south southeast of Hinton, west-central Alberta (see Figure 32). Throughout the April to July period, G040's home range extended to approximately 34 km (east- west), by 16 km (north-south). Elevations within the study area ranged from 2,800 m in the western mountainous areas declining eastwards from the foothills to an elevation of 1,300 m. This elevation gradient resulted in a diversity of habitat types and ecosystems that included barren mountainous regions (17%), alpine and sub-alpine meadow (9%), wet meadow complexes (11%), coniferous dominated forests and mixed wood forests (62%).

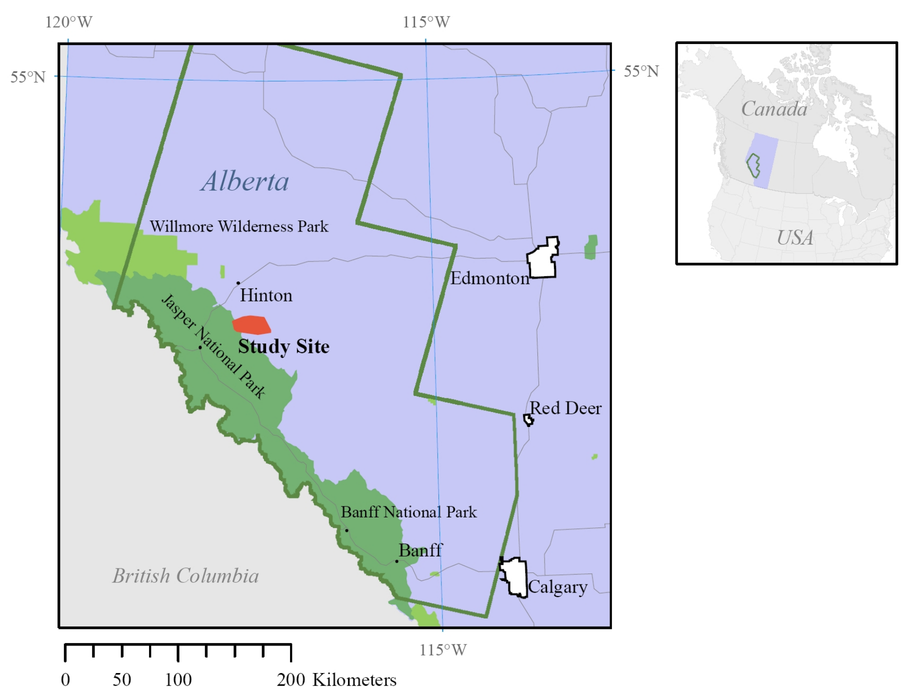


Figure 32: Study area for investigating grizzly bear locomotion in west-central Alberta, Canada

As the study area falls outside the protected area of Jasper National Park, the full range of human activities and disturbances were found, including forest harvesting, oil and gas exploration and development, mining, hunting, trapping, and all-terrain vehicle use. An extensive road network from resource extraction activities as well as seismic lines associated with energy exploration was present.

All capture efforts followed procedures accepted by the Canadian Council on Animal Care for the safe handling of bears. Research protocols were also reviewed and approved by the Animal Care Committee at the Western College of Veterinary Medicine in Saskatoon, Saskatchewan.

The GPS collar was programmed to acquire 26 locations per day[1] at 30-minute intervals during the early morning and early evening, one-hour intervals during the remainder of daylight hours, and two-hour intervals during the night. The hypothesis was that the animal would be more active during the mornings and early evenings, with the least amount of activity during the night. As such we wished to obtain more frequent positions during times with the most activity.

Eighty-one percent of expected locations were obtained during the field trial. The Dilution of Precision (DOP) observed during this period ranged from a minimum of 1 to a maximum of 11, with an average DOP of 3.5 (Mode = 2).

Study Design

For this data set three criterion groups were defined: the first being stationary (Group 1), the second searching (Group 2), and lastly walking (Group 3). A number

---

[1] More positions per day were possible; however, this would also shorten the life of the GPS battery.

of “variables” derived from the accelerometer data have been identified by Ladetto (2000), Ladetto and Merminod (2002), Lee and Mase (2001b), Macheiner et al. (2004) and Watanabe et al. (2005). Initially, following this body of literature[2], Z and Y accelerometer data streams were combined; however, satisfactory results could not be obtained. As such, the initial candidate set selected for this analysis included Z variance (ZV), Y variance (YV), Z range (ZR), Z frequency (ZF), power spectral density (PSD) of the Z axis, roll (R), pitch (P) and slope (S) of the sensor unit. As discussed in Chapter 2, ZV, YV, and ZR all relate directly to the amount of force exerted by the animal as it moves through a stride. ZF is related to movement rate, given the band limit that has been applied to the signal, 0.33 Hz to 4.5 Hz, it is expected that higher frequencies are related to more rapid movement. Roll, Pitch and Slope were included to provide information about the animal’s body posture. For example, given an ideal mounting, when the animal is walking continuously we would expect roll and pitch to be relatively constant and small — given that the NavAid is mounted on the side of the animal we would expect that roll would be larger than pitch, as the collar tends to rotate around the animals neck slightly due to the increased weight on one side of the collar. Whereas when the animal is feeding, we would expect pitch to be higher than roll. If the animal is lying down,

---

2 G. Andrienko et al., "Geovisual analytics for spatial decision support: Setting the research agenda," *International Journal of Geographical Information Science* 21, no. 8 (2007), Ladetto, "On Foot Navigation: Continuous Step Calibration Using Both Complementary Recursive Prediction and Adaptive Kalman Filtering" , Ladetto and Merminod, "Digital Magnetic Compass and Gyroscope Integration for Pedestrian Navigation" , Lee and Mase, "Recognition of Walking Behaviours for Pedestrian Navigation" , Macheiner, Legat, and Hofmann-Wellenhof, "Testing a Pedestrian Navigator" , Watanabe et al., "A new technique for monitoring the detailed behaviour of terrestrial animals: A case study with the domestic cat,"

then we would expect periods of very little variation from moment to moment in roll, pitch and slope, but, all parameters could range between 0° and 360°.

The NavAid has been designed to collect accelerometer data at a rate of 32 Hz, and magnetometer data at a rate of 1 Hz. The acceleration sensors recorded both accelerations related to changes in the movements of the bear, i.e., dynamic accelerations, and, by default, it also records gravitational acceleration $\left(\cong 9.8\text{ms}^{-2}\right)$. The periodic properties of the acceleration signals[3] recorded during dynamic bouts of animal movement (e.g., locomotion, eating, and grooming) allowed us to apply a Fast Fourier Transform (FFT) in order to determine the frequency of the signal at a particular time. Prior to application of the FFT, the signal was band limited using an elliptical filter. The FFT is an expression of a signal as a sum of sinusoids of frequencies $0, \omega_0, 2\omega_0, \ldots, n\omega_0, \ldots$, whose amplitudes are $C_0, C_1, C_2, \ldots, C_n, \ldots$, and whose phases are $0, \theta_1, \theta_2, \ldots, \theta_n, \ldots$, respectively. By plotting $C_n \quad v. \quad \omega$, we produce an amplitude spectrum of the signal (see Figure 33). From the graph we can identify the fundamental frequency for the signal by the frequency with the largest amplitude, C. In the example depicted in Figure 33, the top graph depicts the raw signal, and the bottom is the band-limited spectrum of the signal; the window size used was 4 seconds (128 observations). In this instance the frequency was estimated to be approximately 3.5 Hz. Note the effect of band limiting the signal prior to performing the FFT: all signals outside the pass-band has been removed.

3 See Figure 6: Phases and periods of forelimb activity in Chapter 2.

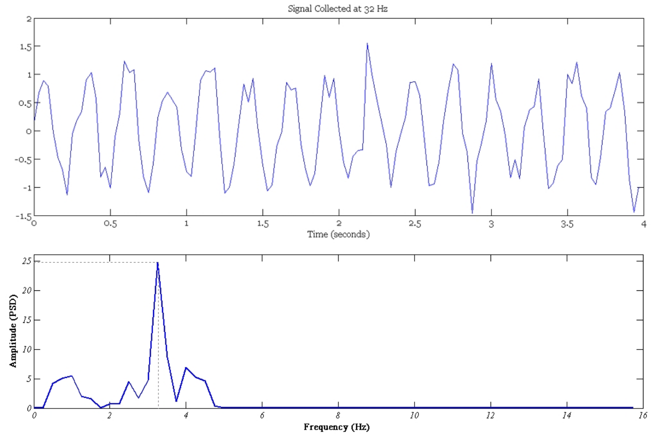


Figure 33: Example of a FFT using a four-second window at 32 Hz

All of the variables discussed above were considered to be non-stationary, in the sense that the range of anticipated values for each of the variables was expected to differ depending on the type of behaviour that the animal was exhibiting. Hence, two windows were selected in order to address the issue of non-stationarity. Following Watanabe et al. (2005) and Tanaka et al. (2001) a four-second window was used for more rapid movements, a fast walk or a trot, and a longer, 16-second window was used to identify slower low frequency movements such as searching and foraging[4] (see Figure 34 to Figure 37).

As physical observation of G040 proved to be unsuccessful, a visual assessment of a portion of the data was undertaken to categorize the different locomotion bouts. A one-hour portion of the data set acquired on April 18, 2006 between 1:15 p.m. and 2:15 p.m. was selected as it was known from the GPS data that the animal covered approximately 1,240 m during this time. It was therefore

4 Hideji Tanaka, Yasuaki Takagi, and Yasuhiko Naito, "Swimming Speeds and Buoyancy Compensation of Migrating Adult Chum Salmon *Oncorhynchus keta* Revealed by Speed/Depth/Acceleration Data Logger," *Journal of Experimental Biology* 204, no. 21 (2001), Watanabe et al., "A new technique for monitoring the detailed behaviour of terrestrial animals: A case study with the domestic cat,"

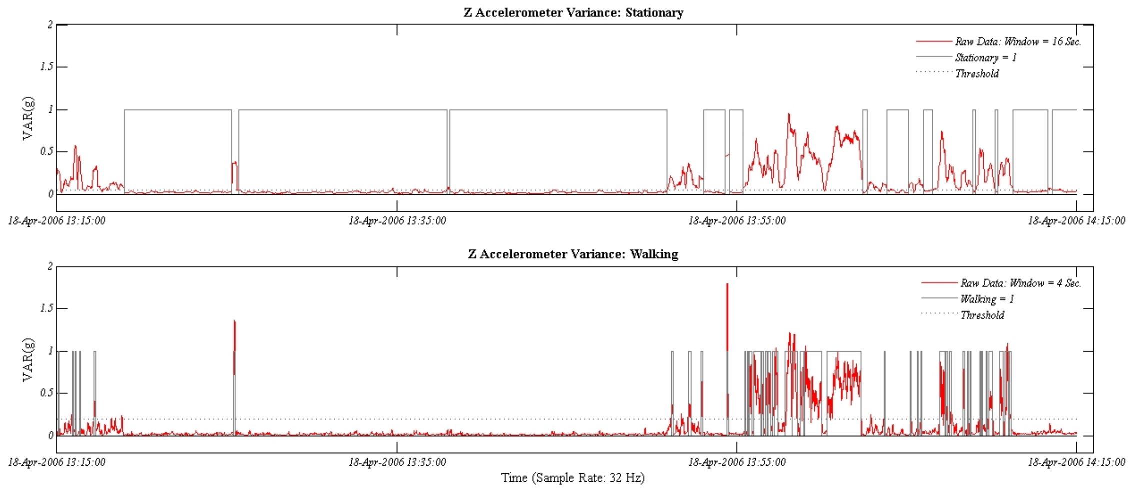


Figure 34: Z variance profile with stationary and walking periods identified

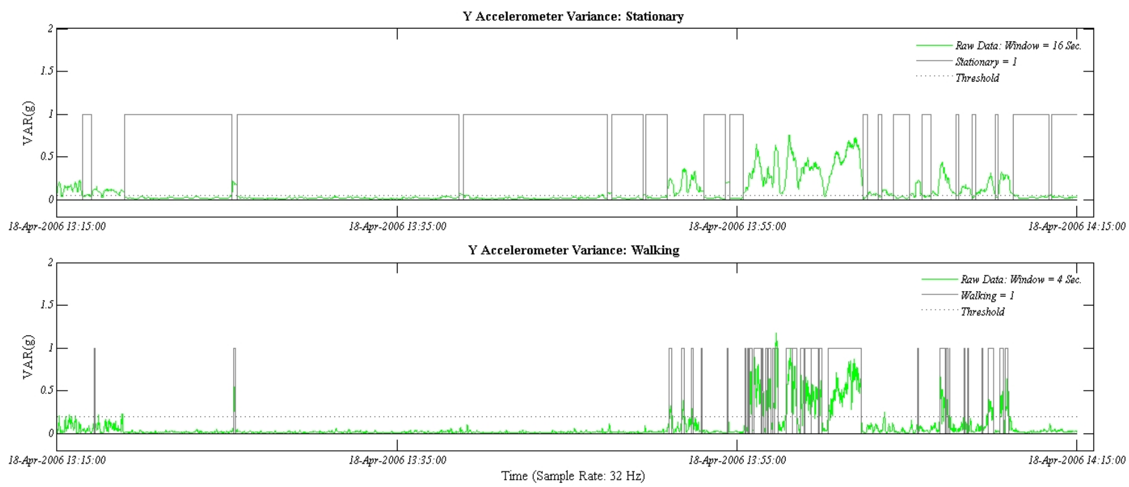


Figure 35: Y variance profile with stationary and walking periods identified

likely that at some point during the hour the animal exhibited a range of gaits.

Thresholds were assessed for Z variance, Y variance, and Z accelerometer range (see dashed lines in Figure 34 to Figure 36) for periods when it appeared that the animal was stationary (all variables exhibited low values), and a second set of

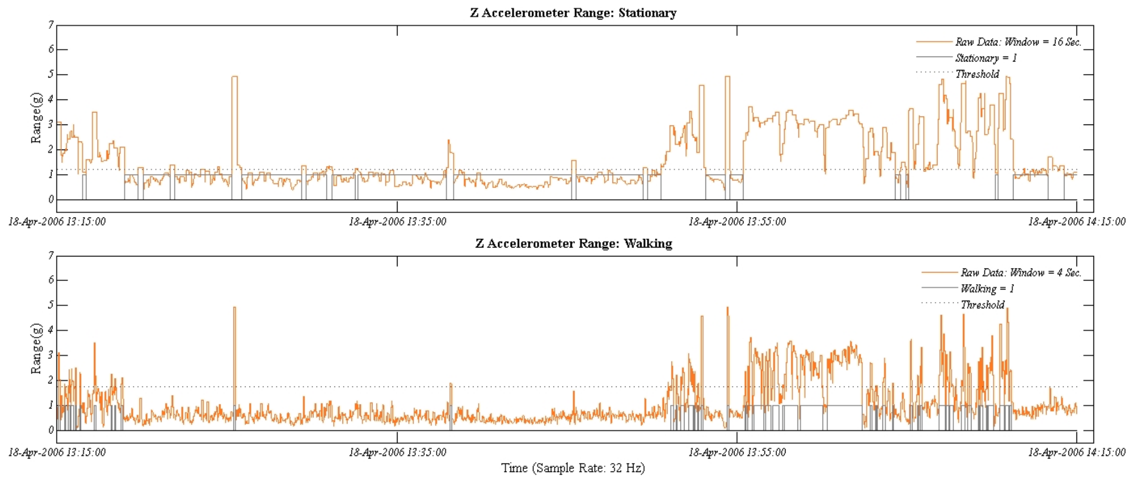


Figure 36: Z range profile with stationary and walking periods identified

thresholds was identified for periods with high values for each of the three variables. When all variables were high, it was assumed that the animal was actively walking, and not searching for food or foraging. All data that fell between the two sets of thresholds was assigned to the searching category. In effect, a low-pass filter was developed for the identification of stationary periods; a high-pass filter for active walking periods; and a band-pass filter for searching or foraging periods.

A (subjective) review of Figure 34 to Figure 36 indicates that there can be relatively long periods of inactivity interspersed with a mix of searching and walking bouts. As expected, the effect of the wider window tends to smooth the data, which ensures that minor movements of only a few seconds are essentially ignored. As evident in Figure 37, during periods of little activity the accelerations are in the range of ±0.5 g, and when the animal appears to be moving the signal will saturate the accelerometer sensors — observe the maximum amplitude of the

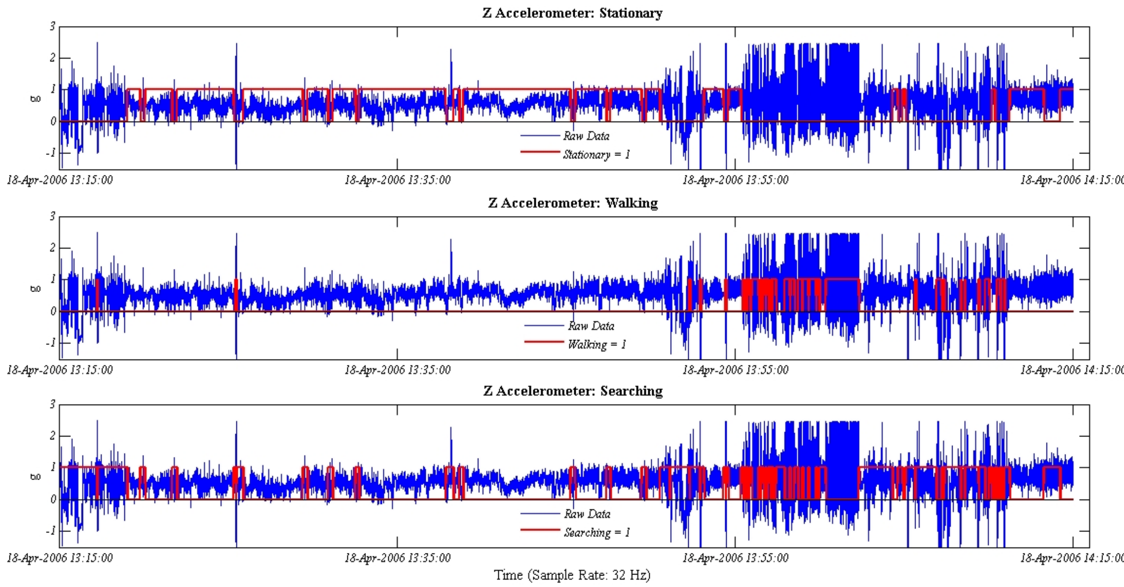


Figure 37: Z accelerometer data categorized into bouts of stationary, searching and walking

signal shortly after 13:55. While it is more beneficial to acquire a complete signal, given the peak detection techniques used to identify steps, saturation of the sensor does not detrimentally affect step identification. This particular issue could be addressed by utilizing the PNI ASIC magnetometer at its ±6 g range; however, we then lose a substantial amount of resolution, which may have a detrimental effect on the quality of the information that can be obtained from the sensor unit.

The data were analyzed using SPSS v. 15 and SAS v. 9.1. Descriptive information for our eight outcome variables is provided in Table 14. The sample size has been based on the following requirement, $\min(n_j) > 5p$, that is, there should be at least five times the number of predictors in smallest group[5], and ideally, sample proportions should reflect actual proportions of population[6].

5 Carl J. Huberty and Stephen Olejnik, *Applied MANOVA and Discriminant Analysis*, 2nd ed. (Hoboken, New Jersey: John Wiley & Sons, Ltd., 2006) pg. 309

6 Ibid. pg. 310

Using the rules identified in Figure 34 to Figure 37 above, it was estimated that G040 spent approximately 12.5 hours per day exhibiting stationary behaviour, 10.5 hours searching or foraging, and slightly less than one hour actively moving. A random selection of each group was then obtained to provide a sample with approximate proportions of 12:10:1 such that group 3 (walking) had 40 samples.

Upon initial review of the data, it was found that the inclusion of frequency (ZF) resulted in a singular variance-covariance matrix of the sum of squares of, and cross products between the variables, and was therefore removed from the candidate set of variables.

A review of univariate Q-Q plots, Skewness and Kurtosis of the individual predictor variables (not included) suggest that multivariate normality is not tenable. In addition the strong patterns evident in the multivariate scatter plot shown in Figure 38 supports the hypothesis that multivariate normality can not be

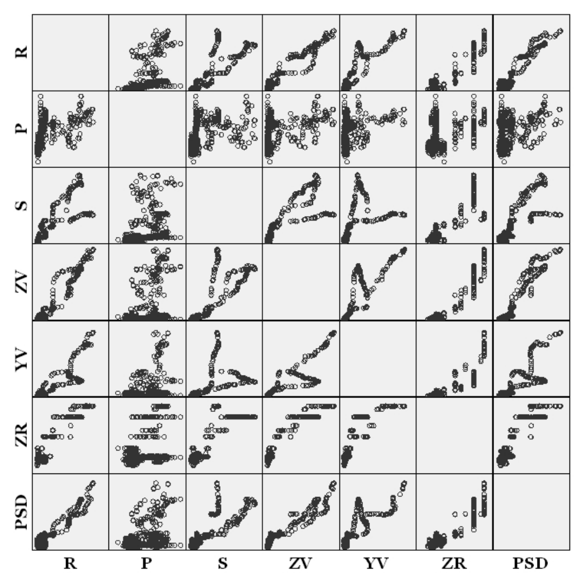


Figure 38: Multivariate scatter plot of predictor variables

expected with this data set.

The results of the univariate hypothesis tests $\left(\mathrm{df}_1 = 2, \mathrm{df}_2 = 862\right)$ indicate that the three populations differ across all seven variables (see Table 14); hence the seven variables would all appear to be capable of distinguishing between the criterion groups.

Table 14: Description information and univariate test for the 3-group accelerometer data

	Group Means (SDs)			**Univariate**	
Var1iable	*Stationary (1)*	*Searching (2)*	*Walking (3)*	$F_{2,862}$	*P*
ZV	0.011 (0.009)	0.080 (0.123)	0.451 (0.109)	497.0	0.000
YV	0.012 (0.009)	0.048 (0.050)	0.394 (0.089)	1,771.5	0.000
ZR	0.824 (0.090)	1.338 (0.847)	3.531 (0.455)	426.0	0.000
PSD	3.749 (2.200)	8.560 (7.742)	29.862 (7.890)	417.3	0.000
R	0.026 (0.019)	0.064 (0.088)	0.328 (0.055)	445.2	0.000
P	-0.006 (0.421)	0.003 (0.451)	0.716 (0.139)	54.2	0.000
S	0.008 (0.005)	0.031 (0.036)	0.057 (0.006)	137.6	0.000
n	437	388	40		

However, a review of the error correlations (Table 15) show that a substantial proportion of the outcome variables are highly correlated. Hence, it is clear that the full set of variables should not be considered together (see red shaded cells — high correlations; orange — medium to high correlations — in Table 15).

Table 15: Error correlations for the 3-group accelerometer data

	YV	*ZR*	*PSD*	*R*	*P*	*S*
ZV	0.81	0.90	0.92	.93	0.35	0.94
YV		0.83	0.78	0.83	0.28	0.77
ZR			0.82	0.87	0.29	0.89
PSD				0.92	0.42	0.91
R					0.52	0.96
P						0.44

The Box test for covariance homogeneity provided little support to indicate that the covariance matrices are similar, given $\left[F\left(56,34037\right)=130.2,\ P\cong 0.000\right]$ and $\left[\chi^2\left(56\right)=7{,}302,\ P\cong 0.000\right]$. Further support for this conclusion is provided by the (natural) logarithms of the determinants of the four covariance matrices: Group 1, -52.4; Group 2, -32.9; Group 3, -41.1; and Error, -34.4. These log determinants[7] are clearly not similar. Based on theoretical considerations, and the nature of the variables studied, we must conclude that the joint distribution of the seven variables within each population is not multivariate normal. Hence, there is ample evidence for the use of a quadratic classification rule.

Huberty and Olejnik (2006) recommend a rank transformation be applied to the data in instances where continuous variables do not meet the assumptions of multivariate normality[8] (see Table 16 and Table 17). While the Box test for

---

7 The determinant represents the generalized variance of a square covariance matrix $\mathbf{S}$. Similar values imply similar variability within a set of data. Taking the natural logarithm of $|\mathbf{S}|$ allows covariance matrices to be assessed using a χ^2 distribution.

8 Huberty and Olejnik, *Applied MANOVA and Discriminant Analysis*

Table 16: Rank transformed description information for the 3-group accelerometer data

	Group Means/(SDs)			Univariate	
Var1iable	*Stationary (1)*	*Searching (2)*	*Walking (3)*	$F_{2,862}$	*P*
ZV	265 (193)	580 (163)	832 (30)	436.6	0.000
YV	279 (189)	563 (194)	845 (12)	338.2	0.000
ZR	348 (168)	487 (281)	837 (31)	107.9	0.000
PSD	294 (202)	549 (204)	827 (23)	252.8	0.000
R	367 (219)	466 (251)	835 (23)	83.8	0.000
P	417 (241)	415 (248)	772 (54)	42.2	0.000
S	298 (212)	550 (206)	779 (14)	216.9	0.000
n	437	388	40		

Table 17: Rank transformed error correlations for the 3-group accelerometer data

	YV	*ZR*	*PSD*	*R*	*P*	*S*
ZV	0.77	0.79	0.83	.74	0.59	0.74
YV		0.68	0.59	0.62	0.41	0.61
ZR			0.67	0.60	0.38	0.56
PSD				0.58	0.46	0.64
R					0.75	0.94
P						0.75

covariance homogeneity did improve, obtaining $\left[F\left(56,34037\right)=49.1,\, P\cong 0.000\right]$ and $\left[\chi^2\left(56\right)=2,755,\, P\cong 0.000\right]$, the data is still clearly not normally distributed. Nor are the logarithms of the determinants of the four covariance matrices: Group 1, 64.7; Group 2, 66.7; Group 3, 33.4; and Error, 67.5, within a similar range. A review of correlation within the Rank transformed variables show that although there are fewer strong correlations, there are an increased number of medium to strong correlations. As with Table 15, red shaded cells indicate high correlations and orange cells, medium to high correlations.

It is clear that transforming the data provided no obvious benefit; hence predictive discriminant analysis (PDA) was applied to the raw data. The trade-off is that the optimality of the results may be in question. That is, the number, or proportion of observations correctly classified cannot be assured to be the maximum possible. Given the high correlations evident in Table 15, the initial candidate set of variables was reduced to 13 combinations of variables that did not exhibit substantial correlation (see Table 18).

Prior probabilities of locomotion prediction were set based on the assessment of the estimated time that G040 spent performing a particular behaviour. We therefore arrived at the following priors for the analysis: 0.50, 0.45, and 0.05 for Groups 1, 2 and 3 respectively.

Table 18: Candidate models for PDA of accelerometer data

Possible Models			
1	ZV	*8*	PSD, P
2	ZV, P	*9*	R
3	YV	*10*	R, P
4	YV, P	*11*	P
5	ZR	*12*	P, S
6	ZR, P	*13*	S
7	PSD		

Classification Results

An external classification rule is generally recommended to estimate group hit rates. Therefore a quadratic Leave-One-Out (LOO) rule was used. Prior to a final examination of the data, a review of the candidate models was undertaken. What was sought was the possibility of deleting one or more of the predictors before performing our final Predictive Discriminant Analysis. Using software developed by J. D. Morris (Florida Atlantic University)[9] an all subsets analysis was undertaken. The results from this analysis indicate that Z accelerometer range (ZR) and pitch (P) produce the highest total group hit rate. The results of the three best models, when using either one or two predictors, are presented in Table 19, with all results reproduced in Figure 39. It is apparent that the total group hit rate increased when a second predictor was added to the model. In both one and two variable subsets, Z accelerometer range provided the highest hit rate levels.

Table 19 summarizes the hit rate for each model, along with $Z_{(i)}$, the transformed hit rate[10], which takes into consideration both the observed and chance hit rate for the data set, and the expected variance of the hit rate, to produce a z score that can be interpreted more objectively. As a result, candidate model 6 (in Table 18), consisting of the variables ZR and P was judged as the best model to be retained. While model 4 (YV P) in Table 18 was given the same rank as model 6 (ZR P) based on its z score, but its hit rate wasn't quite as high.

---

9 See John D. Morris and Alice Meshbane, "Selecting Predictor Variables in Two-group Classification Problems," *Educational and Psychological Measurement* 55, no. 3 (1995)

10 See Carl J. Huberty and Joseph M. Wisenbaker, "Variable Importance in Multivariate Group Comparisons," *Journal of Educational Statistics* 17, no. 1 (1992)

Table 19: Total group LOO hit rates for variable subsets from the 3-group locomotion date

Subset Size	Best Subset	Hit Rate	$Z_{(i)}$	Rank
1	ZR	0.754	17.535	2
	YV	0.729	16.102	3
	S	0.669	12.553	4
2	ZR P	0.845	22.928	1
	YV P	0.816	21.221	1
	P S	0.762	18.013	2

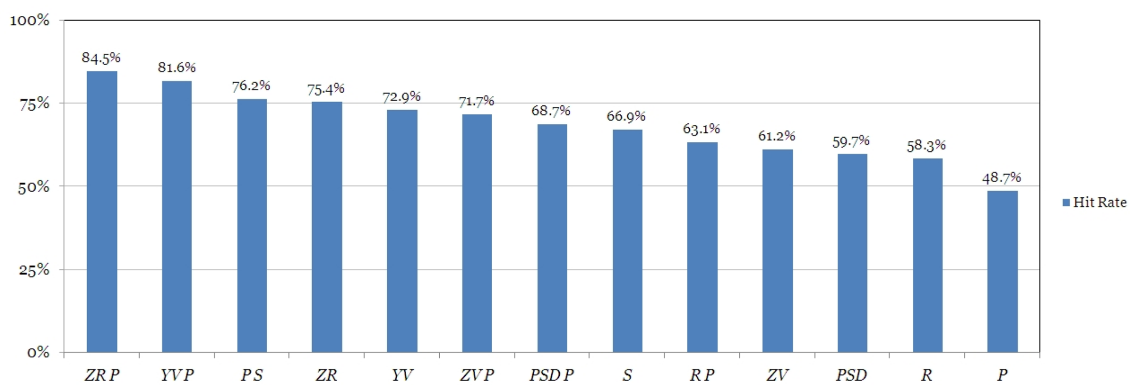


Figure 39: Total group LOO hit rate versus best sub-set for the 3-group locomotion data

Table 20 relates to the multivariate hypothesis implicit in predictive discriminant analysis that the population centroids do not differ. The multivariate hypothesis is similar to the univariate approach, except it examines all variables at one time. The multivariate test has a number of advantages over a group of separate tests, as Type I errors can be controlled, and any underlying structure in the data will also be considered. Clearly the statistics indicate that the population centroids do differ, and that they also differ with respect to the two variables, ZR and P.

Table 20: Multivariate hypothesis tests

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	0.519	151.321	4	1,724	0.000
Hotellings	1.012	217.7414	4	1,720	0.000
Wilks[11]	0.491	183.661	4	1,722	0.000

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	$F_{2,862}$	Sig. of F
ZR	285.943	289.328	142.971	0.336	425.957	0.000
P	19.696	156.730	9.848	.182	54.162	0.000

Table 21 provides a partial list[12] of the results from the individual locomotion prediction results. The table provides a small portion of the quadratic LOO classification results. An examination of the results shows that each unit is assigned to the group with the highest posterior probability. This allows one to examine the closeness in a probabilistic sense, of each unit to the centroid of each of the groups. For example, unit 865 (group 3) is also predicted to be closest to the centroid of group-3, as the posterior probability, $\hat{P}(3|\mathbf{x}) = 0.936$, is clearly the largest. Unit 401 on the other hand, came from group 2, but was predicted to be in group 1, hence this is considered an error or miss.

There are also a number (25) of in-doubt units (≅2.8% of the sample). These are units that are approximately the same distance from two or more group

11 F statistic for Wilks' Λ is exact.

12 A complete tabulation of the locomotion prediction results is provided in Appendix IV.

Table 21: Partial list of cross-validation results using quadratic discriminant function classification – posterior probability of membership to locomotion groups

	Actual	Predicted	*Posterior Probabilities*			*Typicality*	
ID	*Group*	*Group (j)*	$\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	$\hat{P}(\mathbf{x}_u \mid j)$	D^{*2}
79	1	1	0.874	0.127	0.000	0.270	2.699
115	1	1	0.939	0.061	0.000	0.658	1.172
401	2	1	0.943	0.057	0.000	0.001	1.632
411	2	2	0.033	0.968	0.000	0.010	0.743
458	2	2	0.000	1.000	0.000	0.845	3.909
496	2	1	0.518	0.482	0.000	0.909	5.873
800	2	2	0.000	0.557	0.443	0.592	4.485
853	3	3	0.000	0.037	0.963	0.723	1.543
865	3	3	0.000	0.065	0.936	0.808	1.572

centroids, hence their prior probabilities for those groups are similar. Relatively, there are few in-doubt units; therefore we can conclude that an additional group does not exist in the data. All but two of the in-doubt units have similar posterior probabilities for groups 1 and 2. For group 1 units, ZR is close to the group average, but P is much higher than the group average, whereas for group 2 units, ZR is closer to the group 1 average rather than group 2, but P is very low compared to the group average. The majority of the group 2 in-doubt units were predicted to be group 1 members.

In addition to the in-doubt units, there are 18 units (≅2.1%) that may be outliers. These are units that have been assigned to a group, but do not reflect a typical group member. Note the typicality value for unit 411 in Table 21. Typicality is the probability of an observed unit u, consisting of vector $\mathbf{x}_u$, belonging to group j. In essence, it is the proportion of units in group j that have vectors close to $\mathbf{x}_u$. All possible outliers belong to groups 1 or 2, with four of the 18 being misclassified. In all group 1 cases ZR is approximately average for the group, but P is extreme, both negative and positive, suggesting that the NavAid is not in its normal orientation. With regards to the group 2 outliers, ZR is low compared to the group average (typically — close to the group 1 average), as is P. A review of the raw data would suggest that these are probably not searching movements as the signal within a few seconds of these points suggests that the animal is stationary. Perhaps this is the animal having a look around, while it is sitting on the ground.

Locomotion prediction results are presented in Table 22 in the form of a 3 by 3 classification table. The three separate group hit rates are given in parentheses on the main diagonal. The total group hit rate is $(432+264+35)/865=0.846$.

Table 22: Quadratic LOO classification results

		Predicted Groups			
		Stationary (1)	*Searching (2)*	*Walking (3)*	*Total*
Actual Group	*1*	**432 (0.989)**	5	0	437
	2	110	**264 (0.680)**	14	388
	3	0	5	**35 (0.875)**	40
Total		542	274	49	N = 865

In order to assess the quality of the classification, it is worthwhile comparing the classification results with those that could be obtained by chance alone. Table 23 provides a comparison of the PDA classification results with chance classification, where n_{jj} is the number of hits per group, n_j is the size of the group, $e_j = q_j n_j$ is the expected hits given the estimated prior probability $\left(q_j\right)$ of membership for group j. z is a standard normal statistic used to test the null hypothesis that the estimated classification is equal to the chance classification[13] and is calculated using

$$z = \frac{n_{jj} - e}{\sqrt{e\left(N - e\right)/N}}. \quad (5.1)$$

P in Table 23 is the probability of obtaining z. Clearly, the quadratic LOO classification for groups 1 to 3, and the overall group hit rate, are significantly better than would be expected by chance. This is supported by the fact that the lower bound of the 99% confidence interval for each group does not overlap chance

Table 23: Comparison of results with chance classification

Group	n_{jj}	n_j	e_j	Z	P	**Lower Bound for 99% Interval**	I
Staionary (1)	432	437	218.5	20.43	0.000	407.68	98%
Searching (2)	264	388	174.6	9.12	0.000	241.20	42%
Walking (3)	35	40	2	23.94	0.000	31.79	87%
Overall	731	865	395.1	22.93	0.000	696.92	71%

13 See Huberty and Olejnik, *Applied MANOVA and Discriminant Analysis* pp. 316 - 320

classification. Lastly, I provides an indication of the improvement over chance by the classification rule. Hence, the classification rule will predict nearly twice as many group 1 and group 3 units as would be arrived at by chance, and we can therefore consider that the model's effect is meaningful.

An alternative method commonly used within Geomatics to assess the quality of a classification technique is Kappa. Kappa is equivalent to the overall improvement over chance reported in Table 23, i.e., 71%.

The quadratic classification rule developed from the accelerometer sensor data may be used with new data. The quadratic rule is in the form of a quadratic composite of the two predictor variables, ZR and P. The weights and constants for the three groups given in SAS are in Table 24. Given a set of predictor observations

Table 24: SAS output for the quadratic classification functions for locomotion mode

Locomotion Mode	Type	Name	ZR	P
1	QUAD	ZR	-165.998	28.105
1	QUAD	P	28.105	-7.575
1	QUAD	LINEAR	274.071	-46.438
1	QUAD	CONST	-110.053	-110.053
2	QUAD	ZR	-0.778	0.472
2	QUAD	P	0.472	-2.750
2	QUAD	LINEAR	2.078	-1.246
2	QUAD	CONST	-1.169	-1.169
3	QUAD	ZR	-2.421	0.421
3	QUAD	P	0.421	-25.781
3	QUAD	LINEAR	16.489	33.956
3	QUAD	CONST	-41.507	-41.507

for a new time interval, a composite score for each of the groups can be determined using the weights in Table 24; an observation can then be assigned to the group with the largest composite score.

Given the data in Table 24, the quadratic classification functions for groups one to three are, therefore, the following

$$\begin{aligned} Z_{Group1} = & -110.053 + (274.071 \times ZR) + (-46.438 \times P) \\ & + (-165.998 \times ZR^2) + (-7.575 \times P^2) \\ & + (28.105 \times ZR \times P) \end{aligned} \tag{5.2}$$

$$\begin{aligned} Z_{Group2} = & -1.169 + (2.078 \times ZR) + (-1.246 \times P) \\ & + (-0.778 \times ZR^2) + (-2.750 \times P^2) \\ & + (0.472 \times ZR \times P) \end{aligned} \tag{5.3}$$

$$\begin{aligned} Z_{Group3} = & -41.507 + (16.489 \times ZR) + (33.956 \times P) \\ & + (-2.421 \times ZR^2) + (-25.781 \times P^2) \\ & + (0.421 \times ZR \times P) \end{aligned} \tag{5.4}$$

The best subset analysis also identified YV and P as one of the better classification rules (see Table 19). While this rule gave a slightly lower group hit rate (0.816 as opposed to 0.845 for ZR, P) its z score was similar, suggesting that this model may also be equally as good. Hence, it is worthwhile comparing the effectiveness of the two classification rules. Comparison of two rules can be treated in a similar manner to a two level, repeated measures design[14]. As such, the comparison

---

[14] A. Agresti, *Categorical Data Analysis*, 2nd ed. (New York: John Wiley, 2002) pg. 411

results in a 2×2 table that consists of the number of units correctly classified by both rules (upper right, a); the number of units correctly classified by rule 1, but not rule 2 (upper left, b), and so on.

To determine if rule 1 and rule 2 are the same we compare the hit rate proportions of each rule, i.e., rule 1: $HR_1 = (a+b)/N$; and rule 2: $HR_2 = (a+c)/N$, where the null hypothesis is $H_0 : HR_1 = HR_2$. Table 25 summarizes the similarities between the two rules.

Table 25: Comparison of classification rules

		Rule 2 (YV P)		
		Hit	*Miss*	*Total*
Rule 1 (ZR P)	*Hit*	628	103	*731*
	Miss	78	56	*134*
Total		*706*	*159*	*865*

The data results in $HR_1 = 0.846$ and $HR \ \ = 0.816$. Using equation (5.5) produces a standard normal statistic of $z = 1.603\,(P = 0.063)$. Therefore, assuming $\alpha = 0.05$, we must accept H_0, and conclude that the two rules are not statistically different given the observed data, however, the lack of significance is borderline. When comparing rule 1 with the third-ranked rule

$$z = \frac{HR_1 - HR_2}{\sqrt{\dfrac{HR_1\left(1-HR_1\right)}{n_1} \; \dfrac{HR_2\left(1-HR_2\right)}{n_2}}}, \tag{5.5}$$

(model 12 in Table 18), we find that they are significantly different $(z = 4.380, P = 0.000)$; hence model 12 need not be considered further.

From Figure 40(d) (lower left image) it is apparent that group 3 (walking) is clearly separated from the other forms of locomotion, and that accelerometer range accounts for the majority of separation between the three groups. In addition, it is clear that as the rate of locomotion increases so too does the average accelerometer range. It would also appear from the plots that pitch is not related to any specific separation of the groups, however, we must be careful to not place too much reliance on visual inspection, because our perception may not depend on the numerical scales used for the plot axes.

Discussion

The purpose of this study was to determine how well three locomotion behaviours of a grizzly bear might be predicted using predictor scores based on data obtained from an animal NavAid. Of the three groups considered (stationary, searching, and walking) it was found that all groups could be predicted with accuracy of some note. The group hit rate for a stationary animal was nearly 99% accurate, and nearly twice that which could be expected by chance. Walking was also identified at a high level, approaching 88% accuracy, and was 87% better than could be expected by chance. Searching behaviour was the least accurate at 68% (42% better than chance). As a result, a classification rule for use with new data has been developed. The raw observation data did not meet the requirements of multivariate normality; hence the classification rule takes a quadratic form. Prior

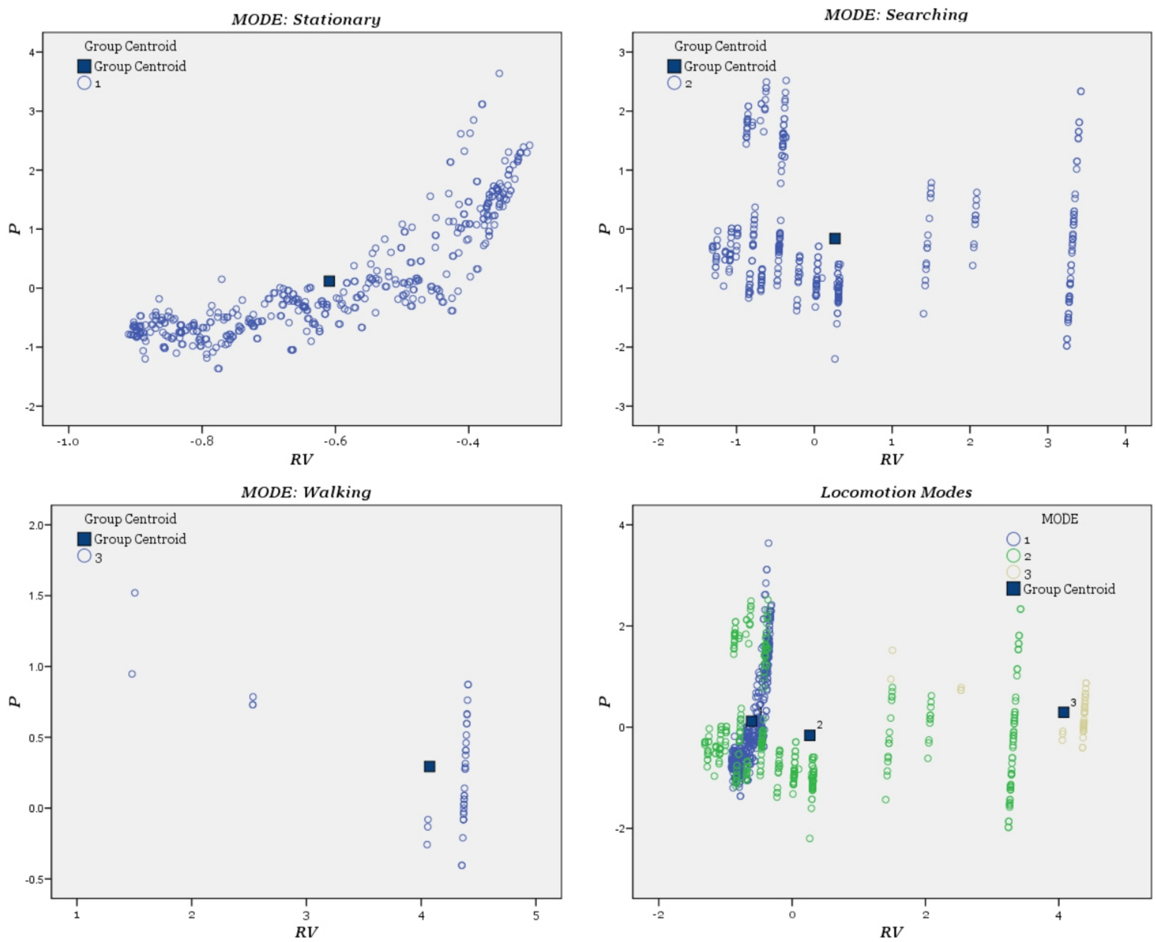


Figure 40: Plots of quadratic LOO group centroids. (a): stationary, (b): searching, (c): walking; and (d) combined

probabilities used to develop the model were derived from a subjective analysis of accelerometer data, however, it should be noted that had equal priors been applied to the model the overall group hit rate would have been only reduced by 1.4%.

The primary limitation of this model is the influence of initially misclassified analysis units on the hit rate estimation; it may be non-consequential, considerable, or of unknown effect. In order to address this adequately, time must be spent in the field. There are relatively few in-doubt units, which suggest that none of the observations belong to a fourth group. Because multivariate normality is untenable, and because we do not know which distribution the data does follow, cluster analysis is a possible solution to group assignment; however, we run the risk of developing a rule that is not easily interpretable. Alternatively, application of a k-nearest neighbour rule may also be of benefit, but there appears to be no consensus regarding validation of this technique[15]. The issue here is to verify, as much as possible, initial group membership.

Given the saturation of the accelerometer signals in Figure 37, a comparison of ±2 g and ±6 g sensors is warranted to determine if saturation of the sensors is having an adverse effect on the analysis. An increase in signal range will however reduce signal resolution, which may adversely affect results. This could be addressed in part by utilizing a 16-bit analogue to digital converter (ADC), as opposed to the 12-bit ADC currently implemented in the NavAid. This would, however, require a design change as the MSP430 microprocessor does not include a 16-bit ADC.

Path Estimation Methodology

The classification process just reviewed was used to identify temporal intervals for each of the three modes of locomotion. The Z axis accelerometer data was then categorized based on those intervals. Additionally, in order to reduce the computation burden it was assumed that the animal was essentially stationary if

---

15 See Huberty and Olejnik, *Applied MANOVA and Discriminant Analysis* pp. 363 – 365 and A. P. White, "Cross Validation of Nearest Neighbour Discriminant Analysis- A Warning to SAS Users," *Journal of Statistical Computation and Simulation* 49, no. 3 (1994)

two or more consecutive GPS positions were not significantly different at a confidence level of 95%, that is

$$s_{a-b} \geq z_{1-\alpha/2}\left(\frac{\mathrm{DOP}_a + \mathrm{DOP}_b}{2}\right)\mathrm{UERE}, \tag{5.6}$$

where s_{a-b} is the distance between the GPS positions, DOP is the Dilution of Precision estimated for each GPS fix, UERE is the user equivalent range error (discussed in Chapter 2), and $z_{1-\alpha/2} = 1.96$.

Following Lee and Mase (2001a), the raw data was then smoothed using an elliptical filter with a pass band from 0.33 Hz to 4.5 Hz. Figure 41 highlights the filter used to smooth the data. Essentially, any signal that had a frequency out side the pass-band limits had the magnitude of its signal attenuated by -20 dB[16].

Upon smoothing, a simple peak detection algorithm was implemented to identify the time at which a step occurred. While trialing this algorithm, by identification of my own steps, similar accuracies to that reported by Lee and Mase

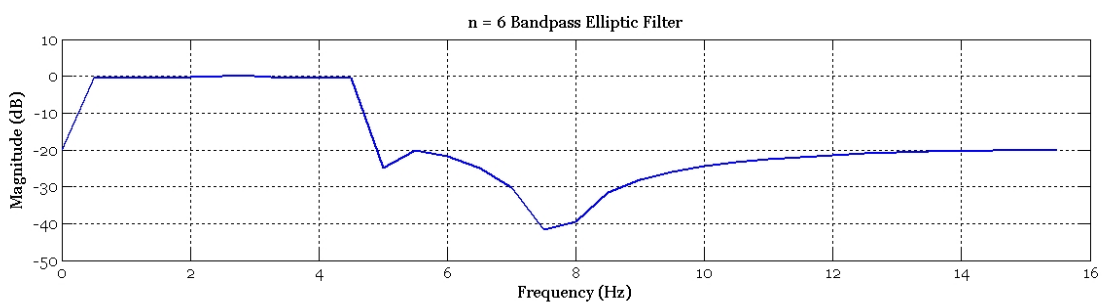


Figure 41: Elliptical filter used to smooth accelerometer data

16 Outside the pass band the signal magnitude will be reduced by a factor of 100 $\text{attenuation}(\text{dB}) = 10 \cdot \log_{10}\left(\frac{\text{Output}}{\text{Input}}\right)$.

(2001a) were noted. That is, over ten repeated trials of 40 strides each on a flat surface, with the NavAid attached to my belt, on my hip, 41.3±1.8 strides were identified at the one sigma level. This would suggest that on average, a route would be approximately 3.2% long, assuming a constant stride length. Studies by Fang et al. (2005), Ladetto (2000) and Lee and Mase (2001a) have shown that stride length is influenced by stride frequency and terrain slope[17]. It is expected that the same effect will be present in animal data, and that factors such as sex, age, and environmental condition may also play a role in stride length. While grizzly bears and humans have a similar heel-toe foot placement action, their fore leg has an elbow, as opposed to a knee, which may also have a detrimental effect on step counting.

The classified Z accelerometer data for groups two and three, from the grizzly bear, was then processed to identify the time at which steps occurred. As a control data set has yet to be acquired, the results have been assessed subjectively, by visual inspection. As shown in Figure 42, it is immediately apparent that for the walking data, the elliptical filter provides a good representation of the raw data.

For this particular sample most peaks have been identified, with possibly four peaks (≅6.3%) of the 63 peaks identified in Figure 42 being identified incorrectly.

---

17 L. Fang et al., "Design of a Wireless Assisted Pedestrian Dead Reckoning System - The NavMote Experience," *IEEE Transactions on Instrumentation and Measurement* 54, no. 6 (2005), Ladetto, "On Foot Navigation: Continuous Step Calibration Using Both Complementary Recursive Prediction and Adaptive Kalman Filtering" , Lee and Mase, "Recognition of Walking Behaviours for Pedestrian Navigation"

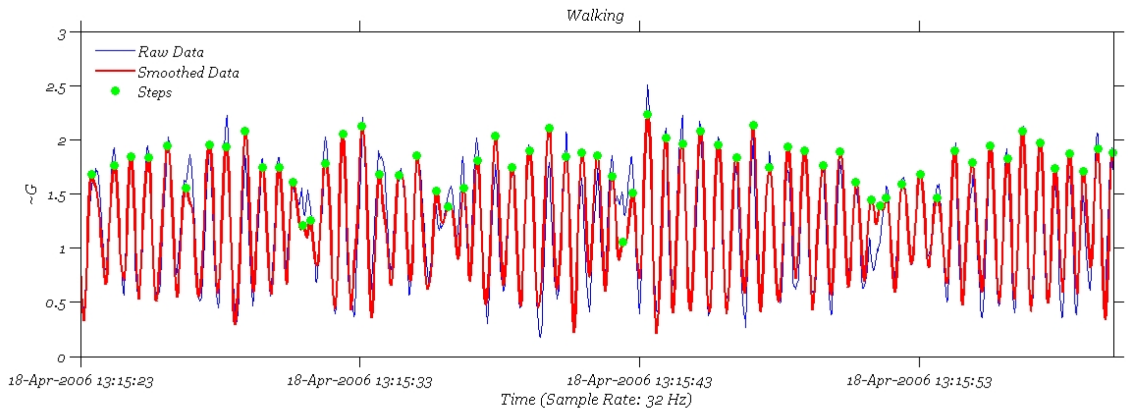


Figure 42: Estimated step for walking behaviour using an elliptical filter

However, when an elliptical filter was applied to the searching data there were typically many more peaks identified than there appeared to be in the raw data.

With regards to the searching data it was generally a lot nosier than the walking data and obviously not as stationary — hence the poor performance of the elliptical filter[18]. If one was to hypothesize as to the source of the extra noise, one might suggest that it could be attributed to head and neck movement of the animal as it searches for food, eats, and surveys its surroundings.

Given the unsatisfactory search results using an elliptical filter a foray into the world of wavelets was embarked upon in an attempt to obtain a better solution. As was outlined in Chapter 2, one of the common applications of wavelets is to denoise a data stream with minimal loss of information. Typically, when undertaking signal processing, it is beneficial to know something about how smooth the signal is, i.e., how sharp are the peaks in the signal; are there sudden jumps in the signal, etc., in order to select an appropriate filter. With wavelets, however, very little

[18] Digital filters assume a stationary signal.

knowledge of the signal is required. Essentially, a wavelet is selected, and the signal is transformed into a set of coefficients, then all coefficients below a certain size are thrown out (set to zero) and the signal is reconstructed. That said, selection of the best basis function could become a time consuming process.

In previous work by Ladetto (2000), a Meyer wavelet was adopted to de-noise the accelerometer data. However, it was found that when applied to the grizzly bear data it produced unsatisfactory results. At one level of decomposition there would be too many peaks, at the next level, there would be too few, or visually, they were clearly not located in the right position. It is speculated that this is due to differences in locomotion between humans and an animal, and the confounding effects of environmental factors such as terrain, vegetation, and a very loose coupling of the NavAid to the grizzly bear, i.e., the collar rotates from side to side around the animal's neck as it moves rather than being fixed, all of which were likely not present in Ladetto's work.

Hence, the Daubechies family of wavelets was selected, primarily for its compact support, which allows the wavelet to efficiently represent signals that contain localized features[19] such as those evident in the raw accelerometer data (see Figure 44). As with the walking data, results of the wavelet analysis have been assessed qualitatively. Using Matlab 7.1 and its Wavelet Toolbox, visual inspection of different orders, and levels of decomposition of the Daubechies wavelets were

19 Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making* and James S. Walker, *A Primer on Wavelets and their Scientific Application* (Boca Raton: CRC Press LLC, 1999)

considered to determine which order/level combination should be used to smooth the data. In essence, the objective of this inspection was to identify an order/decomposition level combination that best represented the times at which steps were considered to have been taken. It was concluded that a Daubechies wavelet of order 8, using the level five coefficients best represented the portion of data that was used during this inspection. The Daubechies order eight wavelet is depicted in Figure 43.

A comparison of level 4 and level 5 decompositions indicated that there was little difference between them in segments where there was a regular occurrence of peaks, however, when the frequency of peaks became quite variable; the level four decomposition appeared to over estimate the number of peaks. At level six decomposition, the resulting wave missed the majority of peaks in the higher frequency portions of the signal. Figure 44 contains the output of a portion of the data during the afternoon of April 18, 2006. When the raw data in Figure 44 is

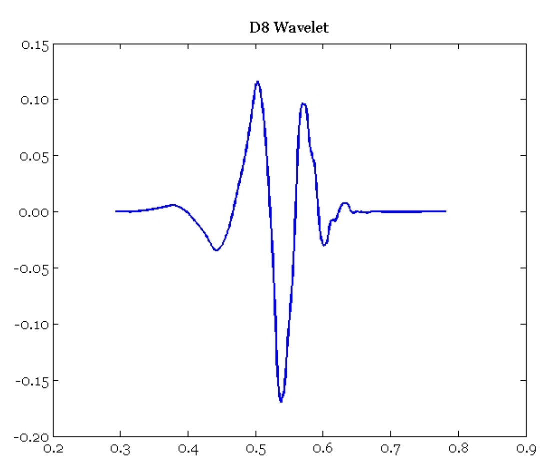


Figure 43: Daubechies wavelet basis function used to denoise accelerometer signal

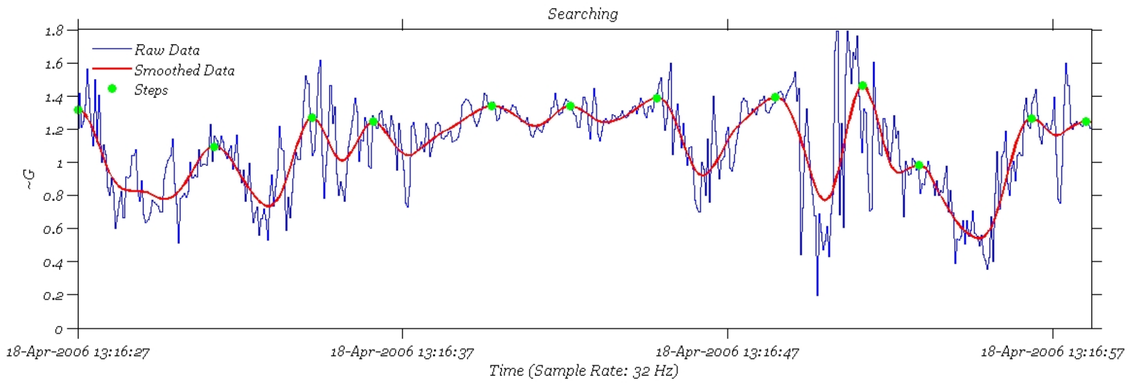


Figure 44: De-noised searching signal with steps identified by peak detection

compared with that in Figure 42 it is apparent that non-stationarity is a much greater issue with the searching data.

Regardless of whether the animal was searching or walking, once a step had been identified, the heading at the time of the step was determined by cubic convolution, using the closest two heading samples either side of the time that the step occurred. It was assumed that the cubic convolution would produce a smoother set of headings, thus minimizing the effect of spurious headings.

Transformation Results

As reported earlier, it has been shown in a number of studies that a variable stride length produces more accurate results than a constant stride length model[20]. Given the results of the Predictive Discriminant Analysis, *Z* accelerometer range was used as a stride length surrogate in the computation of a dead reckoning (DR) route.

---

20 See Fang et al., "Design of a Wireless Assisted Pedestrian Dead Reckoning System - The NavMote Experience," ; Ladetto, "On Foot Navigation: Continuous Step Calibration Using Both Complementary Recursive Prediction and Adaptive Kalman Filtering" ; or Lee and Mase, "Recognition of Walking Behaviours for Pedestrian Navigation" .

Following equation (2.1) a route was calculated for each movement interval defined by the GPS data. Once computed an affine transformation (see equation (2.15)) was computed to place each route in its appropriate geographic space, and to rotate and scale each route such that the routes were forced to fit the GPS data that defined the temporal interval within which they fell. In effect, the scaling process determined estimated stride length, and the rotation was an estimate of the bias between the orientation of the NavAid and the heading of the animal.

A review of the scale factors for the 186 intervals on which movement was deemed to have occurred, nine intervals were identified as outliers. In these instances, the step identification procedure substantially underestimated the number of steps taken; hence the scale factor required to fit the route to the GPS data was large. The scale factor during these intervals ranged from 5.5 to 22.7, with an average of 11.3 ($s = 5.0$). Once these units were removed the average scale factor reduced to 0.88 ($s = 1.2$). A comparison of group means indicates that the outliers do come from a different distribution ($t = 5.29$, $n = 184$, $p = 0.000$), as such, the outliers were not considered in the remainder of the analysis.

A review of rotation factors obtained from the affine transformation indicates that the average rotation was 36° 30' East. That is, each route, on average, had to be rotated 36° 30' anti-clockwise. This implies that the misalignment between the animal frame and the sensor frame (see chapter 2) was 18° 11', given that magnetic declination in the Hinton area was 18° 19' East[21]. Following Yamartino (1984) the dispersion[22] about this mean value was determined to be 103° 40'. This tells us that the heading results are highly variable and we suspect that the cause of the variability is in part due to the fact that the collar is not fixed firmly to the animal, allowing it to roll and pitch much more that we would ideally like. An additional source of error is likely to be due to the accuracy of GPS positions. As depicted in Figure 45, each GPS position has an associated error, the user equivalent range error (UERE). The UERE can be considered a probability distribution around the estimated location, within which the true location falls. We can assume that the UERE is independent and identically distributed, although in reality it is not (Frair, et al., 2004). Hence, there is a chance that an observed line, as depicted by the heavy black line in Figure 45, may in actual fact be the blue line in Figure 45, resulting in a heading error between the two points. If we assume two points are just significantly different in position, that the average DOP for the two positions is 3.2, and the one sigma position error is 8.0 m[23] then the points would be 50.2 m apart, if a confidence level of 95% was used. As such, the error associated with each point would be 25.1 m. Given this geometry, the maximum difference between the observed direction and the true direction would be 45°. As the GPS points move further apart this error will reduce.

---

21 Magnetic declination was estimated using Natural Resources Canada's Magnetic declination calculator for April of 2006, go to http://www.geolab.nrcan.gc.ca/geomag/apps/mdcal_e.php.

22 Yamartino, "A Comparison of Several "Single-Pass" Estimators of the Standard Deviation of Wind Direction,"

23 See Chapter 2 for details regarding these values.

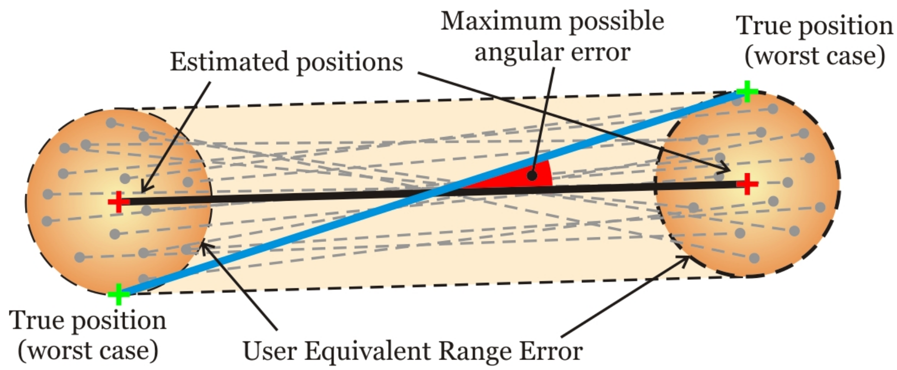


Figure 45: The effect of uncertainty of position on direction

An analysis of stride length shows that the average searching stride was 0.23m (s=0.324, n = 50,499), while the average for walking was 2.62 (s=2.673, n = 1,823). Therefore, as expected, a walking stride is significantly longer than a searching stride $(t = 4.049, P = 0.000)$.

Movement Rates Revisited

Now that we have produced a continuous path for G040, it is worthwhile revisiting the movement rate analysis undertaken in chapter two, except this time a three-process model shall be adopted now that the animal's path is continuous. As stated in chapter two the model is designed to partition grizzly bear movement into locomotion (corridor movement), specialized search movement (patch movement) and feeding movements. Silby et al. (1990) and Johnson et al. (2002) have proposed non-linear curve fitting models to differentiate between the movement bouts[24] using the log transformed frequency distribution of an animal's movement velocity. Major inflections along the curve provide a means of differentiating

---

24 Johnson et al., "Movement Parameters of ungulates and Scale-specific Responses to the Environment," , Silby, Nott, and Fletcher, "Splitting Behaviour into Bouts,"

between the different types of processes, or movements. Once again, the nonlinear regression (NLR) procedure in SPSS 15.0 was used to solve the model

$$y = \ln(N_f \lambda_f e^{-\lambda_f r} + N_s \lambda_s e^{-\lambda_s r} + N_l \lambda_l e^{-\lambda_l r}) \tag{5.7}$$

where N is the total number of counts for each type of movement, λ is the probability that an event occurs in the next movement rate interval, and r is the movement rate. f represents foraging bouts, s represents searching, l represents locomotion bouts, and y is the expected number of movements that occur during each discrete interval of the movement rates.

For G040, 65,536 out of 180,941 movement rates estimated between April 15 and May 1, 2006 were used for the analysis. Movement rates ranged from 0 m per minute to 594 m per minute, with an average movement rate of 24.0±0.3 m[25] per minute.

The raw data is a frequency distribution (histogram) of movement rates as depicted in Figure 46. Some zero counts were observed at high movement rates, and as the analysis was carried out on log-transformed data a constant value of one (1) was added to all counts. The logarithm of the frequencies was then taken, to equalize the variances at different movement rates. Log ~ (frequency) (hereafter called y) is plotted against velocity (metres/minute) in Figure 46.

It is clear from Figure 46 that the $\log_e$ frequency distribution is not a single straight line, as would be expected of a single-process model. To see whether the data might fit a three-process model, the model was fit as follows. The left-hand 80

---

25 95% Confidence Interval

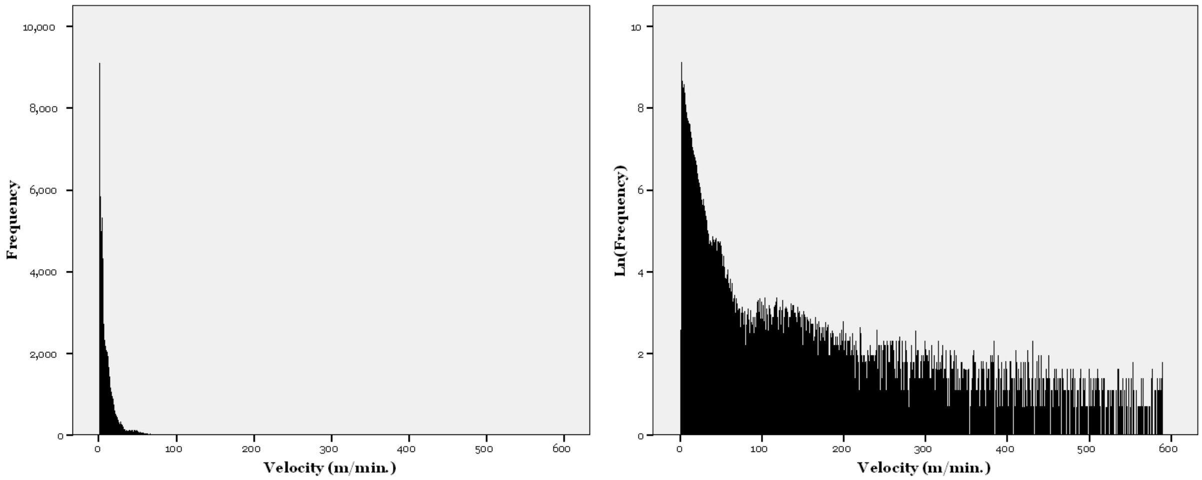


Figure 46: (a) Frequency, (b) and Log$_e$ frequency versus velocity (m/min.) for grizzly bear G040 collected over a two week period

points of Figure 46b were replotted in Figure 47a. The points appear to fall on a straight line with equation $y = 7.854 - 0.069r$ (from linear regression, $F_{1,78} = 383.1,\ p = 0.000$). Hence, $\lambda_f = 0.069$, and $N_f = (1/0.069)e^{7.854} = 37,333$. The middle 190 points of Figure 46b are replotted in Figure 47b. These points appear to follow a straight line with equation $y = 3.726 - 0.007r$ $\left(F_{1,189} = 301.9,\ p = 0.000\right)$.

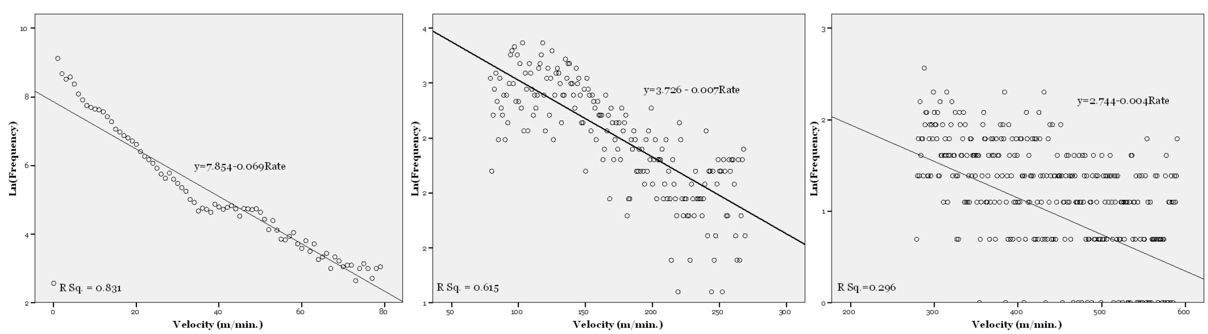


Figure 47: The data of Figure 46 is split into three for further analysis. (a) The left-hand points plotted are for velocities less than 80 m/min. (b) The middle points between movements of 80 m/min and 270 m/min. (c) the points relating to movement rates greater than 270 m/min.

Hence, $\lambda_s = 0.007$, and $N_s = (1/0.007)e^{3.726} = 5,930$. Lastly, the remaining points in Figure 46(b) were replotted in Figure 47(c), and a first order model was fit using linear regression, resulting in a model $y = 2.744 - 0.004r$ ($F_{1,310} = 130.3$, $p = 0.000$), with $\lambda_l = 0.004$, and $N_l = (1/0.004)e^{2.744} = 3,887$.

After 22 iterations, the output from the NLR procedure (in SPSS Release 15.0) provided a set of parameter estimates, the parameter correlation matrix, and the total sum of squares associated with the six-parameter model of equation (5.7) (see Table 26 and Figure 48).

The model accounts for 92.8% $\left(R^2 = 1 - \frac{SSE}{SST_c} \right)$ of the variation in the data. The F-statistic for the three-process model is 4,631.7, with $\mathrm{df}_1 = 5$, $\mathrm{df}_2 = 585$, and $P = 0.000$. The critical value for this test, given a confidence interval of 95%, is 2.114, clearly the model is worth fitting to the data of Figure 46.

As indicated by the upper (UB) and lower (LB) bounds[26] of the coefficients, all are significant, with $N_f = 56,726.9$ $[\mathrm{LB}: 46,575.2, \mathrm{UB}: 66,878.6]$, $\lambda_f = 0.107$ $[\mathrm{LB}: 0.097, \mathrm{UB}: 0.118]$, $N_s = 3,853.9$ $[\mathrm{LB}: 2,062.3, \mathrm{UB}: 5,645.6]$, $\lambda_s = 0.010$ $[\mathrm{LB}: 0.005, \mathrm{UB}: 0.015]$, $N_l = 3,247.4$ $[\mathrm{LB}: 2,020.0, \mathrm{UB}: 4,474.7]$, and $\lambda_l = 0.002$ $[\mathrm{LB}: 0.000, \mathrm{UB}: 0.005]$. The correlation matrix (Table 26) shows moderate correlation between the parameters for foraging and searching behaviour, however, the parameters relating to higher movement rates tend towards quite strong

26 LB and UB represent the confidence interval of an estimated coefficient.

Table 26: Output from the non-linear curve fitting procedure NLR

Parameter	Estimate	Std. Error	95% Conf. Interval	
			Lower Bound	*Upper Bound*
N_f	56,726.887	5,168.794	46,575.197	66,878.576
λ_f	0.107	0.005	0.097	0.118
N_s	3,853.935	912.227	2,062.290	5,645.580
λ_s	0.010	0.003	0.005	0.015
N_l	3,247.358	624.929	2,019.975	4,474.740
λ_l	0.002	0.001	0.000	0.005

Asymptotic Correlation Matrix

	λ_f	N_s	λ_s	N_l	λ_l
N_f	0.637	-0.030	0.079	0.068	0.056
λ_f		-0.202	0.394	0.349	0.299
N_s			-0.877	-0.920	-0.966
λ_s				.982	0.938
N_l					0.949

ANOVA

Source	Sum of Squares	df	Mean Squares	F	P
Regression	4,052.734	5	810.546	4,631.7	0.000
Residual	101.969	584	0.175		
Uncorrected Total	4,154.703	590			
Corrected Total	1,410.366	589			

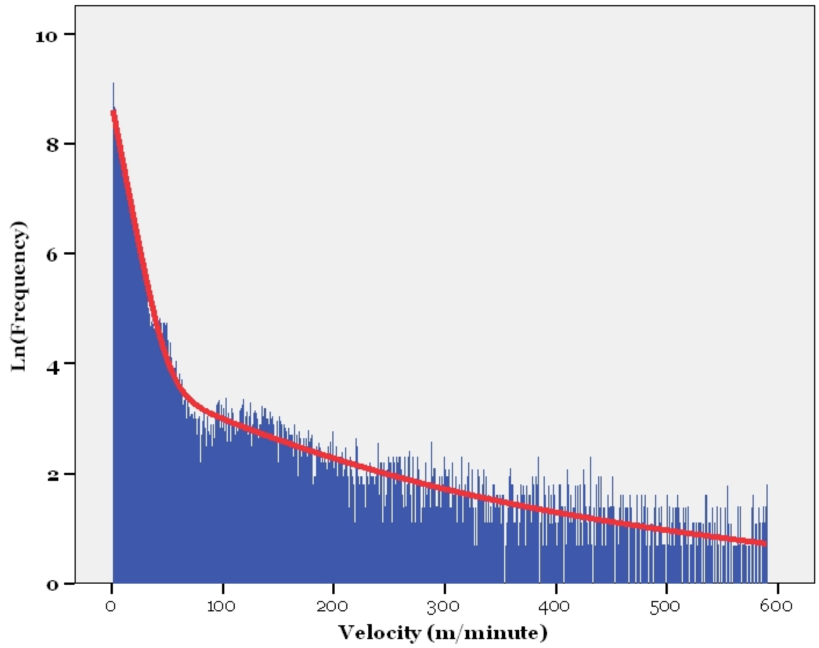


Figure 48: Log$_e$ of frequency of G040's velocity including model

relationships, some of which are excessive. This makes it difficult to attribute effects to one or the other of these parameters; as a consequence their standard errors may be higher than they otherwise might have been.

The movement behaviour criterion, r_c, can now be used to determine the threshold between foraging and searching, and searching and active locomotion using equations (5.8) and (5.9) respectively.

$$r_c^{f-s} = \frac{1}{\lambda_f - \lambda_s} \ln \frac{N_f \lambda_f}{N_s \lambda_s} \tag{5.8}$$

$$r_c^{s-l} = \frac{1}{\lambda_s - \lambda_l} \ln \frac{N_s \lambda_s}{N_l \lambda_l} \tag{5.9}$$

Hence, we can conclude that the threshold between foraging and searching was 52.2 m/minute [LB: 44.0 m/min.; UB: 66.1 m/min.][27], and that the threshold between searching and active locomotion was 222.6 m/minute [LB: 133.1 m/min.; UB: 423.3 m/min.]. Clearly, these intervals do not overlap.

When compared with the model developed from data normally available to researchers and wildlife managers (6.5 m/minute [LB: 5.5 m/min.; UB: 7.7 m/min.]), this model provides a completely different picture with regards to the movement of G040. Given that G040 is female, and G098 was male, we would expect that G098's thresholds would be higher if we had the same type of data to work with, assuming they both spend a similar amount of time feeding, as male grizzly bear home ranges are generally considerably larger than female home ranges.

Following (Slater and Lester, 1982), estimates of the total number of misassigned movements can be obtained by rearranging equations (5.8) and (5.9) to obtain

$$N_f e^{-\lambda_f r_c^{f-s}} + N_s\left(1-e^{-\lambda_s r_c^{f-s}}\right) \tag{5.10}$$

$$N_s e^{-\lambda_s r_c^{s-l}} + N_l\left(1-e^{-\lambda_l r_c^{s-l}}\right). \tag{5.11}$$

Hence, the expected number of movement bouts misassigned between foraging and searching was 1,780, or 2.77% of the total number of movements that fall below the r_c^{s-l} threshold; and the expected number of movement bouts misassigned between

27 95% Confidence Interval

searching and active locomotion was 26, or 0.89% of the total number of movements that fall above r_c^{f-s}.

The classification of the movement data for April 18, 2006 has been depicted in Figure 49. The movement data has been plotted along with IDT land cover data, roads (the two black lines in the north central portion of the figure) and rivers. The initial impression of the route with respect to the IDT data is that there is a possible registration issue between the remotely sensed land cover data, and the GPS data obtained from G040. Assuming the GPS data is correct, it would appear that the land cover data is shifted approximately 150 m north. Alternatively, it may be a limitation of the land cover data, as the alpine transition from shrub through herb to barren it is difficult to distinguish as there is little vegetation[28].

G040 bedded down overnight in the southwest corner of Figure 49 in an area classified as shrubs. She commenced moving shortly after 8:15 a.m. and appears to have spent a large portion of the morning (see Figure 49) in shrub areas. According to the GPS data, there was a pause from 10:46 a.m. until 12:15 p.m. Then around 1:15 p.m. she commenced a period of active locomotion, whereby she headed in a north-westerly direction for one hour, covering close to 1,300 m. During this period it would appear that G040 moved along a barren area into upland trees. Towards the end of the hour, her velocity decreased and it is supposed that she spent more time grazing. After 2:15 p.m., she continued in a north-westerly direction but at a slower pace. A portion of the next hour was spent

---

[28] Dr. Greg McDermid, August 17, 2007, *personal communication*

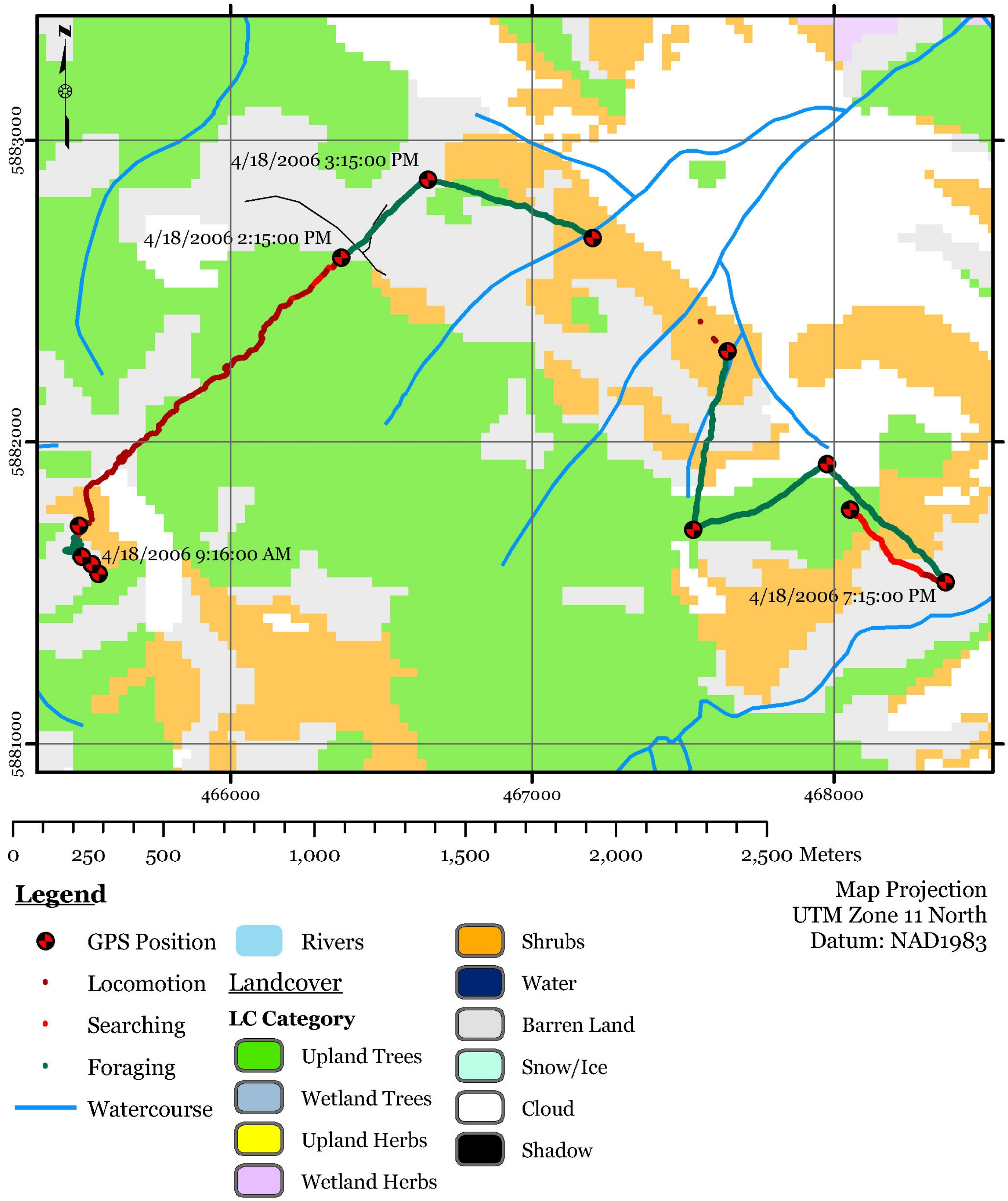


Figure 49: G040 movement on April 18, 2006

in the vicinity of an unclassified road. By 3:15 p.m., G040 had turned in a south-westerly direction within another shrub area and continued to forage until approximately 5:15 p.m. at which time she came across a stream and followed that for an hour. There was also a period between 4:15 p.m. and 5:15 p.m. where very few steps were identified, but according to the GPS data, she had covered 580 m. Once G040 reached the head of the stream she turned in an easterly direction for 30 minutes until she reached a tributary of the stream that she had followed previously. At the end of this stream she headed in a south-west direction for 30 minutes up a gully and over a ridge into what appears to be a barren river valley that is running in a south-east north-west direction. At 7:15 p.m. she turned around and headed back up to the top of the ridge at a much higher velocity, bedding down for the night in a patch of upland trees in the vicinity of the crest of the ridge.

Figure 50 provides an enlarged view of G040's foraging path during the morning of April 18, 2006. What is evident from this figure is that when G040 feeds, her path is considerably more convoluted. This is what would be expected; but, given current data collection techniques, it has not been possible to document the extent of route sinuosity for the grizzly bear. Now that we have a clearer idea of the complete path, it is anticipated that determination of habitat use preference can be greatly refined because we now know where a grizzly bear has been, and where it hasn't. In addition, we can categorize a grizzly bears movement behaviour and examine if different relationships exist between various movement behaviours and the environment that it occupies.

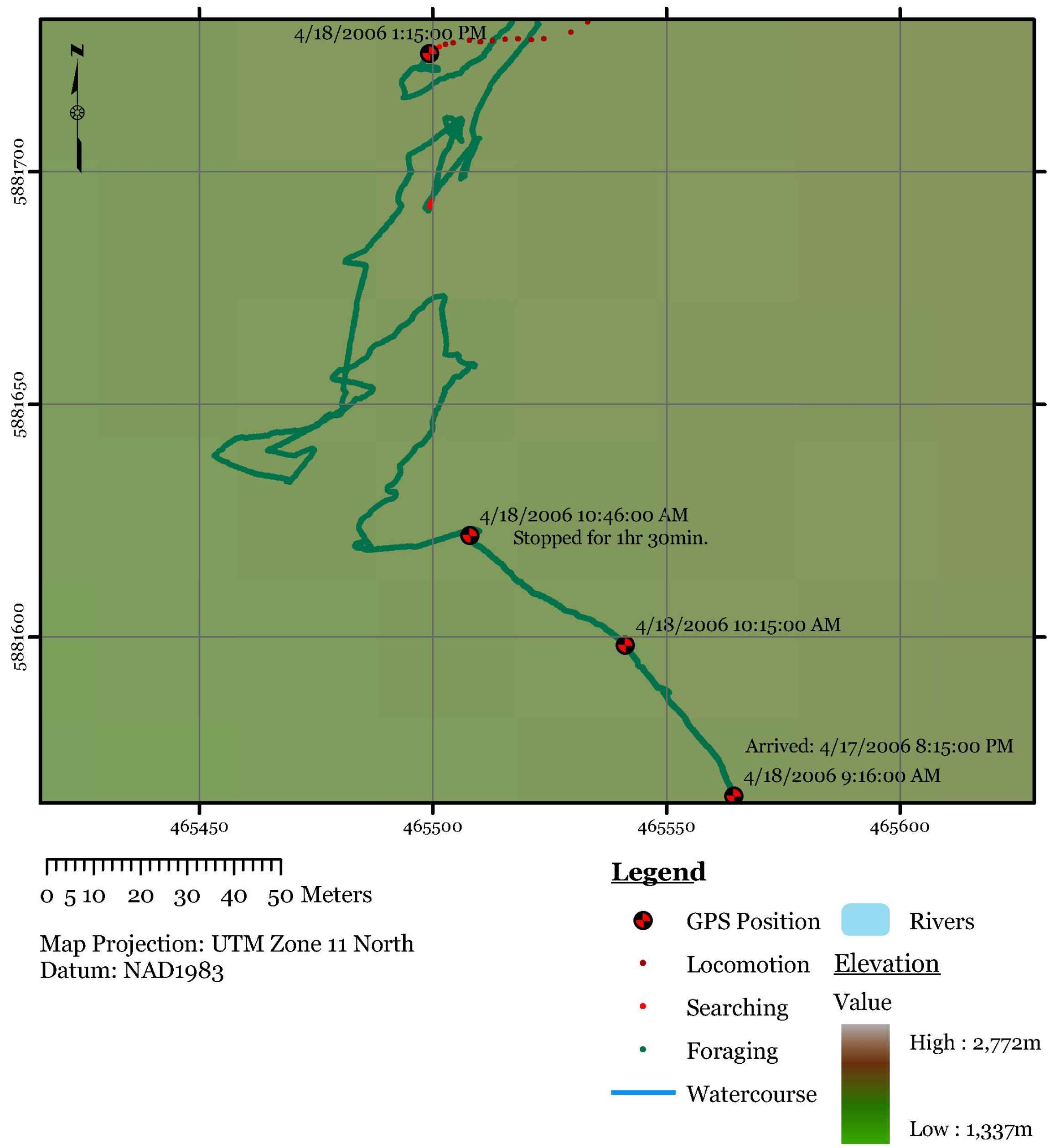


Figure 50: Foraging movement — G040, April 18, 2006

Concluding Remarks

The analysis has shown that the NavAid provides the data necessary to gain insights into animal movement and behaviour. Via two separate approaches we have developed models that allow us to categorize an animal's movements into different types of locomotion behaviour.

At the same time, it has raised a number of issues that warrant further research. Firstly, in order to validate the results of the predictive discriminant analysis, we must obtain some control data from the field. Without this data we are essentially undertaking a data mining exercise, and we run the risk of identifying effects in the data that may not be valid. Secondly, there have been instances where we have not been able to identify steps adequately, which generally results in extremely long strides and excessive movement rates. We need to spend some time investigating why we have been unable to see these steps. Is it an algorithm issue? A hardware issue? A hardware mounting issue? At this stage we do not know. Thirdly, but in the same vein, is the issue of heading variation. In most sensor-based tracking systems, the sensors are considered to be loosely coupled to the body frame of the carrier. This assumes that a simple calibration will resolve any misalignment between the different coordinate frames. With regards to animal tracking, the sensor's frame is "loosely coupled" to the animal frame, but calibration is not feasible unless we can ensure that the NavAid is truly "fixed" to the animal. This issue requires more research if we wish to improve the quality of the heading results.

The elliptical filter works well when the animal is walking, however identification of steps via wavelets when the animal is in a searching mode is subjective at this stage. Are there better wavelets that can be used? Is wavelet choice consequential? Implementation of quantitative methods would seem logical and beneficial.

As the system stands at the moment, it generates an extreme amount of data. Appendix II introduces concepts for minimizing the amount of data that is required to represent a trajectory given certain constraints related to the cost of updating the NavAid's position and the cost of the NavAid being out of position. The NavAid system would benefit from research in this area.

However, regardless of the limitations mentioned, this phase of the research has produced convincing evidence that addresses the first two questions posed of this research. That is, it is possible to take technology used for pedestrian navigation systems and apply them in wildlife tracking applications such as the tracking of grizzly bear. Secondly, analysis of the output from that NavAid has enabled the animal's movement to be classified into general groups consisting of no movement, foraging, searching and locomotion behaviours.

Chapter 6

Contribution, Discussion and Conclusions

This final chapter provides a summary of the key findings and contributions regarding the objectives and questions asked of this research. This works point of origin has been the status of current GPS-based animal tracking techniques and the limitations that they present for wildlife management. Through the adaptation of recent advances in pedestrian navigation systems, an animal tracking tool has been developed to aid the understanding of grizzly bear behaviour and habitat use.

The basic premise behind this research has been that good planning must rest on a foundation of good science that both government and the public trust. The quality of one's scientific work hinges on the quality of the available data. As mentioned in the opening chapter, this is not a criticism. It is a reality that all researchers must face, and as noted by Stauffer (2002), despite having powerful analytical tools at our finger tips, there are limits to the precision of models developed in the wildlife sciences because of the noise inherent in the processes under investigation[1].

The results of this work provide a new type of data that has not been previously possible within the wildlife management domain. With the NavAid that has been developed during this research, we can now begin to provide information about an animal's behaviour as it moves throughout its domain. We are now also able to say where an animal has not been. From a modelling perspective this means that scientists can develop models with greater certainty, if for no other reason than an improved understanding of the boundary conditions of a proposed model. It means that it is now possible to associate animal behaviour with habitat use, which has previously not been possible with GPS alone.

From this perspective, the objective to develop a tool that can assist the scientific community to provide better, more complete data regarding animal movement has been achieved. For example, we have shown that animals do react differently to stimuli when exhibiting different locomotion behaviours (see chapter four), hence a portion of the noise that is observed in the global process can be partitioned off, or accounted for, by each of the behaviours. This will result in a smaller portion of unexplainable information, i.e., noise.

Many ecological, biological, or economic processes operate either through diffusion, exchange and transfer, interaction or dispersal[2] and occur over space and

---

1 Stauffer, "Linking Populations and Habitats: Where have we been? Where are we going?"

2 Haining, *Spatial Data Analysis: Theory and Practice* pg. 21

time. In order to understand their regulation and dynamics we must measure their rates, their results, and the factors that influence them. As stated by Huston (2002), in order for these processes to be meaningful, we must measure them at the appropriate scale[3]. But what is that scale? I believe that the NavAid can assist in providing answers to this question with respect to wildlife research. It provides data that can enable analysis at a very fine scale, both spatially and temporally; hence studies that use this data need not be limited by the spatial or temporal resolution of existing animal tracking methods. This will provide researchers with opportunities to undertake modelling exercises that cover the full range of a species' activity, whether they are at daily, monthly or seasonal scales. It will allow researchers to clearly identify ecological barriers, should they exist, in what otherwise would have to be treated as homogeneous units. An example of this ability to differentiate might be the identification of a patch of habitat that has been classified as a preferred habitat for a particular species, but that falls within the range of a predator of that species. This would likely result in a repulsion effect rather than the expected attraction effect.

So to this end, we can conclude that the primary objective of this research has been met. We have developed the NavAid and undertaken field tests on grizzly bears. The NavAid provides data that has previously not been available to wildlife researcher. Using the data from the NavAid we have generated continuous paths of

3 Huston, "Critical Issues for Improving Predictions,"

an animal via step detection methods using GPS to constrain the extent of each path.

Given the paths derived from the NavAid data, we have been able to categorize the animal's route according to different locomotion behaviours. While we have shown that this is possible with just GPS data, we also have persuasive evidence to support the hypothesis that two- or four-hour sampling intervals do not provide sufficiently detailed information necessary to quantify more than two behaviours, nor does the GPS data provide a range of movement rates that could be expected of a grizzly bear (maximum observed movement rate from GPS only was 59 m/minute. The NavAid data suggests that movement rates of up to 590[4] m/minute (≅35 km/h) are possible, whereas when using GPS data alone, the maximum movement rate was 59 m/minute (≅3.5 km/h). Hence, I believe that there is strong support that the NavAid does enable locomotion behaviour to be associated with an animal's movement path.

With regards to the third question posed of this research regarding grizzly bear selection based on locomotion behaviour, I believe that we have produced convincing evidence that this hypothesis is correct. However, it must be tempered by the fact that the analysis has only been carried out on one animal to date, and it is likely that differences in movement rates will be observed when tracking different animals depending on their age, sex, and reproductive state.

---

4 There were five steps with velocities of 590 m to 595 m per minute.

Recommendations

Given the work that has been completed, the most important recommendation that can be made that will further the quality of the data obtained from the NavAid is the need for field observations for a period of time while the NavAid is on a grizzly bear. A primary assumption of predictive discriminant analysis is that its initial groups are well defined, that is, the group membership of each observation used in the analysis is correct. For reasons related to hardware failure and logistics, field observation was never possible. As such, we run the risk of the accelerometer signal of certain behaviours being assigned to the wrong initial group, which will reduce prediction accuracy.

The influence of initial misclassification can not be readily determined. Although the cost of misclassification when there are only two groupings can generally be estimated by undertaking a sensitivity analysis using different priors, it would not appear from the literature that this is an effective solution when there are more than two groups.

Principal components analysis, cluster analysis, or factor analysis are also options that can be investigated to assess the initial classification, however we run the risk of not being able to easily interpret the constructs that result from such analysis.

The remaining recommendations that I wish to make are more points that I believe others who wish to take on this type of work should be aware of. In all fairness, if I were to start this project over again, I would think twice about using grizzly bears as test animals. From a development perspective, I would recommend using animals that are accessible. The cost of capturing a grizzly bear is high, and the cost to the animal during recapture can also be high. Hence, from a practical, ethical, and economic standpoint, it is likely that any equipment added to the FMFGBP equipment will remain on the animal until all equipment has run its course, or has failed. In addition, the effort that is required to test the NavAid on a grizzly bear is substantial and when combined with a one-year turnaround for testing, development is slow.

Throughout this work we have had excellent support from the FMFGRP and other government and industrial sponsors, and, relative to other research, funding for this work has been quite reasonable, but the cost of development has always outstripped the funding that we have been able to obtain. So, for those who chose to venture down a similar path, ensure that there is sufficient funding to pay for research and development of hardware that is to be deployed, and more importantly, ensure you have sufficient numbers of people with the technical skills necessary to undertake this type of work. Often, with our team of two, we just didn't have enough time in the day to get everything done that was required of a system to be placed in the particular environment that we had chosen. I suspect the lesson here is that when in a university environment, research that requires hardware development is best restricted to the prototype stage. We were, and are still not, set up for production.

Future Work

When I look back at my original plan for this work and compare what has been achieved with what was planned it leads me to think of the often quoted line from John Steinbeck's novel Of Mice and Men[5], the title of which is taken from Robert Burns' poem, To a Mouse[6], "The best-laid plans of mice and men go oft awry …", but given the time it has taken reach this point, the plan suggests there are still avenues to explore for some time to come. Thus, I propose five primary areas of ongoing research.

The first concerns data interpretation. There are clearly a number of signal processing issues that warrant further investigation. I believe that we have proven the general concepts that we set out to investigate regarding the NavAid, but in doing so we have probably raised just as many new questions as have been resolved. To reiterate the conclusions from the analysis of the NavAid data in chapter five, there have been instances where we have not been able to identify steps adequately, which generally results in extremely long strides and excessive movement rates. We need to spend some time investigating why we have been unable to see these steps. Is it an algorithm issue? A hardware issue? A hardware mounting issue? Wavelets have proven to be more effective than traditional digital filters for smoothing the accelerometer signals, but have we found the most efficient wavelet? Equally, we need to investigate methods that can improve

5 John Steinbeck, *Of Mice and Men*, *Penguin Classic* (Penguin, 1938)
6 Robert Burns, "To a Mouse," (John Wilson, 1786)

heading accuracy. This work may be best served by returning to the laboratory to simulate grizzly bear motion.

The second direction concerns hardware development. One of the biggest issues that must be recognized when tracking wildlife is the risk of equipment failure. As stated earlier, wild animals are expensive to capture and monitor. They have a tendency to remove equipment that is put on them, or if not, break it. We need a method of accessing the data, and ideally processed data, without having to first remove the NavAid from an animal. This implies the need for wireless communications. However, in order to add a communication link we must first address the increased demand for power that the addition of a communication system would necessitate. Perhaps this can best be resolved via a solar power addition. We currently use our own GPS receiver for controlling timing of all data streams, but it would be beneficial if this component could be extended so that the NavAid could be made a standalone unit. Lastly, as we have strived to do throughout this work, we must continue to make it smaller and lighter, as this will make it more attractive for use on a wider range of animals.

The third area of research concerns image analysis. At present, we acquire one image every 15 minutes during daylight hours. This implies we could acquire more than 10,000 images per grizzly bear per season. Field trials during the winter of 2006 showed that the NavAid could acquire in excess of 20,000 images before its battery dies. The imagery provides context as to why an animal might be in a certain location, but it is not practical to manually assess the imagery. Hence, research into algorithms for the automatic classification of image structure is required so that we can make use of all the information that we are acquiring.

The fourth area of research relates to data. From a representational perspective what is the most efficient and effective means of presenting the data. We are producing trajectories; this naturally implies two, or three space dimensions, and time. At this stage, processed data is stored as points. It would be better represented as three (or four) dimensional lines. The question then becomes how often does one need to add a new segment to ensure that the representation is correct to within some acceptable tolerance. Appendices I and II provide some background into this issue. The concepts reviewed in these appendices are from the moving object and spatial database literature. That work has focused on network based solutions with known origin and destination points. With wildlife, we may know the origin, the animal's current position, but its destination is unknown. However, the moving object literature shows that when using a network, the amount of data required to know where the object is on the network at any point in the past, or the near future, can be reduced by as much as 85% when first order equations are used to represent trajectories of objects instead of points. It would be of considerable interest to apply these concepts to cost or risk surfaces developed for wildlife.

The final area of research is concerned with spatial analysis. The NavAid produces a lot of data from which valuable information can, and must be drawn. Because of the level of detail that is acquired, it is expected that the effect of spatial dependence will be substantial when using traditional analysis techniques. Hence generalized least squares utilizing a weighting function that describes the spatial dependence will be imperative. With this level of detail it is not unreasonable to suppose that spatial dependence will also differ over time. We already know that grizzly bear feeding habits change on a seasonal basis. To what level does this affect spatial dependence? Does it matter if it does?

Given the results of chapter two there is also an obvious need for further investigation of animal selection preferences based on behaviour. In a sense, we can now say that in terms of an animal's use of its environment, for the period of time that an animal is tracked, we acquire the population set of its movements. Consequently, now that we know where an animal has been and where it hasn't, we feel that a confirmatory approach, using factor analysis and structural equation modelling in order to test behavioural hypotheses, would also be an interesting direction for future research.

Appendices

Appendix I

An Integrated Approach for the Analysis and Visualization of Moving Objects

The body of work contained in Appendix I was presented as a workshop paper at the workshop for Mobile Geospatial Augmented Reality, Banff, 2006, see (Hunter, et al., 2006). Appendix I has been included in the thesis because it originally formed part of the body of the thesis, however, time did not permit the assessment of the concepts reviewed in this section.

Introduction

Traditional Geographic Information Systems (GIS) assume a world that only exists in the present[1]. By including a temporal component in a GIS we may ask additional types of questions regarding spatial data that may benefit from visualization for interpretation. For example, we may wish to understand how a particular species utilizes different types of environments during different periods

---

1 D. J. Peuquet, "Time in GIS and Geographical Databases," in *Geographical Information Systems: Principles, Techniques, Management, and Applications*, ed. P. A. Longley, et al. (New York: John Wiley & Sons, Ltd., 1999)

(i.e., what seasonal changes does a species exhibit) or how human activity might affect the social structure of a species, as any disruption in social behaviour of a species may affect the social structure of a population and ultimately reproduction[2].

There are many reasons why species interactions occur. Interactions may be due to home range overlap[3], such as the common use of travel corridors or high-quality habitats. Of interest also are interactions that involve mutual attraction in which animals are actively interacting for mating, or other biological reasons (family groups), in which the closeness of location cannot be explained by overlap of frequently utilized areas[4]. As is evidenced by many studies of moving objects within domains such as transportation[5], land use simulation[6] and Time-Geography analysis[7], the investigation of movement-behaviour patterns over time is a complex issue that requires the examination of many competing dimensions such as location, time, temporal order and object activity/behaviour[8]. To simplify the

---

2 G. B. Stenhouse et al., "Grizzly Bear Associations along the Eastern Slopes of Alberta," *Ursus* 16, no. 1 (2005)

3 C. P. Doncaster, "Non-parametric Estimates of Interaction from Radio-tracking Data," *Journal of Theoretical Biology* 143 (1990)

4 Stenhouse et al., "Grizzly Bear Associations along the Eastern Slopes of Alberta,"

5 C. R. Bhat, "Work Travel Mode Choice and Number of Non-work Commute Stops," *Transportation Research Part B: Methodological* 31, no. 1 (1997)

6 R. Itami et al., "RBSim 2: Simulating the Complex Interactions between Human Movement and the Outdoor Recreation Environment," *Journal of Nature Conservation* 11 (2003)

7 R. N. Buliung and P. S. Kanaroglou, "On Design and Implementation of an Object-relational Spatial Database for Activity/Travel Behaviour Research," *Journal of Geographic Systems* 6 (2004); T. Hägerstrand, "What about People in Regional Science?" *Papers of the Regional Science Association* 24 (1970); J. Maken, R. G. Healey, and S. Dowers, "Simulation Modelling with Object-oriented GIS: A Prototype Application to the Time Geography of Shopping Behaviour," *Geographical Systems* 4, no. 4 (1997)

8 M-P. Kwan, "Interactive Geovisualization of Activity-Travel Patterns using Three-Dimensional Geographical Information Systems: A Methodological Exploration with a large data set," *Transportation Research Part C* 8 (2000)

analysis process researchers have typically focused on a limited subset of available dimensions[9], or have adopted multivariate methods to develop general activity/behaviour patterns from a large number of indicator variables[10].

As described in Kwan (2000), while these methods have enhanced our understanding of movement and behaviour, they are also limited: few methods can cope with real world problems as the spatial component is typically reduced to a one-dimensional problem (i.e., distance along a network) and many of the analytical methods require the discretization of what is essentially a continuous process[11].

As GeoComputational techniques and technology advance in an effort to better represent real world phenomena, more complexity is being built into proposed models[12]. This complexity requires effective methods for exploring new complex data[13]. GIS based geovisualization techniques such as fly-throughs have

---

9 Bhat, "Work Travel Mode Choice and Number of Non-work Commute Stops," , K. G. Goulias, "Longitudinal Analysis of Activity and Travel Pattern Dynamics using Generalized Mixed Markov Latent Class Models," *Transportation Research Part B: Methodological* 33, no. 8 (1999)

10 L. J. E. Fernández, J. de Cea Ch., and A. Soto O., "A Multi-modal Supply–demand Equilibrium Model for Predicting Intercity Freight Flows," *Transportation Research Part B: Methodological* 37, no. 7 (2003); J. Zhang and A. Fujiwara, "Representing Household Time Allocation Behavior by Endogenously Incorporating Diverse Intra-household Interactions: A Case Study in the Context of Elderly Couples," *Transportation Research Part B: Methodological* 40, no. 1 (2006)

11 Kwan, "Interactive Geovisualization of Activity-Travel Patterns using Three-Dimensional Geographical Information Systems: A Methodological Exploration with a large data set,"

12 E. Olmedo et al., *From Linearity to Complexity: Towards a New Economics* (Monash University, 13 April 2002 [cited Submitted preprint under review); available from http://www.csu.edu.au/ci/draft/olmedo02/; S. Openshaw, "GeoComputation," in *GeoComputation*, ed. S. Openshaw and R. J. Abrahart (London: Taylor and Francis, 2000); R. R. Parwani, *Complexity: An Introduction* (Monash University, 2002 [cited Submitted preprint under review 2006]); available from http://www.csu.edu.au/ci/draft/parwan01/

13 P. Ahonen-Rainio and Kraak. M-J., "Deciding on Fitness for Use: Evaluating the Utility of Sample Maps as an Element of Geospatial Metadata," *Cartography and Geographic Information Science* 32, no. 2 (2005); D. Guo et al., "Multivariate Analysis and Geovisualization with an Integrated Geographic Knowledge Discovery Approach," *Cartography and Geographic Information Science* 32, no. 2 (2005); E. McCormack, "Using a GIS to Enhance the Value of Travel Diaries," *Institute of Transportation Engineers Journal* 69, no. 1 (1999)

been available for sometime. But while the viewer is able to move through an environment, the objects represented within the environment typically remain static, limiting the usefulness of fly-throughs for mobile objects. Kernel-based representations[14] and time-space aquariums[15] have been used to represent dynamic phenomena, but are limited in the number of dimensions that can be modeled clearly at one time. Alternative techniques such as the family of Time Plots[16] used to visualize spatial patterns within a time based frame, rather than a traditional space based framework, work well for the analysis of mobile objects, but require further development to improve ease of interpretation.

Geospatial augmented reality produces composite scenes of real world scenes and computer generated scenes, the objective of which is to aid the viewer's sensory perception of the overall scene. The application of augmented reality to the study of grizzly bear movement and behaviour will enable researchers to gain insight into interactions between mating animals, family groups and between grizzly bears and human beings that are as yet not possible. However, it is my belief that the primary limitation of current visualization techniques is the data structures that are typically used to represent spatial features within a GIS[17]. As such, the remainder of this Appendix will review a data model currently being investigated for the

---

14 A. Gatrell, "Density Estimation and the Visualization of Point Patterns," in *Visualization in Geographical Information Systems*, ed. H. M. Hearnshaw and D. J. Unwin (New York: John Wiley & Sons, Ltd., 1994); Silverman, *Density Estimation for Statistics and Data Analysis*

15 Hägerstrand, "What about People in Regional Science?"

16 S. Imfeld, "Time, Points and Space - Towards a Better Analysis of Wildlife Data in GIS" (University of Zurich, 2000)

17 T. Abraham and J. F. Roddick, "Survey of Spatio-Temporal Databases," *GeoInformatica* 3, no. 1 (1999); Laube, "A Classification of Analysis Methods for Dynamic Point Objects in Environmental GIS"

implementation of dynamic objects within a GIS environment and will simplify the integration of dynamic data with geospatial augmented reality.

Moving Objects Databases

As a result of increased recognition of the relationships between temporal and spatial data, there has been increasing demand for the integration of database support for both spatial and temporal data[18]. Spatio-temporal databases deal with spatial objects that change over time (for example, they move, they grow, or they alter their shape): objects may be cars, planes, people, animals, storms, lakes, forests, etc. Such databases allow queries about objects that change their spatial position and their attributes over time to be addressed.

While database theory exists for temporal databases, and some commercial Database Management Systems (DBMS) are starting to adopt temporal theory, spatial databases typically utilize a static, or snapshot, view of time[19], thus reasoning about motion and change is limited. Most commercial applications that “model” moving objects use point-location management techniques[20], where for each moving object a location-time point is maintained within a database and the coordinates of the point are updated periodically. This method has three primary drawbacks; the first is that if the database is queried regarding a particular time

---

18 M. Erwig and M. Schneider, "A Visual Language for the Evolution of Spatial Relationships and its Translation into a Spatio-Temporal Calculus," *Journal of Visual Languages and Computing* 14 (2003)

19 A. Frank, "Socio-Economic Units: Their Life and Motion," in *Life and Motion of Socio-economic Units*, ed. A. Frank, J Raper, and J-P. Cheylan (New York: Taylor and Francis, 2001)

20 O. Wolfson, "Moving Objects Information Management: The Database Challenge " (paper presented at the 5th Workshop on Next Generation Information Technologies and Systems (NGITS'2002), Caesarea, Israel, June 25 – 26 2002)

and no specific update is available for that time, then interpolation, or perhaps extrapolation for future times, is required, which may return results that do not match reality. Secondly, point-location management techniques generally require some form of precision/resource trade-off, meaning a more precise picture of an object's movement requires more resources, such as wireless bandwidth, data storage capacity and processing power. Lastly, the technique often leads to cumbersome software development[21].

Within spatio-temporal literature there are three main approaches[22] to representing dynamic objects within a database. Güting et al. (2000) have developed a schema based on abstract data types that enables one to model and query histories of movement or evolution of spatial objects over time[23]. In essence, data types and operators that allow for the representation and analysis of both time and space have been developed for a Relational Database Management System (RDBMS). Wolfson (2002) and Sistla et al. (1997) have developed a model for management of current and near future movements for transportation and location based services[24]. The model makes use of *a priori* information about the origin and destination of an object, the network upon which the moving objects navigate, and information about the dynamics of objects within the database. In an effort to

---

21 Ibid.

22 R. H. Güting and M. Schneider, *Moving Objects Databases*, ed. J. Gray, *Data Management Systems* (San Francisco: Morgan Kaufmann Publishers, 2005)

23 Güting et al., "A Foundation for Representing and Querying Moving Objects,"

24 P. Sistla et al., "Modeling and Querying Moving Objects" (paper presented at the Thirteenth International Conference on Data Engineering (ICDE13), Birmingham, UK, 1997), Wolfson, "Moving Objects Information Management: The Database Challenge "

make use of use of standard relational data types and operators, Kanellakis et al. (1990) and Kanellakis et al. (1995) introduced Constraint Databases, an abstract model that uses infinite relations to represent geometric entities in a k dimensional space[25]. Data is represented within an arbitrary dimension that removes the need for specialized data types and operators[26].

Because we are interested in understanding the behaviour of grizzly bears, we are interested in questions such as: when do groups of bears meet up in family groups? Where are their day beds? What distance do they travel during sunrise/sunset/mid afternoon? What was the trajectory of bear G217? When did two bears move into Jasper National Park together? How does an animal's home range change over time? Conceptually these questions are depicted in Figure 51. The bear's movement consists of a continuous trajectory that is determined by instantaneous points obtained from GPS. Instantaneous points constrain both temporal durations along the animal's trajectory and temporal vectors derived from motion sensing equipment on the animal. The time duration vectors can then be interpreted as time duration areas, or home ranges for a specific period. As defined by Trajcevski et al. (2004), a trajectory can be represented by a sequence of points[27] $(E_1, N_1, t_1), (E_2, N_2, t_2), \ldots, (E_n, N_n, t_n), (t_1 < t_2 < \ldots < t_n)$. For a given trajectory,

25 P. C. Kanellakis, G. Kuper, and P. Revesz, "Constraint Query Language," *Journal of Computer and System Sciences* 51, no. 1 (1995), P. C. Kanellakis, G. Kuper, and P. Revesz, "Constraint Query Language" (paper presented at the 9th Symposium on Principles of Database Systems (PODS), Nashville, Tennessee, April 2 - 4 1990)

26 P. Rigaux, M. Scholl, and A. Voisard, *Spatial Databases with Application to GIS* (San Francisco: Morgan Kaufmann Publishers, 2002)

27 Trajcevski et al., "Managing Uncertainty in Moving Objects Databases,"

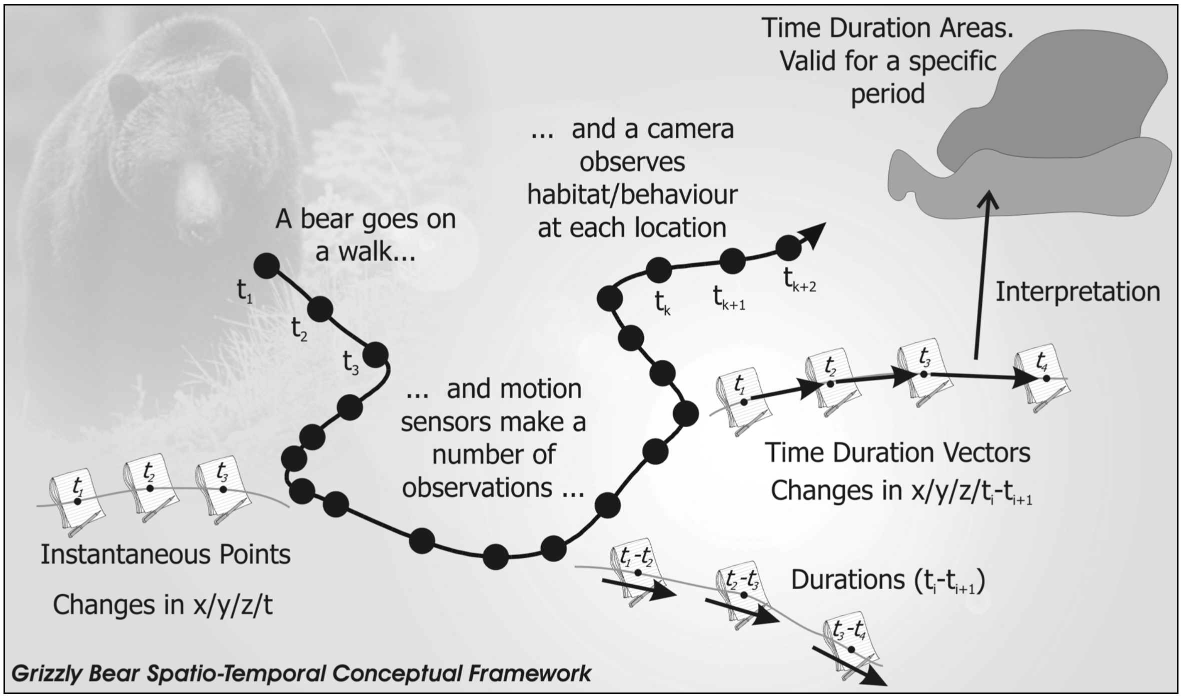


Figure 51: Moving object conceptual framework for grizzly bears

T, its projection into the XY plane defines the route of T. A trajectory defines the moving location of an object as an implicit function of time.

From this conceptual framework a set of Use Case Models[28] describing the primary set of actions required of the system have been developed, two of which are shown in Figure 52. The use case for management of temporal data consists of four actors[29] and three use cases. The actors include the moving object that is being tracked through a spatial domain; motion sensors providing information about the objects locomotion; a user who may be asking questions of the system; and a motion variance controller (MVC) that manages imprecision and uncertainty

[28] G. Booch, J. Rumbaugh, and I. Jacobson, *The Unified Modeling Language User Guide* (Reading: Addison Wesley Longman Ltd., 1999)

[29] A set of roles that a user of use cases plays when interacting with the use cases, Ibid.

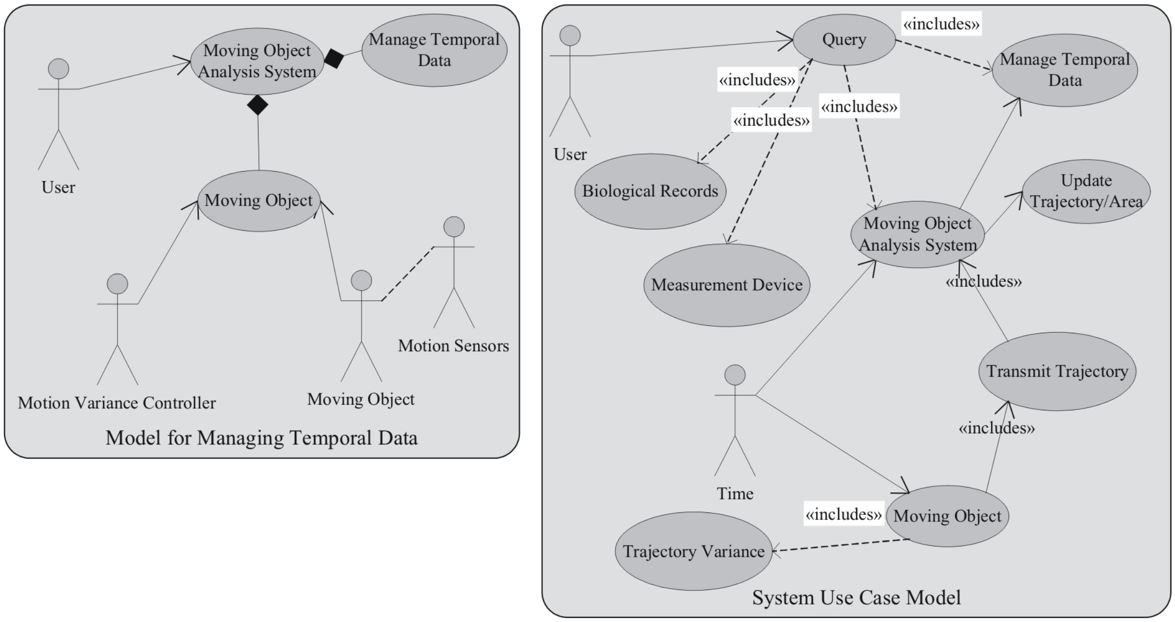


Figure 52: Use case models for grizzly bear space-time data model

within the system. The moving object use case processes the data provided by moving object, motion sensors, and the motion variance controller to provide the moving object analysis system (MOAS) with a trajectory that represents the path of the moving object given the constraints imposed on the system by the MVC. The vector utilizes dynamic attributes[30] that change their value with time automatically. By extending the point data type to include dynamic attributes we are able to represent a moving object with a single point, rather than a series of points, or a set of line segments. Following Sistla et al. (1997), a dynamic attribute[31] A of type T is represented by three sub-attributes, $A.value$, $A.updatetime$, and $A.function$, where $A.value$ is of type T, $A.updatetime$ is a time

30 Sistla et al., "Modeling and Querying Moving Objects" ; Wolfson et al., "Updating and Querying Databases that Track Mobile Units,"

31 Sistla et al., "Modeling and Querying Moving Objects"

value, and $A.function$ is a function such that at time $t=0, f(t) = 0$. For this work $A.value$ is a GPS position, $A.updatetime$ is the time of a GPS fix, and $A.function$ is a formula describing the current trajectory. The value of A at anytime $t \geq A.updatetime$ is:

$$value(A,t) = A.value + A.function(t - A.updatetime) \tag{I.1}$$

The MOAS also consists of a use case Manage Temporal Data that maintains the history of trajectories for each moving object.

The System Use Case Model consists of two actors: a user who accesses the MOAS, and other associated data stores (biological records, measurement devices, etc.), via a query use case; and a time actor that coordinates time throughout the model. Additional use cases in this view of the system include a transmit trajectory use case that is responsible for transmission of data from a moving object to the MOAS, and an update trajectory/area use case which is responsible for updating time duration vectors and their associated time duration areas (refer to Figure 51).

A simplified class diagram, depicted in Figure 53, has been developed from the use case and sequence modelling process. Three components have been developed for data acquisition; locomotion analysis; and representation of space/time trajectories and areas. The moving object class consists of three classes representing the devices that will be used to provide information about the locomotion and behaviour of moving objects, and its associated trajectories.

Similarly, a trajectory is composed of data provided by the locomotion analysis component which includes instantaneous points, from the GPS, which define the limits of each trajectory; motion data describe the movement of an

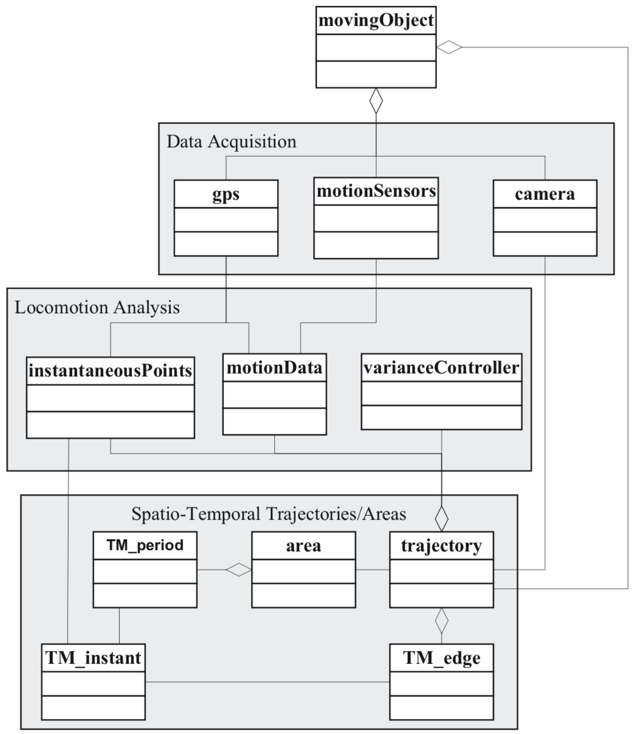


Figure 53: Class diagram of moving object data model

object along the trajectory defined by GPS positions; and a variance controller class that limits the extent of a particular segment within a trajectory. Each trajectory is also composed of a temporal edge, which is a one dimensional topological primitive in time[32]. A temporal edge is defined by an initiation node and a termination node, both of which correspond to temporal instants. A trajectory, or set of trajectories, defines an area, or home range, of a moving object that is limited to a temporal

32 ISO_19108:2002(E), *Geographic Information — Temporal Schema*, vol. ISO 19108:2002(E) (International Organization for Standardization, 2002)

period. A period is an extent in time bounded by beginning and end temporal instants[33]. Two types of behaviour have been identified for describing trajectories. The first behaviour corresponds to the identification of steps taken by an animal along its trajectory. This behaviour will trigger a change in location over time of the animal. The second behaviour represents the inherent imprecision within the system. Each GPS has a specified accuracy based upon the unit's observed single point precision accuracy and the geometry of the satellite constellation at the time of a position fix. In addition the magnetic compass used to determine heading of the animal has a predetermined heading accuracy. Together these error components provide a bounding box on a trajectory, which when traversed will trigger the creation of a new heading within a trajectory. The objective of these behaviours is to represent the sinuosity of a trajectory as accurately as possible, but at the same time minimize the number of headings (line segments) within a trajectory, given the accuracy of the systems.

Summary

This work presents a spatio-temporal data model for moving objects, one goal of which is to improve the visualization of the complex behaviours driving the locomotion of moving objects. The conceptual and logical design of the prototype has been reported for the study of grizzly bear movement and behaviour. The data model has been implemented within an object-oriented framework that utilizes

33 Ibid.

MySQL[34] as the data storage mechanism. By modelling the data in this manner we anticipate that we will be able to more clearly observe and describe the connections between grizzly bear behaviour and the various exogenous inputs that are required to explain behaviour of, and interaction between, grizzly bear and their environment. While validation of this model is still pending, we consider that the incorporation of space and time into the representation of moving objects will open up new opportunities for the integration of spatial data into augmented reality systems.

34 MySQL, "MySQL Reference manual," (2006)

Appendix II

As with Appendix I, this appendix has been included in the thesis because it originally formed part of the body of the thesis, however, time did not permit the assessment of the concepts reviewed in this section.

Location Updates – Balancing Update Frequency and Imprecision

When considering the motion of a NavAid carrier, it is desirable to provide an external observer, a user, with up-to-date information while recognizing the limitations of the NavAid. In effect, the user would like imprecision in the system to be bounded within some limit. However, if high frequency updates, each stride for example, result in limited battery life, the utility of the NavAid may be limited. Hence, one must ask questions about when and how often updates should be provided by the NavAid. High frequency updates may be expensive in terms of cost and overhead performance; infrequent updates result in out-dated information being provided to the user. Consequently, the location of a NavAid is inherently imprecise since the predicted location cannot always be identical to the actual location[1].

Current navigation technology typically uses an ad hoc time-based update threshold. To ensure that imprecision is bounded within predefined limits for a particular application, it is necessary to design the update threshold so that an error does not exceed the design threshold at the high end of the velocity spectrum for that domain. This results in redundant updates whenever the carrier is stationary or moving at a velocity substantially less than the design velocity. In turn, this redundancy results in an increase in processing overhead and power consumption of the NavAid.

An alternative location update policy would be to make use of the deviation (i.e., the difference between the NavAids actual location at time t and its predicted[2] location at time t), and the uncertainty in the location of the NavAid given its current hardware/software/user requirements configuration. For example, if the NavAid was to be carried by a person, uncertainty may be set to 50 m[3] as recommended by the FCC report regarding enhanced 911 emergency calls using handheld communications[4]; if it was to be carried by wildlife, uncertainty may be

1 Güting and Schneider, *Moving Objects Databases* ; D. Pfoser and C. S. Jensen, "Capturing the Uncertainty of Moving-Object Representations" (paper presented at the 6th International Symposium on Advances in Spatial Data (SSD), Hong Kong, China, July 20 - 23 1999)

2 The predicted location is determined from the information provided by the last updated position of the NavAid, and its heading and velocity at that time.

3 The FCC specification states a tighter accuracy standard for handset-based solutions (50 meters for 67 (1 σ) percent of calls) than for network-based solutions (100 meters for 67 percent of calls).

4 Federal Communications Commission, "Revision of the Commission's Rules To Ensure Compatibility with Enhanced 911 Emergency Calling Systems," (Washington, D.C.: Federal Communications Commission, 1999)

set to 250 m. Both deviation and uncertainty can be considered costs in terms of incorrect decision-making[5].

Following Wolfson et al. (1998) and Wolfson et al. (1999), if we assume that the penalty for each unit of deviation during a time unit is weighted by a constant unit of one, then the deviation cost can be defined as[6]

$$\mathrm{COST}_d\left(t_2,t_2\right)=\int_{t_1}^{t_2} d\left(t\right)dt\,, \tag{II.1}$$

where $d\left(t\right)$ describes the deviation as a function of time. If we let C_1 be the update cost (i.e., the cost to recalibrate the NavAid and update its heading and position) defining it as the ratio between the update cost and the cost of a unit of deviation per unit of time, then the NavAid requires $1/\mathrm{C}_1$ updates to reduce the deviation by one during one unit of time.

The cost of uncertainty depends on the size of uncertainty and on the duration for which it lasts. As with (II.1), cost uncertainty, C_2, is a penalty for each unit of uncertainty during a unit of time and can be defined as

$$\mathrm{COST}_u\left(t_1,t_2\right)=\int_{t_1}^{t_2} \mathrm{C}_2 u\left(t\right)dt\,, \tag{II.2}$$

where $u\left(t\right)$ is the uncertainty of the NavAid as a function of time. Now we can define the information cost of a trip, COST_I, over the interval $\left[t_1,t_2\right]$ as

$$\mathrm{COST}_I\left[t_1,t_2\right]=C_1+\mathrm{COST}_d\left[t_1,t_2\right]+\mathrm{COST}_u\left[t_1,t_2\right] \tag{II.3}$$

---

5 O. Wolfson et al., "Cost and Imprecision in Modeling the Position of Moving Objects" (paper presented at the Fourteenth International Conference on Data Engineering (ICDE14), Orlando, FL, 1998)

6 Ibid, Wolfson et al., "Updating and Querying Databases that Track Mobile Units,"

The total information cost of a trip is the sum of COST_I s for every pair of consecutive updates, t_1 and t_2. Hence, if there are $t_0, t_1, t_2, \ldots, t_n$ updates on a route, or trip, then the total information cost of a trip is

$$\mathrm{COST}_I\left[t_0, t_n\right] = \mathrm{COST}_d\left[t_0, t_1\right] + \mathrm{COST}_u\left[t_0, t_1\right] + \sum_{i=1}^{n} COST_I\left[t_i, t_{i+1}\right]. \qquad \text{(II.4)}$$

Cost Optimization for Dead Reckoning

A DR update policy requires that there be some threshold th set within the NavAid. The threshold is checked against the distance between the location of the NavAid and its predicted position. When the deviation of the NavAid exceeds th, the NavAid updates its navigation solution. The objective of a dead reckoning policy is to set a deviation threshold, K, such that the total information cost is minimized[7]. A typical strategy requires that the NavAid predicts the future behaviour and direction of the deviation. This prediction is used to compute the average cost per unit of time between now and the next update as a function, f, of the new threshold, K. Then K is set to minimize f.

As above, let C_1 denote the update cost and C_2 denote the uncertainty cost. If we assume that t_1 and t_2 are consecutive update instances, then the deviation $d(t)$ between t_1 and t_2 is given by some function $a(t-t_1)$ with $t_1 \le t \le t_2$, where a is a positive function, and the uncertainty is fixed at K between t_1 and t_2. If we let $d(t) = a(t-t_1)$ and $u(t) = K$ then from (II.3) we obtain

7 Wolfson et al., "Cost and Imprecision in Modeling the Position of Moving Objects"

$$COST_I\left[t_1,t_2\right]=C_1+\int_{t_1}^{t_2} a\left(t-t_1\right)dt+\int_{t_1}^{t_2} C_2Kdt$$
$$=C_1+\frac{a}{2}\left(t_2-t_1\right)^2+C_2K\left(t_2-t_1\right) \tag{II.5}$$

Let

$$f\left(t_2\right)=\frac{COST_I\left[t_1,t_2\right]}{\left(t_2-t_1\right)} \tag{II.6}$$

denote the average information cost per unit time between t_1 and t_2 at update time t_2. As t_1 and t_2 are consecutive, and because the deviation is equal to the uncertainty threshold at t_2, $d\left(t\right)=u\left(t\right)$, then $K=a\left(t_2-t_1\right)$, by replacing t_2 in $f\left(t_2\right)$ with $\left(K/a\right)+t_1$ we obtain

$$f\left(K\right)=\frac{a\mathrm{C}_1}{K}+\left(0.5+\mathrm{C}_2\right)K\ . \tag{II.7}$$

By taking the derivative of (II.7), the minimum of $f\left(K\right)$ is obtained when

$$K=\sqrt{\frac{2a\mathrm{C}_1}{2\mathrm{C}_2+1}} \tag{II.8}$$

Dead Reckoning Update Policy

Implementation of an adaptive DR policy allows some flexibility in terms of when an update will be performed. For example, when the NavAid is moving at very low velocity or is stationary, the system will require few, if any, updates. However, if the NavAid starts moving rapidly, the update frequency can increase to ensure that a user defined precision threshold is maintained. An adaptive policy would start with an initial (arbitrary) deviation threshold, th_1. The NavAid then tracks the actual deviation and updates the navigation solution when th_1 is exceeded. The DR updates would consist of the location at the time of the update, the velocity and heading of the NavAid, and a new threshold th_2. The new threshold would be computed as follows: let t_1 be the number of time units since the beginning of a trip until the deviation exceeds th_1 for the first time. I_1, the deviation cost, can be calculated from (II.4) for the interval. If we assume that $a_2 = \frac{2I_1}{t_{1-2}}$, then

$th_2 = \sqrt{\frac{2a_1 C_1}{2C_{2+1}}}$, where C_1 is the update cost and C_2 is the uncertainty cost. When th_2 is reached, a similar update is determined where, $th_3 = \sqrt{\frac{2a_2 C_1}{2C_{2+1}}}$, $a_3 = \frac{2I_2}{t_{2-3}}$, and I_2 is the deviation cost from the first update to the second.

Summary

The benefit of developing an update policy such as the one described is twofold. Firstly, it extends the life of the NavAid by conserving battery power due to fewer transmissions of data to an external user, and fewer onboard computations; secondly, it reduces the volume of data that must be stored by an information system in order to recreate the route of a NavAid, by recognizing that dead reckoning solutions contain errors and redundant information.

It is recommended that in order to implement an update policy such as that described above it is necessary to investigate the relative costs C_1, the update cost, and C_2, the uncertainty cost. C_1 will essentially be determined by power consumption of the NavAid to recalibrate the NavAid and update its heading and position. More frequent updates imply higher power consumption and shorter battery life, which relates directly to user satisfaction and utility. C_2 represents the cost to the user of receiving incorrect information, and is largely an operational cost resulting from the user making an incorrect decision, i.e., a Park Ranger opens a trail to tourists when there is a grizzly bear in the vicinity that is unknown to the Ranger. Appropriate weights for C_1 and C_2 are required such that the cost of updating the system is balanced by the cost of providing incorrect information to the NavAid user.

Appendix III

A Review of Predictive Discriminant Analysis

Introduction

Discriminant Analysis is a well known research technique with a long history[1], especially in the biological and medical sciences. But within the Geomatics world, most empirical research is only concerned with outcomes that contain a single variable, and no grouping within the variable, therefore it is not a technique that is discussed regularly. As such the purpose of this appendix is to provide sufficient background information to allow an understanding and implementation of Predictive Discriminant Analysis. Much of the content found in this appendix has been summarized from Huberty and Olejnik (2006).

---

1 See authors such as T. W. Anderson, S. Das Gupta, and G. P. Styan, *A Bibliography of Multivariate Statistical Analysis* (Edinburgh: Olivier and Boyd, 1972) and more recently N. C. Giri, *Multivariate Statistical Analysis* (New York: Dekker, 2004) for a historical perspective.

Predictive Discriminant Analysis

The notions of explanation and prediction are closely aligned. It can be argued that explanation – identification of patterns or of structure – is a prerequisite of prediction. It can equally be argued that the converse is also true. For this work prediction is viewed as a means of enhancing an explanation and as a practical solution to the problem of locomotion identification based on the output of a set of accelerometer sensors.

A common approach to making empirical, or statistical, predictions is multiple regression. This technique is appropriate in situations involving a set of p predictor variables, $X_1, X_2, \ldots, X_p$, that may be random or fixed, and a single criterion (random) variable, Y. Hence, in this situation we are dealing with a single group of N observed units for which there are $p+1$ response measures. The basic goal of multiple regression is to set up a rule, based on an $N \times (p+1)$ data matrix, to be used in predicting, or estimating, an outcome variable, given the observations on the p predictors. In essence, one determines a set of regression weights $b_1, b_2, \ldots, b_p$ corresponding to the set of p predictor variables to give a linear composite value that represents a predicted value of the outcome variable.

Following Montgomery (2001) the model for a particular outcome unit (u) may be represented as[2]:

$$\hat{Y}_u = b_0 + b_1 X_{1u} + b_2 X_{2u} + \ldots + b_p X_{pu} + e_u, \qquad \text{(III.1)}$$

2 D. Montgomery, *Design and Analysis of Experiments* (New York: John Wiley & Sons, Ltd., 2001)

where b_0 is the regression constant. The model may be expressed more compactly as:

$$\hat{Y}_u = b_0 + \sum_{i=1}^{p} b_i X_{iu} + e_u \,, \tag{III.2}$$

or as

$$\hat{Y}_u = b_0 + \mathbf{b}'\mathbf{x}_u + e_u \,, \tag{III.3}$$

where $\mathbf{b}'$ is the $1\times p$ row vector of regression weights, and $\mathbf{x}_u$ is the $p\times 1$ column vector of predictor variable measures for unit u.

When the X measures are based on different metrics, or measurement scales, one can mitigate the effect of varying metrics on regression weights through standardization of the regression weights[3]:

$$b_i^* = b_i \frac{s_i}{s_Y} \,, \tag{III.4}$$

where s_i and s_Y are the estimated standard deviations of X_i and Y. These weights can then be used to predict a standard Y measure,

$$\hat{Z}_Y = \mathbf{b}^{*\prime}\mathbf{z}_u \tag{III.5}$$

where $\mathbf{z}_u$ is a $p\times 1$ vector of standardized X measures.

An alternative approach for making empirical predictions is a form of discriminant analysis called predictive discriminant analysis (PDA)[4]. PDA is appropriate when the outcome variable consists of multiple groups. In this instance we have p X measures for each unit belonging to one of J groups. It is assumed that the J groups of n_j units represent J meaningful populations. With

3 Huberty and Olejnik, *Applied MANOVA and Discriminant Analysis* p. 256
4 Ibid.

such models, the outcome variable may be dichotomous or polytomous. A goal of PDA is to set up a rule based on $J\ n_j \times p$ data matrices that will predict the population membership of a unit. It is assumed that a unit does in fact belong to one of the J outcome populations.

Classification Rules

A PDA classification rule may take one of three forms; it may be a composite of the predictor it measures; it may take the form of an estimated probability of population membership; or it may take the form of a distance from the estimated centroid of a population.

The concept of distance is either explicit or implicit to each of the forms of classification. The Pythagorean Theorem is the standard measure of distance between two points, $A{:}\left(X_{1A}, X_{2A}\right)$ and $B{:}\left(X_{1B}, X_{2B}\right)$, in a Euclidean metric space:

$$\begin{aligned} \tilde{d}_{AB}^2 &= \left(X_{1A} - X_{1B}\right)^2 + \left(X_{2A} - X_{2B}\right)^2 \\ &= \sum_{i=1}^{2} \left(X_{iA} - X_{iB}\right)^2 \end{aligned} \tag{III.6}$$

that may also be expressed as:

$$\left[\mathbf{x}_A - \mathbf{x}_B\right]' \left[\mathbf{x}_A - \mathbf{x}_B\right], \tag{III.7}$$

where $\mathbf{x}_A$ and $\mathbf{x}_B$ are 2×1 column vectors of scores, and $\mathbf{x}_A - \mathbf{x}_B$ is a 2×1 column vectors of differences. This index is appropriate if two conditions are assumed[5]: the first is that X_1 and X_2 are uncorrelated, (i.e., $\rho_{AB} = 0.0$); and that the measures

5 P. C. Mahalanobis, "On the Generalized Distance in Statistics," *Proceedings of the National Institute of Sciences of India* 2, no. 1 (1936) and Carl J. Huberty, "Mahalanobis Distance," in *Encyclopedia of Statistics in Behavioural Science*, ed. B. Everitt and D. C. Howell (London: John Wiley & Sons, Ltd., 2005)

on X_1 and X_2 have unit variances (i.e., $\sigma_A^2 = \sigma_B^2 = 1.0$). This Euclidean metric concept may be easily extended into a p variate (dimensional) space[6]:

$$\tilde{d}_{AB}^2 = \sum_{i=1}^{p} \left(X_{iA} - X_{iB} \right)^2 , \tag{III.8}$$

or

$$\left[\mathbf{x}_A - \mathbf{x}_B \right]' \left[\mathbf{x}_A - \mathbf{x}_B \right], \tag{III.9}$$

where $\mathbf{x}_A$ and $\mathbf{x}_B$ are $p \times 1$ vectors.

Similar to the bivariate case, the p variate case assumes all variables are uncorrelated, with unit variances. That is the $p \times p$ covariance matrix, Σ, is an identity matrix[7]:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1p}\sigma_1\sigma_p \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2p}\sigma_2\sigma_p \\ \vdots & & & \vdots \\ \rho_{p1}\sigma_p\sigma_1 & & \cdots & \sigma_p^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & & \cdots & 1 \end{bmatrix}, \tag{III.10}$$

The basic requirement when comparing distances is that the same metric is used in the computation of the distances. This is typically achieved by ensuring that all standard deviations, or variances, are equal. If this is not the case then the unequal variances must be taken into consideration. In general, because empirical studies deal with variables that are inter-correlated (albeit modestly), these inter-correlations must also be taken into consideration when assessing distances. This

6 Mícheál Ó Searcóid, *Metric Spaces* (London: Springer-Verlag Ltd., 2007) (p. 4) and A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Quantitative Geography. Perspectives on Spatial Data Analysis* (London: Sage Publications Ltd., 2000) (p. 20)

7 Huberty, "Mahalanobis Distance,"

is accomplished by introducing the covariance matrix into the distance measure as follows:

$$D_{AB}^2 = [\mathbf{x}_A - \mathbf{x}_B]' \Sigma^{-1} [\mathbf{x}_A - \mathbf{x}_B] \tag{III.11}$$

which is P.C. Mahalanobis' generalized distance index[8]. D may be used to determine the distance between two population centroids, μ_1 and μ_2, such that equation (III.11) becomes

$$D_{12} = \left([\mu_1 - \mu_2]' \Sigma^{-1} [\mu_1 - \mu_2] \right)^{\frac{1}{2}}. \tag{III.12}$$

In this instance, Σ is the covariance matrix common to both populations, which are assumed to be equal. In addition, D may be used to determine the distance of a measurement unit from one, or all, group centroids. In this instance (III.11) becomes

$$D_{uj} = \left([\mathbf{x}_u - \mu_j]' \Sigma_j^{-1} [\mathbf{x}_u - \mu_j] \right)^{\frac{1}{2}} \tag{III.13}$$

where $\mathbf{x}_u$ is the observed vector for unit u, μ_j is the population centroid, and Σ_j is the covariance matrix for population j.

To summarize, there are three types of distances that one could consider in an analysis: unit to unit (III.11); centroid to centroid (III.12); or unit to centroid (III.13). It is this third type that is of most importance for PDA[9].

8 Mahalanobis, "On the Generalized Distance in Statistics,"

9 Huberty and Olejnik, *Applied MANOVA and Discriminant Analysis* (p.260)

Decision Rules

The basic purpose of PDA is to determine which population a particular observed unit belongs to. Decision rules are commonly based on the principle of maximum likelihood (ML), that is, assign a unit to the population in which the observed vector has the greatest likelihood of occurrence[10]. Mathematically, this may be expressed as:

Assign unit u to population j if

$$P(\mathbf{x}_u \mid j) > P(\mathbf{x}_u \mid j') \quad \text{(III.14)}$$

for $j' \neq j$.

In general, the adequacy of a ML rule is dependent upon the quality of the probability distribution assigned to a population $P(\mathbf{x}_u \mid j)$, and the representativeness of the training samples on which the estimates are based[11]. Therefore, following Huberty and Olejnik (2006) we must consider the relative size of each population when estimating the probability of an observation unit belonging to a particular population. If we let π_j equal the prior probability[12], or the proportion of the units in the total universe that is in population j, we can calculate a posterior probability, $P(j \mid \mathbf{x}_u)$, of membership in population j using Bayes' Theorem. Hence, the posterior probability of unit u belonging to population j, given an observed vector $\mathbf{x}_u$ is

---

10 Kenneth P. Burnham and David R. Anderson, *Model Selection and Multimodal Inference: A Practical Information-Theoretic Approach* (New York: Springer, 2002)

11 Huberty and Olejnik, *Applied MANOVA and Discriminant Analysis* (p.263)

12 D. Wrinch and H. Jeffreys, "On Certain Fundamental Principles of Scientific Inquiry," *Philosophical Magazine* 42, no. 6th Series (1921)

$$P(j \mid \mathbf{x}_u) = \frac{\pi_j \cdot P(\mathbf{x}_u \mid j)}{\sum_{j'=1}^{J} \pi_{j'} \cdot P(\mathbf{x}_u \mid j')} \tag{III.15}$$

Rule (III.14) can therefore be extended to a maximum Bayesian probability rule as follows:

Assign unit u to population j if

$$P(j \mid \mathbf{x}_u) > P(j' \mid \mathbf{x}_u) \tag{III.16}$$

for $j \neq j'$, where $P(j \mid \mathbf{x}_u)$ is defined as in (III.15).

This rule will minimize the total number of misclassification errors[13]. However, in order to utilize rule (III.16) we need to know the distribution parameters, Σ and μ, which generally are unknown and require estimation. Three approaches for determining distribution parameters are possible. The first approach is to specify a theoretical distribution, assume the data fits the distribution, estimate the model parameters using the data, and then construct the classification rule using those estimates. The second approach is to estimate probability density values directly from the data with no prior model specification and construct the rule using those estimates. The third approach is a combination of the first and second using a Bayesian framework to estimate probability density values.

As with many statistical methods, the theoretical distribution often assumed is the normal probability distribution. This distribution is easily extended to the multivariate case as follows

---

13 Huberty and Olejnik, *Applied MANOVA and Discriminant Analysis* (p. 269)

$$f\left(\mathbf{x} \mid j\right)=\frac{1}{\sqrt{\left(2\pi\right)^{p}}\sqrt{\left|\Sigma_{j}\right|}}\exp\left[-\frac{1}{2}\left(\mathbf{x}-\mu_{j}\right)'\Sigma_{j}^{-1}\left(\mathbf{x}-\mu_{j}\right)\right]. \tag{III.17}$$

By substituting D_{uj}^{2} for the distance between an observed unit and the centroid of population J (see (III.13) above), and the estimated mean $\left(\overline{\mathbf{x}}_{j}\right)$ and covariance $\left(\mathbf{S}_{j}\right)$ of population J, (III.17) for unit u can be stated as

$$\hat{f}\left(\mathbf{x}_{u} \mid j\right)=\left(2\pi\right)^{-p/2}\cdot\left|\mathbf{S}_{j}^{-1/2}\right|\exp\left(-\frac{1}{2}D_{uj}^{2}\right). \tag{III.18}$$

If we then substitute $q_{j}=\pi_{j}$ for the prior probabilities (q_{j} is the estimated prior probability), and cancel out $\left(2\pi\right)^{-p/2}$ the posterior probability using the maximum probability rule (III.15), based on a normal distribution, becomes:

$$\hat{P}\left(j \mid \mathbf{x}_{u}\right)=\frac{q_{j}\cdot\left|\mathbf{S}_{j}\right|^{-1/2}\cdot\exp\left(-\frac{1}{2}D_{uj}^{2}\right)}{\sum_{j'=1}^{J}q_{j'}\cdot\left|\mathbf{S}_{j'}\right|^{-1/2}\cdot\exp\left(-\frac{1}{2}D_{uj'}^{2}\right)} \tag{III.19}$$

and the maximum probability rule for the p variate normal case can be expressed as

Assign unit u to population j if

$$\hat{P}\left(j \mid \mathbf{x}_{u}\right)>\hat{P}\left(j' \mid \mathbf{x}_{u}\right) \tag{III.20}$$

for $j \neq j'$, where $\hat{P}\left(j \mid \mathbf{x}_{u}\right)$ is defined as in (III.19).

In terms of classification, the denominator of (III.19) may be ignored such that assignment to a population can be achieved by maximizing

$$q_{j}\cdot\left|\mathbf{S}_{j}\right|^{-1/2}\cdot\exp\left(-\frac{1}{2}D_{uj}^{2}\right) \tag{III.21}$$

which, can be accomplished by taking the natural logarithm of (III.21) to give

$$Q_{uj} = \ln\left(q_j\right) - \frac{1}{2}\ln\left|\mathbf{S}_j\right| - \frac{1}{2}D_{uj}^2. \tag{III.22}$$

Equation (III.22) is quadratic in $\mathbf{x}_u$. Therefore, the maximum probability rule for the p variate normal case may now be expressed as:

Assign unit u to population j if

$$Q_{uj} > Q_{uj'} \tag{III.23}$$

for $j \neq j'$, where Q_{uj} is defined as in (III.22).

An equivalent alternative to (III.19) is to use a distance-based classifier. In this instance a unit u would be assigned to the population centroid that is closest according to (III.13), or $D_{uj}^2 = \left[\mathbf{x}_u - \overline{\mathbf{x}}_j\right]' \mathbf{S}_j^{-1}\left[\mathbf{x}_u - \overline{\mathbf{x}}_j\right]$. Maximizing (III.22) is equivalent to minimizing

$$\begin{aligned} d_{uj} &= -2Q_{uj} \\ &= \ln\left|\mathbf{S}_j\right| + D_{uj}^2 - 2\ln\left(q_j\right) \end{aligned} \tag{III.24}$$

If the J population covariance matrices are equal, $\Sigma_1 = \Sigma_2 = \cdots = \Sigma_j = \Sigma$, then the estimator for Σ is the error covariance matrix, $\mathbf{S}_e$. In this instance, $\left|\mathbf{S}_j\right|$ and $\left|\mathbf{S}_{j'}\right|$ from (III.18), (III.19), (III.21) and (III.22) are equal, and they cancel out of these equations. Lastly, if equal prior probabilities are also imposed, (III.18), (III.19), (III.21) and (III.22) can be simplified further by the removal of q_j and $q_{j'}$.

Hence, to summarize classification statistics, when restricted to normal based rules, the general statistic involves estimating the posterior probability $\hat{P}\left(j \,|\, \mathbf{x}\right)$ in (III.23). Two special cases may also be considered: if the J Population covariance matrices are not equal, then the quadratic rule (III.20), must be used, otherwise a linear rule may be employed using $\ln\left(q_j\right)-\frac{1}{2}D_{uj}^{*2}$, where $D_{uj}^{*2}=\left(\mathbf{x}_u-\overline{\mathbf{x}}_j\right)'\mathbf{S}_e^{-1}\left(\mathbf{x}_u-\overline{\mathbf{x}}_j\right)$; the second case refers to prior probabilities — if the priors are considered to be equal then q_j or $q_{j'}$ need not be considered. These conditions are summarized in Table 27 following.

Table 27: Alternative forms of classification statistics

	Covariance Matrices[14]	
Prior Probabilities	*Unequal (Quadratic Rule)*	*Equal (Linear Rule)*
Unequal	$\hat{P}\left(j\mid\mathbf{x}_u\right)=\dfrac{q_j\cdot\left\lvert\mathbf{S}_j\right\rvert^{-1/2}\cdot\exp\left(-\frac{1}{2}D_{uj}^2\right)}{\sum_{j'=1}^{J}q_{j'}\cdot\left\lvert\mathbf{S}_{j'}\right\rvert^{-1/2}\cdot\exp\left(-\frac{1}{2}D_{uj'}^2\right)}$ $Q_{uj}=\ln\left(q_j\right)-\frac{1}{2}\ln\left\lvert\mathbf{S}_j\right\rvert-\frac{1}{2}D_{uj}^2$ $d_{uj}=\ln\left\lvert\mathbf{S}_j\right\rvert+D_{uj}^2-2\ln\left(q_j\right)$	$\hat{P}\left(j\mid\mathbf{x}_u\right)=\dfrac{q_j\cdot\exp\left(-\frac{1}{2}D_{uj}^{*2}\right)}{\sum_{j'=1}^{J}q_{j'}\cdot\exp\left(-\frac{1}{2}D_{uj'}^{*2}\right)}$ $Q_{uj}=\ln\left(q_j\right)-\frac{1}{2}D_{uj}^{*2}$ $d_{uj}=D_{uj}^{*2}-2\ln\left(q_j\right)$
Equal	$\hat{P}\left(j\mid\mathbf{x}_u\right)=\dfrac{\left\lvert\mathbf{S}_j\right\rvert^{-1/2}\cdot\exp\left(-\frac{1}{2}D_{uj}^2\right)}{\sum_{j'=1}^{J}\left\lvert\mathbf{S}_{j'}\right\rvert^{-1/2}\cdot\exp\left(-\frac{1}{2}D_{uj'}^2\right)}$ $Q_{uj}=-\frac{1}{2}\ln\left\lvert\mathbf{S}_j\right\rvert-\frac{1}{2}D_{uj}^2$ $d_{uj}=\ln\left\lvert\mathbf{S}_j\right\rvert+D_{uj}^2$	$\hat{P}\left(j\mid\mathbf{x}_u\right)=\dfrac{\exp\left(-\frac{1}{2}D_{uj}^{*2}\right)}{\sum_{j'=1}^{J}\exp\left(-\frac{1}{2}D_{uj'}^{*2}\right)}$ $Q_{uj}=-\frac{1}{2}D_{uj}^{*2}$ $d_{uj}=D_{uj}^{*2}$

14 Adopted from Ibid. , pp. 271 - 278

Predictive Discriminant Analysis Assumptions

When undertaking PDA two data conditions are of concern. The first is that multivariate normality of the observation vectors exists. A number of authors (Cox and Small, 1978, Fan, 1996, Rencher, 2002) discuss numerical and graphical techniques for assessing multivariate normality[15]. The most common techniques for determining univariate normality are the Q-Q plot, skewness and kurtosis. Multivariate normality can be assessed by development of a Q-Q plot of D^2, scatter plots of all pairs of variables, or generalizations of skewness and kurtosis to multiple dimensions. The second condition, or assumption, relates to the equality of covariance matrices. This condition can be tested using the Box test for covariance homogeneity. The Box test compares the log-transformed determinants of the covariance matrices. The determinant represents the generalized variance of a covariance matrix $\mathbf{S}$. Similar values imply similar variability within a set of data. Taking the natural logarithm of $|\mathbf{S}|$ allows covariance matrices to be assessed using a χ^2 distribution. If sample sizes, and or covariance matrices are not equal, the resulting estimates of the effects of measured covariates may be inefficient or biased[16] and reported P values may underestimate, or overestimate the actual P

15 D. R. Cox and N. J. H. Small, "Testing Multivariate Normality," *Biometrica* 65, no. 2 (1978), Xitao Fan, "An SAS Program for Assessing Multivariate Normality," *Educational and Psychological Measurement* 56, no. 4 (1996), A. C. Rencher, *Methods of Multivariate Analysis* (New York: John Wiley & Sons, Ltd., 2002)

16 See Jay Teachman et al., "Covariance Structure Models for Fixed and Random Effects," Sociological Methods and Research 30, no. 2 (2001); William H. Greene, Econometric Analysis, 5th ed. (Upper Saddle River, New Jersey: Prentice Hall, 2002); or Cheng Hsiao, Analysis of Panel Data, 2nd ed. (New York: Cambridge University Press, 2003) (p. 11, 72)

value. In reality, this does not affect PDA, as the statistics described in Table 27 allow us to take these conditions into consideration.

If equal covariances are untenable, it is recommended that a quadratic rule be adopted as this should lead to higher cross-validation hit accuracy[17]. However, it has been reported by Meshbane and Morris (1995) that differences between linear and quadratic rules are often not statistically different, in particular when the sample size is relatively low compared to the number of groups being analyzed[18]. But if the groups are multivariate normal and there are numerous groups and or observations, a quadratic rule will tend to perform better[19].

Once the first and second conditions have been assessed, one should undertake tests to determine if the set of predictor variables are capable of identifying groups within the outcome variable. The most common approach for assessing individual predictor variables is by means of a one-way Analysis of Variance. Multivariate assessment is undertaken using tests such as Wilks Criterion, the Bartlet-Pillai Criterion, Roy Criterion, and the Hotelling-Lawley Criterion.

Selection of prior probability should reflect how likely it is for an observation to come from a particular group. Huberty and Olejnik (2006) suggest that priors

---

17 See T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd ed. (New York: John Wiley & Sons, Ltd., 1984) (p.235)) and Alice Meshbane and John D. Morris, "A Method for Selecting Between Linear and Quadratic Classification Models in Discriminant Analysis," *Journal of Experimental Education* 63, no. 3 (1995)

18 Meshbane and Morris, "A Method for Selecting Between Linear and Quadratic Classification Models in Discriminant Analysis,"

19 Paul A. Rubin, "A Comparison of Linear Programming and Parametric Approaches to the Two-Group Discriminant Problem," *Decision Sciences* 21, no. 2 (1990)

should correspond to relative sample size only if a proportional sampling plan has been utilized, or if equal sample sizes are involved.

Classification Results

The predominant method of group assignment used by software packages is the estimated posterior probability of group membership. For example, the SAS[20] software package outputs the J normal based posterior estimates, $\hat{P}(j\,|\,\mathbf{x})$, for each observation unit. By examining the J $\hat{P}(j\,|\,\mathbf{x})$ values, one can assess, probabilistically, the closeness of each observation to the centroid of each of the J groups.

When the largest estimated posterior probability is assigned to the correct group, the assignment can be considered a “hit”, if not it is a “miss”, or error. An in-doubt observation is an observation in which two or more groups are assigned approximately equal $\hat{P}(j\,|\,\mathbf{x})$ values. This implies that an observed vector is approximately the same distance from the centroids of the groups. Fence sitters may sometimes be used to identify why some group members may resemble typical members of other groups. If there are large numbers of in-doubt units, this may suggest the actual existence of an additional group in between existing groups.

Some software packages, for example SPSS[21], also output a typicality probability, $\hat{P}(\mathbf{x}_u\,|\,j)$, which is the typicalness, or probability, of observation u belonging to group j. These probabilities are derived from a (Linear) Discriminant

20 SAS/STAT® is a statistical software package, SAS Institute, 100 SAS Campus Drive, Cary, North Carolina 27513-2414, U.S.A. See http://www.sas.com.

21 Typicality probabilities are produced by the SPSS DISCRIMINANT procedure (http://www.spss.com/).

Function (LDF) determined from the data. An LDF is a linear function that is essentially a set of weights (similar to β coefficients in linear regression) that maximize the correlation between the LDF (Z), consisting of multiple predictor variables, and the outcome variable. The purpose of a LDF is to maximize the ratio of the between-groups sum of squares to the within-group sum of squares[22].

The general model is as follows:

$$Z = b_1Y_1 + b_2Y_2 + \ldots + b_pY_p \,. \qquad \text{(III.25)}$$

An LDF describes the squared distance between a particular unit and the centroid of a group, which in a LDF space has a chi-squared distribution with $\mathrm{df} = J - 1$[23]. This statistic can then be used to help identify potential outliers.

Hit Rate Estimation

Hit rate estimation can be undertaken either internally, or externally. An internal classification organizes the observed units based on the parameters obtained from the study samples. As with least squares regression equations, the LDFs are fit to the data such that they maximize the predictive power of the classification rule for the particular sample being studied, consequently, internal hit rate estimation will tend to be overly optimistic[24].

---

22 Peter A. Lachenbruch, "Discriminant Diagnostics," *Biometrics* 53, no. 4 (1997)

23 See R. A. Fisher, "The Use of Multiple Measures in Taxonomic Problems," *Annals of Eugenics* 7 (1936) and R. J. McKay and N. A. Campbell, "Variable Selection Techniques in Discriminant Analysis: Part I Description," *British Journal of mathematical and Statistical Psychology* 35, no. 1 (1982)

24 See Ronald E. Frank, William F. Massy, and Donald G. Morrison, "Bias in Multiple Discriminant Analysis," *Journal of Marketing Research* 2, no. 3 (1965)and G. J. McLachlan, "The Bias of Sample Based Posterior Probabilities," *Biometrical Journal* 19, no. 6 (1977)

With external hit rate estimation, the classification rule is determined with a subset of the data, and then used to classify the remaining data. Typically, results are presented in the form of a classification table. The training subset may consist of 25-35 percent of the total sample[25]. There are, however, several drawbacks associated with this technique[26]. In order to withhold data from classification rule estimation a large sample is required, which may not be possible. The rule that is used to classify the data is not based on the complete set of data, and, therefore may not represent the full data set adequately. If the test subset is large, a good assessment of the classification rule will be obtained, even though the rule itself may be poor, whereas a small test sample should result in a better classifier, but results may be quite variable. Lastly the method can be economically impractical if the cost of obtaining a sample is high, as the method requires a sample size that is larger than that which is required to develop a good classification rule.

Alternatively, a Leave-one-out (LOO) method may be used in which $N-1$ observations are used to estimate an LDF, and then the LDF is used to classify the remaining observation. The process is then repeated N times, and the proportion of units correctly classified are used as the hit rate estimate[27]. A limitation of this technique is that it has been shown to produce somewhat variable

25 W. Schaafsma and G. N. van Vark, "Classification and Discrimination Problems with Applications - Part IIa," *Statistica Neerlandica* 33 (1979)

26 See Peter A. Lachenbruch and M. Ray Mickey, "Estimation of Error Rates in Discriminant Analysis," *Technometrics* 10, no. 1 (1968)

27 See Peter A. Lachenbruch, "An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis," *Biometrics* 23, no. 4 (1967)

results[28], which is attributable to the high reuse of the original data — N sets of J LDFs derived from nearly identical data.

A third option is the Maximum Posterior Probability (MPP) method[29]. MPP is essentially the mean of the estimated posterior probabilities for all observation units assigned to population j by the classification rule used. The MPP estimator for $P_j^{(a)}$ is

$$\hat{P}_j^{(a)} = \frac{1}{N \cdot q_j} \sum_{j'=1}^{J} \left\{ \sum_{u=1}^{nj'} \begin{bmatrix} \text{post. prob. for all } \mathbf{x_u} \text{ in} \\ \text{Group } j' \text{ assigned to Group } j \end{bmatrix} \right\}, \qquad \text{(III.26)}$$

And the total group hit rate, $P^{(a)}$, is

$$\begin{aligned} \hat{P}^{(a)} &= \sum_{j=1}^{J} q_j \hat{P}_j^{(a)} \\ &= \frac{1}{N} \sum_{u=1}^{N} \max \begin{bmatrix} \hat{P}(1 \mid \mathbf{x_u}), \hat{P}(2 \mid \mathbf{x_u}), \ldots, \\ \hat{P}(j \mid \mathbf{x_u}), \ldots, \hat{P}(J \mid \mathbf{x_u}) \end{bmatrix}, \end{aligned} \qquad \text{(III.27)}$$

i.e., $P^{(a)}$ is calculated from the mean of the maximum estimated posterior probabilities for each unit.

According to Hora and Wilcox (1982) and Glick (1978), assuming normality assumptions are met, MPP combined with cross validation (LOO) is the preferred method of estimating the hit rate, as it has low bias, and is not sensitive to sampling variability[30]. If normality is untenable then a quadratic LOO hit rate

---

28 See Ned Glick, "Additive Estimators for Probabilities of Correct Classification," *Pattern Recognition* 10 (1978) and Stephen C. Hora and James B. Wilcox, "Estimation of Error Rates in Several-population Discriminant Analysis," *Journal of Marketing Research* 19, no. 1 (1982)

29 See W.R. Dillon and M Goldstein, *Multivariate Analysis: Methods and Applications* (New York: John Wiley & Sons Ltd., 1984) pp. 406 - 409

30 Glick, "Additive Estimators for Probabilities of Correct Classification," Hora and Wilcox, "Estimation of Error Rates in Several-population Discriminant Analysis,"

estimator is preferred[31]. If covariance heterogeneity is suspected, then it is preferable that the smallest group have at least five times the number of observations as there are predictors[32].

31 M. Connally, "Identifying Covariance Differences in Comparison of Linear versus Quadratic Classification Rule" (University of Georgia, 2004)

32 Anil K. Jain, Robert P. W. Duin, and Jianchang Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, no. 1 (2000)

Appendix IV

Classification Results for 3-Group Locomotion Data

Table 28: Cross-validation results using quadratic discriminant function classification – posterior probability of membership to locomotion groups

			Posterior Probabilities				
ID	*Actual Group*[1]	*Predicted Group (j)*[2]	$\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	*Typicality*[3]	D^{*2}
1	1	1	0.949	0.051	0.000	0.723	0.607
2	1	1	0.949	0.051	0.000	0.725	0.725
3	1	1	0.940	0.060	0.000	0.688	0.850
4	1	1	0.947	0.053	0.000	0.600	0.910
5	1	1	0.936	0.064	0.000	0.621	0.941
6	1	1	0.947	0.053	0.000	0.632	0.906
7	1	1	0.945	0.055	0.000	0.593	1.184
8	1	1	0.845	0.155	0.000	0.633	2.851
9	1	1	0.945	0.055	0.000	0.602	1.228
10	1	1	0.948	0.053	0.000	0.559	0.529
11	1	1	0.941	0.059	0.000	0.627	0.912
12	1	1	0.839	0.162	0.000	0.400	3.128
13	1	1	0.945	0.055	0.000	0.596	0.664
14	1	1	0.864	0.136	0.000	0.560	3.105
15	1	1	0.877	0.123	0.000	0.243	2.561

1 1 refers to data classified as being stationary, 2 as searching data, and 3 as walking data. Red cells are possible in-doubt units, and green cells identify possible outliers.

2 Red cells indicate in-doubt, or fence sitter units. These are units with approximately equal probabilities for two or more groups. That is, if their probabilities are close then they are a similar distance from two or more group centroids. Fence-sitters can provide insight into why some units resemble the typical member of one group – reflected by the group centroid – about as much as the typical member of another group.

3 Typicality refers to how typical an observation is to its group. It can be considered the proportion of units in group *j* that have vectors, $\mathbf{x}_v$, close to unit *u*. As such, it typicality is low for unit u, it is possible that u is an outlier, or perhaps a member of a group outside of the original groups specified. Cells with low typicality (< 0.05) have been highlighted in green.

ID	*Actual Group*[1]	*Predicted Group (j)*[2]	*Posterior Probabilities* $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	*Typicality*[3]	D^{*2}
16	1	1	0.878	0.122	0.000	0.548	2.530
17	1	1	0.944	0.056	0.000	0.549	0.634
18	1	1	0.838	0.162	0.000	0.740	3.353
19	1	1	0.862	0.138	0.000	0.602	2.834
20	1	1	0.939	0.061	0.000	0.249	0.730
21	1	1	0.855	0.145	0.000	0.600	3.140
22	1	1	0.940	0.060	0.000	0.735	0.930
23	1	1	0.945	0.055	0.000	0.217	0.679
24	1	1	0.840	0.160	0.000	0.393	3.324
25	1	1	0.837	0.163	0.000	0.691	3.343
26	1	1	0.874	0.126	0.000	0.192	2.617
27	1	1	0.862	0.138	0.000	0.678	3.132
28	1	1	0.935	0.065	0.000	0.687	2.150
29	1	1	0.920	0.080	0.000	0.697	1.382
30	1	1	0.857	0.144	0.000	0.187	3.113
31	1	1	0.854	0.146	0.000	0.254	3.254
32	1	1	0.798	0.202	0.000	0.250	3.796
33	1	1	0.896	0.105	0.000	0.258	2.420
34	1	1	0.826	0.174	0.000	0.313	3.516
35	1	1	0.945	0.055	0.000	0.698	1.076
36	1	1	0.893	0.107	0.000	0.167	2.386
37	1	1	0.930	0.070	0.000	0.220	1.144
38	1	1	0.943	0.057	0.000	0.671	0.628
39	1	1	0.885	0.115	0.000	0.170	2.773
40	1	1	0.836	0.164	0.000	0.469	3.165
41	1	1	0.886	0.114	0.000	0.605	2.749
42	1	1	0.931	0.069	0.000	0.185	1.564
43	1	1	0.829	0.171	0.000	0.594	3.464
44	1	1	0.943	0.057	0.000	0.681	0.947
45	1	1	0.893	0.108	0.000	0.708	2.291
46	1	1	0.937	0.063	0.000	0.169	1.695
47	1	1	0.819	0.181	0.000	0.168	3.570
48	1	1	0.822	0.178	0.000	0.246	3.539
49	1	1	0.921	0.079	0.000	0.184	1.623
50	1	1	0.846	0.154	0.000	0.316	3.197
51	1	1	0.894	0.106	0.000	0.502	2.263
52	1	1	0.776	0.225	0.000	0.394	3.963
53	1	1	0.945	0.055	0.000	0.476	0.707
54	1	1	0.899	0.101	0.000	0.188	2.085
55	1	1	0.900	0.100	0.000	0.399	2.075
56	1	1	0.869	0.131	0.000	0.173	3.014
57	1	1	0.781	0.220	0.000	0.134	3.905
58	1	1	0.922	0.078	0.000	0.704	1.595
59	1	1	0.857	0.143	0.000	0.137	3.003
60	1	1	0.937	0.063	0.000	0.268	1.711
61	1	1	0.927	0.073	0.000	0.554	1.700
62	1	1	0.944	0.057	0.000	0.153	0.726
63	1	1	0.877	0.123	0.000	0.558	2.815
64	1	1	0.885	0.115	0.000	0.274	2.655

ID	*Actual Group*[1]	*Predicted Group (j)*[2]	*Posterior Probabilities* $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	*Typicality*[3]	D^{*2}
65	1	1	0.884	0.116	0.000	0.140	2.684
66	1	1	0.899	0.101	0.000	0.533	2.468
67	1	1	0.900	0.100	0.000	0.701	2.442
68	1	1	0.876	0.124	0.000	0.222	2.818
69	1	1	0.888	0.112	0.000	0.188	2.293
70	1	1	0.932	0.068	0.000	0.159	1.446
71	1	1	0.872	0.128	0.000	0.225	2.969
72	1	1	0.889	0.111	0.000	0.676	2.397
73	1	1	0.867	0.133	0.000	0.585	2.701
74	1	1	0.932	0.068	0.000	0.176	1.424
75	1	1	0.846	0.154	0.000	0.526	3.219
76	1	1	0.849	0.151	0.000	0.159	3.169
77	1	1	0.880	0.120	0.000	0.422	2.783
78	1	1	0.934	0.067	0.000	0.272	1.345
79	1	1	0.874	0.127	0.000	0.270	2.699
80	1	1	0.871	0.129	0.000	0.427	2.745
81	1	1	0.925	0.075	0.000	0.158	2.184
82	1	1	0.883	0.117	0.000	0.590	2.780
83	1	1	0.880	0.120	0.000	0.290	2.822
84	1	1	0.932	0.068	0.000	0.142	1.323
85	1	1	0.937	0.063	0.000	0.399	1.407
86	1	1	0.912	0.089	0.000	0.463	2.272
87	1	1	0.883	0.118	0.000	0.215	2.740
88	1	1	0.934	0.066	0.000	0.150	1.223
89	1	1	0.933	0.067	0.000	0.152	1.367
90	1	1	0.816	0.185	0.000	0.184	3.410
91	1	1	0.945	0.055	0.000	0.410	1.110
92	1	1	0.712	0.288	0.000	0.181	4.461
93	1	1	0.949	0.051	0.000	0.294	0.726
94	1	1	0.931	0.070	0.000	0.196	1.296
95	1	1	0.932	0.068	0.000	0.125	1.291
96	1	1	0.831	0.169	0.000	0.472	3.247
97	1	1	0.941	0.059	0.000	0.200	1.090
98	1	1	0.934	0.067	0.000	0.671	1.407
99	1	1	0.945	0.055	0.000	0.325	1.120
100	1	1	0.945	0.055	0.000	0.326	1.296
101	1	1	0.943	0.057	0.000	0.196	1.611
102	1	1	0.942	0.058	0.000	0.129	1.389
103	1	1	0.941	0.060	0.000	0.287	1.110
104	1	1	0.941	0.059	0.000	0.309	1.020
105	1	1	0.941	0.059	0.000	0.314	1.000
106	1	1	0.933	0.067	0.000	0.416	1.433
107	1	1	0.942	0.058	0.000	0.331	1.374
108	1	1	0.933	0.067	0.000	0.200	1.558
109	1	1	0.857	0.143	0.000	0.395	2.782
110	1	1	0.921	0.079	0.000	0.270	1.635
111	1	1	0.933	0.068	0.000	0.392	1.269
112	1	1	0.933	0.067	0.000	0.221	1.241
113	1	1	0.922	0.078	0.000	0.408	1.601

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
114	1	1	0.932	0.068	0.000	0.204	1.583
115	1	1	0.939	0.061	0.000	0.658	1.172
116	1	1	0.931	0.069	0.000	0.273	1.464
117	1	1	0.930	0.070	0.000	0.320	1.492
118	1	1	0.946	0.054	0.000	0.664	1.141
119	1	1	0.936	0.064	0.000	0.218	1.703
120	1	1	0.875	0.125	0.000	0.237	2.528
121	1	1	0.771	0.229	0.000	0.233	3.624
122	1	1	0.763	0.237	0.000	0.397	3.704
123	1	2	0.313	0.687	0.000	0.260	0.294
124	1	2	0.301	0.700	0.000	0.264	0.293
125	1	1	0.722	0.278	0.000	0.218	4.051
126	1	1	0.946	0.054	0.000	0.414	1.252
127	1	1	0.946	0.054	0.000	0.292	1.249
128	1	1	0.605	0.395	0.000	0.366	5.012
129	1	1	0.946	0.054	0.000	0.450	1.050
130	1	1	0.935	0.065	0.000	0.360	1.725
131	1	1	0.947	0.054	0.000	0.201	1.043
132	1	1	0.539	0.461	0.000	0.495	5.551
133	1	1	0.727	0.273	0.000	0.274	4.006
134	1	1	0.736	0.264	0.000	0.280	3.923
135	1	1	0.799	0.201	0.000	0.237	3.194
136	1	1	0.892	0.108	0.000	0.455	1.782
137	1	1	0.940	0.060	0.000	0.179	1.823
138	1	1	0.561	0.439	0.000	0.225	5.396
139	1	1	0.943	0.057	0.000	0.232	1.601
140	1	1	0.947	0.054	0.000	0.183	1.172
141	1	1	0.915	0.085	0.000	0.221	1.960
142	1	1	0.548	0.452	0.000	0.475	5.494
143	1	1	0.949	0.051	0.000	0.234	0.661
144	1	1	0.787	0.213	0.000	0.229	3.318
145	1	1	0.743	0.257	0.000	0.304	3.806
146	1	1	0.824	0.176	0.000	0.309	2.863
147	1	1	0.818	0.182	0.000	0.221	2.935
148	1	1	0.778	0.222	0.000	0.216	3.413
149	1	1	0.658	0.342	0.000	0.294	4.580
150	1	1	0.772	0.228	0.000	0.475	3.487
151	1	1	0.669	0.331	0.000	0.481	4.491
152	1	1	0.923	0.077	0.000	0.483	1.796
153	1	1	0.742	0.258	0.000	0.461	3.797
154	1	1	0.895	0.105	0.000	0.288	1.726
155	1	1	0.949	0.051	0.000	0.226	0.599
156	1	1	0.950	0.050	0.000	0.172	0.444
157	1	1	0.949	0.051	0.000	0.507	0.612
158	1	1	0.950	0.050	0.000	0.543	0.526
159	1	1	0.949	0.051	0.000	0.494	0.616
160	1	1	0.950	0.050	0.000	0.469	0.525
161	1	1	0.950	0.050	0.000	0.466	0.449
162	1	1	0.950	0.050	0.000	0.099	0.483

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
163	1	1	0.950	0.050	0.000	0.167	0.516
164	1	1	0.950	0.050	0.000	0.513	0.485
165	1	1	0.902	0.098	0.000	0.551	1.577
166	1	1	0.943	0.057	0.000	0.103	1.520
167	2	2	0.000	1.000	0.000	0.689	1.226
168	1	1	0.957	0.043	0.000	0.488	0.921
169	1	1	0.957	0.043	0.000	0.489	0.353
170	1	1	0.771	0.229	0.000	0.180	3.853
171	1	1	0.957	0.043	0.000	0.546	0.350
172	1	1	0.775	0.225	0.000	0.175	3.813
173	1	1	0.957	0.043	0.000	0.502	0.774
174	1	1	0.950	0.050	0.000	0.426	0.477
175	1	1	0.954	0.046	0.000	0.459	1.192
176	1	1	0.950	0.050	0.000	0.472	0.474
177	2	2	0.000	1.000	0.000	0.547	1.226
178	2	2	0.000	1.000	0.000	0.505	1.252
179	1	1	0.939	0.061	0.000	0.429	1.224
180	1	1	0.957	0.043	0.000	0.472	0.259
181	1	1	0.957	0.043	0.000	0.466	0.780
182	2	2	0.000	1.000	0.000	0.540	1.213
183	2	2	0.000	1.000	0.000	0.567	1.225
184	1	1	0.941	0.059	0.000	0.573	0.503
185	1	1	0.920	0.080	0.000	0.453	1.141
186	1	1	0.945	0.055	0.000	0.476	1.314
187	2	2	0.000	1.000	0.000	0.533	1.095
188	1	1	0.905	0.095	0.000	0.424	1.525
189	1	1	0.939	0.061	0.000	0.231	1.240
190	2	2	0.000	1.000	0.000	0.408	1.434
191	2	2	0.000	1.000	0.000	0.526	1.434
192	2	2	0.000	1.000	0.000	0.495	1.138
193	2	2	0.000	1.000	0.000	0.516	1.613
194	2	2	0.000	1.000	0.000	0.502	1.258
195	2	2	0.000	1.000	0.000	0.415	1.138
196	2	2	0.000	1.000	0.000	0.418	1.070
197	1	1	0.945	0.055	0.000	0.522	0.373
198	2	2	0.000	1.000	0.000	0.445	1.329
199	2	2	0.000	1.000	0.000	0.438	1.258
200	1	1	0.673	0.327	0.000	0.546	10.975
201	1	1	0.943	0.057	0.000	0.412	2.846
202	1	1	0.953	0.047	0.000	0.406	0.100
203	1	1	0.880	0.121	0.000	0.396	6.462
204	1	1	0.941	0.059	0.000	0.336	0.809
205	2	2	0.000	1.000	0.000	0.266	1.292
206	1	1	0.953	0.047	0.000	0.260	0.082
207	1	1	0.946	0.055	0.000	0.134	0.360
208	1	1	0.953	0.047	0.000	0.169	0.091
209	1	1	0.950	0.050	0.000	0.163	0.369
210	1	1	0.953	0.047	0.000	0.861	0.116
211	1	1	0.953	0.047	0.000	0.861	0.087

ID	*Actual Group*[1]	*Predicted Group (j)*[2]	*Posterior Probabilities* $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	*Typicality*[3]	D^{*2}
212	1	1	0.881	0.119	0.000	0.139	6.427
213	1	1	0.953	0.047	0.000	0.522	0.117
214	1	1	0.949	0.051	0.000	0.086	0.414
215	1	1	0.552	0.448	0.000	0.523	13.080
216	1	1	0.949	0.051	0.000	0.091	0.406
217	1	2	0.326	0.674	0.000	0.571	9.715
218	1	1	0.951	0.049	0.000	0.391	2.676
219	1	1	0.958	0.042	0.000	0.574	1.210
220	1	1	0.953	0.047	0.000	0.068	0.112
221	1	1	0.860	0.140	0.000	0.111	2.330
222	1	1	0.953	0.047	0.000	0.071	0.116
223	1	2	0.333	0.667	0.000	0.823	9.713
224	1	1	0.946	0.054	0.000	0.605	1.145
225	1	1	0.952	0.048	0.000	0.855	0.163
226	1	1	0.729	0.271	0.000	0.152	10.502
227	1	1	0.899	0.101	0.000	0.159	1.621
228	1	1	0.861	0.139	0.000	0.221	7.534
229	1	2	0.051	0.949	0.000	0.855	13.046
230	1	1	0.904	0.096	0.000	0.229	1.512
231	1	1	0.903	0.097	0.000	0.107	1.537
232	1	1	0.948	0.052	0.000	0.609	0.829
233	1	1	0.939	0.061	0.000	0.424	0.786
234	1	1	0.948	0.052	0.000	0.415	0.829
235	1	1	0.857	0.143	0.000	0.077	2.389
236	1	1	0.948	0.052	0.000	0.089	0.878
237	1	1	0.948	0.052	0.000	0.456	0.878
238	1	1	0.928	0.072	0.000	0.546	0.925
239	1	1	0.938	0.062	0.000	0.547	0.802
240	1	1	0.947	0.053	0.000	0.484	1.013
241	1	1	0.943	0.057	0.000	0.484	0.422
242	1	1	0.861	0.139	0.000	0.455	2.325
243	1	1	0.918	0.082	0.000	0.822	1.189
244	1	1	0.943	0.057	0.000	0.342	0.621
245	1	1	0.909	0.092	0.000	0.074	1.414
246	1	1	0.911	0.089	0.000	0.121	1.364
247	1	1	0.948	0.052	0.000	0.116	0.380
248	1	1	0.944	0.056	0.000	0.709	0.592
249	1	1	0.851	0.149	0.000	0.358	2.486
250	1	1	0.946	0.054	0.000	0.205	1.019
251	1	1	0.956	0.044	0.000	0.165	0.721
252	1	1	0.920	0.080	0.000	0.254	2.819
253	1	1	0.950	0.050	0.000	0.159	0.247
254	1	1	0.919	0.082	0.000	0.245	2.850
255	1	1	0.949	0.051	0.000	0.093	0.618
256	1	1	0.951	0.050	0.000	0.203	0.374
257	1	1	0.950	0.050	0.000	0.198	0.497
258	1	1	0.838	0.162	0.000	0.413	2.800
259	1	1	0.950	0.050	0.000	0.195	0.272
260	1	1	0.949	0.051	0.000	0.350	0.292

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
261	1	1	0.948	0.052	0.000	0.113	0.340
262	1	1	0.948	0.052	0.000	0.188	0.329
263	1	1	0.950	0.050	0.000	0.118	0.280
264	1	1	0.949	0.051	0.000	0.181	0.302
265	1	1	0.948	0.052	0.000	0.710	0.286
266	1	1	0.950	0.050	0.000	0.250	0.263
267	1	1	0.949	0.052	0.000	0.366	0.308
268	1	1	0.948	0.052	0.000	0.244	0.275
269	1	1	0.932	0.068	0.000	0.157	0.797
270	1	1	0.924	0.076	0.000	0.372	1.007
271	1	1	0.950	0.050	0.000	0.163	0.255
272	1	1	0.945	0.055	0.000	0.436	0.424
273	1	1	0.939	0.061	0.000	0.477	0.600
274	1	1	0.938	0.062	0.000	0.780	0.601
275	1	1	0.949	0.051	0.000	0.737	0.288
276	1	1	0.948	0.052	0.000	0.737	0.336
277	1	1	0.952	0.048	0.000	0.783	0.158
278	1	1	0.956	0.045	0.000	0.734	0.727
279	1	1	0.945	0.055	0.000	0.760	0.398
280	1	1	0.923	0.077	0.000	0.781	1.050
281	1	1	0.952	0.048	0.000	0.733	0.145
282	1	1	0.932	0.068	0.000	0.759	0.778
283	1	1	0.908	0.092	0.000	0.783	1.498
284	1	1	0.939	0.062	0.000	0.778	0.578
285	1	1	0.859	0.141	0.000	0.761	3.046
286	1	1	0.951	0.049	0.000	0.761	0.216
287	1	1	0.941	0.059	0.000	0.778	0.503
288	1	1	0.866	0.134	0.000	0.467	3.193
289	1	1	0.946	0.054	0.000	0.781	0.309
290	1	1	0.910	0.090	0.000	0.476	1.452
291	1	1	0.952	0.049	0.000	0.539	0.165
292	1	1	0.938	0.062	0.000	0.781	5.706
293	1	1	0.864	0.136	0.000	0.674	3.224
294	1	1	0.930	0.070	0.000	0.856	0.868
295	1	1	0.931	0.069	0.000	0.171	0.832
296	1	1	0.875	0.126	0.000	0.857	2.394
297	1	1	0.878	0.122	0.000	0.711	2.338
298	1	1	0.863	0.137	0.000	0.174	3.241
299	1	1	0.860	0.140	0.000	0.104	2.333
300	1	1	0.862	0.138	0.000	0.720	2.302
301	1	1	0.894	0.106	0.000	0.783	1.738
302	1	1	0.919	0.081	0.000	0.105	1.241
303	1	1	0.802	0.198	0.000	0.896	3.171
304	1	1	0.959	0.041	0.000	0.605	2.687
305	1	1	0.861	0.139	0.000	0.291	3.272
306	1	1	0.959	0.041	0.000	0.709	2.205
307	1	1	0.959	0.041	0.000	0.602	2.686
308	1	1	0.952	0.048	0.000	0.676	0.150
309	1	1	0.947	0.053	0.000	0.784	0.771

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
310	1	1	0.959	0.041	0.000	0.532	2.205
311	1	1	0.947	0.053	0.000	0.539	4.777
312	1	1	0.952	0.048	0.000	0.532	0.169
313	1	1	0.958	0.042	0.000	0.564	1.064
314	1	1	0.953	0.047	0.000	0.894	0.095
315	1	1	0.952	0.048	0.000	0.718	0.144
316	2	2	0.012	0.988	0.000	0.539	0.968
317	1	1	0.959	0.041	0.000	0.295	2.379
318	1	1	0.953	0.047	0.000	0.542	0.119
319	1	1	0.952	0.048	0.000	0.542	0.149
320	1	1	0.934	0.066	0.000	0.549	0.958
321	1	1	0.935	0.065	0.000	0.539	0.926
322	1	1	0.954	0.046	0.000	0.549	0.031
323	1	1	0.945	0.055	0.000	0.788	0.417
324	1	1	0.954	0.046	0.000	0.782	0.023
325	1	1	0.944	0.056	0.000	0.408	0.466
326	1	1	0.951	0.049	0.000	0.589	0.170
327	1	1	0.945	0.055	0.000	0.880	0.485
328	1	1	0.949	0.052	0.000	0.884	0.313
329	1	1	0.954	0.046	0.000	0.522	0.027
330	1	1	0.945	0.055	0.000	0.575	0.465
331	2	2	0.013	0.987	0.000	0.479	0.981
332	1	1	0.947	0.053	0.000	0.595	0.307
333	2	2	0.013	0.987	0.000	0.560	0.985
334	1	1	0.959	0.041	0.000	0.410	2.379
335	2	2	0.013	0.987	0.000	0.443	0.981
336	1	1	0.944	0.056	0.000	0.485	0.443
337	1	1	0.909	0.092	0.000	0.450	1.867
338	1	1	0.947	0.053	0.000	0.450	0.325
339	1	1	0.959	0.041	0.000	0.485	2.032
340	2	2	0.013	0.987	0.000	0.563	0.990
341	1	1	0.956	0.044	0.000	0.443	0.632
342	1	1	0.957	0.043	0.000	0.017	0.246
343	1	1	0.958	0.042	0.000	0.575	1.852
344	1	1	0.959	0.042	0.000	0.581	1.763
345	1	1	0.957	0.043	0.000	0.530	0.371
346	2	2	0.021	0.979	0.000	0.563	1.073
347	2	2	0.012	0.988	0.000	0.581	0.968
348	2	2	0.013	0.987	0.000	0.839	0.990
349	1	1	0.951	0.049	0.000	0.008	0.153
350	1	1	0.951	0.049	0.000	0.512	0.145
351	1	1	0.932	0.068	0.000	0.594	0.824
352	2	2	0.019	0.981	0.000	0.530	1.057
353	1	1	0.956	0.044	0.000	0.512	0.108
354	1	1	0.931	0.069	0.000	0.008	0.883
355	2	2	0.037	0.963	0.000	0.290	1.220
356	1	1	0.910	0.090	0.000	0.293	1.475
357	2	2	0.019	0.981	0.000	0.017	1.052
358	2	2	0.006	0.995	0.000	0.521	0.879

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
359	1	1	0.921	0.079	0.000	0.944	1.402
360	2	2	0.019	0.981	0.000	0.591	1.052
361	1	1	0.952	0.048	0.000	0.944	0.096
362	2	2	0.044	0.956	0.000	0.947	1.276
363	1	1	0.956	0.044	0.000	0.692	0.074
364	1	1	0.952	0.048	0.000	0.060	0.103
365	2	2	0.014	0.986	0.000	0.687	1.727
366	1	1	0.944	0.056	0.000	0.521	0.421
367	1	1	0.886	0.114	0.000	0.948	2.342
368	1	1	0.923	0.077	0.000	0.844	1.324
369	2	2	0.001	0.999	0.000	0.947	1.042
370	2	2	0.002	0.998	0.000	0.826	1.139
371	2	2	0.001	0.999	0.000	0.856	1.026
372	2	2	0.000	1.000	0.000	0.935	3.029
373	2	2	0.001	0.999	0.000	0.945	1.026
374	2	2	0.001	0.999	0.000	0.854	1.042
375	1	1	0.887	0.113	0.000	0.322	2.314
376	1	1	0.957	0.043	0.000	0.627	1.747
377	2	2	0.001	0.999	0.000	0.825	0.837
378	2	2	0.000	1.000	0.000	0.002	3.029
379	2	2	0.002	0.998	0.000	0.061	1.138
380	1	1	0.899	0.102	0.000	0.935	2.020
381	2	2	0.003	0.997	0.000	0.816	0.846
382	2	2	0.001	0.999	0.000	0.870	0.837
383	1	1	0.928	0.072	0.000	0.565	1.061
384	2	2	0.002	0.998	0.000	0.003	0.835
385	2	2	0.002	0.998	0.000	0.002	0.835
386	1	1	0.929	0.071	0.000	0.818	1.042
387	2	2	0.000	1.000	0.000	0.521	2.715
388	1	1	0.918	0.082	0.000	0.001	1.497
389	1	1	0.938	0.063	0.000	0.313	0.699
390	2	2	0.000	1.000	0.000	0.870	3.558
391	1	1	0.938	0.062	0.000	0.589	0.684
392	1	1	0.917	0.083	0.000	0.589	1.519
393	1	1	0.941	0.059	0.000	0.729	0.544
394	1	1	0.942	0.058	0.000	0.622	0.530
395	2	2	0.000	1.000	0.000	0.003	3.074
396	2	1	0.942	0.058	0.000	0.933	1.470
397	2	1	0.942	0.058	0.000	0.814	1.470
398	2	1	0.944	0.056	0.000	0.332	1.750
399	2	1	0.943	0.057	0.000	0.735	1.632
400	2	1	0.943	0.057	0.000	0.930	1.568
401	2	1	0.943	0.057	0.000	0.001	1.632
402	2	2	0.000	1.000	0.000	0.004	3.229
403	2	1	0.943	0.057	0.000	0.315	1.559
404	2	2	0.131	0.869	0.000	0.564	0.835
405	2	1	0.943	0.057	0.000	0.911	1.559
406	2	2	0.000	1.000	0.000	0.919	3.463
407	1	1	0.955	0.045	0.000	0.910	1.546

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
408	2	1	0.936	0.065	0.000	0.923	1.350
409	2	1	0.936	0.065	0.000	0.004	1.350
410	2	1	0.942	0.058	0.000	0.010	1.376
411	2	2	0.033	0.968	0.000	0.010	0.743
412	2	2	0.033	0.968	0.000	0.467	0.743
413	1	1	0.957	0.043	0.000	0.461	1.723
414	2	2	0.102	0.898	0.000	0.038	0.796
415	1	1	0.957	0.043	0.000	0.001	1.725
416	2	1	0.920	0.081	0.000	0.037	1.803
417	2	2	0.053	0.947	0.000	0.001	0.744
418	2	2	0.053	0.947	0.000	0.484	0.744
419	2	2	0.000	1.000	0.000	0.478	0.041
420	2	1	0.942	0.058	0.000	0.838	1.376
421	2	2	0.045	0.955	0.000	0.651	0.741
422	2	1	0.939	0.061	0.000	0.680	1.365
423	2	2	0.000	1.000	0.000	0.652	4.043
424	2	2	0.045	0.955	0.000	0.324	0.741
425	1	1	0.959	0.041	0.000	0.636	2.114
426	1	1	0.959	0.041	0.000	0.635	2.551
427	1	1	0.959	0.041	0.000	0.642	2.115
428	1	1	0.959	0.041	0.000	0.675	2.550
429	2	2	0.000	1.000	0.000	0.603	4.007
430	1	1	0.959	0.041	0.000	0.580	2.515
431	2	2	0.000	1.000	0.000	0.820	4.840
432	2	2	0.000	1.000	0.000	0.334	4.324
433	1	1	0.958	0.042	0.000	0.222	1.089
434	2	2	0.000	1.000	0.000	0.567	4.840
435	2	2	0.000	1.000	0.000	0.655	4.007
436	2	2	0.183	0.817	0.000	0.596	0.910
437	2	1	0.844	0.156	0.000	0.739	2.901
438	2	1	0.844	0.156	0.000	0.509	2.901
439	2	1	0.913	0.087	0.000	0.521	1.571
440	1	1	0.959	0.041	0.000	0.834	2.514
441	2	1	0.920	0.081	0.000	0.750	1.803
442	2	1	0.892	0.108	0.000	0.447	2.012
443	2	1	0.863	0.137	0.000	0.311	2.569
444	2	1	0.934	0.066	0.000	0.265	1.146
445	2	1	0.945	0.055	0.000	0.219	1.479
446	2	1	0.933	0.067	0.000	0.435	1.158
447	2	1	0.877	0.123	0.000	0.301	2.300
448	2	1	0.775	0.225	0.000	0.602	3.925
449	2	1	0.933	0.067	0.000	0.714	1.158
450	2	1	0.892	0.108	0.000	0.890	2.012
451	2	1	0.820	0.180	0.000	0.273	3.277
452	2	1	0.924	0.076	0.000	0.887	1.605
453	2	1	0.915	0.085	0.000	0.269	1.514
454	2	1	0.915	0.085	0.000	0.721	1.514
455	2	1	0.912	0.088	0.000	0.722	1.581
456	2	1	0.934	0.066	0.000	0.840	1.132

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
457	2	1	0.914	0.086	0.000	0.809	2.033
458	2	2	0.000	1.000	0.000	0.845	3.909
459	2	1	0.915	0.085	0.000	0.766	1.976
460	2	1	0.915	0.085	0.000	0.274	2.000
461	2	1	0.894	0.106	0.000	0.834	2.728
462	2	1	0.915	0.085	0.000	0.852	1.530
463	2	1	0.925	0.075	0.000	0.813	1.555
464	2	1	0.896	0.104	0.000	0.869	1.920
465	2	1	0.925	0.075	0.000	0.852	1.555
466	2	2	0.000	1.000	0.000	0.855	3.827
467	2	1	0.899	0.101	0.000	0.723	2.556
468	1	1	0.959	0.041	0.000	0.849	2.826
469	2	1	0.899	0.101	0.000	0.764	2.556
470	2	2	0.000	1.000	0.000	0.866	3.827
471	2	2	0.220	0.780	0.000	0.874	0.635
472	2	1	0.909	0.091	0.000	0.656	2.221
473	2	1	0.935	0.065	0.000	0.857	1.108
474	2	2	0.498	0.503	0.000	0.655	0.318
475	2	2	0.379	0.622	0.000	0.841	0.418
476	2	2	0.279	0.721	0.000	0.879	0.539
477	2	2	0.149	0.851	0.000	0.844	0.798
478	2	2	0.149	0.851	0.000	0.741	0.798
479	2	1	0.899	0.101	0.000	0.833	2.571
480	2	1	0.927	0.073	0.000	0.691	1.476
481	2	1	0.715	0.285	0.000	0.621	4.109
482	2	1	0.594	0.406	0.000	0.827	5.217
483	2	1	0.715	0.285	0.000	0.868	4.109
484	2	1	0.661	0.339	0.000	0.816	4.617
485	2	1	0.933	0.067	0.000	0.877	1.189
486	2	1	0.815	0.185	0.000	0.751	3.038
487	1	1	0.959	0.041	0.000	0.762	1.864
488	1	1	0.958	0.042	0.000	0.753	1.346
489	2	1	0.706	0.294	0.000	0.852	4.194
490	1	1	0.958	0.042	0.000	0.718	1.348
491	1	1	0.950	0.050	0.000	0.861	1.029
492	1	1	0.954	0.046	0.000	0.850	1.052
493	2	2	0.000	1.000	0.000	0.834	0.531
494	1	1	0.959	0.041	0.000	0.818	1.868
495	1	1	0.953	0.047	0.000	0.824	3.234
496	2	1	0.518	0.482	0.000	0.909	5.873
497	2	1	0.690	0.310	0.000	0.712	4.346
498	1	1	0.953	0.047	0.000	0.694	3.247
499	2	1	0.573	0.427	0.000	0.828	5.400
500	1	1	0.949	0.051	0.000	0.608	1.357
501	1	1	0.952	0.048	0.000	0.688	1.081
502	1	1	0.935	0.065	0.000	0.596	1.385
503	2	1	0.573	0.427	0.000	0.915	5.400
504	1	1	0.935	0.065	0.000	0.697	1.361
505	2	2	0.001	1.000	0.000	0.502	3.798

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
506	1	1	0.958	0.042	0.000	0.763	1.822
507	1	1	0.958	0.042	0.000	0.752	1.826
508	1	1	0.956	0.044	0.000	0.359	1.674
509	2	1	0.532	0.468	0.000	0.896	5.753
510	2	2	0.000	1.000	0.000	0.253	0.476
511	1	1	0.951	0.049	0.000	0.904	1.484
512	1	1	0.952	0.048	0.000	0.355	1.091
513	2	1	0.673	0.328	0.000	0.803	4.513
514	1	1	0.958	0.043	0.000	0.793	1.613
515	1	1	0.852	0.148	0.000	0.237	2.519
516	2	2	0.458	0.542	0.000	0.997	0.347
517	2	1	0.673	0.328	0.000	0.995	4.513
518	1	1	0.952	0.048	0.000	0.690	1.240
519	2	2	0.112	0.888	0.000	0.680	0.921
520	2	1	0.532	0.468	0.000	0.867	5.753
521	1	1	0.916	0.084	0.000	0.513	1.597
522	2	1	0.723	0.277	0.000	0.858	4.028
523	2	1	0.569	0.431	0.000	0.908	5.431
524	2	1	0.569	0.431	0.000	0.853	5.431
525	1	1	0.951	0.049	0.000	0.081	1.495
526	2	1	0.711	0.289	0.000	0.234	4.151
527	1	1	0.952	0.048	0.000	0.844	1.230
528	2	1	0.597	0.403	0.000	0.627	5.186
529	2	1	0.711	0.289	0.000	0.669	4.151
530	2	2	0.498	0.503	0.000	0.680	0.318
531	2	1	0.558	0.442	0.000	0.336	5.527
532	1	1	0.954	0.046	0.000	0.073	1.173
533	2	1	0.602	0.398	0.000	0.071	5.144
534	1	1	0.915	0.085	0.000	0.690	1.568
535	2	1	0.584	0.416	0.000	0.616	5.306
536	1	1	0.954	0.047	0.000	0.345	1.181
537	2	2	0.000	1.000	0.000	0.232	0.932
538	1	1	0.916	0.085	0.000	0.329	1.544
539	2	2	0.000	1.000	0.000	0.334	0.711
540	2	2	0.000	1.000	0.000	0.428	0.897
541	2	2	0.000	1.000	0.000	0.577	0.711
542	2	2	0.000	1.000	0.000	0.434	0.635
543	2	2	0.000	1.000	0.000	0.566	0.897
544	2	2	0.000	1.000	0.000	0.899	0.635
545	2	1	0.541	0.459	0.000	0.082	5.674
546	2	2	0.000	1.000	0.000	0.225	0.580
547	2	2	0.000	1.000	0.000	0.229	0.444
548	2	2	0.000	1.000	0.000	0.299	0.490
549	2	2	0.000	1.000	0.000	0.229	0.490
550	2	2	0.000	1.000	0.000	0.366	0.907
551	2	1	0.541	0.459	0.000	0.135	5.674
552	2	2	0.000	1.000	0.000	0.299	0.583
553	2	2	0.000	1.000	0.000	0.915	0.907
554	2	2	0.001	0.999	0.000	0.685	3.610

ID	*Actual Group*[1]	*Predicted Group (j)*[2]	*Posterior Probabilities* $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	*Typicality*[3]	D^{*2}
555	2	2	0.000	1.000	0.000	0.273	1.296
556	2	1	0.547	0.454	0.000	0.367	5.628
557	2	1	0.547	0.454	0.000	0.617	5.628
558	1	1	0.959	0.041	0.000	0.133	2.420
559	2	2	0.000	1.000	0.000	0.122	0.587
560	2	2	0.000	1.000	0.000	0.122	0.583
561	2	2	0.000	1.000	0.000	0.908	0.521
562	2	1	0.584	0.416	0.000	0.169	5.306
563	2	2	0.000	1.000	0.000	0.259	0.587
564	2	2	0.000	1.000	0.000	0.273	0.626
565	2	2	0.000	1.000	0.000	0.616	0.626
566	2	2	0.000	1.000	0.000	0.909	0.122
567	2	2	0.000	1.000	0.000	0.166	0.122
568	2	2	0.000	1.000	0.000	0.945	0.494
569	1	1	0.953	0.047	0.000	0.945	1.580
570	2	2	0.000	1.000	0.000	0.946	0.199
571	2	2	0.000	1.000	0.000	0.926	0.199
572	2	2	0.000	1.000	0.000	0.421	0.494
573	2	2	0.000	1.000	0.000	0.613	0.311
574	2	2	0.000	1.000	0.000	0.930	0.900
575	2	2	0.000	1.000	0.000	0.904	0.887
576	2	2	0.000	1.000	0.000	0.340	0.834
577	1	1	0.959	0.041	0.000	0.945	2.420
578	2	2	0.000	1.000	0.000	0.198	0.041
579	2	2	0.001	0.999	0.000	0.932	3.610
580	2	2	0.000	1.000	0.000	0.924	0.302
581	2	2	0.000	1.000	0.000	0.646	0.437
582	1	1	0.951	0.049	0.000	0.832	1.574
583	2	2	0.000	1.000	0.000	0.416	0.531
584	1	1	0.943	0.057	0.000	0.975	1.529
585	1	1	0.947	0.053	0.000	0.655	1.575
586	1	1	0.958	0.042	0.000	0.978	2.120
587	1	1	0.958	0.042	0.000	0.898	2.123
588	1	1	0.959	0.041	0.000	0.824	2.433
589	1	1	0.958	0.042	0.000	0.899	2.180
590	2	2	0.000	1.000	0.000	0.964	0.255
591	2	2	0.000	1.000	0.000	0.915	0.622
592	2	2	0.007	0.993	0.000	0.983	4.893
593	2	2	0.007	0.993	0.000	0.807	4.893
594	2	2	0.027	0.973	0.000	0.919	3.820
595	1	1	0.956	0.044	0.000	0.798	3.801
596	1	1	0.956	0.044	0.000	0.862	3.798
597	1	1	0.959	0.041	0.000	0.980	2.627
598	2	2	0.009	0.991	0.000	0.805	6.485
599	2	2	0.000	1.000	0.000	0.610	0.622
600	2	2	0.014	0.986	0.000	0.867	6.011
601	2	2	0.014	0.986	0.000	0.608	6.011
602	2	2	0.000	1.000	0.000	0.608	0.255
603	2	2	0.019	0.981	0.000	0.967	5.731

ID	*Actual Group*[1]	*Predicted Group (j)*[2]	*Posterior Probabilities* $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	*Typicality*[3]	D^{*2}
604	2	2	0.063	0.937	0.000	0.340	4.560
605	2	2	0.063	0.937	0.000	0.610	4.560
606	2	2	0.070	0.930	0.000	0.856	4.466
607	2	2	0.241	0.759	0.000	0.815	3.169
608	2	2	0.035	0.966	0.000	0.429	5.149
609	1	1	0.958	0.042	0.000	0.793	2.255
610	2	2	0.000	1.000	0.000	0.860	3.796
611	2	2	0.000	1.000	0.000	0.784	0.926
612	2	2	0.000	1.000	0.000	0.393	0.926
613	2	2	0.000	1.000	0.000	0.393	1.066
614	2	2	0.000	1.000	0.000	0.423	0.896
615	2	2	0.000	1.000	0.000	0.200	0.719
616	2	2	0.000	1.000	0.000	0.905	0.843
617	2	2	0.000	1.000	0.000	0.607	0.896
618	2	2	0.000	1.000	0.000	0.768	0.699
619	2	2	0.000	1.000	0.000	0.906	0.737
620	2	2	0.000	1.000	0.000	0.958	0.753
621	2	2	0.000	1.000	0.000	0.423	1.005
622	2	2	0.000	1.000	0.000	0.425	0.479
623	2	2	0.000	1.000	0.000	0.765	0.753
624	2	2	0.000	1.000	0.000	0.958	0.923
625	2	2	0.000	1.000	0.000	0.458	0.635
626	2	2	0.000	1.000	0.000	0.859	0.747
627	2	2	0.000	1.000	0.000	0.582	1.070
628	2	2	0.000	1.000	0.000	0.613	1.024
629	2	2	0.000	1.000	0.000	0.869	0.968
630	2	2	0.000	1.000	0.000	0.607	0.457
631	2	2	0.000	1.000	0.000	0.955	0.476
632	2	2	0.000	1.000	0.000	0.874	0.169
633	2	2	0.000	1.000	0.000	0.932	0.737
634	2	2	0.000	1.000	0.000	0.540	0.999
635	2	2	0.000	1.000	0.000	0.679	0.707
636	2	2	0.000	1.000	0.000	0.936	1.455
637	2	2	0.000	1.000	0.000	0.685	1.455
638	2	2	0.000	1.000	0.000	0.861	0.923
639	1	1	0.955	0.045	0.000	0.442	1.925
640	2	2	0.000	1.000	0.000	0.587	0.572
641	2	2	0.000	1.000	0.000	0.957	1.895
642	2	2	0.000	1.000	0.000	0.627	0.561
643	2	2	0.000	1.000	0.000	0.627	0.561
644	2	2	0.000	1.000	0.000	0.525	0.775
645	2	2	0.000	1.000	0.000	0.667	0.775
646	2	2	0.000	1.000	0.000	0.534	0.572
647	2	2	0.000	1.000	0.000	0.540	0.898
648	2	2	0.000	1.000	0.000	0.600	0.898
649	2	2	0.000	1.000	0.000	0.508	0.395
650	1	1	0.954	0.047	0.000	0.673	1.863
651	2	1	0.942	0.058	0.000	0.587	4.662
652	1	1	0.930	0.070	0.000	0.514	1.769

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
653	1	1	0.929	0.071	0.000	0.588	1.787
654	1	1	0.958	0.042	0.000	0.686	3.397
655	2	1	0.959	0.041	0.000	0.641	1.802
656	2	1	0.959	0.041	0.000	0.529	1.802
657	2	2	0.000	1.000	0.000	0.599	0.193
658	1	1	0.898	0.102	0.000	0.599	2.169
659	2	1	0.893	0.107	0.000	0.588	8.086
660	2	1	0.959	0.041	0.000	0.475	1.319
661	2	1	0.955	0.046	0.000	0.600	3.067
662	2	2	0.000	1.000	0.000	0.691	0.207
663	2	1	0.955	0.046	0.000	0.475	3.067
664	2	1	0.958	0.042	0.000	0.582	2.387
665	2	1	0.959	0.041	0.000	0.819	1.302
666	2	1	0.958	0.042	0.000	0.957	2.387
667	2	1	0.957	0.043	0.000	0.525	2.479
668	2	1	0.932	0.068	0.000	0.973	5.607
669	2	1	0.957	0.043	0.000	0.974	2.479
670	1	1	0.953	0.048	0.000	0.818	1.929
671	2	1	0.957	0.043	0.000	0.813	2.487
672	1	1	0.957	0.043	0.000	0.954	2.286
673	1	1	0.937	0.064	0.000	0.563	1.784
674	1	1	0.902	0.098	0.000	0.417	2.136
675	1	1	0.958	0.042	0.000	0.824	2.451
676	1	1	0.919	0.081	0.000	0.417	1.933
677	2	1	0.957	0.043	0.000	0.346	2.487
678	1	1	0.958	0.042	0.000	0.548	3.395
679	1	1	0.919	0.081	0.000	0.591	1.933
680	1	1	0.958	0.042	0.000	0.563	2.452
681	1	1	0.956	0.044	0.000	0.596	2.113
682	1	1	0.956	0.044	0.000	0.204	2.113
683	2	1	0.959	0.041	0.000	0.596	2.151
684	2	1	0.935	0.065	0.000	0.641	5.357
685	2	2	0.000	1.000	0.000	0.607	0.214
686	1	1	0.959	0.041	0.000	0.527	3.065
687	1	1	0.950	0.050	0.000	0.607	5.223
688	1	1	0.954	0.046	0.000	0.591	4.208
689	1	1	0.944	0.057	0.000	0.527	6.030
690	2	1	0.921	0.079	0.000	0.351	6.452
691	1	1	0.957	0.043	0.000	0.441	2.123
692	1	1	0.950	0.050	0.000	0.395	4.948
693	1	1	0.948	0.052	0.000	0.653	1.711
694	2	2	0.000	1.000	0.000	0.616	0.361
695	1	1	0.949	0.052	0.000	0.204	1.707
696	1	1	0.947	0.053	0.000	0.563	5.321
697	1	1	0.955	0.045	0.000	0.400	1.914
698	2	2	0.000	1.000	0.000	0.458	1.365
699	2	2	0.000	1.000	0.000	0.543	1.365
700	2	2	0.000	1.000	0.000	0.609	1.164
701	2	2	0.000	1.000	0.000	0.651	1.559

ID	*Actual Group*[1]	*Predicted Group (j)*[2]	*Posterior Probabilities* $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	*Typicality*[3]	D^{*2}
702	1	1	0.947	0.053	0.000	0.651	5.316
703	2	1	0.952	0.048	0.000	0.653	3.532
704	2	2	0.000	1.000	0.000	0.563	0.658
705	1	1	0.959	0.041	0.000	0.240	2.705
706	1	1	0.959	0.042	0.000	0.616	2.304
707	1	1	0.953	0.047	0.000	0.616	4.451
708	1	1	0.949	0.051	0.000	0.654	1.665
709	1	1	0.953	0.047	0.000	0.609	4.454
710	2	1	0.952	0.048	0.000	0.654	3.532
711	1	1	0.942	0.058	0.000	0.622	5.963
712	2	2	0.000	1.000	0.000	0.240	0.628
713	1	1	0.952	0.048	0.000	0.507	4.656
714	2	1	0.959	0.041	0.000	0.726	2.184
715	1	1	0.947	0.053	0.000	0.154	5.340
716	1	1	0.952	0.048	0.000	0.731	4.659
717	2	1	0.960	0.040	0.000	0.199	1.669
718	2	1	0.960	0.040	0.000	0.502	1.697
719	2	1	0.960	0.040	0.000	0.779	1.697
720	2	2	0.000	1.000	0.000	0.505	0.429
721	2	2	0.000	1.000	0.000	0.784	0.429
722	2	1	0.957	0.043	0.000	0.199	1.073
723	2	1	0.958	0.042	0.000	0.500	2.355
724	2	2	0.000	1.000	0.000	0.479	0.467
725	2	1	0.946	0.054	0.000	0.453	0.943
726	2	1	0.954	0.046	0.000	0.465	0.969
727	2	1	0.960	0.040	0.000	0.453	1.667
728	2	1	0.960	0.040	0.000	0.402	1.812
729	2	1	0.954	0.046	0.000	0.465	1.469
730	2	2	0.000	1.000	0.000	0.402	0.524
731	2	1	0.954	0.046	0.000	0.423	1.469
732	2	1	0.960	0.040	0.000	0.184	2.089
733	2	1	0.960	0.040	0.000	0.435	2.152
734	2	2	0.000	1.000	0.000	0.423	0.170
735	2	2	0.000	1.000	0.000	0.435	0.170
736	2	2	0.000	1.000	0.000	0.184	0.212
737	2	2	0.000	1.000	0.000	0.436	0.212
738	2	2	0.000	1.000	0.000	0.656	0.321
739	2	2	0.000	1.000	0.000	0.154	0.321
740	2	2	0.000	1.000	0.000	0.436	0.635
741	2	2	0.000	0.998	0.002	0.162	1.383
742	2	2	0.000	0.999	0.001	0.474	1.270
743	2	2	0.000	0.998	0.002	0.162	1.383
744	2	2	0.000	0.997	0.003	0.132	1.555
745	2	2	0.000	0.996	0.004	0.478	1.711
746	2	2	0.000	1.000	0.000	0.474	0.863
747	2	2	0.000	1.000	0.000	0.474	1.891
748	2	2	0.000	1.000	0.000	0.474	0.983
749	2	2	0.000	1.000	0.000	0.456	0.664
750	2	2	0.000	1.000	0.000	0.132	0.664

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
751	2	2	0.000	1.000	0.000	0.474	0.764
752	2	2	0.000	1.000	0.000	0.120	1.087
753	2	2	0.000	0.968	0.032	0.685	2.132
754	2	2	0.000	1.000	0.000	0.685	0.746
755	2	2	0.000	1.000	0.000	0.656	0.812
756	2	2	0.000	1.000	0.000	0.445	0.663
757	2	2	0.000	1.000	0.000	0.668	1.568
758	2	2	0.000	1.000	0.000	0.472	1.501
759	2	2	0.000	1.000	0.000	0.668	0.691
760	2	2	0.000	1.000	0.000	0.445	1.495
761	2	2	0.000	0.949	0.052	0.441	2.337
762	2	2	0.000	0.992	0.008	0.685	1.792
763	2	2	0.000	0.992	0.008	0.685	1.792
764	2	2	0.000	0.995	0.005	0.979	1.717
765	2	2	0.000	1.000	0.001	0.474	4.548
766	2	2	0.000	1.000	0.001	0.686	4.548
767	2	2	0.000	0.997	0.003	0.472	1.655
768	2	2	0.000	1.000	0.000	0.120	4.777
769	3	2	0.000	1.000	0.000	0.686	3.601
770	3	2	0.000	1.000	0.001	0.378	2.011
771	2	2	0.000	1.000	0.000	0.456	5.233
772	2	2	0.000	0.995	0.005	0.314	1.728
773	2	2	0.000	1.000	0.000	0.686	5.233
774	3	2	0.000	0.726	0.274	0.378	3.448
775	2	2	0.000	1.000	0.000	0.314	4.794
776	2	2	0.000	1.000	0.000	0.686	4.794
777	2	2	0.000	1.000	0.000	0.122	4.911
778	2	2	0.000	0.980	0.020	0.319	1.987
779	2	2	0.000	1.000	0.000	0.103	4.911
780	3	2	0.000	0.726	0.274	0.631	3.448
781	3	2	0.000	0.710	0.290	0.482	3.564
782	2	2	0.000	1.000	0.000	0.079	5.347
783	2	2	0.000	1.000	0.000	0.103	5.347
784	2	2	0.000	1.000	0.000	0.605	6.171
785	2	2	0.000	1.000	0.000	0.604	6.171
786	2	2	0.000	0.763	0.237	0.079	4.402
787	2	2	0.000	0.995	0.005	0.651	4.416
788	2	2	0.000	1.000	0.000	0.122	4.830
789	2	2	0.000	1.000	0.000	0.631	4.569
790	2	3	0.000	0.276	0.725	0.237	1.941
791	2	2	0.000	1.000	0.000	0.314	4.569
792	2	2	0.000	0.888	0.112	0.434	4.365
793	2	3	0.000	0.140	0.860	0.200	0.906
794	3	3	0.000	0.088	0.912	0.237	1.205
795	2	2	0.000	1.000	0.000	0.434	6.484
796	3	3	0.000	0.106	0.894	0.319	1.560
797	2	2	0.000	1.000	0.000	0.441	6.484
798	3	3	0.000	0.183	0.817	0.355	2.619
799	2	3	0.000	0.146	0.854	0.276	0.871

ID	Actual Group[1]	Predicted Group (j)[2]	Posterior Probabilities $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	Typicality[3]	D^{*2}
800	2	2	0.000	0.557	0.443	0.592	4.485
801	3	3	0.000	0.042	0.958	0.312	0.815
802	2	2	0.000	0.815	0.185	0.314	4.386
803	2	3	0.000	0.366	0.634	0.592	2.659
804	2	3	0.000	0.366	0.634	0.464	2.659
805	2	2	0.000	0.667	0.333	0.276	4.437
806	2	3	0.000	0.230	0.770	0.464	1.545
807	2	3	0.000	0.311	0.689	0.150	2.233
808	2	2	0.000	0.667	0.333	0.526	4.437
809	2	3	0.000	0.315	0.685	0.312	2.266
810	2	3	0.000	0.230	0.770	0.150	1.545
811	3	3	0.000	0.042	0.958	0.526	0.815
812	2	2	0.000	0.985	0.015	0.355	4.377
813	3	3	0.000	0.031	0.970	0.129	0.309
814	2	2	0.000	0.985	0.015	0.200	4.377
815	2	2	0.000	0.672	0.328	0.366	4.435
816	2	3	0.000	0.296	0.704	0.482	4.214
817	2	3	0.000	0.296	0.704	0.432	4.214
818	3	3	0.000	0.031	0.970	0.446	0.309
819	2	2	0.000	1.000	0.000	0.446	5.426
820	2	2	0.000	0.774	0.226	0.370	8.036
821	2	2	0.000	0.884	0.116	0.432	8.454
822	2	2	0.000	0.884	0.116	0.292	8.454
823	2	2	0.000	1.000	0.000	0.595	5.426
824	3	3	0.000	0.027	0.973	0.595	0.182
825	3	3	0.000	0.025	0.975	0.398	0.212
826	2	2	0.000	0.966	0.034	0.398	9.090
827	2	2	0.000	0.917	0.083	0.129	4.359
828	3	3	0.000	0.025	0.975	0.279	0.212
829	2	2	0.000	1.000	0.000	0.408	4.853
830	2	2	0.000	0.966	0.034	0.408	9.090
831	3	3	0.000	0.033	0.967	0.404	0.411
832	2	2	0.000	1.000	0.000	0.292	4.913
833	2	2	0.000	1.000	0.000	0.442	5.513
834	2	3	0.000	0.203	0.797	0.442	1.307
835	2	3	0.000	0.203	0.797	0.493	1.307
836	2	2	0.000	0.591	0.409	0.370	4.469
837	2	2	0.000	1.000	0.000	0.648	11.531
838	2	2	0.000	0.999	0.001	0.478	4.491
839	3	3	0.000	0.034	0.966	0.493	0.486
840	2	2	0.000	0.761	0.239	0.478	7.997
841	2	2	0.000	1.000	0.000	0.134	11.531
842	2	3	0.000	0.214	0.787	0.443	2.949
843	3	3	0.000	0.032	0.968	0.443	0.397
844	3	3	0.000	0.037	0.963	0.404	0.585
845	3	3	0.000	0.029	0.971	0.366	0.255
846	3	3	0.000	0.029	0.971	0.229	0.255
847	3	3	0.000	0.045	0.955	0.315	2.010
848	3	3	0.000	0.026	0.974	0.279	0.430

ID	*Actual Group*[1]	*Predicted Group (j)*[2]	*Posterior Probabilities* $\hat{P}(1\|\mathbf{x})$	$\hat{P}(2\|\mathbf{x})$	$\hat{P}(3\|\mathbf{x})$	*Typicality*[3]	D^{*2}
849	3	3	0.000	0.065	0.935	0.229	2.923
850	3	3	0.000	0.037	0.963	0.315	1.543
851	3	3	0.000	0.042	0.958	0.329	0.817
852	3	3	0.000	0.046	0.954	0.134	2.064
853	3	3	0.000	0.037	0.963	0.723	1.543
854	3	3	0.000	0.027	0.973	0.366	0.545
855	3	3	0.000	0.025	0.975	0.602	0.276
856	3	3	0.000	0.027	0.973	0.602	0.545
857	3	3	0.000	0.029	0.972	0.329	0.761
858	3	3	0.000	0.025	0.975	0.508	0.232
859	3	3	0.000	0.031	0.969	0.508	1.038
860	3	3	0.000	0.038	0.962	0.058	0.642
861	3	3	0.000	0.159	0.841	0.393	3.256
862	3	3	0.000	0.111	0.889	0.366	4.106
863	3	3	0.000	0.159	0.841	0.065	3.256
864	3	3	0.000	0.111	0.889	0.058	4.106
865	3	3	0.000	0.065	0.936	0.808	1.572

References

AB, Televilt TVP Positioning. *GPS-Simplex™, Televilt TVP Positioning AB*. Lindesburg: Televilt TVP Positioning AB, 2001.
Abraham, T., and J. F. Roddick. "Survey of Spatio-Temporal Databases." *GeoInformatica* 3, no. 1 (1999): 61 - 99.
Agresti, A. *Categorical Data Analysis*. 2nd ed. New York: John Wiley, 2002.
Ahonen-Rainio, P., and Kraak. M-J. "Deciding on Fitness for Use: Evaluating the Utility of Sample Maps as an Element of Geospatial Metadata." *Cartography and Geographic Information Science* 32, no. 2 (2005): 101-13.
Aminian, K., B. Najafi, C. Bula, P. F. Leyvraz, and Ph Robert. "Spatio-temporal Parameters of Gait Measured by an Ambulatory System using Miniature Gyroscopes." *Journal of Biomechanics* 35, no. 5 (2002): 689 - 99.
Anderson, T. W. *An Introduction to Multivariate Statistical Analysis*. 2nd ed. New York: John Wiley & Sons, Ltd., 1984.
Anderson, T. W., S. Das Gupta, and G. P. Styan. *A Bibliography of Multivariate Statistical Analysis*. Edinburgh: Olivier and Boyd, 1972.
Andrienko, G., N. Andrienko, P. Jankowski, D. Keim, M. J. Kraak, A. MacEachren, and S. Wrobel. "Geovisual analytics for spatial decision support: Setting the research agenda." *International Journal of Geographical Information Science* 21, no. 8 (2007): 839 - 57.
Anselin, L. *Spatial Econonmetrics: Methods and Models*. New York: Kluwer Academic, 1988.
Backus, Vickie. "Comprehensive Conservation Modeling: A Spatially Explicit Individual-Based Approach using Grizzly Bears as a Case Study." University of Utah, 2006.
Bédard, Y., C. Caron, Z. Maamar, B. Moulin, and D. Vallière. "Adapting Data Models for the Design of Spatio-Temporal Databases." *Computer, Environment and Urban Systems* 20, no. 1 (1996): 19 - 41.
Bélisle, M., and C. C. St. Clair. "Cumulative Effects of Barriers on the Movements of Forest Birds." *Conservation Ecology* 5, no. 2 (2001): 9.
Bender, D. J., L. Tischendorf, and L. Fahrig. "Using Patch Isolation Metrics to Predict Animal Movement in Binary Landscapes." *Landscape Ecology* 18, no. 1 (2003): 17 - 39.
Bertazzon, S. "Spatial and Temporal Autocorrelation in Innovation Diffusion Analysis." Paper presented at the International Conference on Computational Science and its Applications, Montreal, Canada, May 18 - May 21 2003.

Bethke, R., M. Taylor, S. Amstrup, and F. Messier. "Population Delineation of Polar Bears Using Satellite Collar Data." *Ecological Applications* 6, no. 1 (1996): 311 - 17.

Bhat, C. R. "Work Travel Mode Choice and Number of Non-work Commute Stops." *Transportation Research Part B: Methodological* 31, no. 1 (1997): 41 - 54.

Biewener, Andrew A. *Animal Locomotion*. Oxford: Oxford University Press, 2003.

Bigham, B. Oran. *Fast Fourier Transform and Its Applications*. Edited by Alan V. Oppenheim. Upper Saddle River, NJ: Prentice Hall, 1988.

Booch, G., J. Rumbaugh, and I. Jacobson. *The Unified Modeling Language User Guide*. Reading: Addison Wesley Longman Ltd., 1999.

Bouten, C. V. C., K. T. M. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen. "A Triaxial Accelerometer and Portable Data Processing Unit for the Assessment of Daily Physical Activity." *IEEE Transactions on Biomedical Engineering* 44, no. 3 (1997): 136 – 47.

Boyce, Mark S., Pierre R. Vernier, Scott E. Nielsen, and Fiona K. A. Schmiegelow. "Evaluating resource selection functions." *Ecological Modelling* 157, no. 2-3 (2002): 281.

Britting, Kenneth R. *Inertial Navigation Systems Analysis*. New York: Wiley - Interscience, 1971.

Buliung, R. N., and P. S. Kanaroglou. "On Design and Implementation of an Object-relational Spatial Database for Activity/Travel Behaviour Research." *Journal of Geographic Systems* 6 (2004): 237 - 62.

Burnham, Kenneth P., and David R. Anderson. *Model Selection and Multimodal Inference: A Practical Information-Theoretic Approach*. New York: Springer, 2002.

Burns, Robert. "To a Mouse." John Wilson, 1786.

Campbell, W. H. ""Magnetic" Pole Locations on Global Charts are Incorrect." *EOS Transactions, AGU* 77, no. 36 (1996): 345 - 47.

Canada, Environment. 2004. Canadian Climate Normals 1971-2000. In, Environment Canada, http://www.climate.weatheroffice.ec.gc.ca/climate_normals/index_e.html. (accessed 17 October, 2005).

Cardillo, Marcel. "Biological Determinants of Extinction Risk: Why are Smaller Species Less Vulnerable?" *Animal Conservation* 6, no. 01 (2003): 63 - 69.

Guidelines On: The Care and Use of Wildlife, Canadian Council on Animal Care (2003).

Carroll, Carlos, Reed F. Noss, and Paul C. Paquet. "Carnivores as Focal Species for Conservation Planning in the Rocky Mountain Region." *Ecological Applications* 11, no. 4 (2001): 961 - 80.

Caruso, M.J. "Applications of Magnetic Sensors for Low Cost Compass Systems." Paper presented at the Positioning, Location and Navigation Symposium (PLANS) 2000, San Diego, CA 2000.

Caruso, M.J., and L. S. Withanawasan. "Vehicle Detection and Compass Applications using AMR Magnetic Solutions." Paper presented at the Sensors Expo, Baltimore, MD 1999.

Charrier, Sophie, Sandrine Petit, and Francoise Burel. "Movements of Abax parallelepipedus (Coleoptera, Carabidae) in woody habitats of a hedgerow network landscape: a radio-tracing study." *Agriculture, Ecosystems & Environment* 61, no. 2-3 (1997): 133.

Cho, S. Y., K. W. Lee, C. G. Park, and J. G. Lee. "A Personal Navigation System using Low-Cost MEMS/GPS/Fluxgate." Paper presented at the ION 59th Annual Meeting/ CIGTF 22nd Guidance Test Symposium, Albuquerque, NM, 23 - 25 June 2003 2003.

Claramunt, C., C. Parent, S. Spaccapietra, and M. Thériault. "Database Modelling for Environmental and Land Use Change." In *Geographical Information and Planning: European Perspectives*, edited by S. Geertman, S. Openshaw and J. Stillwell. New York: Springer-Verlag Ltd., 1999.

Clark, P. J., and F. C. Evans. "Distance - Nearest Neighbour as a Measure of Spatial Relations in Populations." *Ecology* 35 (1954): 445 - 53.

Commission, Federal Communications. "Revision of the Commission's Rules To Ensure Compatibility with Enhanced 911 Emergency Calling Systems." Washington, D.C.: Federal Communications Commission, 1999.

Committee on the Status of Endangered Wildlife in Canada, COSEWIC. "COSEWIC Species Database: Bear, Grizzly." Government of Canada, 2002.

Connally, M. "Identifying Covariance Differences in Comparison of Linear versus Quadratic Classification Rule." University of Georgia, 2004.

Cox, D. R., and N. J. H. Small. "Testing Multivariate Normality." *Biometrica* 65, no. 2 (1978): 263 - 72.

Crist, Eric. P., and Richard C. Cicone. "A Physically Based Transformation of Thematic Mapper Data - The TM Tasseled Cap." *IEEE Transactions on Geoscience and Remote Sensing* GE-22, no. 3 (1984): 256 - 63.

D'Eon, Robert G., Robert Serrouya, Graham Smith, and Christopher O. Kochanny. "GPS Radiotelemetry Error and Bias in Mountainous Terrain." *Wildlife Society Bulletin* 30, no. 2 (2002): 430.

Daubechies, Ingrid. *Ten Lectures of Wavelets*. Philadelphia: SIAM, 1992.

Dettmers, R., D. A. Buehler, and J. B. Bartlett. "A Test and Comparison of Wildlife Habitat Modeling Techniques for Predicting Bird Occurrence at a Regional Scale." In *Predicting Species Occurrences: Issues of Accuracy and Scale*, edited by J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall and F. B. Samson. Washington, USA: Island Press, 2002.

Di Orio, Aaron P., Richard Callas, and Robert J. Schaefer. "Performance of Two GPS Telemetry Collars under Different Habitat Conditions." *Wildlife Society Bulletin* 31, no. 2 (2003): 372.

Dillon, W.R., and M Goldstein. *Multivariate Analysis: Methods and Applications*. New York: John Wiley & Sons Ltd., 1984.

Dobson, Annette. J. *An Introduction to Generalized Linear Models*. Boca Raton: Chapman and Hall, 1999.
Doncaster, C. P. "Non-parametric Estimates of Interaction from Radio-tracking Data." *Journal of Theoretical Biology* 143 (1990): 431 - 43.
Donoho, David L., Iain M. Johnstone, Gerard Kerkyacharian, and Dominique Picard. "Wavelet Shrinkage: Asymptopia?" *Journal of the Royal Statistical Society. Series B (Methodological)* 57, no. 2 (1995): 301 - 69.
Donovan, Therese M., and Frank R. Thompson Iii. "Modeling The Ecological Trap Hypothesis: A Habitat and Demographic Analysis for Migrant Songbirds." *Ecological Applications* 11, no. 3 (2001): 871 - 82.
Dussault, C., R. Courtois, J-P. Ouellet, and J. Huot. "Evaluation of GPS Telemetry Collar Performance for Habitat Studies in the Boreal Forest." *Wildlife Society Bulletin* 27, no. 4 (1999): 965 - 72.
Edwards, T. C., G. G. Moisen, and D. R. Cutler. "Assessing Map Accuracy in a Remotely Sensed Ecoregion-scale Cover Map." *Remote Sensing of Environment* 63 (1998): 73 - 83.
El-Sheimy, Naser. "Inertial Techniques and INS/GPS Integration." Calgary: The Department of Geomatics Engineering, The University of Calgary, 2003.
Emlen, S. T. In *Avian Biology*, edited by D. S. Farner and J. R. King. New York: Academic Press, 1975.
Erwig, M., R. H. Güting, M. Schneider, and M. Vazirgiannis. "Abstract and Discrete Modeling of Spatio-Temporal Data Types." Paper presented at the 6th ACM Symposium on Advances in Geographic Information Systems, Washington D.C., November, 1998 1998.
———. "Spatio-Temporal Data Types: An Approach - Modeling and Querying Moving Objects in Databases." *GeoInformatica* 3, no. 3 (1999): 269 - 96.
Erwig, M., and M. Schneider. "A Visual Language for the Evolution of Spatial Relationships and its Translation into a Spatio-Temporal Calculus." *Journal of Visual Languages and Computing* 14 (2003): 181 - 211.
Fan, Xitao. "An SAS Program for Assessing Multivariate Normality." *Educational and Psychological Measurement* 56, no. 4 (1996): 668 -75.
Fang, L., P. J. Antsaklis, L. A. Montestruque, M. B. McMickell, M. Lemmon, Y. Sun, H. Fang, I. Koutroulis, M. Haenggi, M. Xie, and X. Xie. "Design of a Wireless Assisted Pedestrian Dead Reckoning System - The NavMote Experience." *IEEE Transactions on Instrumentation and Measurement* 54, no. 6 (2005): 2342 - 58.
Fernández, L. J. E., J. de Cea Ch., and A. Soto O. "A Multi-modal Supply–demand Equilibrium Model for Predicting Intercity Freight Flows." *Transportation Research Part B: Methodological* 37, no. 7 (2003): 615 - 40.
Ferrier, Simon, Graham Watson, Jennie Pearce, and Michael Drielsma. "Extended Statistical Approaches to Modelling Spatial Pattern in Biodiversity in Northeast New South Wales. I. Species-Level Modelling." *Biodiversity and Conservation* 11, no. 12 (2002): 2275.

Fisher, R. A. "The Use of Multiple Measures in Taxonomic Problems." *Annals of Eugenics* 7 (1936): 179 - 88.

Foody, G. M. "Status of Land Cover Classification Accuracy Assessment." *Remote Sensing of Environment* 80 (2002): 185 - 201.

Forman, R. T. T., and M. Gordon. *Landscape Ecology*. New York: John Wiley & Sons, Ltd., 1986.

Fotheringham, A. S., C. Brunsdon, and M. Charlton. *Quantitative Geography. Perspectives on Spatial Data Analysis*. London: Sage Publications Ltd., 2000.

Fotheringham, A. S., and M. Wegener. *Spatial Models and GIS: New Potential and New Models*. London: Taylor and Francis, 2000.

Frair, J., S. Nielsen, E. Merrill, S. Lele, M. Boyce, R Munro, G. Stenhouse, and H. Beyer. "Removing GPS Collar Bias in Habitat Selection Studies." *Journal of Applied Ecology* 41 (2004): 201 – 12.

Frank, A. "Socio-Economic Units: Their Life and Motion." In *Life and Motion of Socio-economic Units*, edited by A. Frank, J Raper and J-P. Cheylan. New York: Taylor and Francis, 2001.

Frank, Ronald E., William F. Massy, and Donald G. Morrison. "Bias in Multiple Discriminant Analysis." *Journal of Marketing Research* 2, no. 3 (1965): 250 - 58.

Franklin, Alan B., David R. Anderson, R. J. Gutierrez, and Kenneth P. Burnham. "Climate, Habitat Quality, and Fitness in Northern Spotted Owl Populations in Northwestern California." *Ecological Monographs* 70, no. 4 (2000): 539 - 90.

Franklin, S. E., M. B. Lavigne, M. A. Wulder, and G. B. Stenhouse. "Change Detection and Landscape Structure Mapping Using Remote Sensing." *The Forestry Chronicle* 78, no. 5 (2002): 618-25.

Franklin, S. E., G. B. Stenhouse, M. J. Hansen, C. C. Popplewell, J. A. Dechka, and D. R. Peddle. "An Integrated Decision Tree Approach (IDTA) - Classification of Landcover in support of Grizzly Bear Habitat Analysis in the Alberta Yellowhead Ecosystem." *Canadian Journal of Remote Sensing* 27, no. 6 (2001): 579 - 93.

Gambaryan, P. P. *How Mammals Run: Anatomical Adaptations*. Translated by Hilary Hardin. New York: John Wiley & Sons, Ltd., 1974.

Gatrell, A. "Density Estimation and the Visualization of Point Patterns." In *Visualization in Geographical Information Systems*, edited by H. M. Hearnshaw and D. J. Unwin. New York: John Wiley & Sons, Ltd., 1994.

Gau, R. J., R. Mulders, L. J. Ciarniello, D. C. Heard, C. L. B. Chetkiewicz, M. S. Boyce, R. H. M. Munro, G. B. Stenhouse, B. Chruszcz, M. L. Gibeau, B. Milakovic, and K. L. Parker. "Uncontrolled Field Performance of Televilt GPS-Simplex (TM) Collars on Grizzly Bears in Western and Northern Canada." *Wildlife Society Bulletin* 32, no. 3 (2004a): 693-701.

Gau, Robert J., Robert Mulders, Lana M. Ciarniello, Douglas C. Heard, Cheryl-Lesley B. Chetkiewicz, Mark Boyce, Robin Munro, Gordon Stenhouse, Bryan Chruszcz, Michael L. Gibeau, Brian Milakovic, and Katherine L.

Parker. "Uncontrolled Field Performance of Televilt GPS-Simplex™ Collars on Grizzly Bears in Western and Northern Canada." *Wildlife Society Bulletin* 32, no. 3 (2004b): 693 - 701.
Getis, A. "Spatial Statistics." In *Geographical Information Systems: Principles, Techniques, Management and Applications*, edited by P. A. Longley, M. F. Goodchild, D. J. Maguire and D. W. Rhind. New York: John Wiley & Sons, Ltd., 1999.
Gibeau, Michael L., Anthony P. Clevenger, Stephen Herrero, and Jack Wierzchowski. "Grizzly Bear Response to Human Development and Activities in the Bow River Watershed, Alberta, Canada." *Biological Conservation* 103, no. 2 (2002): 227 - 36.
Giri, N. C. *Multivariate Statistical Analysis*. New York: Dekker, 2004.
Glick, Ned. "Additive Estimators for Probabilities of Correct Classification." *Pattern Recognition* 10 (1978): 211 - 27.
Goodwin, B. J., and L. Fahrig. "How does Landscape Structure influence Landscape Connectivity?" *Oikos* 99, no. 3 (2002): 552 - 70.
Gordon, Ian R. "The Ewe's Oestrous Cycle and Seasonal Breeding Activity." In *Controlled Breeding in Farm Animals*, 169 - 79. Oxford, UK: Pergamon, 1983.
Gould, James L. "The Locale Map of Honey Bees: Do Insects Have Cognitive Maps?" *Science* 232, no. 3 (1986): 861 - 69.
Goulias, K. G. "Longitudinal Analysis of Activity and Travel Pattern Dynamics using Generalized Mixed Markov Latent Class Models." *Transportation Research Part B: Methodological* 33, no. 8 (1999): 535 - 58.
Graps, Amara. "An Introduction to Wavelets." *IEEE Computational Science and Engineering* 2, no. 2 (1995): 50 - 61.
Greene, William H. *Econometric Analysis*. 5th ed. Upper Saddle River, New Jersey: Prentice Hall, 2002.
Greig-Smith, P. "The use of Random and Contiguous Quadrats in the Study of the Structure of Plant Communities." *Annals of Botany* 16, no. 62 (1952): 293 - 316.
Griffith, D. A., and L. J. Layne. *A Casebook for Spatial Statistical Data Analysis: A Compilation of Analyses of Different Thematic Data Sets*. New York: Oxford University Press, 1999.
Guo, D., M. Gahegan, A. M. MacEachren, and B. Zhou. "Multivariate Analysis and Geovisualization with an Integrated Geographic Knowledge Discovery Approach." *Cartography and Geographic Information Science* 32, no. 2 (2005): 113 - 33.
Güting, R. H., M. H. Böhlen, M. Erwig, C.S. Jensen, N.A. Lorentzos, M. Schneider, and M. Vazirgiannis. "A Foundation for Representing and Querying Moving Objects." *ACM Transactions on Database Systems* 25, no. 1 (2000): 1 - 42.
Güting, R. H., and M. Schneider. *Moving Objects Databases*. Edited by J. Gray, *Data Management Systems*. San Francisco: Morgan Kaufmann Publishers, 2005.

Hägerstrand, T. "What about People in Regional Science?" *Papers of the Regional Science Association* 24 (1970): 7 - 21.
Haining, R. *Spatial Data Analysis: Theory and Practice*. Cambridge: Cambridge University Press, 2003.
Hamer, J. D. W. "Feeding Ecology of Grizzly Bear in the Cascade and Panther Valleys of Banff National Park, Alberta." University of Calgary, 1985.
Hasse, P. "Spatial Pattern Analysis in Ecology based on Ripley's K-Function: Introduction and Methods of Edge Correction." *Journal of Vegetation Science* 6, no. 4 (1995): 575 - 82.
Heglund, P. J. "Foundations of Species-Environment Relations." In *Predicting Species Occurrences: Issues of Accuracy and Scale*, edited by J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall and F. B. Samson. Washington, USA: Island Press, 2002.
Herrero, S. "Introduction." In *Bears: Status Survey and Conservation Action Plan*, edited by C. Servheed, S. Herrero and B. Peyton. Gland, Switzerland: IUCN, 1999.
Hodge, G. *Planning Canadian Communities: An Introduction to the Principles, Practice, and Participants*. 4 ed. Scarborough: Thompson Nelson, 2003.
Hofmann-Wellenhof, B., H. Lichteneggar, and J. Collins. *Global Positioning System: Theory and Practice*. New York: Springer-Verlag Wien, 1997.
Holt, D., D. G. Steel, and M. Tranmer. "Area Homogeneity and the Modifiable Area Unit Problem." *Geographical Systems* 3 (1996): 181 - 200.
Hora, Stephen C., and James B. Wilcox. "Estimation of Error Rates in Several-population Discriminant Analysis." *Journal of Marketing Research* 19, no. 1 (1982): 57 - 62.
Hornsby, K., and M. J. Egenhofer. "Qualitative Representation of Change." Paper presented at the Conference on Spatial Information Theory, COSIT '97, Laurel Highlands, PA 1997.
Hosmer, David W., and Stanley Lemeshow. *Applied Logistic Regression*. 2 ed: John Wiley & Sons, Ltd., 2000.
Hsiao, Cheng. *Analysis of Panel Data*. 2nd ed. New York: Cambridge University Press, 2003.
Hubbard, Barbara B. *The World According to Wavelets: The Story of a Mathematical Technique in the Making*. 2nd ed. Natick, Massachusetts: A K Peters, 1998.
Huberty, Carl J. "Mahalanobis Distance." In *Encyclopedia of Statistics in Behavioural Science*, edited by B. Everitt and D. C. Howell, 1110 - 11. London: John Wiley & Sons, Ltd., 2005.
Huberty, Carl J., and Stephen Olejnik. *Applied MANOVA and Discriminant Analysis*. 2nd ed. Hoboken, New Jersey: John Wiley & Sons, Ltd., 2006.
Huberty, Carl J., and Joseph M. Wisenbaker. "Variable Importance in Multivariate Group Comparisons." *Journal of Educational Statistics* 17, no. 1 (1992): 75 - 91.

Hulbert, I. A. R. "GPS and its Use in Animal Telemetry: The Next Five Years." Paper presented at the International Conference on Tracking Animals with GPS, Aberdeen, Scotland, 12 - 13 March 2001.

Hunter, A. J. S., J Cranston, N. El-Sheimy, and G. B. Stenhouse. "An Integrated Approach for the Analysis and Visualization of Moving Objects." Paper presented at the First International Workshop on Mobile Geospatial Augmented Reality, Banff, Canada, May 29 - 30 2006.

Huston, Michael A. "Critical Issues for Improving Predictions." In *Predicting Species Occurrences: Issues of Accuracy and Scale*, edited by J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall and F. B. Samson. Washington, USA: Island Press, 2002.

Hutchins, R. B., R. L. Blevins, J. D. Hill, and E. H. White. "The Influence of Soils and Microclimate on Vegetation of Forested Slopes in Eastern Kentucky." *Soil Science* 121, no. 4 (1976): 234 - 41.

Imfeld, S. "Time, Points and Space - Towards a Better Analysis of Wildlife Data in GIS." University of Zurich, 2000.

MSP430x15x, MSP430x16x, MSP430x161x Mixed Signal Microcontroller Version E. Texas Instruments Inc., Dallas, Texas.

ISO_19108:2002(E). *Geographic Information — Temporal Schema*. Vol. ISO 19108:2002(E): International Organization for Standardization, 2002.

Itami, R., R. Raulings, G. MacLaren, K. Hirst, R. Gimblett, and D. Zanon. "RBSim 2: Simulating the Complex Interactions between Human Movement and the Outdoor Recreation Environment." *Journal of Nature Conservation* 11 (2003): 278 - 86.

J. J. Spiker, Jr. "GPS Signal Structure and Performance Characteristics." In *Global Positioning System*, edited by P. M. Janiczek, 29 - 54. Alexandria, VA: The Institute of Navigation, 1980.

Jain, Anil K., Robert P. W. Duin, and Jianchang Mao. "Statistical pattern recognition: a review." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, no. 1 (2000): 4 - 37.

Jelinski, D. E., and J. Wu. "The Modifiable Areal Unit Problem and Implications for Landscape Ecology." *Landscape Ecology* 11, no. 3 (1996): 129 - 40.

Jensen, John R. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Upper Saddle River, NJ.: Prentice Hall, 2005.

Jerde, Christopher L., and Darcy R. Visscher. "GPS Measurement Error Influences on Movement Model Parameterization." *Ecological Applications* 15, no. 3 (2005): 806 – 10.

Jirawimut, R., P. Ptasinski, V. Garaj, F. Cecelja, and W. Balachandran. "A Method for Dead Reckoning Parameter Correction in Pedestrian Navigation Systems." Paper presented at the 18th IEEE Instrumentation and Measurement Conference 2001.

Johnson, C. J., K. L. Parker, D. C. Heard, and M. P. Gillingham. "Movement Parameters of ungulates and Scale-specific Responses to the Environment." *Journal of Animal Ecology* 71 (2002): 225 - 35.

Johnson, C., and P.Ward. "GPS Application to Seismic Oil Exploration." In *Global Positioning System*, edited by P. M. Janiczek, 195 - 203. Alexandria, VA: The Institute of Navigation, 1980.

Johnson, R. A. *Miller & Freund's Probability and Statistics for Engineer*. 6th ed. New Jersey: Prentice-Hall, 2000.

Johnston, F. R., and P. J. Harrison. "Discount Weighted Moving Averages." *The Journal of the Operational Research Society* 35, no. 7 (1984): 629 - 35.

Jonsen, I. D., R. A. Myers, and J. Mills Flemming. "Meta-Analysis of Animal Movement using State-Space Models." *Ecology* 84, no. 11 (2003): 3055 - 63.

Jonsen, I. D., and P. D. Taylor. "Fine-scale Movement Behaviours of Calopterygid Damselfies are Influenced by Landscape Structure: An Experimental Manipulation." *Oikos* 88, no. 3 (2000): 553 - 62.

Kalkhan, M. A., R. M. Reich, and T. J. Stohlgren. "Assessing the Accuracy of Landsat Thematic Mapper Classification using Double Sampling." *International Journal of Remote Sensing* 19, no. 2049 - 2060 (1998).

Kanellakis, P. C., G. Kuper, and P. Revesz. "Constraint Query Language." *Journal of Computer and System Sciences* 51, no. 1 (1995): 26 - 52.

———. "Constraint Query Language." Paper presented at the 9th Symposium on Principles of Database Systems (PODS), Nashville, Tennessee, April 2 - 4 1990.

Käppi, J., J. Syrjärinne, and J. Saarinen. "MEMS-IMU Based Pedestrian Navigator for Handheld Devices." Paper presented at the ION GPS 2001, Salt Lake City, UT, 11-14 September 2001 2001.

Kareiva, P. M., and N. Shigesada. "Analyzing Insect Movement as a Correlated Random Walk." *Oecologia* 56, no. 234 - 238 (1983).

Keating, K. A., and S. Cherry. "Use and Interpretation of Logistic Regression in Habitat-Selection Studies." *Journal of Wildlife Management* 64, no. 4 (2004): 774 - 89.

Khan, Ashfaq. 2005. Digital Signal Processing Fundamentals. In, Da Vinci Engineering Press, http://www.ebrary.com.ezproxy.lib.ucalgary.ca/corp/index.jsp. (accessed 23 June, 2007).

Kim, J. W., H. J. Jang, D.-H. Hwang, and C. Park. "A Step, Stride and Heading Determination for the Pedestrian Navigation System." *Journal of Global Positioning Systems* 3, no. 1 - 2 (2004): 273 - 79.

Krebs, John R., Jonathan T. Erichsen, Michael I. Webber, and Eric L. Charnov. "Optimal Prey Selection in the Great Tit (*Parus major*)." *Animal Behaviour* 25, no. 1 (1977): 30 - 38.

Kwan, M-P. "Interactive Geovisualization of Activity-Travel Patterns using Three-Dimensional Geographical Information Systems: A Methodological Exploration with a large data set." *Transportation Research Part C* 8 (2000): 185 - 203.

Lachenbruch, Peter A. "An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis." *Biometrics* 23, no. 4 (1967): 639 - 45.

———. "Discriminant Diagnostics." *Biometrics* 53, no. 4 (1997): 1284 - 92.

Lachenbruch, Peter A., and M. Ray Mickey. "Estimation of Error Rates in Discriminant Analysis." *Technometrics* 10, no. 1 (1968): 1 - 11.

Ladetto, Q. "On Foot Navigation: Continuous Step Calibration Using Both Complementary Recursive Prediction and Adaptive Kalman Filtering." Paper presented at the ION GPS 2000, Salt Lake City, Utah, September 19 - 22. 2000.

Ladetto, Q., V. Gabaglio, and B. Merminod. "Two Different Approaches for Augmented GPS Pedestrian Navigation." Paper presented at the International Symposium on Location Based Services for Cellular Users: Locellus, Munich, Germany 2001.

Ladetto, Q., and B. Merminod. "Digital Magnetic Compass and Gyroscope Integration for Pedestrian Navigation." Paper presented at the 9th Saint Petersburg International Conference on Integrated Navigation Systems, Saint Petersburg, Russia, May 27-29 2002.

Lambeck, R. J. "Focal Species: A Multi-species Umbrella for Nature Conservation." *Conservation Biology* 11 (1997): 849 - 56.

Langran, G., and N. R. Chrisman. "A Framework for Temporal Geographic Information." *Cartographica* 25, no. 3 (1988): 1 - 14.

Lathi, B. P. *Linear Systems and Signals*. Berkeley: Cambridge University Press, 1992.

Laube, P. "A Classification of Analysis Methods for Dynamic Point Objects in Environmental GIS." Paper presented at the 4th AGILE Conference - GI in Europe: Integrative, Interoperable, Interactive, Brno, Czech Republic, April 19 – 21 2001.

Laube, P., S. Imfeld, and R. Weibel. "Discovering Relative Motion Patterns in Groups of Moving Point Objects." *International Journal of Geographical Information Science* 19, no. 6 (2005): 639 - 68.

Lee, S-W., and K. Mase. "Recognition of Walking Behaviours for Pedestrian Navigation." Paper presented at the IEEE Conference on Control Applications (CCA01), Mexico City, Mexico, September 2001a.

Lee, S.-W., and K. Mase. "Recognition of Walking Behaviors for Pedestrian Navigation." Paper presented at the Proceedings of 2001 IEEE Conference on Control Applications (CCA01), Mexico City, Mexico 2001b.

Legendre, P., and M. - J. Fortin. "Spatial Pattern and Ecological Analysis." *Vegetatio* 80 (1989): 107 - 38.

Leopold, Aldo. *A Sand Country Almanac and Sketches Here and There*. New York: Oxford University Press, 1949.

Levi, R. W., and T. Judd. "Dead Reckoning Navigational System using Accelerometers to Measure Foot Impacts." United States, 1996.

Lewis, M. A., and J. D. Murray. "Modelling Territoriality and Wolf-Deer Interactions." *Nature* 366, no. 23 (1993): 738 - 40.

Lowes, F.J. 2005. IAGA Division V-MOD Geomagnetic Field Modeling: IGRF Proper Use. In, IAGA Working Group VMOD, http://www.ngdc.noaa.gov/IAGA/vmod/igrfhw.html. (accessed May 3, 2006).

Lynch, J. P., A. Partridge, K. H Law, T. W. Kenny, A. S. Kiremidjian, and E. Carryer. "Design of Piezoresistive MEMS-Based Accelerometer for Integration with Wireless Sensing Unit for Structural Monitoring." *Journal of Aerospace Engineering* 16, no. 3 (2003): 108 - 14.

MacArthur, R. H., and E. O. Wilson. *The Theory of Island Biogeography*. Princeton: Princeton University Press, 1967.

Mace, Richard D., John S. Waller, Timothy L. Manley, Katherine Ake, and William T. Wittinger. "Landscape Evaluation of Grizzly Bear Habitat in Western Montana." *Conservation Biology* 13, no. 2 (1999): 367.

Mace, Richard D., John S. Waller, Timothy L. Manley, L. Jack Lyon, and Hans Zuuring. "Relationships Among Grizzly Bears, Roads and Habitat in the Swan Mountains, Montana." *Journal of Applied Ecology* 33, no. 6 (1996): 1395.

Macheiner, K., K. Legat, and B. Hofmann-Wellenhof. "Testing a Pedestrian Navigator." Paper presented at the European Navigation Conference, GNSS 2004, Rotterdam, The Netherlands 2004.

MacIntyre, N. E., and J. A. Wiens. "Interactions between Landscape Structure and Animal Behaviour: the Roles of Heterogeneously Distributed Resources and Food Deprivation on Movement Patterns." *Landscape Ecology* 14 (1999): 437 - 47.

Mahalanobis, P. C. "On the Generalized Distance in Statistics." *Proceedings of the National Institute of Sciences of India* 2, no. 1 (1936): 49 - 55.

Maken, J., R. G. Healey, and S. Dowers. "Simulation Modelling with Object-oriented GIS: A Prototype Application to the Time Geography of Shopping Behaviour." *Geographical Systems* 4, no. 4 (1997): 397 - 429.

Mallet, Stéphane. *A Wavelet Tour of Signal Processing*. 2nd ed. New York: Academic Press, 1999.

Manly, B. F. J., L. L. McDonald, D. L. Thomas, T. L. McDonald, and W. P. Erickson. *Resource Selection by Animals: Statistical Design and Analysis for Field Studies*. 2nd ed. Dordrecht, the Netherlands: Kluwer Academic Press, 2002.

Marei, E. J. "Mechanics of the Animal Organisms *(Russian)*." *St.Petersburg State University* (1875).

Martin, E. H. "GPS User Equipment Error Models." In *Global Positioning System*, edited by P. M. Janiczek, 109 -18. Alexandria, VA: The Institute of Navigation, 1980.

Martin, Paul, and Patrick Bateson. *Measuring Behaviour: An Introductory Guide*. Cambridge: Cambridge University Press, 1986.

McCormack, E. "Using a GIS to Enhance the Value of Travel Diaries." *Institute of Transportation Engineers Journal* 69, no. 1 (1999): 38 - 43.

McDermid, Dr. Greg. August 17, 2007. *personal communication*

McDermid, Greg J. "Remote Sensing for Large-Area, Multi-Jurisdictional Habitat Mapping." University of Waterloo, 2005.
McKay, R. J., and N. A. Campbell. "Variable Selection Techniques in Discriminant Analysis: Part I Description." *British Journal of mathematical and Statistical Psychology* 35, no. 1 (1982): 1 - 29.
McLachlan, G. J. "The Bias of Sample Based Posterior Probabilities." *Biometrical Journal* 19, no. 6 (1977): 421 - 26.
McLellan, B. N., and D. M. Shackleton. "Grizzly Bears and Resource-Extraction Industries: Effects of Roads on Behaviour, Habitat Use and Demography." *Journal of Applied Ecology* 25, no. 2 (1988): 451.
McLellan, Bruce N., Fred W. Hovey, Richard D. Mace, John G. Woods, Daniel W. Carney, Michael L. Gibeau, Wayne L. Wakkinen, and Wayne F. Kasworm. "Rates and Causes of Grizzly Bear Mortality in the Interior Mountains of British Columbia, Alberta, Montana, Washington, and Idaho." *The Journal of Wildlife Management* 63, no. 3 (1999): 911.
Mech, L. D., and S. M. Barber. "Critique of Wildlife Radio-Tracking and its Use in National Parks: A Report to the U.S. National Park Service." edited by Northern Prairie Wildlife Research Center U.S. Geological Survey. Jamestown, N.D.: Northern Prairie Wildlife Research Center Online, 2002.
Mennis, J. L, D. J. Peuquet, and L. Qian. "A Conceptual Framework for Incorporating Cognitive Principles into Geographical Database Representation." *International Journal of Geographical information Science* 14, no. 6 (2000): 501 - 20.
Meshbane, Alice, and John D. Morris. "A Method for Selecting Between Linear and Quadratic Classification Models in Discriminant Analysis." *Journal of Experimental Education* 63, no. 3 (1995): 263.
Mikhail, Edward M., James S. Bethel, and J. Chris McGlone. *Introduction to Modern Photogrammetry*. New York: John Wiley and Sons Inc., 2001.
Moen, Ron, John Pastor, Yosef Cohen, and Charles C. Schwartz. "Effects of Moose Movement and Habitat Use on GPS Collar Performance." *The Journal of Wildlife Management* 60, no. 3 (1996): 659 - 68.
Montgomery, D. *Design and Analysis of Experiments*. New York: John Wiley & Sons, Ltd., 2001.
Mooij, Wolf M., and Donald L. DeAngelis. "Error Propagation in Spatially Explicit Population Models: a Reassessment." *Conservation Biology* 13, no. 4 (1999): 930 - 33.
Moorcroft, P. R., M. A. Lewis, and R. L. Crabtree. "Home Range Analysis Using a Mechanistic Home Range Model." *Ecology* 80, no. 5 (1999): 1656 - 65.
Moore, Ian D. "Terrain Analysis Programs for the Environmental Sciences (TAPES)." *Agricultural Systems and Information Technology* 4 (1992): 37–9.
Moore, Ian D., T. W. Norton, and Jann E. Williams. "Modelling Environmental Heterogeneity in Forested Landscapes." *Journal of Hydrology* 150, no. 2 - 4 (1993): 717 - 47.

Morris, John D., and Alice Meshbane. "Selecting Predictor Variables in Two-group Classification Problems." *Educational and Psychological Measurement* 55, no. 3 (1995): 438 - 41.
Morrison, Michael L. "A Proposed Research Emphasis to Overcome the Limits of Wildlife-Habitat Relationship Studies." *The Journal of Wildlife Management* 65, no. 4 (2001): 613.
Munro, R. H. M., S. E. Nielsen, M. H. Price, G. B. Stenhouse, and M. S. Boyce. "Seasonal and Diel Patterns of Grizzly Bear Diet and Activity in West-Central Alberta." *Journal of Mammalogy* 87, no. 6 (2006): 1112 - 21.
MySQL. "MySQL Reference manual." 2006.
Nagy, J. A., W. L. Hawley, M. W. Barret, and J. W. Nolan. *Population Characteristics of Grizzly and Black Bears in West-Central Alberta*. Edited by Alberta Environment Centre. Vol. AECV88-R1. Vegrevill, Alberta, Canada: Government of Alberta, 1989.
Naves, Javier, Thorsten Wiegand, Eloy Revilla, and Miguel Delibes. "Endangered Species Constrained by Natural and Human Factors: the Case of Brown Bears in Northern Spain." *Conservation Biology* 17, no. 5 (2003): 1276.
Nellemann, Christian, and Raymond D. Cameron. "Effects of Petroleum Development on Terrain Preferences of Calving Caribou." *Arctic* 49, no. 1 (1996): 23 - 28.
Nielsen, Scott E., M. S. Boyce, and G. B. Stenhouse. "Grizzly Bears and Forestry I. Selection of Clearcuts by Grizzly Bears in West-central Alberta, Canada." *Forest Ecology and Management* 199, no. 1 (2004a): 51-65.
Nielsen, Scott E., M. S. Boyce, G. B. Stenhouse, and R. H. M. Munro. "Modeling Grizzly Bear Habitats in the Yellowhead Ecosystem of Alberta: Taking Autocorrelation Seriously." *Ursus* 13 (2002): 45 - 56.
Nielsen, Scott E., and Alan Haney. "Gradient Responses for Understory Species in a Bracken Grassland and Northern Dry Forest Ecosystem of Northeast Wisconsin." In *Transactions of the Wisconsin Academy of Sciences, Arts and Letters*, edited by William J. Urbrock, 149 - 66, 1998.
Nielsen, Scott E., S. Herrero, M. S. Boyce, R. D. Mace, B. Benn, M. L. Gibeau, and S. Jevons. "Modelling the Spatial Distribution of Human Caused Grizzly Bear Mortalities in the Central Rockies Ecosystem of Canada." *Biological Conservation* 120, no. 1 (2004b): 101 - 13.
Nielsen, Scott Eric, Gordon B. Stenhouse, and Mark S. Boyce. "A Habitat-based Framework for Grizzly Bear Conservation in Alberta." *Biological Conservation* 130, no. 2 (2006): 217.
Nolet, B. A., and W. M. Mooj. "Search Paths of Swans Foraging on Spatially Autocorrelated Tubers." *Journal of Animal Ecology* 71 (2002): 451 - 62.
Noss, R. F. "Indicators for Monitoring Biodiversity: A Hierarchical Approach." *Conservation Biology* 4, no. 4 (1990): 355 - 64.
Olmedo, E., R. Mateos, M. Sancho, and J.M. Valderas. 2002. From Linearity to Complexity: Towards a New Economics. In *Complexity International*, Monash University, http://www.csu.edu.au/ci/draft/olmed002/. (accessed Submitted preprint under review.

OmniVision. "OV7640 Color CMOS VGA (640 x 480) Camera Chip." Sunnyvale, CA: OmniVision, 2005.
Openshaw, S. "GeoComputation." In *GeoComputation*, edited by S. Openshaw and R. J. Abrahart, 1 - 31. London: Taylor and Francis, 2000.
———. *The Modifiable Areal Unit Problem. CATMOG 38*. Norwich, England.: GeoBooks, 1984.
Openshaw, S., and P.J. Taylor. "The Modifiable Areal Unit Problem." In *Quantitative Geography: A British View*, edited by Neil Wrigley and R. Bennett, 60 - 69. London: Routledge and Kegan Paul, 1981.
Openshaw, S., I. Turton, and J. MacGill. "Using the Geographical Analysis Machine - Analyze Limiting Long-term Illness Census Data." *Geographical and Environmental Modelling* 3, no. 1 (1999): 83 - 99.
Parwani, R. R. 2002. Complexity: An Introduction. In *Complexity International*, Monash University, http://www.csu.edu.au/ci/draft/parwan01/. (accessed Submitted preprint under review, 2006).
Peuquet, D. J. "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, no. 3 (1994): 441 - 61.
———. "Time in GIS and Geographical Databases." In *Geographical Information Systems: Principles, Techniques, Management, and Applications*, edited by P. A. Longley, M. F. Goodchild, D. J. Maguire and D. W. Rhind. New York: John Wiley & Sons, Ltd., 1999.
Peuquet, D. J., and N. Duan. "An Event-based Spatiotemporal Data Model (ESTDM) for Temporal Analysis of Geographical Data." *Journal of Geographical information Science* 9, no. 7 - 24 (1995).
Pfoser, D., and C. S. Jensen. "Capturing the Uncertainty of Moving-Object Representations." Paper presented at the 6th International Symposium on Advances in Spatial Data (SSD), Hong Kong, China, July 20 - 23 1999.
Pilkington. 2006. Glass that cleans itself! In, Pilkington Activ, http://www.pilkington.com/about+pilkington/education/glass+that+cleans+itself.htm. (accessed 24 October, 2006).
3-Axis Magneto-Inductive Sensor Driver and Controller with SPI Serial Interface. PNI Corporation, Santa Rosa, CA.
Preisler, H. K., D. R. Brillinger, A. A. Ager, J. G. Kie, and R. P. Akers. "Stochastic Differential Equations: A Tool for Studying Animal Movement." Paper presented at the IUFRO 4.11 Conference on Forest Biometry, Modelling and Information Science, Greenwich, UK, June 25-29 2001.
Purvis, Andy, John L. Gittleman, Guy Cowlishaw, and Georgina M. Mace. "Predicting Extinction Risk in Declining Species." *Proceedings of the Royal Society of London - B* 267, no. 1456 (2000): 1947 - 52.
Rencher, A. C. *Methods of Multivariate Analysis*. New York: John Wiley & Sons, Ltd., 2002.
Rettie, W. J., and P. D. McLoughlin. "Overcoming Radiotelemetry Bias in Habitat Selection Studies." *Canadian Journal of Zoology* 77, no. 8 (1999): 1175 - 84.

Ridley, Mark. *Animal Behaviour: A Concise Introduction*. Oxford: Blackwell Scientific Publications, 1986.

Rigaux, P., M. Scholl, and A. Voisard. *Spatial Databases with Application to GIS*. San Francisco: Morgan Kaufmann Publishers, 2002.

Ripley, B. D. "The Second Order Analysis of Stationary Point Processes." *Journal of Applied Probability* 13 (1976): 255 - 66.

Roever, Carrie L. "Grizzly Bear (Ursus Arctos L.) Selection of Roaded Habitats in a Multi-Use Landscape." University of Alberta, 2007.

Romesburg, H. C. *Cluster Analysis for Researchers*. Belmont: Life Long Learning Publications, 1984.

Rubin, Paul A. "A Comparison of Linear Programming and Parametric Approaches to the Two-Group Discriminant Problem." *Decision Sciences* 21, no. 2 (1990): 373 - 87.

Russell, R. E., R. K. Swihart, and Z. Feng. "Population Consequences of Movement Decisions in a Patchy Landscape." *Oikos* 103, no. 1 (2003): 142 - 52.

Schaafsma, W., and G. N. van Vark. "Classification and Discrimination Problems with Applications - Part IIa." *Statistica Neerlandica* 33 (1979): 91 - 126.

Schooley, R. L., and J. A. Wiens. "Finding Habitat Patches and Directional Connectivity." *Oikos* 102, no. 3 (2003): 559 - 70.

Searcóid, Mícheál Ó. *Metric Spaces*. London: Springer-Verlag Ltd., 2007.

Shannon, C. E. "A Mathematical Theory of Communication." *The Bell System Technical Journal*, 27 (1948): 379 – 423, 623-56.

Shin, E-H., and N. El-Sheimy. "A New Calibration Method for Strapdown Inertial Navigation Systems." *Zeitschrift fur Vermessungswesen* 127, no. 1 (2002): 1 - 10.

Shin, S. H., and H. S. Hong. "MEMS Based Personal Navigator Equipped on User's Body." Paper presented at the ION GNSS 18th International Technical Meeting of the Satellite Division, Long Beach, CA 2005.

Silby, R. M., H. M. R. Nott, and D. J. Fletcher. "Splitting Behaviour into Bouts." *Animal Behaviour* 39 (1990): 63 - 69.

Silverman, B. W. *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall, 1986.

Sistla, P., O. Wolfson, S. Chamberlain, and S. Dao. "Modeling and Querying Moving Objects." Paper presented at the Thirteenth International Conference on Data Engineering (ICDE13), Birmingham, UK 1997.

Slater, P. J. B., and N. P. Lester. "Minimising Errors in Splitting Behaviour into Bouts." *Behaviour* 79 (1982): 153 - 61.

Smith, J. N. M. "The Searching Behaviour of Two European Thrushes: I. Description and Analysis of Search Paths." *Behaviour* 48 (1974): 276 - 302.

Stauffer, D. E. "Linking Populations and Habitats: Where have we been? Where are we going?" In *Predicting Species Occurrences: Issues of Accuracy and Scale*, edited by J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler,

M. G. Raphael, W. A. Wall and F. B. Samson. Washington, USA: Island Press, 2002.

Stehman, S. V. "Basic Probability Sampling Designs for Thematic Map Accuracy Assessment." *International Journal of Remote Sensing* 20 (1999): 2423 - 41.

Steinbeck, John. *Of Mice and Men*, *Penguin Classic*: Penguin, 1938.

Stenhouse, G. B., J. Boulanger, J.L. Lee, K. Graham, J. Duval, and J Cranston. "Grizzly Bear Associations along the Eastern Slopes of Alberta." *Ursus* 16, no. 1 (2005): 31 - 40.

Stenhouse, G.B., M.S. Boyce, and J. Boulanger. "Report on Alberta Grizzly Bear Assessment of Allocation." Hinton: Alberta Sustainable Resource Development, Fish and Wildlife Division, Alberta, 2003.

Stenhouse, Gordon. March 28, 2005. *personal communication*

Sterling, R. "Development of a Pedestrian Navigation System using Shoe Mounted Sensors." University of Edmonton, 2004.

Sterling, R., K. Fyfe, and G. Lachapelle. "Evaluation of a New Method of Heading Estimation for Pedestrian Dead Reckoning using Shoe Mounted Sensors." *The Journal of Navigation* 58, no. 1 (2005): 31 – 45.

LIS2L02AS4 - MEMS Inertial Sensor:2-Axis - ±2g/±6g Linear Accelerometer. STMicroelectronics, Geneva, Switzerland.

Stockwell, D. R., and A. T. Peterson. "Bias in Biodiversity Data." In *Predicting Species Occurrences: Issues of Accuracy and Scale*, edited by J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall and F. B. Samson, 537 – 46. Washington, USA: Island Press, 2002.

Strong, W. L. *Ecoregions and Ecodistricts of Alberta*. Edmonton: Alberta Forestry, Lands and Wildlife, 1992.

Sugden, Andrew, and Elizabeth Pennisi. "When to Go, Where to Stop." *Science* 313, no. 5788 (2006): 775.

Sustainable Resource Development, Alberta. "Species Currently Listed under the Wildlife Act and New Species Assessed by the ESCC." Edmonton: Alberta Government, 2007.

Tanaka, Hideji, Yasuaki Takagi, and Yasuhiko Naito. "Swimming Speeds and Buoyancy Compensation of Migrating Adult Chum Salmon *Oncorhynchus keta* Revealed by Speed/Depth/Acceleration Data Logger." *Journal of Experimental Biology* 204, no. 21 (2001): 3895 - 904.

Taylor, M. K., S. Akeeagok, D. Andriashek, W. Barbour, E. W. Born, W. Calvert, H. D. Cluff, S. Ferguson, J. Laake, A. Rosing-Asvid, I. Stirling, and F. Messier. "Delineating Canadian and Greeenland Polar Bear (Ursus maritimus) Populations by Cluster Analysis of Movements." *Canadian Journal of Zoology* 79 (2001): 690 - 709.

Taylor, P. D., L. Fahrig, K. Henein, and G. Merriam. "Connectivity is a Vital Element of Landscape Structure." *Oikos* 68, no. 3 (1993): 571 - 73.

Teachman, Jay, Greg J. Duncan, W. Jean Yeung, and Dan Levy. "Covariance Structure Models for Fixed and Random Effects." *Sociological Methods and Research* 30, no. 2 (2001): 271 - 88.

Theberge, Jeanette. C. "Scale-dependent Selection of Resource Characteristics and Landscape Pattern for Female Grizzly Bear in the Eastern Slopes of the Canadian Rocky Mountains." University of Calgary, 2002.

Tischendorf, L., and L. Fahrig. "How Should we Measure Landscape Connectivity?" *Landscape Ecology* 15, no. 7 (2000a): 633 - 41.

———. "On the Usage and Measurement of Landscape Connectivity." *Oikos* 90, no. 1 (2000b): 7 - 19.

Trajcevski, G., O. Wolfson, K. Hinrichs, and S. Chamberlain. "Managing Uncertainty in Moving Objects Databases." *ACM Transactions on Database Systems (TODS)* 29, no. 3 (2004): 463 - 507.

Turchin, P. *Quantitative Analysis of Movement: Measuring and Modeling Population Redistribution in Animals and Plants*. Sunderland, Massachutsetts: Sinauer Associates Inc., 1998.

———. "Translating Foraging Movements in Heterogeneous Environments into Spatial Distribution of Foragers." *Ecology* 72, no. 4 (1991): 1253 - 66.

Valeo, Dr. Caterina. August 28, 2007. *personal communication*

van Horne, B. "Approaches to Habitat Modelling: The Tension Between Pattern and Process and Between Specificity and Generality." In *Predicting Species Occurrences: Issues of Accuracy and Scale*, edited by J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall and F. B. Samson. Washington, USA: Island Press, 2002.

Wagenaar, R. C., and W. J. Beek. "Hemiplegic Gait: A Kinematic Analysis using Walking Speed as a Basis." *Journal of Biomechanics* 25, no. 9 (1992): 1007 - 15.

Walker, James S. *A Primer on Wavelets and their Scientific Application*. Boca Raton: CRC Press LLC, 1999.

Wang, J-H., and Y. Gao. "A New Magnetic Compass Calibration Algorithm using Neural Networks." *Measurement of Science and Technology* 17, no. 1 (2006): 153 - 60.

Watanabe, Shinichi, Masako Izawa, Akiko Kato, Yan Ropert-Coudert, and Yasuhiko Naito. "A new technique for monitoring the detailed behaviour of terrestrial animals: A case study with the domestic cat." *Applied Animal Behaviour Science* 94, no. 1-2 (2005): 117.

Wentz, E. A., A. F. Campbell, and R. Houston. "A Comparison of Two Methods - Create Tracks of Moving Objects: Linear Weighted Distance and Constrained Random Walk." *International Journal of Geographical Information Science* 17, no. 7 (2003): 623 - 45.

White, A. P. "Cross Validation of Nearest Neighbour Discriminant Analysis- A Warning to SAS Users." *Journal of Statistical Computation and Simulation* 49, no. 3 (1994): 129 - 40.

White, Joseph D., Steven W. Running, Ramakrishna Nemani, Robert E. Keane, and Kevin C. Ryan. "Measurement and Remote Sensing of LAI in Rocky Mountain Montane Ecosystems." *Canadian Journal of Forest Research* 27, no. 11 (1997): 1714 - 27.

Wiens, J. A., R. L. Schooley, and Jr. R. D. Weeks. "Patchy Landscapes and Animal Movements: Do Beetles Percolate?" *Oikos* 78, no. 2 (1997): 257 - 64.

Williams, P. H., and K. J. Gaston. "Measuring more of Biodiversity: Can Higher-Taxon Richness Predict Wholesale Species Richness?" *Biological Conservation* 67, no. 3 (1994): 211 - 17.

Wilson, John P., and John C. Gallant. *Terrain Analysis: Principles and Applications*. New York: John Wiley & Sons, Ltd., 2000.

Wolfson, O. "Moving Objects Information Management: The Database Challenge." Paper presented at the 5th Workshop on Next Generation Information Technologies and Systems (NGITS'2002), Caesarea, Israel, June 25 – 26 2002.

Wolfson, O., S. Chamberlain, S. Dao, L. Jiang, and G. Mendez. "Cost and Imprecision in Modeling the Position of Moving Objects." Paper presented at the Fourteenth International Conference on Data Engineering (ICDE14), Orlando, FL 1998.

Wolfson, O., P. Sistla, S. Chamberlain, and Y. Yesha. "Updating and Querying Databases that Track Mobile Units." *Distributed and Parallel Databases* 7, no. 3 (1999): 257 - 57.

Worboys, M. F. "A Unified Model of Spatial and Temporal Information." *The Computer Journal* 37, no. 1 (1994): 26 -34.

Wrinch, D., and H. Jeffreys. "On Certain Fundamental Principles of Scientific Inquiry." *Philosophical Magazine* 42, no. 6th Series (1921): 369 - 90.

Yamartino, R.J. "A Comparison of Several "Single-Pass" Estimators of the Standard Deviation of Wind Direction." *Journal of Climate and Applied Meteorology* 23, no. 1362 - 1366 (1984).

Zhang, J., and A. Fujiwara. "Representing Household Time Allocation Behavior by Endogenously Incorporating Diverse Intra-household Interactions: A Case Study in the Context of Elderly Couples." *Transportation Research Part B: Methodological* 40, no. 1 (2006): 54 - 74.

Zollner, P. A., and S. L. Lima. "Search Strategies for Landscape-level Interpatch Movements." *Ecology* 80, no. 3 (1999): 1019 - 30.