Title
Space, Time, and Energy in Dismounted Navigation

Author
Irmischer, Ian Joseph

Publication Date
2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Space, Time, and Energy in Dismounted Navigation

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Geography

by

Ian Joseph Irmischer

Committee in charge:

Professor Keith C. Clarke, Chair

Professor Daniel R. Montello

Dr. Benjamin C. Trumble

Dr. Michael D. Hendricks

June 2016

The dissertation of Ian J. Irmischer is approved.

Daniel R. Montello

Benjamin C. Trumble

Michael D. Hendricks

Keith C. Clarke, Committee Chair

June 2016

Space, Time, and Energy in Dismounted Navigation

by

Ian J. Irmischer

ACKNOWLEDGEMENTS

I would like to express deepest appreciation to my committee chair Dr. Keith Clarke, for his years of helpful advice and patience. Keith has been an all-star mentor and role model. Without his dedicated guidance and support this dissertation would not have been possible.

I would also like to thank committee members, Dr. Daniel Montello and Dr. Benjamin Trumble. Dan's expertise in navigation and research design was instrumental in developing my methodology. Ben's advice with regards to energy expenditure and his instruction on statistical techniques was paramount.

In addition, COL(R) Michael Hendricks was responsible for sparking my interest in this topic and guiding me through the initial phases of project development. Much of this study relied on contributions he made as an Academy Professor at West Point.

I would like to thank Frank Davenport and Kevin Mwenda for their assistance with my statistical analysis. Frank introduced me to R and was an essential confidant as I prepared for my exams. Kevin was critical in my learning of Mixed Effects Modeling.

My sanity and this dissertation were contingent on the company and counsel of my friend, Greg Husak. The importance of our daily discussions, lunchtime recreation, and our families' evening and weekend get-togethers cannot be understated.

Most importantly, I would like to thank my wife Jennifer Bernstein. She supported me throughout and let me focus on my dissertation while taking care of the needs of our family. To my daughters, Ava and Dylan: I love you.

Lastly, I would like acknowledge the funding support provided by the U.S. Army's Geospatial Research Lab. This generous grant was essential to purchase equipment and fund required travel and contributed greatly to the success of this project.

Ian J. Irmischer
Curriculum Vitae

Personal:

Work: U.S. Army Student Detachment
University of California, Santa Barbara
Department of Geography
808-312-7047
irmischer@geog.ucsb.edu

Home: 7448 San Carpino Drive,
Goleta, CA, 93117
808-312-7047
ian.irmischer@gmail.com

Status: Married with two children
Security Clearance: Top Secret/SCI

Education:

PhD Geography (2016), **University of California, Santa Barbara**
Primary course work: Geographic Information Systems, Computational Modeling
Thesis: "Space, Time, and Energy of Dismounted Navigation"
Thesis Advisor: Dr. Keith C. Clarke

MA Geography (2006), **University California Santa Barbara**
Primary course work: Geographic Information Systems
Thesis: "An Examination of Line of Sight Algorithm Error"
Thesis Advisor: Dr. Keith C. Clarke

M.Sc. Engineering Management (2001) **University Missouri Rolla**

B.Sc. Chemical Engineering (1996) **Tulane University**

Appointments and Awards:

Omar N. Bradley Research Fellowship in Mathematics FY 2016

Advanced Civil Schooling Appointment, U.S. Army—PhD in Geographic Information Science, 2013

Bronze de Fleury Medal, 2013

Geospatial Engineer Officer Additional Skill Identifier (W2), 2013

Space Activities Additional Skill Identifier (3Y), 2009

Assistant Professor, Dept. of Geog & Environmental Eng, United States Military Academy, 2008

Instructor, Dept. of Geography & Environmental Engineering, United States Military Academy, 2006

Fully Funded Masters Degree (University of California, Santa Barbara), U.S. Army, 2004

Bronze Star (1 OLC), Meritorious Service Medal (1 OLC), Army Commendation Medal (1 OLC)

Distinguished Military Graduate, U.S. Army ROTC, Tulane University, 1996

Three Year Scholarship, U.S. Army ROTC, 1993

Military Experience:

Division Engineer, 25th Infantry Division, Schofield Barracks, HI 2013

Deputy Division Engineer, 25th Infantry Division, Schofield Barracks, HI 2012

Brigade Engineer and Facilities Manager, 3BCT, 25th Infantry Division, Jalalabad, Afghanistan 2012

Operations Officer and Executive Officer, 3BSTB, 25th ID, Jalalabad, Afghanistan, 2011–2012

Resident Command and General Staff Course (CGSC), 2010–2011

Instructor and Assistant Professor, United States Military Academy, West Point New York, 2006–2009

Bravo Company Commander, 588th Engineers, 4th Infantry Division, Baqubah, Iraq, 2003–2004

Geospatial Analysis Officer, 2nd BCT, 4th Infantry Division, Fort Hood, Texas 2003

Commander, 610th Engineer Detachment (Topo), Fort Hood, Texas, 2002.

Executive Officer, B Company, 40th Engineers, 1st Armor Division, Camp Bedrock, Bosnia, 1998–1999

Terrain Analysis Officer, 2-6 Infantry Battalion, 2nd Brigade, 1st Armor Division, Tuzla, Bosnia, 1998

Academic Experience:

Three years teaching experience with the Department of Geography and Environmental Engineering, United States Military Academy:

Course Director, EV378, Cartography

Instructor, EV398, Geographic Information Systems

Instructor, EV203, Physical Geography

Instructor, EV489A, Independent Study

Grants:

$35,800 funded by Army Geospatial Center (AGC), "Review of Current Algorithm for Energy Expenditures," 2014

$20,000 funded by the Topographic Engineer Center (TEC), "Terrain Visualization," 2007

Service:

Guest Lecture, "Military Applications of GIS," Harvard University, 2013

Instructor—Functional Academic Skills Training (FAST), FOB Fenty, Jalalabad, 2011–2012

Tutor—Remedial Mathematics, Leavenworth High School, 2009–2010

Supervisor —Advanced Individual Academic Development to US Army Topographic Center, 2008

Supervisor—Advanced Individual Academic Development to ESRI, 2008

Supervisor—Advanced Individual Academic Development to Panama, 2007 and 2008

Undergraduate Academic Counselor (2006–2009)—Responsible academic progress of 15 students/yr

Undergraduate Club Officer Representative (2006–2009)—Cadet Mountaineering Team.

Disaster Relief Assistance, Hurricane Katrina, FEMA, Baton Rouge, LA, 2005

Professional Activities:

Member, Army Engineer Association, 1996–present

Member, Association of the United States Army, 1996–present

Member, Society of Military Engineers, 2010–present

Member, Cartography and Geographic Information Society, 2012

Member, Gamma Theta Upsilon, International Geographical Honor Society

Member, Association of American Geographers, 2004–present

Publications and Professional Articles:

Schmidt, D.W., **I. Irmischer,** C. Jensen, A. Khan-Hernandez, M. Peeler, M. Shafer. 2010. *Technical Report for Secretary of Defense—National Security Implications of Climate Change.* U.S. Army Command and General Staff College, Fort Leavenworth, Kansas.

Fleming, S., **I. Irmischer,** D. Roux, B. Boxler, A. Moomin, A. Paulo, M. Seese, L. Summerlin, and S. Vaughn. 2009. *Field Technical Report—TRTC and YTC Test Site Characterization, Mapping and Imaging,* Center for Environmental and Geographical Sciences, Department of Geography and Environmental Engineering, United States Military Academy, West Point, NY.

Fleming, S., T. Baker, H. Didier, N. Dieter, **I. Irmischer,** M. Lewis, E. McAllister and L. McBride. 2008. *Field Technical Report—Test Site Characterization, Mapping and Imaging,* Department of Geography and Environmental Engineering, United States Military Academy, West Point, NY and GIS Cell, G3—Operations, Yuma Proving Ground, Yuma, AZ.

Cartographic Illustration:

Krakowka, Amy R. and Hummel, Laurel J. (2009) Understanding Africa: A Geographic Approach. Carlisle, Pennsylvania: Center for Strategic Leadership United States Army War College.

Conference Presentations:

Irmischer, I. (2015) *Space, Time, and Energy in Dismounted Navigation.* Human Mobility, Cognition, and Geographic Information Science, Copenhagen, DE, Nov, 9, 2015.

Irmischer, I. (2015) *Individual Differences That Affect Navigational Performance.* Annual Meeting of American Association of Geographers, Chicago, IL, April, 20, 2015.

Irmischer, I., (2014) *The Energy of Dismounted Navigation.* Colloquium Lecture Series, University of California Santa Barbara, Santa Barbara, California, December, 2014.

Irmischer, I., (2014) *Google's Geo-Tools for Research & the Classroom.* Spatial Technology Lecture Series, University of California Santa Barbara, Santa Barbara, California, November, 2014.

Irmischer, I., Lytle, D., (2009) *Web-based Thematic Cartography Instruction: Methods and Assessment.* Annual Meeting of American Association of Geographers, Las Vegas, NV, March, 2009.

Irmischer, I., (2008) *Teaching Cartography through Distance Learning.* ESRI Education User Conference, San Diego, CA, August, 2008.

Irmischer, I., Hendricks, M.., (2008) *A Faculty Development Model for Using Geographic Tools in the Classroom.* Annual Meeting of American Association of Geographers, Boston, MA, April, 2008.

Irmischer, I., (2007) *Volunteered Geographic Information Uses for National Security.* Workshop on Volunteered Geographic Information, Santa Barbara, California, December, 2007.

Irmischer, I., Hendricks, M., Harrison, J., Galgano, F., (2007) *GIS Data Preparation and Model Development to Support Analysis in Military Geography.* Annual Meeting of American Association of Geographers, San Francisco, CA, April, 2007.

Irmischer, I., (2006) *Line of Sight Algorithm Development.* Colloquium Lecture Series, UCSB Department of Geography, 2006.

Irmischer, I., (2005) *Modeling Improvised Explosive Devise Attacks in a Geographic Information System.* Graduate Student Seminar, UCSB Department of Geography, 2005.

Irmischer, I., (2005) *Line of Sight Error Assessment.* Colloquium Lecture Series, UCSB Department of Geography, 2005.

Irmischer, I., (2004) *GIS in the United States Military.* Graduate Student Seminar, University of California, Santa Barbara, 2004.

Software Skills:

ESRI ArcGIS 10.2, ArcSDE, ArcIMS, ERDAS Imagine, VB, Python, R, Matlab, VBA, HTML

Professional Development:

Human mobility, cognition, and GISc Workshop, University of Copenhagen, Denmark, 2015.

Google for Higher Ed Summit, Mountain View, CA, 2014.

Master Teacher Program, USMA, 2006–2007.

Presenting Data and Information, Edward Tufte, Manhattan, NY, 2007.

New Instructor Training, USMA 2006.

ABSTRACT

Space, Time, and Energy in Dismounted Navigation

by

Ian Joseph Irmischer

Navigation, defined as goal-related movement through space and time to reach a destination, is a fundamental human activity. Geographers, physiologists, archaeologists, anthropologists, and psychologists have long been interested in the spatial, temporal, and energy expenditure aspects of navigation. Hikers, search and rescue teams, firefighters, the military, and others navigate on foot through rugged terrain, and their success depends on understanding how the dynamics of foot-based navigation affect individual capabilities, caloric requirements, and risk potential.

This research project modeled energy expenditure and speed of movement of human beings engaged in foot-based navigation in wooded environments with varied terrain. The models were developed using spatiotemporal analysis of a subject's movement trajectories and biometrics. Energy expenditure estimates were collected via biosensors and Global Navigation Satellite Systems (GNSS) from subjects while they engaged in foot-based navigation through undeveloped, forested landscapes. Trajectory data from 200 subjects were merged with a land cover data set to analyze characteristics of human navigation over varying slopes and terrain. Generalizing these characteristics provided a model of energy expenditure and navigational speed from an origin to a destination along an unknown route. The equation developed to model energy expenditure of a human's route during navigation

uses terrain slope, land cover, body mass index (BMI), sex, and traveled distance to predict Calorie consumption with an accuracy of 89 percent. The model of navigation speed accurately predicts route completion time within 10 percent. These models help to explain the human dynamics of navigation.

Contents

TABLE OF FIGURES

CHAPTER 1

INTRODUCTION

One of the most basic tasks of human existence is to move from an origin to a destination. The coordinated and goal-directed movement through an environment is called navigation (Montello 2005). The task is accomplished many times over during the course of a day. This research helps us understand the energy required for human navigation. It also investigates the speed at which the task is accomplished.

An important subset of origin-to-destination movement is that which is done in the natural wilderness. Navigation on foot through the wilderness is done by hikers, search and rescue personnel, firefighters, and the military, among many others. It is often referred to as dismounted navigation. The human dynamics of dismounted navigation are critical to understanding individual capabilities, requirements, and risk.

Current models of energy expenditure for wilderness navigation are not sufficient to assist users during route planning and logistical estimation. Most models are crude when considering land cover implications. None have been validated with field data. Improving our understanding of Calorie consumption during navigation will prove to be beneficial: Hikers will be able to estimate caloric needs based on trail choice; search and rescue will improve load planning and routing; and societal uses of an energy expenditure model for on-foot navigation are many.

Understanding human movement rates during navigation is also important. Certainly, search and rescue workers would like to understand the time-based range capabilities of a lost person. Hikers would benefit to know how long it would take to navigate through an unknown portion of their journey. Military planning often requires estimating how long a patrol would take when navigating through unfamiliar terrain, which is especially important when an operation is contingent upon precise timing. Just like a model of energy expenditure, knowing how fast people can move while navigating has far-reaching implications.

When moving from an origin to a destination, people must constantly make decisions on what route to take. These decisions are based on an infinite number of variables. Increasingly, computational aids assist humans during this process. Devices and applications support human decision-making by assessing and analyzing their route choice prior to the execution of their movement. It is quite common for an individual to use Google Maps or other web-based systems when planning how to accomplish navigational tasks. These web-based analytical tools can provide route planners with critical advice. Simple input of an origin and a destination can yield routing information to be used to select a course. Routing algorithms, in conjunction with Geographic Information Systems (GIS), are capable of prediction methods to help navigators plan courses. These route planning aids can compute the shortest route, the fastest route, and other valuable metrics. The metric of energy expenditure (EE), specifically its effect on navigation, is critically under-researched. Additionally, speed of human movement when people are navigating in the wilderness has been somewhat taken for granted. Much work has been done in the fields of geography, exercise physiology, and kinesiology to predict human walking speed. However, very little has been done to account for speed decreases due to the cognitive cost of navigation.

This dissertation research models the human energy expenditure of foot-based navigation in the outdoor environment. Further, it investigates the speed at which humans conduct dismounted navigation. It uses Geographic Information theory and systems to create mathematical representations of EE and speed and investigates how both of these models are based on environmental and terrain characteristics. Additionally, it explores how individual physiological characteristics can contribute to EE models. Collection of spatio-temporal information has provided a means for statistical modeling of this important human process.

Pinpointing the focus, one of the specific motivations of this research is to better understand military land navigation. It is a basic skill required of every Soldier and Marine, and military land navigation is taught at each form of basic training as well as at all U.S. Army leadership schools. Improved understanding of the human dynamics of dismounted navigation is of great interest to the armed services.

One of the major components of military land navigation is route selection (U.S. Army 2013). Route selection is an element of every military operation involving dismounted troops. Conservation of human energy expenditure is a vital consideration when planning routes, since physical exhaustion has been shown to cause degraded cognitive ability and decision making (Fleury and Bard 1987). Fatigue has also been shown to lead to lower than normal physical performance in tasks such as marksmanship (Tharion et al. 1997). Military field manuals emphasize that energy consumption prediction and exhaustion avoidance are critical when planning military operations and patrols (U.S. Army 1990). There are many geospatial tools available to assist Soldiers and Marines during route selection (See ArcGIS Military Analyst among others). However, no validated tools address the variable of consumed energy. The omission of tools that provide route planning assistance based on potential exhaustion leaves a gap in current U.S. military capability.

Likewise, calculation of human movement speed during navigation is essential for military operations. Many missions require adjacent units to coordinate meeting in space and time. Estimation of human navigational speed is necessary to successfully arrive at a certain location at a specified time. Understanding dismounted movement rates during long-range reconnaissance planning is essential to maximizing enemy detection capability. It also minimizes the risk of capture. The integrated use of dismounted troops and aviation assets demands detailed movement speed planning and analysis. Current navigation models do not adequately account for slowed movement due to wayfinding.

This research effort improves our understanding of geospatial considerations available to our military, specifically those that focus on route planning and navigation. Improving geospatial-related technologies will support the U.S. Army Warfighter through enhanced knowledge of the environment and his/her individual capabilities.

Components of Modeling Energy Expenditure and Speed of Navigation

Computational assessment of the best route has been studied since the 1950s, when the Bellman-Ford algorithm was formulated (Ford and Fulkerson 1962). Geospatial-based routing tools have been used by military planners for more than 25 years (Reynolds and Taylor 1988). However, these tools have disregarded the need for planning routes based on physical exertion. Dismounted operations in the mountainous terrain of Afghanistan have emphasized a need to assist mission planners with tools that can compute human energy expenditure of dismounted navigation. This study has partially filled this needed gap. Additionally, this study betters our understanding of human movement rates. It uses empirical data to model movement speeds for individuals conducting dismounted navigation in wooded and hilly terrain.

Energy expenditure in humans due to physical activity varies in accordance with the individual characteristics of the navigator. If two humans navigated between an origin and a destination along the same exact route, at the same exact time, they would expend a different quantity of Calories. These differences are based on factors such as Body Mass Index (BMI), fitness, and sex. This study models how these individual differences cause variance in energy expenditure.

The energy cost of dismounted navigation also varies over different terrain conditions. If other factors are held constant, steep uphill navigation is more energy-expensive than navigation over level ground. The vegetation, or more generally land cover, through which a navigator must traverse also affects human energy expenditure. For example, navigating over vegetated terrain costs more time and energy than navigating over a paved road; navigating in snow requires more human energy than navigating over a dirt path.

Similar to EE, predicted speed of navigation is dependent on terrain conditions. This research shows how changing land cover can slow the movement rates of human navigation. Additionally, the grade at which an individual travels up or down also limits the speed of travel. Most importantly, however, this examination attempts to show how movement rates are slower when individuals are required to continually assess their current position and make new routing decisions.

Scale of the Model Development

Most dismounted military navigation between an origin and a destination occurs at the temporal scale of hours, not minutes, days, or weeks. From a spatial perspective, the majority of military dismounted navigation covers distances in the range of 3–15 kilometers.[1]

[1] The spatial and temporal scales of military dismounted navigation were developed from the investigator's 18 years of U.S. Army route planning experience.

Therefore, the research concentrates on navigation that occurs uninterrupted for approximately 2–4 hours. It studies navigators who are traversing routes between 4–8 kilometers. This scale is also deemed appropriate since it is in keeping with normal human eating habits, such that navigators do not need to consume Calories during the navigation event.

Research Contributions and Products

The core contributions and the research products that have been developed from this work are specifically focused on the dismounted route planning. Implications of human energy expenditure and movement speeds of navigation have been examined: (1) The research has assigned relative weights to individual characteristics that contribute to energy expenditure; (2) the significance of environmental and terrain effects to energy expenditure have been determined; (3) some current energy expenditure models have been assessed; and (4) the analysis has developed a model for human navigational speeds considering slope and land cover as the independent variables.

The products of this research are fourfold: (1) The dissertation develops a model of human energy expenditure due to dismounted cross-country navigation through wooded terrain; (2) it develops a quadratic function that can be used to predict the dismounted navigational speed of human beings, based on slope percentage and land cover; (3) it provides a quantitative assessment of the routing algorithm developed by U.S Army's Geospatial Research Lab (GRL); and (4) analytic techniques for handling and visualizing large quantities of mobility data are developed.

Research Questions and Hypotheses

This research strives to answer quantitative questions about space, time, and energy used during human navigation. The first inquiry into navigation is focused on individual differences, posing the following query:

1. What are the contributing weights of BMI, fitness[2], and sex (individual characteristics) to an energy expenditure model of dismounted navigation?

I anticipate that these factors do not contribute equally to the amount of energy a person expends during dismounted navigation. This research will quantify and measure the effects of individual differences on the energy expenditure of human beings.

Similarly, the second inquiry involves understanding the energy expenditure of navigation, although it differs in that it investigates the environmental effects of energy expenditure during navigation. Specifically, this research asks:

2. What are the contributing weights of slope, land cover, and distance traveled from origin (environmental/terrain variables) to energy expenditure while navigating?

Here I postulate that environmental and terrain variables do not contribute equally to the amount of energy a person expends during dismounted navigation.

---

[2] BMI and fitness are known to be correlated (Leyk et al. 2006). Care was taken during data analysis to address this known association.

The project also investigates a previously un-validated routing algorithm used by the U.S. Military to assist in route planning. The algorithm's accuracy is assessed in this question and further answers:

3. How well do current energy expenditure models used by the U.S. Army match the experimental energy expenditure data?

I postulate that current energy expenditure models do not accurately predict caloric expenditure in navigators. This research will provide critical insight to the correct weighting of model parameters. It will also quantitatively assess the error associated with GRL's energy expenditure model by using field-based empirical data.

Finally, the relationship between slope and movement speed is examined. This area of study will be used to better understand how quickly (or slowly) humans move while they are navigating in hilly woods, posing the following question:

4. How does arduous wayfinding affect human movement speed when navigating in hilly wooded terrain?

I postulate that a change in terrain elevation will affect a navigator's speed. It is also proposed that the act of navigating affects movement speed. Specifically, it is hypothesized that people move at a slower speed when they are constantly planning movement, conducting location assessment, and selecting/adjusting routes.

This research collected energy expenditure estimates of Soldiers navigating on foot over hilly, wooded terrain. The estimates provided an opportunity to model energy expenditure during dismounted navigation. The model defines the contribution of individual and terrain factors to human energy expenditure. Knowledge of the influences contributing to energy expenditure provides a basis for assessing and improving current GIS routing tools, which in turn leads to improved routing algorithms. Further, it advances geospatial tools for the mission planner and improves current capabilities. These improvements will undoubtedly increase the operational safety of woodland firefighters, the success rates of search and rescue teams, and it will increase military mission accomplishment and countless other applications. Finally, it provides better tools to access information about how the environment and human capabilities affect navigation.

The remainder of this document is organized to communicate the study background, methods, data, results, and conclusions. Chapter 2 reviews the significant literature in the fields of energy expenditure, navigation, movement rates, GIS, and spatial analysis as they relate to this research effort. Chapter 3 describes the methods used for data collection, data analysis, visualization, and communication. Chapter 4 explains the data, striving to provide enough detail to future researchers to validate findings and continue work with published data. Chapter 5 describes the results of the dissertation research. Finally, Chapter 6 describes the conclusions and recommends avenues for future work.

CHAPTER 2

PREVIOUS RESEARCH

Navigation

Navigation is defined as the coordinated and goal-directed movement through an environment (Montello 2005). It involves the physical act of moving and the cognitive aspects of deciding on and following a route. These two components of navigation are classified as *locomotion* and *wayfinding*.

Human locomotion during navigation can be accomplished in many ways, such as in a vehicle, assisted by bicycle, in an airplane, or by foot. Each of these forms of locomotion requires energy. This research will focus on unassisted human locomotion by walking. Locomotion via walking involves body movement through muscular contraction, which of course, requires human energy. One of the primary goals of this research is to develop a predictive model of energy expenditure due to navigational locomotion.

The second component, wayfinding, is a more cognitively centered aspect of navigation. Wayfinding involves planning movement, route selection, continued reassessment of location, and the constant decision-making process to adjust the route. The use of the brain power and sensory processes require the expenditure of some energy to wayfind. Navigators continually update their location by recognizing landmarks and terrain features. Humans primarily use vision to accomplish these tasks, but also use hearing and the vestibular senses. However, the expenditure of energy for these is insignificant in comparison to walking or muscular contraction (Clarke and Sokoloff 1999; Lennie 2003). Still, a navigator uses muscular energy while wayfinding to assist in the sensory processes - to look around, turn around, or look behind oneself. Investigation of these activities involving muscular movement are included in our modeling effort.

Individual differences have been shown to impact navigation in many ways. Variances in a person's body mass index (BMI), fitness level, and sex have been shown to affect locomotion technique, efficiency, and energy expenditure during physical activity (Wolinsky and Driskell 2008a). These factors will be investigated in this study.

The cognitive contributions to energy expenditure during navigation is fascinating, but lie beyond the scope of this research. Instead, the focus is on the impacts of physiological differences among individuals, and how these affect energy expenditure. A detailed review of how energy expenditure in humans is affected by BMI, fitness level, and sex will be discussed in the following sections.

The research presented here will also investigate how movement speeds are impacted by cognition. The next research question looks specifically at how cognition affects navigation rates. Using Montello's theoretical framework of navigation (Montello 2005), we can devise an equation:

$$\text{Navigation} = \text{Wayfinding} + \text{Locomotion}$$

Then a framework can be designed to study the cognitive cost of navigation (wayfinding) if navigation speed and locomotion speeds are known.

Military Land Navigation

This investigation will narrow the focus to military dismounted navigation. The research will also focus on U.S. Army navigation, as opposed to naval or aeronautical navigation.

Sometimes Soldiers run from point to point, although this is less common than walking. Therefore, this research mostly studies Soldiers walking rather than running. Military navigation uses maps, lensatic compasses, and GPS devices as the chief tools for navigation (U.S. Army 2013).

The Soldier's *Manual of Common Tasks* describes navigation as a complex task that involves many subtasks, such as identifying topographic symbols on a military map and measuring distances on a map (U.S. Army 2012). It also requires the individual to be able to orient a map and determine a magnetic azimuth using a lensatic compass, as well as to determine location on the ground by terrain association. Terrain association is the process by which an individual senses the environment (usually visually), and accurately compares the real world with the map. The process involves accurately identifying terrain features that are in the navigator's field of view (e.g., hilltops, ridges, valleys, roads, streams, etc.) on the map (U.S. Army 2012). Map orientation, distance and direction finding, and terrain association are instrumental components of wayfinding.

Currently Used Models of Human Dismounted Movement (Locomotion)

One of the oldest models of estimating human walking speed over cross-country terrain was devised by William W. Naismith. On May 2, 1892, Naismith set out alone for a hike through the hills and mountains near Crainlarich, Scotland. After completing a 10-mile hike and climbing 6,300 feet in altitude, Naismith concocted his simple formula. The excursion had taken Naismith 6.5 hours to complete. He writes at the end of his journal entry: "This tallies exactly with a simple formula that may be found useful in estimating what time men in fair condition should allow for easy expeditions, namely, 1 hour for every 3 miles on the map, with an additional hour for every 2,000 feet of ascent" (Naismith 1892). This rule of thumb has been published many times over, and is still used today to estimate the time required for a given walk cross country (Aitken 1977; Langmuir 1984; Clarke 2014).

Since its inception in the 19th century, there have been several efforts to improve Naismith's model. Naismith originally focused on the horizontal and vertical component of the terrain. Others recognized the need to include individual fitness, terrain type, and a downhill slope correction. In 1965 a correction, shown in Figure 1, was created by Scottish Mountaineer Phillip Tranter. Tranter's adjustment corrected Naismith's rule, based on a hiker's fitness level (Langmuir 1969). In 1977 Aitken refined the model to include terrain conditions. He noted that a man can walk at 5 kph (~3mph) on paths, tracks, and road but this speed is reduced to 4 kph over all other types of terrain. Later, Langmuir offered further adjustments. He recommended a downhill adjustment, such that speed is a variable based on the steepness of the terrain. Langmuir adjusted Naismith's rule by subtracting 10 minutes/hour for descents between 5 and 12 percent. Finally, he recommended adding 10 min/hour when descending slopes greater than 12 degrees since climbing down steep slopes is very slow (Langmuir 1984).

Individual fitness in minutes	Time taken in hours estimated using Naismith's rule															
	2	3	4	5	6	7	8	9	10	12	14	16	18	20	22	24
15 (very fit)	1	1.5	2	2.75	3.5	4.5	5.5	6.75	7.75	10	12.5	14.5	17	19.5	22	24
20	1.25	2.25	3.25	4.5	5.5	6.5	7.75	8.75	10	12.5	15	17.5	20	23		
25	1.5	3	4.25	5.5	7	8.5	10	11.5	13.3	15	17.5					
30	2	3.5	5	6.75	8.5	10.5	12.5	14.5								
40	2.75	4.25	5.75	7.5	9.5	11.5						Too much to be attempted!				
50 (unfit)	3.3	4.8	6.5	8.5												

Figure 1: Tranter's fitness correction to Naismith's rule. Individual fitness is measured in minutes by how fast a hiker covers 800 meters while ascending 300 meters.

A second model of hiking was described by Waldo Tobler. Using empirical data provided by Eduard Imhof in 1950, Tobler formulated an equation for walking speed based on terrain slope. The equation yields that walking velocity, W, can be approximated by:

$$W = 6e^{-3.5|S+.05|} \qquad Where\ S = \frac{\Delta\,\text{elevation}}{\Delta\,\text{distance}} = Tan\ \theta\ when\ \theta\ is\ degrees\ slope$$

Tobler used a correction factor of 0.6 to describe velocity when not walking on a path. Despite the equation's simplicity and its omission of obvious variables, it remains commonly used in GIS travel time computations (Richards-Rissetto and Landau 2014). Remarkably, the walking speed along flat terrain matches Naismith's original rule—5 km/hour. The function shows a maximum hiking speed at approximately 6 km/ hour. This speed occurs at a slight downhill slope, at approximately 3 degrees. Figure 2 further illustrates the comparison of walking speed as a measure of slope, based on Tobler's Hiking function (Tobler 1993).

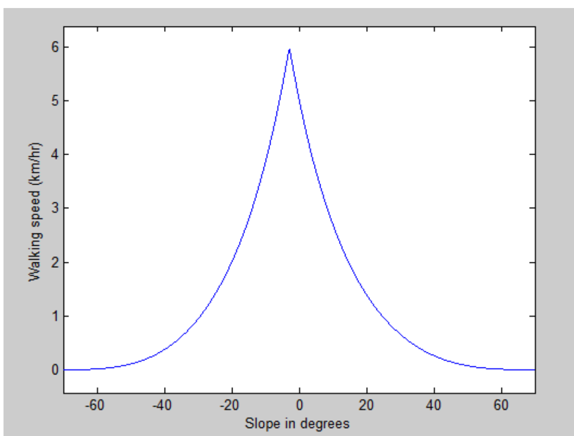

Figure 2: Tobler's hiking function

Naismith's rule and Tobler's hiking function are two popular methods of modeling the movement of humans across variable terrain on foot. The models are similar in conceptualization: (1) Both representations consider human movement over space (distance), time, and slope; (2) each model considers downhill movement slightly different than uphill movement. Tobler's function is not symmetric around zero degrees slope, but rather 3 degrees downhill. Naismith's rule has been modified by Langmuir to include a downhill correction. (3) The models recognize that movement is different based on terrain type. Each categorizes path vs non-path movement and adjusts accordingly (recognizing the Aitken adjustment for Naismith's rule). One difference in the models is the additional correction table created by Tranter, which allows Naismith's rule to account for individual fitness. The variables considered by Naismith and Tobler act as a starting point for further development of human dismounted movement.

Human Energy Expenditure

Human energy expenditure can be defined as the amount of energy used over a given time. This is usually measured in kilocalories or Metabolic Equivalent of Task (MET). A kilocalorie in terms of physical activity and the energy stored in foods, is the heat energy required to raise the temperature of 1 kilogram of water by 1 degree Celsius. A MET is defined as the rate of energy created per surface area of an average human while at rest. It is essentially a rating of activity intensity. If a task has a high MET, then it requires more human energy expenditure than an activity with a lower MET. If an activity has a MET of 2, then it requires twice as much energy as resting (Wolinsky and Driskell 2008a).

Human beings require Calories for a number of different reasons. Over 99 percent of energy used by the body is due to (www.fao.org):

- Basal metabolism—Functions that are essential for life, such as cell function
- Metabolic response to food—Energy for the ingestion and digestion of food.
- Physical activity—Movement and other activities
- Immune response to fight parasites and pathogens (Muehlenbein et al. 2010)
- Growth—Energy needed to synthesize and support growing tissues
- Pregnancy—Extra energy is needed for growing the fetus
- Lactation—The energy cost of lactation

This research focuses on the energy expenditure due to physical activity.

Measurement of Energy Expenditure

There are a number of ways to measure human energy expenditure. The measurement and estimation of human energy expenditure over a period of time is known as calorimetry. Calorimetry is based on the principle that energy expended in a human can be calculated if the amount of heat transfer from the body over a given time is known. There are four main methods for determining a value for energy expended: (1) Direct calorimetry (DC) measures the actual heat loss of an individual. (2) Indirect calorimetry (IC) measures respiratory gases such as oxygen consumption or carbon dioxide production to estimate energy expenditure. (3) Yet another method of measuring EE is using physical activity monitors[3]. These devices employ a number of different modeled variables such as body acceleration and heart-rate (HR) to estimate the amount of energy expended. (4) A technique known as Doubly Labeled

[3] Sometimes the use of physical activity monitors and doubly labeled water can be considered forms of indirect calorimetry but most literature separate these techniques from indirect calorimetry.

Water[3] (DLW) can be used to determine energy expenditure over long periods of time (4–21 days)[4] (Pettee, Tudor-Locke, and Ainsworth 2008; McMurray and Ondrak 2008).

Direct calorimetry calculates the energy used by the body directly from measuring the heat given off by a human (Wolinsky and Driskell 2008b). This technique is most often conducted in a small sealed metabolic chamber, equipped with specialized sensors, to measure changes in the air temperature. Energy expended by the subject can be calculated very precisely based on the principles of heat transfer. However, the applications are somewhat restricted due to the high cost of the equipment, confinement to a chamber, and the complexity of the equipment.

Indirect calorimetry is based on measuring the amount of oxidation in the body (Wolinsky and Driskell 2008b). It is important to understand that indirect and direct calorimetry do not measure the same energy. Indirect calorimetry estimates the energy expended based on consumed oxygen and produced carbon dioxide. Research has found, that the amount of energy that the subject is using, is proportional to the differences in the amount of O_2 and CO_2 inhaled and exhaled[5] (Leonard 2012), which can be measured by a variety of bags/hoods, metabolic carts, or portable facemasks. Examples are shown in Figures 3 and 4 (www.cosmedusa.com):

---

[4] Doubly labeled water is considered the gold standard in energy expenditure determination but it will not be described in detail since it cannot be used for studies such as my research project that have a duration less than 96 hours.

[5] Occasionally, researchers include a urinary nitrogen measurement in their predictive model but most often it is excluded since it contributes less than 2% to total energy expenditure.

Figure 3: COSMED Fitmate PRO indirect calorimetry device

Figure 4: COSMED Fitmate GS indirect calorimetry device

There are generally two types of indirect calorimetry. One uses a closed circuit system and the other uses an open circuit. The closed circuit system uses an airtight cylinder filled with oxygen, which measures the amount of oxygen consumed over time to estimate expended energy. In this method, the person must breathe only through the mouthpiece connected to the oxygen supply. Closed circuit indirect calorimetry severely limits the mobility of the subject, since he must continually be connected to the oxygen supply. A subject may use as much as 100 liters of oxygen over an hour-long test (Levine 2005).

Using the open circuit system, a subject's volume of oxygen consumption (VO_2), is measured by comparing a subject's inhaled air composition to the exhaled air composition. This comparison can reveal the amount of expired O_2 and CO_2. This method is measuring room air inhalation, and therefore the subject need not be connected to an oxygen canister. There are several systems that can be used to compare the inhaled and exhaled air. These include a computerized cart, a bag, or a portable device. Advantages of this IC system are that they can be lightweight and mobile, and don't require oxygen tanks.

Doubly Labeled Water is a non-invasive method of estimating energy expenditure over 4–24 days. The procedure involves drinking water with a known concentration of naturally occurring isotopes of hydrogen and oxygen. Differences between the isotope elimination rates of oxygen and hydrogen are measured periodically throughout the experiment by testing urine, saliva, or blood. These differences allow the investigator to calculate the amount of carbon dioxide and water produced by the body after ingestion. Known amounts of carbon dioxide and water produced are finally used to estimate energy expended (Ainslie, Reilly, and Westerterp 2003).

Physical activity monitors are wearable devices that collect biological and physiological information about an individual. The most common wearable physical activity monitors used to estimate energy expenditure are accelerometers and heartrate monitors. Accelerometers are biosensors that measure the acceleration of the body along one, two, or three axes. A predictive model can be used to estimate energy expended, since acceleration is proportional to force applied (McMinn et al. 2013). Similarly, heart-rate monitors have been shown to be capable of predicting energy expenditure after adjusting for age, sex, and body mass (Keytel et al. 2005). The accelerometer is typically worn on a subject's hip or arm, and heart rate monitors are usually worn around the chest. Several examples of typical research grade accelerometers and heart-rate monitors are shown in Figures 5, 6, and 7 (Source: www.actigraphcorp.com):

Figure 5: Actigraph GT3X on waist.

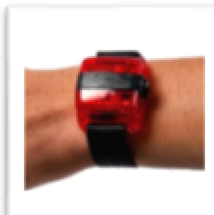

Figure 6: Actigraph GT3X on wrist.

Figure 7: Heart-rate monitor on chest.

Most attempts to estimate energy expenditure from accelerometers use regression techniques. Models of energy expenditure are developed from accelerometer count information (Crouter, Clowers, and Bassett 2006; Crouter et al. 2010; Tapia, 2008). There have been models devised using linear regression (Freedson, Melanson, and Sirard 1998; Swartz et al. 2000) as well as non-linear regression (Chen and Sun 1997; Crouter, Clowers, and Bassett 2006). The equations are developed by relating the accelerometer counts over a given time with energy expenditure information, collected with a closed circuit IC device. The most accurate equations are particular to a specific activity, such as walking, running, or mopping the floor (Crouter, Churilla, and Bassett Jr 2006).

A novel model of energy expenditure for persons walking was developed in 2006 by Crouter and his research team (Crouter, Clowers, and Bassett 2006). His previous research tested the validity of published energy expenditure models created from accelerometers. This study found that a single regression equation for estimating energy expenditure from accelerometer counts tended to overestimate the energy expended during walking and running (Crouter, Churilla, and Bassett Jr. 2006).

Crouter subsequently created a new algorithm to estimate EE. First, he determined if the activity was indeed walking/running, or some other type of activity. He accomplished this by analyzing the variation of counts every 10-second interval for one entire minute. If there was little variation in counts/epoch, then the entire minute was considered walking/running. If there was significant variation, then the minute was classified as another activity. Two regression equations were created: one for walking/running and another for all other activities. This was an improvement, because there is a significant difference between EE/count when walking/running compared to other activities. The new algorithm was more accurate than all others Crouter had tested in his previous study.

In 2010 Crouter and his colleagues further refined this energy expenditure model to better classify activities based on superior temporal analysis. The refined method examines each 10-second epoch, in comparison to all arrangements of the neighboring five 10-second epochs. This is in contrast to the 2006 model, which analyzed each 10-second epoch versus other epochs in each minute. The method essentially created a sliding temporal range for each epoch, instead of a minute-by-minute assessment. Again, they used the variation of counts/epoch to determine if the activity constitute walking/ running, or something else.

Finally, if the counts/epoch are less than 8, then the subject is considered at rest during that period, and by definition, is given a MET=1 for that time period.

This work has led to a state-of-the-art algorithm to estimate energy expenditure listed below (Crouter et al. 2010):

If the counts˙10 sec−1 are > 8

(a) CV of the counts per 10 sec are ≤ 10, then energy expenditure (METS) = 2.294275 * (exp(0.00084679 * ActiGraph counts˙10 sec−1)) (R2 = 0.739; SEE = 0.250),

(b) CV of the counts per 10 sec are > 10, then energy expenditure (METS) = 0.749395 + (0.716431 * (Ln(ActiGraph counts˙10 sec−1))) – (0.179874 * (Ln(ActiGraph counts˙10 sec−1))2) + (0.033173 * (Ln(ActiGraph counts˙10 sec−1))3) (R2 = 0.840; SEE = 0.863)

If the counts˙10 sec−1 are ≤ 8, energy expenditure = 1.0 MET

Where:
CV = (Standard Deviation of Data) / (Mean of the Data)

Heart rate monitors are also used as energy expenditure estimation devices, but studies have shown that accelerometers produce more accurate energy expenditure estimates. Research completed by Keytel and his colleagues, remarkably demonstrated a regression model that can predict energy expenditure from heart-rate that explains 73 percent of the variance (Keytel et al. 2005). This accuracy is, however, significantly lower than the above mentioned ~90 percent accuracy of the accelerometer (Crouter et al. 2010; Kuffel et al. 2011; Brandes et al. 2012). Also, using heart-rate monitors to estimate energy expenditure is somewhat challenging because there is variation within individuals due to emotion, nicotine, and digestion (Freedson and Miller 2000). One final drawback of using a heart-rate device, is that heart-rate variations are based on a delayed reaction from physical activity. Thus, it would be somewhat difficult to study changes that occur on a high resolution temporal scale.

Metabolic Cost of Walking

Mathematical models that predict metabolic cost of walking are at the center of tools created to assist military planners with route selection. Current models of metabolic cost over terrain involve knowledge of the environment as well as individual characteristics such as sex, height, weight, BMI, fitness, and load carried. Generally speaking, these equations are devised by affixing indirect calorimeters to subjects, while they perform walking tasks over different types of terrain. Then regression analysis is used to fit a model that explains the variation in metabolic cost with a set of predictors. One of the most commonly used models of energy expenditure was devised by Givoni and Goldman in 1971 and refined by Pandolf in 1977 (Potter et al. 2013).

The Givoni and Goldman equation was derived by studying 26 subjects walking on treadmills at different speeds and grades. The research unsurprisingly found that metabolic rates increase with walking speed. Similarly, they proved that energy expenditure also rose with increasing grade in a near-linear fashion. Further, the examination found that the metabolic cost of walking was linearly related to the summation of the person's body weight and the load (up to 30kg) with which they were burdened. The devised equation is listed below (Givoni and Goldman 1971). The proposed equation touted a correlation coefficient of .97.

$$MW = \eta\,(W+L) \bullet [2.3 + 0.32 \bullet (V-2.5)^{1.65} + G \bullet (0.2 + 0.07 \bullet (V - 2.5))]$$

Mw = metabolic cost of walking (in watts);
η = terrain factor (terrain for this equation was only considered as 1.0 as it accounted for treadmill surfaces only);
W = body mass (kilograms);
L = load mass (kilograms);
V = velocity or walk rate (kph); and
G = slope or grade (%)

The Pandolf equation was a refinement of the work done by Givoni and Goldman (Pandolf, Givoni, and Goldman 1977). This equation is often used as a benchmark for validating other research relating to energy expenditure (Hall et al. 2004; Duggan and Haisman 1992; Kramer and Sylvester 2011). Pandolf used six subjects, who walked for 15 minutes with backpacks. A second study was conducted with 10 subjects standing still with a loaded backpack to determine the energy expenditure of standing with different loads. The cost of different terrain factor coefficients were taken from a previous study at the same laboratory (Soule and Goldman 1972). A correction factor for downhill walking at a pace of 1.12 m/sec was added by Santee (Santee et al. 2003). Pandolf's original equation was devised with a correlation coefficient of .96 and the correction factor had an r > .90. Listed below is the finalized outcome (Pandolf, Givoni, and Goldman 1977; Santee et al. 2003):

<u>1977 Equation</u>

$MW = 1.5 \cdot W + 2.0 \cdot (W + L) \cdot (L / W)^2 + ŋ \cdot (W + L) \cdot (1.5 \cdot V2 + 0.35 \cdot V \cdot G)$

Where:

Mw = metabolic cost of walking (or standing) (in watts)
W = body mass (kilograms)
L = load mass (kilograms)
ŋ = terrain factor
V = velocity or walk rate (m/s)
G = slope or grade (%)

The terrain factor categories are: 1.0 = black top road or treadmill; 1.1 = dirt road; 1.2 = light brush; 1.5 = heavy brush; 1.8 = swampy bog; 2.1 = loose sand; 2.5 = soft snow, 15 cm depth; 3.3 = soft snow 25 cm deep; 4.1 = soft snow, 35 cm depth (12).

<u>2003 Correction</u>

Mw = PE – CF

Where PE is the 1977 Pandolf Equation and CF is the correction factor listed below.

$$CF = \eta \bullet [(G \bullet (W + L) \bullet V) / 3.5 - ((W + L) \bullet (G + 6)^2) / W) + (25V^2)]$$

Where:

η = terrain factor
G = grade (%)
W = body wt (kg)
L = load wt (kg)
V = velocity (m/s)

Spatial Analysis of Terrain

The evolution of computer storage techniques in the 1950s led to the representation of the landscape in a digital map form. The advent of the personal computer, the high speed Central Processing Unit (CPU), and the introduction of software products for digital map analysis, have allowed the general public a means of conducting computations on digital geographic data. These advancements of computer technology have permitted the use of computational methods to be applied to digital terrain data that can solve geographic problems in a more efficient and accurate way (Li, Zhu, and Gold 2010).

Digital Elevation Data

The desire to digitally represent the topographic surface has developed into a discipline known as digital terrain modeling. A digital terrain model (DTM) is easily defined as a representation of the terrain in digital form. The terrain can be viewed mathematically as a bivariate function defined over a domain in the Euclidean plane. The formation of the terrain surface with respect to this definition is to associate elevation values (Z) to specific geographic locations (X,Y) in the plane such that Z=f(X,Y). The DTM is the necessary data input for the computation of slope across a surface. Two common ways to represent and store the DTM in digital form are the Digital Elevation Model (DEM) and the Triangulated Irregular Network (TIN). The important difference in these two data structures is the distribution and storage of the modeled elevation points (Clarke 1995).

The TIN is a vector-based data structure for storing geographic elevation data. In a TIN, the locations of the geographic points used for the terrain surface representation are scattered and form no regular pattern. They are convenient places for elevation measurement. These elevation points are connected into are a network of triangles to approximate the landscape. The triangles share edges and vertices, forming a continuous surface in which the corners of the triangles match the elevation values exactly if the source data are precise. The TIN was a commonly used data structure before increases in computer processing speed, advances in computer storage space, and improvements in remote sensing data collection techniques. Today, most digital terrain models store elevation information using a DEM.

The most commonly used model of the terrain is a simplified version of the DTM, the DEM, which is a representation of the elevation at a specific geographic location in digital form. The data storage structure of a DEM is a matrix with elevation values in each cell. In a DEM, the locations used for terrain representation of the elevation points are evenly distributed in the form of rectangles. It is possible for the model to form a discontinuous (Figure 8) or continuous (Figure 9) reproduction of the surface in this type of terrain replication. The latter is the preferred representation since we generally consider the terrain to be a continuous surface. To form a continuous surface, sample points are modeled by a polynomial function to form a bilinear surface in which elevation values are determined at regular intervals, specifically, the four corners of a regular square (Li, Zhu, and Gold 2010).

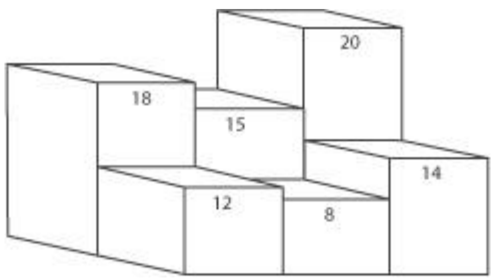

Figure 8: Discontinuous DEM with elevation values of the grid cells.

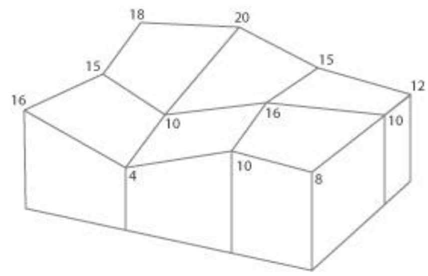

Figure 9: Continuous DEM with elevation values at the grid nodes.

There are many advantages to using the DEM, specifically, the gridded data structure and its simplicity. Gridded representation has long been used in cartography and is how most terrain attributes are stored (Robinson et al. 1995). This structure allows for computation, statistical analysis, and interpolation of elevation values to be performed more easily than terrain data stored in a vector format such as in the TIN. Many computer languages are based on data stored in the form of matrices. Additionally, common algorithms exist to perform operations on this data structure. Images used in computer display are predominately stored in raster format, which allows the terrain data in a DEM to be associated easily with these images. Finally, the most commonly used source data for creating models of the topography are collected by remote sensing. This form of collection easily acquires data directly in raster format (Clarke 2010).

However, there are several disadvantages of the DEM data structure as well. Most importantly, storage of gridded data is much less efficient than the TIN model. There is significant redundancy in the collected points, especially in areas of smooth terrain. There is a loss of feature clarity due to the adherence to storing data at only the corner points of the grid. Terrain features are overlooked if they fall beyond the edges of the grid cell. The smoothing that occurs due to the gridded-data model is a major source of error when conducting analysis of the terrain surface (Clarke 1995). Despite the disadvantages of the DEM, it remains the most common method for geographical elevation data storage.

Calculation of Slope

The computation of elevation change over a distance is referred to as slope—the steepness of the surface traveled. Slope or grade along which a navigator travels directly affects the energy expenditure, as stated in the discussion of the Pandolf equation. Slope is sometimes described as the rise over run. In this explanation, rise is defined by the change of elevation in the Z plane and run is the change in distance over the X,Y plane. It is computed using the equation:

$$Slope = \frac{\text{Rise}}{\text{Run}} \quad \text{or} \quad Slope = \frac{\Delta Z}{\sqrt{\Delta X^2 + \Delta Y^2}}$$

DEMs are readily available to perform this calculation in their raster format. The general method of calculating slope is to compute a surface of best fit through neighborhood points and measure the change of elevation per unit distance (Clarke 1995). A new output raster can be developed with each grid represented by the resulting computation. However, in the case of computing slope along a linear feature, such as a trajectory, only three cells are needed. This calculation is referred to as longitudinal slope. In this computation, we need only consider the elevation values of grid cells that are traversed by the navigator, as shown in Figure 10.

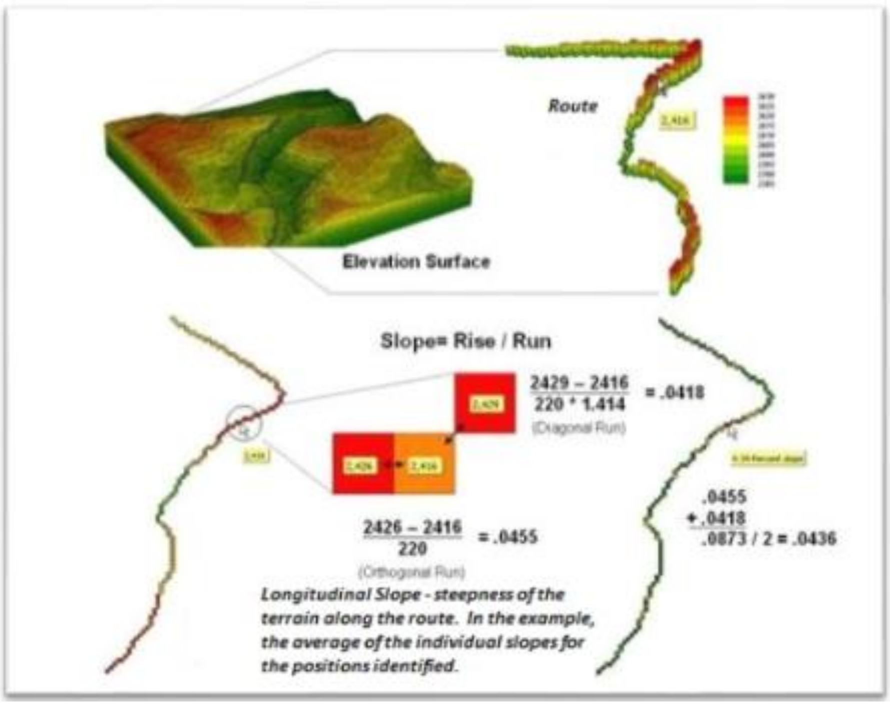

Figure 10: Longitudinal slope calculation adapted from Map Analysis by Joseph Berry (Berry 2007).

Classification of Terrain (Land Cover)

Land cover can be defined as the material at the surface of the earth. It is what humans and animals see on the ground, walk through, or pass over. Land cover can be discerned visually, and therefore can be sensed remotely through aerial photography or satellite images (Fisher 2005). It also affects our locomotion across the terrain. Land cover is often described as different classes, such as trees, water, bare ground, boulders, paths, roads, and many others. A list representing most common classes of land cover can be found in the Land Cover Classification System, produced by the Environment and Natural Resources Service of the

Food and Agriculture Organization. However, much of land cover and its classification is user-dependent based on analytical needs (Di Gregorio and Jansen 2000).

Land cover classification is an abstract representation of the local condition. It represents the process of sorting or arranging entities into groups or categories. This process can be accomplished in several ways: (1) visual inspection and collection of empirical field measurements; (2) use of remotely sensed data, which can consist of simple visual inspection of aerial photography or other images to classify the terrain, or by automated computational image processing based on statistical pattern recognition techniques applied to multispectral data (Jensen 2005).

A Review of Spatio-Temporal Analytics

The data used for the modeling energy and speed of navigation is spatio-temporal in nature. It has both a temporal and a geographic component. The combined analysis of space and time is a rapidly developing research effort in geography and GIScience. Location-based technologies, increased geo-temporal data availability, and improved computational opportunity has allowed the communities to store, integrate, analyze, and visualize this type of data in fantastic ways. Space-time research was mentioned as a major cross-cutting theme by the Steering Committee on New Research Directions for the National Geospatial-Intelligence Agency (National Research Council (U.S.) et al. 2010). Another telling indicator that spatio-temporal analytics is at the forefront of scientific research is the emphasis it received at recent conferences and dedicated special issues in reputable journals. A space-time research panel at the GIScience 2012 International Conference and another at the Association of American Geographers (AAG) in 2014 highlighted the state of space-time research. The *Annals of the AAG*, and the *International Journal of Geographical Information Science*, have recently published special editions dedicated to space-time research.

The roots of studying geographical systems in space and time have been documented throughout the past century of geographical thought. As early as 1941 Sauer was emphasizing time in geographical analysis. He encouraged historical and cultural geographers, to build a "retrospective science" of cultural and historical processes in geography, which enables us to acquire an ability to look forward (Sauer 1941). However, it was the conceptual framework described by Torsten Hagerstrand in his presidential address to the Regional Science Association in 1970 that has directed present theory with respect to human activity and accessibility in time and space (Hagerstrand 1970).

This space and time theory in geography originated with Hagerstrand's models of diffusion and time geography. During his address to the Regional Science Association, Hagerstrand conceptualized his integration of time and space to analyze socio-environmental mechanisms by describing potential human paths defined by constraints. He introduced the concept of space-time prisms to visualize potential paths (Hägerstraand 1970). Allan Pred further defined uses and conceptual structures to solve geographical problems in time and space. He defended Hagerstrand's time geography, and provided a framework for applying time geography to analyze human geographical problem sets, mostly focused on planning, modeling, and accessibility analysis (Pred 1977). Much of his theory is used today when analyzing movement and trajectory data (Andrienko and Andrienko 2013).

The integration of computation into geographical analysis, has offered increased quantitative and analytical opportunities to the study of time in geographical systems. The work of Tobler in the 1960s and 1970s, provided some of the earliest applied examples of integrating time, space, and computation (Tobler 1970). The introduction of digital GIS, during the late 1970s and 1980s led to increased attempts to incorporate space, time, and attributes for geographical analysis (Langran 1992; Peuquet 1994). The progress and innovation of GIS provided a deeper opportunity for analytics using time geography as the theoretical basis.

An early contribution to the field of accessibility and individual movement was published by Miller in 1991 in the *International Journal of Geographical Information Systems* (Miller 1991). Miller provides a comprehensive explanation of space-time prisms, defines a general procedure for creating space-time prisms in GIS, and uses the space-time prism as the analytical basis for researching the urban transportation network. Finally, he offers other potential applications of these constructs in a GIS. In essence, he operationalized the Hagerstrand theories for use in a GIS, and paved the way for future applied research using space-time.

In the early 1990s the U.S. National Center for Geographic Information and Analysis (NCGIA) promoted initiatives to continue research pertaining to space and time in GIS. The effort brought together experts in the fields of spatial reasoning, computer science, and geographic analysis for a series of conferences to further methods for spatio-temporal modeling and reasoning. Papers and results of these sessions are summarized in *Spatial and Temporal Reasoning in Geographic Information Systems* (Egenhofer and Golledge 1998).

Miller’s seminal work with space-time prisms, mentioned above, led to many applied research efforts in accessibility and movement using GIS during the 1990s. A significant contributor in this domain has been Mei-Po Kwan, who has published extensively on the subject of the spatial constraints and opportunities of different individuals and groups, for more than two decades. In 1998 she compared integral methods of calculating individual accessibility with a GIS-based approach using space-time prisms (Kwan 1998). In this well-cited publication, the author found that space-time measures for computing accessibility are better suited to capturing interpersonal differences in individual accessibility than previously used integral methods. The study is also significant because it defines operational procedures for using space-time prisms with computational algorithms in a GIS to solve problems involving human movement.

The analytical methods of modeling geographical systems and processes continued to be on the forefront of GIScience research objectives into the new millennium. A broader conceptualization of geographical, spatio-temporal analysis is to consider a changing world over time. Known as geographic dynamics, the scope encompasses all time dependent aspects of physical and human systems on or near the surface of the earth. Generic theories pertaining to the subject of geographic dynamics use spatio-temporal activities, events, and processes to create geographic change and movement. Development of theories, data models, visualization, analytical techniques, and modeling can be found in *Computation and Visualization for Understanding Dynamics in Geographic Domains* (Yuan and Hornsby 2007).

One of the major outcomes of the work done by the University Consortium for Geographic Information Science (UCGIS), described in the paragraph above, was a defined necessity for the GIS community to discover capabilities of visual representation that can explore multi-dimensional data simultaneously (Yuan and Hornsby 2007). Focusing specifically on spatio-temporal data, the research group emphasized creating opportunities to visualize space-time data. Working from a definition and research agenda proposed by

Thomas and Cook (2005), the community shaped a way forward for the GIS discipline to incorporate visual analytics into space-time analysis.

Visual analytics is defined as the science of analytical reasoning facilitated by interactive visual interfaces. It revolves around an analyst being able to manipulate, identify patterns, and obtain knowledge from large quantities of spatio-temporal data. It concerns itself with analytical reasoning to discern deep insights for assessment, planning, and decision making. It also involves data representation and transformation, to resolve conflicting and dynamic data, for visualization and interaction, and it enables multidimensional data representation and interaction in the visual form. Further, it provides effective information production, presentation, dissemination, and communication of analytical results (Thomas and Cook 2006).

A practical application of employing these fundamentals can be seen in recent work by Shaw and Yu studying virtual communication implications. Innovation and development of information and communication technologies (ICT) during the past decade has changed human activity and travel. Shaw and Yu examined ways to grow Hagerstrand’s time geography to include virtual activities and communications conducted via information and communications technologies (ICT) (Shaw and Yu 2009). The paper outlines a space-time GIS that is robust enough to handle complex activity and interaction data, and it provides a GIS platform and framework to empirically study the intricate and constantly evolving effects of ICTs on individual activities and relations in a hybrid physical-virtual space.

Due to the rise of location-based services (LBS), large data sets of spatio-temporal data are now commonplace. LBS are loosely defined as services that deliver data and information that is tailored to the current or projected location of an object or person. LBS incorporates GIS, wireless technologies, positioning systems, and human computer use (Brimicombe and Li 2009), which is precisely the type of data that is collected and analyzed in this investigation of navigation.

The ubiquitous use of LBS has created a need to be able to analyze large spatio-temporal databases, such as those derived in this study. One way to analyze and visualize this "big data" is to display each record individually. The analyst can then detect space-time patterns within the inherent data. However, the amount of data in a data set may overwhelm the analyst to a point where patterns are not visible (Andrienko et al. 2010).

Three alternative approaches are commonly used to better depict space-time data and are used in this research. One method involves data aggregation and summarization prior to analysis or graphical representation. This is a common method used when dealing with data representing movement of things, animals, or humans. The computational extraction of specific data types prior to analysis or visualization is another method of working with big space-time data. Another approach involves projecting the data to separate it, which creates a visualization that actually alters features in geographic locations to fill available visualization space more efficiently. These three methods are described in detail by a research collaboration led by Gennady Andrienko (Andrienko et al. 2010). Other significant work in the realm of analyzing and visualizing large datasets of trajectories is summarized by Andrienko and Andrienko described below (Andrienko and Andrienko 2013).

Visual Analytics of Trajectory Data

A way to understand the factors that affect dismounted navigators in geographic space, is to collect and analyze space-time trajectories. A trajectory is defined as a path of a moving entity through space. This research focuses on the movement of objects, rather than processes or events, but all can be represented as trajectories. The discipline of visual analytics provides many general methods to analyze data that are represented by a sequence of time-referenced locations.

The ubiquitous use of location-based services has served to jumpstart interest in trajectory analysis in both academia and industry. Data about where and when people/objects move is being collected at an amazing rate. However, visual analysis techniques and treatment of this data remains challenging. Andrienko and Andrienko provide a thorough review of methods, tools, and procedures for handling trajectory data (Andrienko and Andrienko 2013).

One of the most common forms of visualizing trajectories as they pass through time and space is the animated map. Animation of movement dates back to Tobler's migration studies (Tobler 1970). This method of visualization uses a sequence of static maps that are framed as a movie, based on a desired time interval, making it possible to view continuous change of location, attributes, or entities. A limitation of the animated map is that it does not provide the user with the ability to see the entire trajectory as a single view and it is poor at showing large quantities of trajectory paths simultaneously. Certainly, many trajectories can be displayed concurrently, but limitations of human comprehension generally prevent knowledge acquisition.

A second way to visualize trajectory data is with a simple static map, which merely plots all the locations along the trajectory. There is some utility to visualizing data in this manner, as all locations can be visited in one snapshot. Unfortunately, this method totally disregards the temporal component of the data, and similar to animation, it struggles to handle large quantities of data. All trajectories can be plotted as a visualization, but the knowledge presented to the user will likely be limited. Nonetheless, this technique cannot be said to be useless; the technique has been shown beneficial in looking for a specific location that is not visited by more than 20,000 trajectories, or in attempting to identify spatial outliers. Plotting all data will identify these areas.

A third method of viewing trajectory data makes use of time geography theory (Hägerstraand 1970). Hagerstrand's space-time path concept provides a mechanism for mapping space on the X and Y axis, and plotting time on the Z axis. This three-dimensional representation was operationalized by Miller and later, Kwan, as a visualization technique for movement data (Miller 1991). Today this method seems to have the greatest traction, as it is seen in most mainstream GIS software packages. For example, ArcGIS has a Space-Time Pattern Mining toolbox available in its newest release. There are several limitations of this technique: its functionality to deal with many trajectories is poor and occlusion of data increases with trajectory quantity. However, this technique has been shown to be important in searching for trends in the clustering of space-time data.

Another common technique for handling large amounts of trajectory data is clustering. Data can be clustered in many ways to support a research objective. Clustering is a way to group or categorize large quantities of data, based on user-defined properties. Trajectory data can be aggregated by the shape of the trajectory, relative to time, speed, movement direction, and other attributes. This reduces the amount of data displayed on the visualization. An advantage to clustering is that it reduces the density of data and allows grouping of similar data. Often this technique will illuminate patterns within the data.

Research has shown that this data-manipulation technique can be used to reduce the number of plotted trajectories to make manageable visualizations when communicating results.

A final technique that is in common practice for immense quantities of trajectory data is generalization. With a large dataset of trajectories, it is necessary to find ways to simplify calculations, identify trends in the data, and reduce the complexity of the visualizations. Generalization of presence and density are a few of the many ways to generalize trajectory data. Generalization of presence allows discovery of locations where objects or processes visit. Techniques are also available to discover the density of trajectories by location, which are further outlined by Willems (Willems 2009, 2011).

This review of available methodology has provided examination of five different techniques that can be used to analyze and visualize space-time data. Each technique has been shown to have advantages and disadvantages. Spatio-temporal visual analytic techniques are similar to map use in cartography—techniques used by the map maker are defined by the purpose. Cartographers use maps to communicate information and discover trends in data, and the same is true for the visual analytic user. Research into dismounted movement, navigation, and route planning require a mixture of these spatio-temporal visualization methods.

Global Navigation Satellite Systems (GNSS) as a Spatiotemporal Data Collection Device

Visualization of trajectory data requires an LBS to track and record an individual's dismounted movement. The past decades have seen a remarkable leap forward in technologies that enable the measurement and collection of this type of space-time data. Positioning Systems (PS) are a subset of LBS, providing location to a device or user. A common PS device used to collect trajectory data is a Global Navigation Satellite System (GNSS). These technologies use a constellation of satellites to geometrically determine a device's location. These are typically thought of as a Global Positioning System (GPS) device, but nearly any data-collection device can be equipped with a location receiver and can be employed to collect spatiotemporal data.

The GPS is one of the most important technological innovations of the 20th century. GPS was originally developed by the Department of Defense to assist in navigation and weapon targeting. This technology became fully operational in 1993 and much innovation has ensued in the realm of positioning and navigation since then. Additional GNSS have been launched by Russia, China, Japan, and Europe. Receiver pricing and development has evolved to a point where size and price allow for nearly ubiquitous use of positioning technology, and it is now changing nearly every facet of our lives.

Advances in positioning systems have allowed GPS devices to be miniaturized to the point where the receivers are smaller than a dime. These systems and technologies have driven space-time data availability into the realm of "big-data," in which a user can possess high-resolution location information about processes, persons, or events at short time intervals. The ability to collect information at such a high temporal interval allows for precise analytics and modeling of dynamic phenomena (Jiang and Yao 2006).

GNSS uses a constellation of satellites broadcasting a radio signal at known transmission times to provide users with geographical location. The basics of the satellite based location systems revolve around four components:

(1) Known location of at least four satellites.

(2) Known time the radio wave transmission leaves each satellite.

(3) Known speed of radio wave transmission from each satellite.

(4) Known time of radio wave arrival at receiver.

These variables can be combined to solve for a user's location. The location equation uses the principle that distance can be calculated based on distance=speed*time. Reviewing the listed variables shows that we know time (time signal arrived at receiver, time signal left the satellite) and the speed of the radio wave. Simultaneous measurements of three different satellites can narrow the positional estimate to two locations by computing the intersection of three derived distances. Usually, one of these locations is easily rejected as "not possible" by the receiver. A fourth satellite is used to eliminate receiver errors. The calculations yield an exact location, provided that exact values for the listed variables are known and that the radio signal travels directly to the receiver. Unfortunately, there are constraints and errors inherent in the process that limit the GNSS application, accuracy, and repeatability (Brimicombe and Li 2009).

A constraint of the system is that line of sight is required between the receiver and the satellites. Radio signals from the navigational satellites are generally not received by the receiver underground through the human body, in deep canyons, inside buildings, or underwater. Additionally, tree foliage, the outside of buildings, and other elements of the local environment near the receiver can cause the signal to bounce around before arrival. This is called multipath error (shown in the table below), and causes distance calculations to be inaccurate. These factors and other errors can significantly degrade positional accuracy and receiver performance (Brimicombe and Li 2009).

Error affecting the accuracy of positioning technology can be summarized in an error budget. An error budget is simply a table listing error sources and the typical error associated with each miscalculation. An error budget clipped from www.trimble.com is listed below. The following paragraphs will generally describe each error type.

Summary of GPS Error Sources (www.trimble.com)

Typical Error in meters:

(per satellites)	Standard GPS
Satellite Clocks	1.5m
Orbit Errors	2.5m
Ionosphere	5.0m
Troposphere	0.5m
Receiver Noise	0.3m
Multipath	0.6m

Time is an important variable in distance calculation. It can be measured as the time difference between the time the radio wave leaves the satellite and the time it is received at the receiver. Error is introduced when either clock is wrong. Despite the very accurate atomic clocks used on satellites, error can be introduced, as seen in the table above. The time measured by the receiver clock is considerably less accurate. Fortunately, trigonometry can be used by considering the fourth (or furthest) satellite in the constellation. This is known as receiver clock offset and is used to eliminate the receiver clock error.

Positional accuracy of the satellites in space is essential to the distance equation. The locations of satellites at any given time are stored in the receiver. If a satellite drifts off its orbit it causes a receiver to use an inaccurate satellite location in its distance calculation, which is known as an orbit or ephemeris error. The ionosphere and the troposphere cause delays in the radio waves as they travel from the satellite to the receiver. This delay affects the time calculation in our distance equation. Incorrect distance causes our positional accuracy to be degraded.

Receivers also have inherent error. Any object that is not at absolute zero temperature emits electromagnetic energy. The receiver's components and antenna emit considerable electromagnetic energy. These emissions interact with the incoming radio signals, and some of this is at the GNSS frequencies, which will contribute a range error to the measurement. This is called receiver noise.

Many of these errors can be reduced by a technique known as differential GPS, which uses a known stationary receiver with an exact location. This receiver uses the same satellites as the user's receiver. Comparison of the radio signal calculations used by the reference receiver and the user's receiver eliminates systematic errors. These calculations are possible because the reference location is known exactly. Below is a revised error budget that shows the significance of applying differential GPS to improve receiver accuracy. Notice the significant improvement in error minimization in all types except receiver noise and multipath error.

Summary of GPS Error Sources (www.trimble.com)

<u>Typical Error in Meters</u>

(per satellites)	Standard GPS	Differential GPS
Satellite Clocks	1.5m	0m
Orbit Errors	2.5m	0m
Ionosphere	5.0m	0.4m
Troposphere	0.5m	0.2m
Receiver Noise	0.3m	0.3m
Multipath	0.6m	0.6m

A final consideration when discussing error in GNSS is the positioning of satellites in the sky relative to the user. A wider angle between the receiver and the satellite receivers used in the geometric distance calculations will yield a higher accuracy. Generally speaking, it is optimal to have satellites spread out across the sky. If all the radio signals come from the same direction, it is difficult for the geometric calculations to provide accurate location information to the user. This factor makes it obvious that it is important to have as many satellites in the sky as possible (Brimicombe and Li 2009). Terrain with steep slopes, cliffs, and valleys will likely limit the line of sight between some satellites in the constellation and receivers, forcing location calculations to be made using less than optimal satellite locations, and narrower satellite-receiver-satellite angles.

An area characterized by vegetation consisting of oak, hickory, and maple trees has significant limitations when collecting satellite-based location information. The foliage of these tree types is significant in that the canopy of branches and leaves will affect GNSS accuracy in several ways. (1) The trees sometimes block satellite signals from the receiver, limiting receiver accuracy (as described in discussing satellite positioning and wide angle desirability). (2) Foliage causes the radio signal to bounce around off leaves, tree limbs, and the ground before reaching the receiver. These multipath signals significantly affect the distance calculation, and reduce the locational accuracy. Additionally, the tree trunks produce scattering, deflection, and diffraction. The overall effect of the constraints of GNSS technology in hilly and wooded terrain will be inaccurate locational data. Positional accuracies in the range of 4–6 meters are typical in wooded environments (Rodríguez-Pérez et al. 2007).

Summary of Previous Work

Navigation involves two components: locomotion and wayfinding. The energy expenditure of navigation is mostly due to muscular contraction during locomotion and wayfinding. During locomotion, muscular contraction moves the navigator through his/her environment. During wayfinding, the navigator uses muscular contraction to assist in the sensory processes to look around, turn around, or look behind. A model can therefore be derived by studying how individual and terrain factors change EE during cross-country dismounted navigation.

This investigation requires collection of EE estimates, by a number of possible measures: direct calorimetry, indirect calorimetry, Doubly Labeled Water (DLW), and physical activity monitors. Researchers must weight each device's advantages and disadvantages to determine the most expedient collection method.

Energy expenditure of navigation can be modeled using individual attributes of the subjects and characteristics of the terrain surface. Individual characteristics such as BMI, sex, fitness, age, and load have been shown to directly affect the metabolic cost of walking (Potter et al. 2013). The characteristics of the terrain surface that most affect energy expenditure and speed of navigation are slope and land cover (Pandolf, Givoni, and Goldman 1977). Slope can be calculated using a high-resolution digital elevation model in the gridded data structure, using a longitudinal slope calculation. It also can be computed by finding the elevation difference between two sequential GNSS points. Land-cover classification can be created by visually inspecting aerial photography, maps, or by automated computational image processing. These variables can be represented in a computationally efficient gridded data structure for further spatiotemporal analysis and modeling.

Space-time analytics of trajectories provide a method for modeling the energy expenditure and speed of navigation. Several approaches of investigating and visualizing trajectory data have been reviewed above. Techniques from time geography, accessibility, clustering, generalization, and visual analytics are commonly used for analyzing large quantities of trajectory data.

CHAPTER 3

METHODS

Introduction

This chapter will chronicle the subjects, materials, and procedures used to answer the research questions, outlining the implementation of methods employed during the research. It explains the observational study used to collect the empirical data that was required for analysis of each research question. The equipment used for testing is described, explaining the rationale for the selection of said equipment. Further, this chapter provides information pertaining to the software used for data organization, analysis, and visualization, and concludes by discussing constraints and limitations of the methods used.

This research collects energy expenditure and speed estimates of Soldiers while they navigated on-foot, over hilly, wooded terrain. The estimates provide an opportunity to model the contribution of both individual and terrain factors to energy expenditure and speed during dismounted navigation. The research is comprised of several studies that gathered information about how much humans move, whereby data on movement, location, and time were collected for varying terrain conditions that included trees, swamps, roads, trails, brush, hills, and cliffs. This was accomplished by affixing wearable devices to cadets at the USMA while they participated in a navigational test that was part of their summer military training. Collected data also provided a basis for assessing and improving current GIS routing tools.

Cases

Approximately 1000 cadets participate in Cadet Basic Training each summer in West Point, NY, at USMA. They are divided into eight subgroups called companies; each company is comprised by roughly 125 cadets. Each company participates in land navigation testing on different days in July and August, which allowed the investigator to collect data from eight different groups on different days. It allowed maximum use of the limited number of accelerometers that were purchased (a more detailed description of the accelerometers is provided in the materials section of this chapter). The investigator obtained data on individual differences, such as age, sex, height, weight, and fitness score from empirical data collection and the USMA training management office.

The observational study consisted of asking cadets from the academy to volunteer to wear accelerometers and heart-rate monitors while they were tested on their ability to navigate through the woods at West Point. The population of this research study can be defined as human beings who navigate via walking. The sampling frame was junior military officers between the ages of 17–25. This is representative of the segment of the military most directly involved with route planning for operations. The sample can be described as:

- 200 cadets at USMA who participated in land navigation testing during Cadet Basic Training (CBT) in July and August, 2015.
- Consisting of both sexes at an approximate ratio of 80.5 percent male and 19.5 percent female which is similar to the sampling frame referenced above.
- Cadets wore military field uniforms with a pistol belt and camelback hydration system.
- Cadets who volunteered had varying heights, weights, and fitness levels.
- Age of cadets is fell between 17 and 23.
- Other specifics pertaining to the subjects can be found in Chapter 4: Data.

Materials

The objective of this research was to measure the energy expended by the subjects and speed at which navigators moved. Here we describe the equipment used for data collection: the scale, stadiometer, GNSS devices, accelerometer, and the heart-rate monitor.

The scale used in the study was developed by Tanita. A high precision device that could quickly measure an individual's weight was required. Speed of measurement and ease of use were important during the data collection. It was imperative that the researcher weigh the subjects and get them back to their navigation training quickly. Figure 11 shows the scale used. The scale product name is Tanita BC-548 Ironman and it is well reviewed by the medical and physical training communities.

Figure 11: Scale used in research to obtain weight measurements.

Height measurements were taken by a Detecto stadiometer, which was attached to a standard army scale. This device had been calibrated within the three months prior to the testing. These devices are common devices used by the military to take height and weight measurements for Army physical fitness assessments. See Figure 12 for a visual representation of the device.

Figure 12: Detecto Stadiometer used to measure heights of subjects.

Two different GNSS devices were used to collect location data, the Qstarz BT-1300ST Sports Recorder (see Figure 13) and the DeLorme inReach Explorer GNSS (see Figure 15). The primary device used was a Qstarz BT-1300ST Sports Recorder. A picture of the device is shown in Figure 13. This device was placed in the subject's shoulder pocket so that it would not be lost during the navigational exercise. Specifications of the device are shown in Figure 14. The GNSS was able to record data for 4 hours, which exceeded the duration of the navigational test. These devices were used for three primary reasons: (1) availability; they were the only devices under the direct control of the researcher and were provided by USMA free of charge; (2) the devices had direct interoperability with the accelerometers used, which was essential for merging the location and movement data; (3) they were small enough to easily fit inside a shoulder pocket of the subject's shirt.

Figure 13: Qstarz BT-1300ST GNSS device.

General	
GPS Chip	MTK II GPS Module
Frequency	L1, 1575.42MHz
C/A Code	1.023MHz chip rate
Channels	66-CH Performance
Antenna (Internal)	Built-in patch antenna with LNA
Sensitivity	Tracking -165dBm
Datum	WGS84
Performance Characteristic	
Position	Without aid: 3.0m 2D-RMS
Accuracy	<3m CEP(50%) without SA (horizontal) DGPS (WAAS, ENGOS, MSAS): 2.5m
Velocity	Without aid: 0.1m/s, DGPS (WAAS, ENGOS, MSAS): 0.05m/s
Time	50 ns RMS
Cold/Warm/Hot Start	35/33/1 sec, average
Dynamic Condition	
Altitude	<18,000m
Velocity	<515m/sec
Acceleration	<4g
Protocol	
GPS Output Data	NMEA 0183 (V3.01) -GGA, GSA, GSV, RMC (Default)
	VTG, GLL(Optional)
Baud Rate	115,200 bps
Power	
Built-in rechargeable Li-ion battery	
Bluetooth	
Standard	Fully compliant with *Bluetooth* V1.2
Bluetooth Profile	Serial Port Profiles (SPP), Up to 10 meters
Others	
Size / Weight	62 (L) X 38 (W) X 7 (H) mm/ 22g
Operating Temperature	- 10 ℃ to + 60 ℃
Storage Temperature	- 20 ℃ to + 60 ℃
Charging	0 ℃ to + 45 ℃

Figure 14: Qstarz Specifications.

The second GNSS device, the DeLorme inReach Explorer GNSS, was used for redundancy, in the event that the primary GNSS receivers did not work properly. This device was issued to all navigating cadets whether or not they had volunteered for the research study. The redundancy device was provided by contractors supporting the navigational training exercise. While these devices were not controlled by the researcher, he was able to access and save the data if needed. As it turned out, there were no instances in which the back-up data was required. However, this description is included since it was part of the original study design. Specifications can be found for the DeLorme at the following web link: http://www.inreachdelorme.com/product-info/inreach-explorer-rugged-communication-kit.php.

Figure 15: DeLorme inReach Explorer.

The first three research questions are predicated on energy expenditure. There are many factors that must be addressed when choosing a method for estimating energy expenditure. The differences among direct calorimetry, indirect calorimetry, and accelerometers or heart-rate monitors are significant. Figures 16, 17, and 18 describe some advantages and disadvantages of each type of device, as summarized from the following studies (Levine 2005; Wolinsky and Driskell 2008a; McMurray and Ondrak 2008; Pettee, Tudor-Locke, and Ainsworth 2008).

Direct Calorimeter

Major Characteristics	Advantages	Disadvantages	Typical Uses
• Chamber based measurement of heat lost by humans or animals	• Very Accurate	• Not Mobile • Expensive to operate • Cannot be used for short time scale. It takes 30 min to equilibrate before and after heat settings • Difficult to maintain • Limited to exercises that can be evaluated inside a room or chamber	• Lab based studies • Measure energy expenditure of sleeping, walking, etc. But not good for sports related tests • Metabolism experiments in animals and humans

Figure 16: Characteristics, advantages, disadvantages and uses of direct calorimeters.

Indirect Calorimeter

Major Characteristics	Advantages	Disadvantages	Typical Uses
• Estimates heat production by measuring Oxygen consumption and CO2 production • Hoods, bags and facemasks • Oxidation can be measured with either an open circuit (most common) or closed circuit spirometers	• Can be a portable device to use in field studies • Less expensive than direct calorimeter and almost as accurate • Good for measuring metabolic rate over short periods of time	• Less accurate than Direct Calorimeters • Much more expensive than Accelerometers • More limiting in the amount of unconstrained physical activity that can be tested.	• Exercise physiology experiments involving energy expenditure • Sports related studies involving fitness and metabolism

Figure 17: Characteristics, advantages, disadvantages and uses of indirect calorimeters.

Accelerometers or Heart-Rate Monitors

Major Characteristics	Advantages	Disadvantages	Typical Uses
• Small inexpensive biosensors • Attach to hip, wrist and chest	• Very mobile • Technologically easy to understand • Durable • Quick to set up • Lowest cost	• Least accurate of all three types	• Physical Activity estimation of free living activities • Personal health monitoring • Research studies where funding makes Indirect - Calorimeters impossible • Research studies over long periods of time • Studies where hindering subjects with an indirect calorimeter would negatively affect research design

Figure 18: Characteristics, advantages, disadvantages and uses of accelerometers and heart rate monitors.

This study primarily focused on accelerometers to measure energy expenditure due to their mobility, unobtrusive nature, cost, and simplicity. The most accurate energy expenditure estimates come from indirect or direct calorimetry, but the high cost, intricacy, invasiveness, and human resources required, made these devices infeasible for this study. Accelerometers can easily be transported to a field-collection site for use. They are easily affixed to subjects, and they do not restrict movement. Additionally, the cost of accelerometers is significantly lower than indirect calorimetry. Research-grade accelerometers cost less than $300 per unit, which is minimal in comparison to the ca. $30,000 expense of indirect calorimetry devices (Actigraph 2014a). Finally, accelerometers are relatively simple to calibrate, set up, and use for physical activity data collection (Wolinsky and Driskell 2008).

Accelerometers have been shown to predict energy expenditure to greater than 90 percent accuracy while walking (Crouter et al. 2010; Kuffel et al. 2011; Brandes et al. 2012). They are a common method for assessing energy expenditure in field-based research (Welk et al. 2004). Technological advances in micro-computing and data storage have led to an increased variety of research and consumer-grade accelerometers. Some of the more common devices on the market are Actiheart, Actical, Actigraph, Body Media Fit, Fitbit Zip, Jaw Bone Up, and Nike Fuel Band. Most smartphones are now equipped with an accelerometer, which has given rise to a new form of energy expenditure estimation device (Pande et al. 2013). This research used research-grade Actigraph accelerometers as data-collection devices, which are commonly used as instruments in scientific studies to estimate energy expenditure (Lee at al. 2014). They also use mostly open-source algorithms to estimate energy expenditure, which is necessary for our scientific understanding. Finally, these devices are paired with a software package that allows researchers to easily download acceleration data and analyze results.

The study collected individual movement data using the Actigraph GT3X (Figure 19) (Actigraph 2014a). Thirty devices were procured with funding provided from the U.S. Army Geospatial Research Laboratory.

This sensor determines acceleration values using a triaxial Microelectro-Mechanical-System accelerometer. The biosensor is capable of detecting static (from force of gravity) and dynamic accelerations (from movement) in three directions. The accelerometer detects change in acceleration through measuring variations in the sensor’s electric charge storage. The Actigraph GT3X has 256 MB of onboard memory to record and store data in the time scale of days, which was sufficient for our study. The GT3X can be worn around the wrist or waist (John and Freedson 2012), based on experimental design preferences. In our study, the GT3X was worn on the right hip. The data from the accelerometers is used to estimate each cadet’s expended energy every 10 seconds.

Figure 19: Actigraph GT3X Accelerometer

The study used the water-resistant Polar H6 heart-rate monitor (Figure 20), which is a combined chest strap and heart-rate monitor used to measure the subject’s heart rate and send data to a compatible devise. The H6 uses wireless Bluetooth data transmission to pass the data to the GT3X. The strap is connected around the user's breastbone, below the chest muscles. . The heart-rate data was not used in the modeling effort pertaining to Research Questions 1–4, but it is mentioned here because the data was collected and remains part of the final dataset.

Figure 20: Polar H6 heart rate monitor.

Computation Materials and Software

The data was downloaded and analyzed on a DELL laptop computer. The observational study design allowed for collection of location, motion, and heart-rate data of 200 subjects while they navigated through hilly, wooded terrain. A computer with significant hardware capabilities was required to handle geospatial and statistical analytics of a dataset of this magnitude. Additionally, mobility to a field site for data capture and data download was essential. Figure 21 below describes the hardware specifications of the computation equipment used.

Dell Precision M4800 Base
39.6cm (15.6") HD(1366x768) Anti-Glare LED-backlit display
Intel Core i7-4810MQ Processor (Quad Core 2.80GHz, 3.80GHz Turbo)
32GB (4x8GB) 1600MHz DDR3L RAAM
NVIDIA Quadro K2100M w/2GB GDDR5 (490-BBLD)
6MB 47W, w/HD Graphics 4600) (338-BFIB) 1
512GB Solid State Drive Full Mini Card
2nd 2.5 inch 1TB Solid State Hybrid Drive (401-AAEV) 1
8X DVD+/-RW Drive Tray Load
Windows 7 Pro 64-bit English
180W AC Adapter (450-AATJ) 1
6-cell (65Wh) Primary Battery

Figure 21: Computer hardware specifications

The data collected by the wearable devices is complex and requires specialized software for data analysis. Four main software packages were used to download, analyze, and visualize the information collected. Software created by the Actigraph Corporation, called Actilife 6, was used with the accelerometer hardware. Q-Travel software was used to initialize and download data from the Qstarz GNSS. Spatial analysis and visualization was conducted with ArcGIS 10.3 for Desktop from the Environmental Systems Research Institute

(ESRI) Corporation. Statistical analysis and visualization was accomplished using the R Project for Statistical Computing. Other general purpose operations were accomplished using the Microsoft Office 2013 suite.

Actilife 6 is a software package used to estimate human energy expenditure. The software platform was created by the Actigraph Corporation to accompany the Actigraph GT3X accelerometers and other wearable biosensors they offer. This is the only software package that is compatible with the Actigraph GT3X sensors for initialization of the devices, data download, and subsequent data analysis. The software package requires a license and product key for use.

The Actilife 6 software has several algorithms for estimating EE. There are five physical activity algorithms and twelve MET based algorithms that can be used to estimate EE within the program. Williams and Freedson have derived equations used in the software to directly calculate EE from the GT3X counts (ActiGraph Support 2014; Freedson et al. 1998; Sasaki et al. 2011). Similarly, Freedson, Crouter, Hendelman and others created models to calculate METS from accelerometer counts which can be converted to energy expenditure (Actigraph 2015b; Freedson et al. 1998; Crouter et al. 2006; Crouter et al. 2010; Hendelman et al. 2000).

The investigator used the Crouter Refined 2-Regression Model within the Actilife 6 software to estimate METS consumed per ten second interval (Crouter et al. 2010; Actigraph Software Department 2012). The review, background, and specifics of this model were discussed in Chapter 2. The 2010 Crouter Refined 2-Regression Model has been validated in reputable research studies (Kuffel 2011). Additionally, this algorithm is built into the Actilife 6 data analysis software program that can be used to analyze accelerometer data (Actigraph

Software Department 2012). Finally, it is built specifically for the two activities that occur during navigation: walking and rest.

The Geospatial software used for data organization, spatial analysis, and visualization was ArcGIS 10.3 for Desktop. This software platform is a state-of-the-art program used by many professionals analyzing spatial data. The researcher chose this software because of his 15 years of familiarity with the program, its longstanding reputation in the research community, and its ability to handle data originating from a GNSS device (Kennedy 2009).

The mathematical modeling and statistical program chosen to organize, analyze, and visualize the energy, motion, and spatial data was R Project for Statistical Computing. R is an open-source statistical programing language with a wide range of libraries to support the user community. It was used for this project because of its wide range of capabilities, its open-source nature, vast user community, and interoperability with ArcGIS (R Development Core Team 2015).

Data Collection Procedures

The primary data-collection strategy of the research was to collect movement information (that was later turned into energy expenditure estimates) from different people as they navigated over different types of terrain. A general framework for the data collection is shown below:

Day 1 of the empirical data collection

- Ask for volunteers to participate in the research study
- Subjects sign a consent form
- Subject demographic information collected such as name, company, sex, etc.
- Measure subject's height and weight without carried load

Day 2 of the empirical data collection

- Affix GNSS, accelerometers and heart rate monitors to the subjects
- Measure subjects loaded weight (usually about 25 pounds more than previous weight)
- Subjects conduct navigation test.
    - 2–3 hours long
    - They were tested on their ability to find 7 control points in the woods
    - Cadets used a map and compass as navigational aids.
    - Each subject was carrying a Qstarz GNSS device. The GNSS device is not used by the cadets to navigate but were carried by each subject to record their spatiotemporal position. An additional GNSS was given to each cadet for redundancy but was not needed for data analysis.
    - The per-second trajectory information of each cadet was recorded by the GNSS device.
    - More detailed discussion of the event is described in the upcoming sections
- Subjects finish navigation test
- GNSS, accelerometers and heart rate monitors are removed from the subjects
- Researcher downloads GNSS, heart rate and motion data from biosensors to computer

Since the data collection was conducted as part of the Cadet Summer Training at USMA, the research was careful not to interfere with or significantly alter the training. Hence, data collection for each subject was split into two days. The first day's data collection occurred in a classroom while the cadets received map-reading instructions from a lecturer and an overhead projector. The second day of data collection occurred in the woods where cadets participated in the navigational test.

The military training area where the navigational testing took place fell within the following bounding box (Figure 22):

Upper Left MGRS Coordinate: 18TWL7573578270

Lower Right MGRS Coordinate: 18TWL7848976111

The red line in Figure 22 represents a typical path that a cadet may take during the navigational test.

Figure 22: USMA Navigation Training Area

As mentioned earlier, the investigator ran the study with eight different companies (groups), spanning two days for each group. Figure 23 shows the dates of data collection. Column 4 shows the number of volunteers for each part of the study. A difference between the number of Day 1 and Day 2 subjects illustrated in Column 4 reflects instances where a cadet initially agreed to participate but later changed his mind about using the wearable devices and withdrew from the testing. The fifth column shows the number of subjects used for data analysis. A difference between Columns 4 and 5 can represent several problems with the data: malfunctioning equipment that led to poor or no data; improper use of the devise; improper collection of data. In other cases, it represents a loss of data after the test occurred.

Date	Company	Day 1 or Day 2	Number of Volunteers	Number of subjects used for data analysis.
21 July2015	D	Day 1	30	6
22 July2015	D	Day 2	27	6
23 July2015	E	Day 1	30	29
24 July2015	E	Day 2	30	29
25 July2015	F	Day 1	27	25
26 July2015	F	Day 2	26	25
27 July2015	G	Day 1	30	30
28 July2015	G	Day 2	30	30
29 July2015	H	Day 1	30	26
30 July2015	H	Day 2	28	26
31 July2015	A	Day 1	27	26
1 Aug2015	A	Day 2	26	26
2 Aug2015	B	Day 1	30	30
3 Aug2015	B	Day 2	30	30
4 Aug2015	C	Day 1	29	28
5 Aug2015	C	Day 2	28	28
				Total: 200

Figure 23: Data-collection schedule.

On Day 1 of the study, the researcher asked the company of ~125 cadets if they would like to volunteer to be part of the research study. This was done while they were participating in classroom training. Documentation of the consent and institutional review board approval can be found in

APPENDIX A—Institutional Review Board (IRB) Documentation. Willing individuals were excused from the training for approximately 10 minutes while the researcher documented: Date, Last Name, First Name, Middle Name, Company, Platoon, Squad, measured each subject's height and weight, and asked each to sign a consent form.

Height measurements were taken using a stadiometer; weight measurements were taken using a digital scale, both of which have been described above. The subject's height was measured in socks, without boots. The subjects were weighed with the standard Army Combat Uniform cargo pants, belt, socks, and t-shirt. All objects were removed from pants pockets prior to the weight measurement. Figure 24 shows the typical gear worn during the measurements. After the measurements were taken, the subjects redressed and returned to the classroom training.

Figure 24: Subject during height measurement.

During the classroom training, the subjects planned the routes they would navigate the next day. They were given Military Grid Reference System (MGRS) coordinates of seven different locations to find in the woods located approximately 400 meters apart. These locations are hereby referred to as control points in this document. The navigation training area had a total of 36 control points which were identified by an orienteering bag as shown in Figure 25. No subjects, however, had the same groupings of the seven points. Also on the bag was an electric scoring device similar to the one displayed in Figure 26. After mapping the locations of the points on a 1:25000 scale topographic map, the subjects planned the course they would take to find the control points during their navigation test the following day.

Figure 25: Orienteering bag similar to those used to define control points.

Figure 26: Electronic scoring device

On Day 2 of the observational study, the researcher arrived at the testing site at ~5am to prepare the wearable devices. Between 5 and 6am, the researcher organized and turned on the ~30 sets of equipment. This allowed the Qstarz GNSS to acquire satellites before the data collection and it allowed for quick distribution of the equipment, which was essential so as not to alter the cadet training timeline in a manner that would disrupt other mandatory events.

At approximately 6am, the researcher issued each subject an Actigraph GT3X accelerometer, a Qstarz GNSS device, a polar heart-rate monitor and a DeLorme GNSS device (for redundancy). The subjects were instructed to attach the accelerometer to their right hip using a belt strap. The accelerometers were worn over their undershirts and under their Army Combat Uniform blouses as shown in Figure 27. The polar heart-rate monitors were worn against the skin, across the subject’s breast bones as shown in Figure 28. The Qstarz GPS were worn inside the subjects right shoulder pocket. Finally the DeLorme GNSS were affixed to the subjects’ Army Fighting Load Carrier.

Figure 27: Location of accelerometer during the study.

Figure 28: Location of heart-rate monitor during the study.

The subjects (and all cadets participating in the navigational test) were also assigned a safety partner. This safety partner was another cadet who was instructed to follow the cadet while they completed the navigational test. The safety partner was instructed not to assist the navigator in their task completion. They were simply present to ensure the subject was not alone in the woods and that they were moving safely.

After the subjects had donned the wearable devices, they were inspected by the researcher for proper fit, and were then weighed again on a digital scale with the entire load. Subjects carried water in canteens, a camelback, ammunition, and other various items in their pockets while conducting the navigational test and observational study. Figure 29 is a typical outfit worn by the subjects.

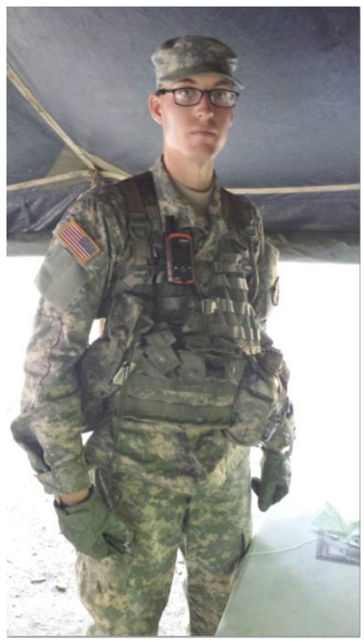

Figure 29: Subject with gear worn during navigation study.

Following inspection and weigh-in, the subjects and their safety partners were directed to the starting control point with a temporal spacing of 1–5 min between individual participating subjects. No subjects started closer than 1 min apart and there were approximately 30 minutes between the starting time of the first subject and the last subject.

All subjects began the navigation study from the starting control point. Subjects carried a card equipped with a radio-frequency identification (RFID) chip. Placing this card within 6 inches of an electronic scoring device (Figure 26) at the starting control point, recorded the time the navigator had begun his test. Subjects then began navigating toward their first control point. Upon finding the first control point, or what they thought was the control point, subjects again scanned their RFID cards on the electronic scoring device that was attached to the orienteering bag. Then the navigator proceeded to find his remaining control points and return to the start point. Navigators had 150 minutes to complete the course. Some navigators failed to find all points in the allotted time and some navigators found wrong control points. Nevertheless, all subjects eventually returned to the start point, where the cadets removed the wearable equipment and were scored using the information on the RFID card. Subsequently, the researcher downloaded the logged movement and location data for later analysis.

The observational study protocol and design was approved by an institutional review board (IRB) conducted by the Keller Army Community Hospital (KACH) at USMA. Expedited review of the KACH Protocol 15-020, IRBNet#413717-1 titled “Space, Time and Energy in Dismounted Navigation” occurred 9 July, 2015. The IRB found there to be no human subjects’ protection issues. The study was assessed as Minimal Risk. The University of California, Santa Barbara (UCSB) has entered into an Institutional Review Board (IRB) authorization agreement with the USMA. This agreement has allowed UCSB to rely on the designated IRB review from USMA and continuing oversight of its human subjects’ research. All documentation concerning the IRB request and approval can be found in Appendix A.

GIS Methods

Multiple GIS analysis methods were used during the research. They were used to investigate the spatial data for the purpose of understanding the dynamics of navigation. The data was organized by creating feature classes from GNSS data. Then methods such as intersection, selection, and buffer were used to clean the data and remove unnecessary portions for analysis. GIS was used to assist in classifying the terrain over which the subjects navigated. Elevation models were created using ArcGIS. It was also used to extract elevation values from DEMs to use in initial slope calculations. Finally, the GIS was used to visualize trajectories and communicate results. These will all be described more fully below.

Feature Class Creation

The first data manipulation method was to transform raw, downloaded GNSS data to a format understandable by the GIS. The raw GNSS data was in a text format that included latitude and longitude. The “Make XY Event Layer” tool in the data management toolbox of ArcGIS was used to create point features based on coordinates from the source table.

Also, GIS analytics were used to remove unwanted data from the dataset, which was a significant step. Primarily, intersection, selection, and buffer techniques were used to organize and clean the data. A detailed description of the steps taken to prepare the data for analysis is provided in Chapter 4: Data.

Land Cover Classification

This research requires an understanding of the terrain over which the subjects are navigating. This is imperative to answering each of the research questions. Classification was used to describe the land cover of the terrain. In essence, the entire land area was categorized into generalized groups. Orienteering maps, foundation data, and inspection of 1-meter satellite imagery from 2013 were used to generalize the study test site into eight classes: on road, open woods, rocky terrain, light vegetation, moderate vegetation, heavy vegetation, swamp, and water. The orienteering maps were obtained from the USMA orienteering team; the value of these maps for classification purposes is in that they have been ground validated by members of the orienteering team. The feature datasets used included roads, wetlands, lakes, and forests. This data was provided by the USMA Department of Public Works and Land Management. The next chapter provides details pertaining to the data used during classification.

Trajectory Analysis

One of the most valuable methods for analyzing the movement of the subjects while they navigated was the visualization of the trajectories. Several methods were used to understand the navigational behavior of the subjects. Trajectory visualization was used to get a basic understanding of how people navigated over different terrain types and for a general comparative understanding of different navigational tendencies. Most importantly, visualization of the navigator's trajectories allowed for identifying flawed data and anomalies in navigational performance. It allowed for identification of data that was later removed from the dataset before the spatio-temporal analysis.

The research initially used Google Earth (GE) to get a basic understanding of the data quality being produced by the wearable sensors. Google Earth has a simple interface to read and display GNSS data. During the data download process, the researcher validated data presence and quality. Figure 30 illustrates the ease with which a researcher can validate data existence and value. Immediately after downloading the data from the wearable sensors, it was easily loaded into GE to validate its presence and quality. By selecting the “tools” menu item, a user can choose “GPS” and then “import from file.” This loads the data into the rudimentary GIS. Quick visual inspection validated that data was present, that it was in the proper geographic location, and that time stamps were accurately depicting when the subject actually navigated. This visualization shows Bull Pond and Lake Georgina, two prominent water bodies around the study site. The red line depicts the navigator’s trajectory. This visualization validates the presence of the data and that it is in generally the appropriate location. Careful examination of the time slider at the top of the window shows that the data was collected between ~5 and ~9 am on 22 July, 2015. This review validates that the data being downloaded from the device was collected during the time of the subject’s navigational test.

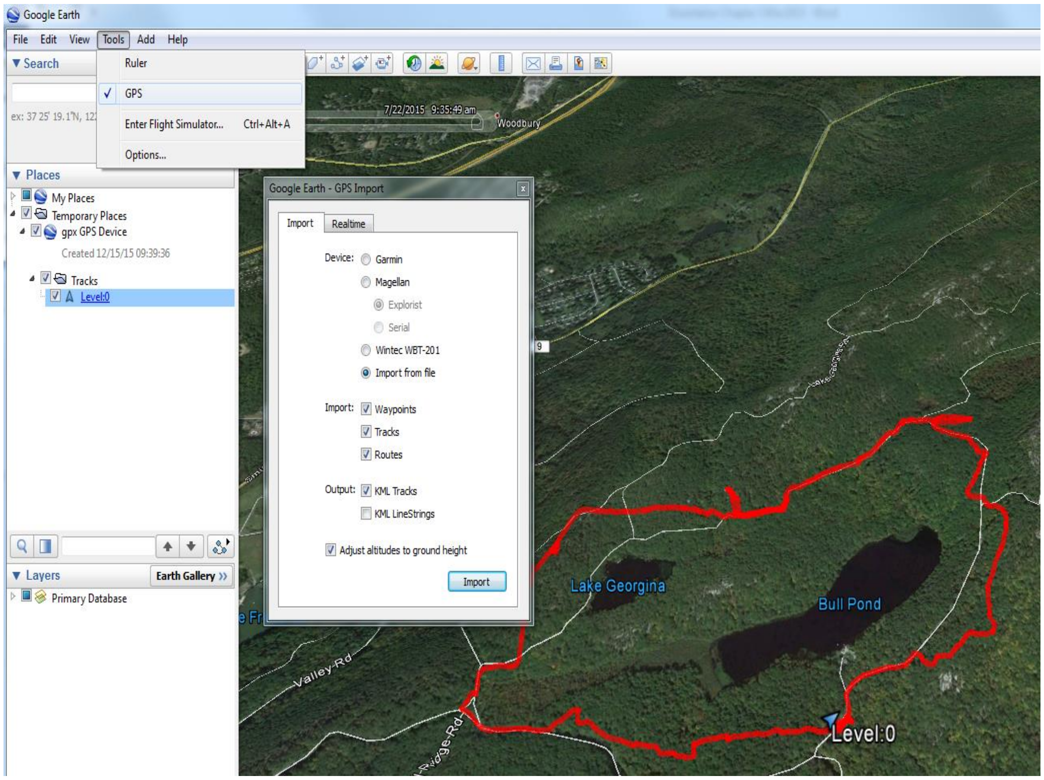

Figure 30: Google Earth data visualization and validation.

A second manner in which visualization was used to indirectly answer the research questions was to identify data to remove prior to analysis. As previously explained, the Qstarz GNSS devices were turned on prior to the time that navigation started and subjects did various things before beginning the navigational test (e.g., linger near the start point, make trips to the bathroom, return to their sleeping area, etc.) Removal of this data was essential to modeling actual navigation. Using ArcGIS Desktop, visualization of trajectories allowed for removal of this data and isolated only navigational movement. First, all the points before 6:30am were selected. Then trajectories from the start point to the bathroom and back were identified. Figure 31 below shows tracks from a single subject. Those points are locations of the subject every 10 seconds. In the figure, the cyan represents locations of the subject before 6:30am. Locations in red are locations visited by the subject after 6:30am. The magenta dot is the starting point. The figure shows some lingering around the start point. These are the cyan dots around the start point. Then the subject visits the bathroom. These are shown by the cyan dots in a line to the northeast. After using the bathroom, the subject returns to the start point. Finally, the subject begins navigating to the southwest. This is represented by the series of cyan dots from the start point toward the bottom-left of the figure. This visualization method allows removal of the irrelevant movement before analysis. The red dots around the start point and leading to the bathroom show that the subjects also lingered and used the restroom after they finished navigating. These data were also removed.

Figure 31: Selection of pre-navigation bathroom trips

GIS visualization was also used to compare navigational behavior between two or more individuals. Often during the data analysis it was necessary to compare the trajectories of several subjects. This was accomplished by standardizing the start time of all navigators to a common reference frame. This allowed comparison of all navigators in space and time, regardless of which day the data was collected.

The initial step was to standardize the start time of each navigator. First, an arbitrary date/time in the future, such as March 29, 2016 at 6:38am, was chosen. This date was subtracted from the start date/time for an individual and the difference was added to each subsequent track dot of that navigator. The process was done for all subjects. This made it appear that each navigator started at March 29, 2016 at 6:38am and allowed for visual comparison of any two individuals in time and space. Figure 32 shows an example of this technique. In this image, one subject's movement is depicted in blue and another in red. The navigator in red was part of the study conducted on 24 July. The navigator in blue was from 3 August. The technique mentioned here allowed for visualizations of synchronized starting times. The code for the Python script used to execute this task can be found in APPENDIX B – Computer Programming Titled: Python Script to synchronize subject's navigation time.

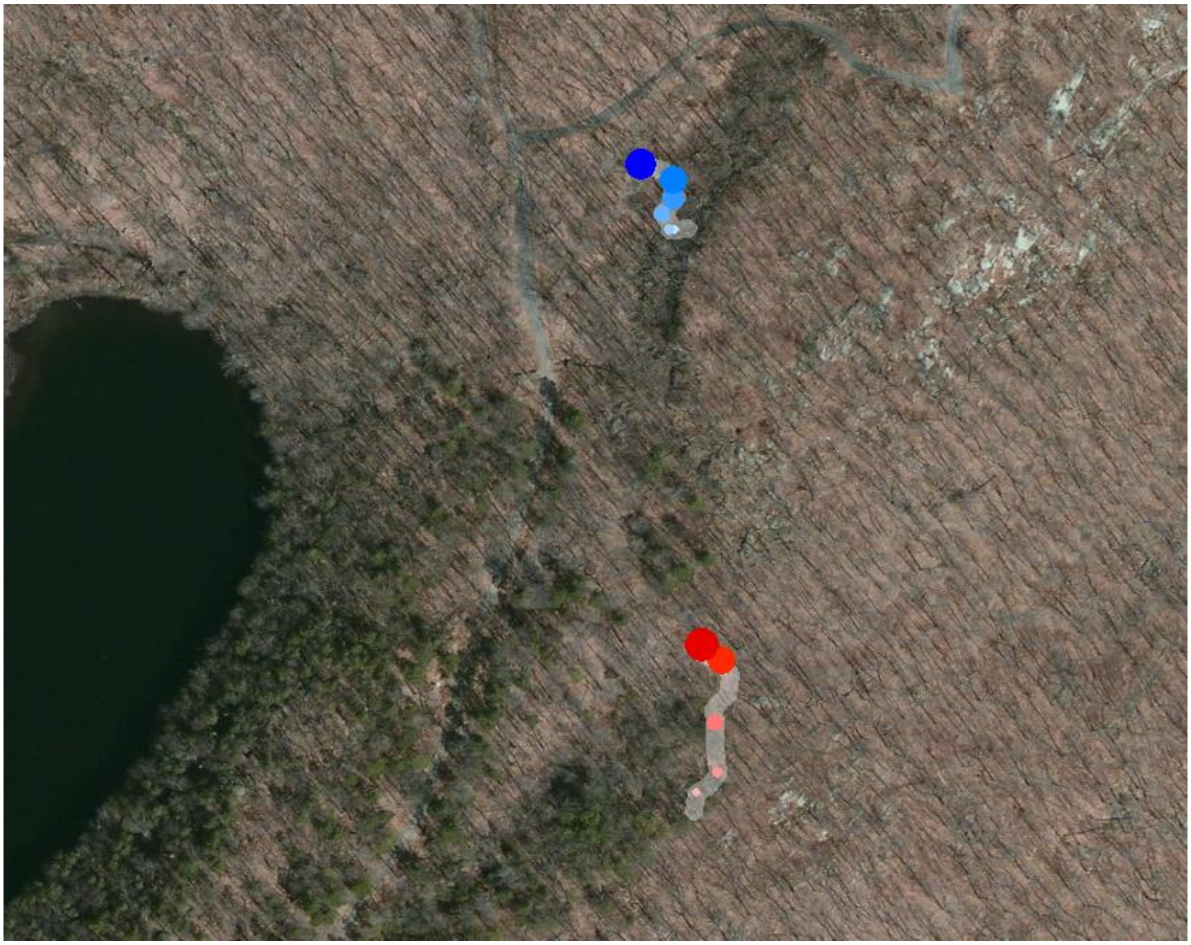

Figure 32: Example of synchronized subject start time

After this preprocessing step, visualizations were created to investigate and compare two or more navigators to inspect navigation behavior (e.g., how a heavy navigator moved vs a lighter navigator). This technique was also used to visualize how fast navigators tended to travel as opposed to slower navigators. Many comparisons of navigators were visualized in this manner and it was an excellent method to get a general understanding of trends, outliers, and basic navigational behavior.

This research used three basic methods of visual analytics to break down the routes taken by navigators: (1) The trajectories were reviewed and studied using static map analysis techniques; (2) forms of generalization were applied to the data and graphics to simplify understanding of the enormous dataset; and (3) animation was applied to visualize movement and spatiotemporal trends of the dataset.

Static maps are visual representations of geographic areas, showing the spatial relationship between entities (Wade and Sommer 2006). This technique was used to analyze the trajectories recurrently throughout the data exploration and data analysis. It is difficult to describe in detail every instance where this technique was applied due to its voluminous uses. The technique was also used to understand the areas visited by individual navigators. It was used to show areas where navigators loitered. It was used to investigate the land cover features in the navigational area and the slopes over which subjects traveled. The technique is further used extensively throughout the communication of the research and results. This can be seen throughout the document but especially in Chapters 4 and 5.

Generalization is the reduction or simplification of the data or visualization. It can involve the process of reducing the number of points or trajectories shown (Wade and Sommer 2006). A significant generalization process that was applied to the spatiotemporal dataset involved down sampling. The GNSS data was originally collected every second. This data was generalized to a temporal scale of 10 seconds for ease of computation as well as the fact that the energy expenditure estimates were made every 10 seconds.

Another application of generalization was the land cover classification. Here the complexity of the land cover was reduced. It allowed for easy visualization of the trajectories and the land cover. Quick identification of whether a subject was on the road or in the woods could be seen. Similar to static maps, generalization occurred frequently throughout the investigation.

The most revealing visual analytic process applied was animation, which is the technique of displaying successive tracks to create the illusion of the navigator's movement. Animation is especially effective in displaying change over time. It follows that animation would be a logical analytical method for navigational movement since it is based on the change of location over time.

Animation of movement trajectories was used for many applications during the processes of data exploration and data analysis. A salient example is using animation to find similarities between navigators. Animation was used in conjunction with processing described above that led to Figure 32. Secondly, it was used to assist in finding anomalies in data as described earlier during the discussion of data cleaning. Using animation to show trips to the bathroom before the navigational event was highlighted during data exploration. Animation for visual inspections was used regularly during our analysis of the data. These examples are two applications that highlight the usefulness of the technique.

Statistical and Modeling Methods

A critical method used to represent EE (Research Questions 1 and 2) and speed (Research Question 4) as a relationship between predictor variables was mathematical modeling. In conjunction with this effort, statistical methods were used to organize, analyze, and understand the data and models developed. The significant methods are described below:

Table of Statistical and Modeling Methods Used

All accomplished using R (R Development Core Team 2015).

METHOD (Research Question Applied)	DESCRIPTION OF THE METHOD AND USE

General Statistics (All)	Initial investigation of energy expenditure, speed, and the variables of slope, land cover, cumulative distance, BMI, sex, and fitness score was accomplished through the computation of basic statistics. Analysis of mean, median, max, min, variance, and standard deviation of each of the variables occurred. Additionally, sum totals of energy expenditure over an entire route were examined. Finally, average individual speed, slope traveled and energy expenditure over the total navigational trip were inspected.
Histogram (All)	All variables were analyzed using histograms, which are to organize and display the frequency of data values. The histogram is a vertical bar chart that shows the underlying data of the distribution. Many types of statistical analysis and modeling efforts assume a Gaussian distribution. Observation of a dataset's histogram allows for validation of that assumption. For example, the distribution of EE was non-Gaussian. This understanding prompted transformation of that variable for further analysis and modeling.
Scatterplot – analysis (All)	A scatterplot is used to show the association between two quantitative variables. Each of the predictor variables was graphed against EE and speed, primarily to validate the linearity of the modeling effort. It also helped to discover potential outliers in the data. Finally, it assisted in teasing out unexpected patterns. One difficulty with the scatter plots was the vast number of points. In many cases, identifying trends in the data was difficult because of noise in the data, since the dataset contains over 150,000 points. In some cases the data was generalized before plotting. This was an effective method of minimizing the noise and improving the ability to discern trends. This methodology is discussed further in the section on kernel density.
Box-Cox Transformation (RQ 1,2,3)	As briefly described above, a normal or Gaussian data distribution is an assumed element of some statistical modeling. The Box-Cox transformation (Box and Cox 1964) is a method used to predict how to transform data that is not normally distributed into a dataset that is Gaussian. This method was used to predict an appropriate transformation of the Energy Expenditure dataset. The method was applied in the R statistical package using the MASS package (Venables and Ripley 2002).
Mathematical modeling (All)	Mathematical modeling is a process of expressing the real world in mathematical terms. It attempts to use mathematics to find solutions to the problem and describe reality. A mathematical model can be considered a generalization or a simplification of a complex, real-world situation into a mathematical form. The mathematical problem can then be solved using a variety of statistical and mathematical techniques. This solution is then interpreted and translated back into real terms. Mathematical modeling was used in Research Questions 1, 2, and 3, to

	understand how human navigators expend energy as they move about different types of terrain. Similarly, for Research Question 4, how quickly humans navigate across the environment is investigated.
Multiple Regression (RQ 1,2,3)	The first modeling step was to develop an understanding of whether the predictor variables would effectively model energy expenditure in a linear model. Multiple regression was used to get a general understanding of the relationships between EE and the predictor variables. Regression analysis is one of the most widely used techniques for analyzing multifactor data. Given the dataset of energy expenditure and predictor variables, coefficients were estimated that minimized the sum of squared errors. This was done by using the lm function in the basic stats package in R (R Development Core Team 2015). The output of this function was used to understand the significance of each predictor variable. Finally, it provided a basic understanding of the applicability of the model.
Linear mixed effects modeling (RQ 1,2,3)	Linear mixed effects modeling was used because the data is grouped by individual. The data is therefore correlated by group based in how each individual moves. Mixed effects modeling provides a flexible procedure to model the phenomenon despite this clear violation of the independence assumption associated with multiple regression. The data represents 200 subjects, each with approximately 1000 data points (more on the dataset in the next chapter). This likely leads to varying intercepts and varying regression coefficients based on the individual. The employment of this method gives an intercept and a series of coefficients for a linear model. It also describes the variance associated with the random effects created by the grouped data. Specifically, the lme function in the nlme package is used to model the data (Pinheiro et al. 2015).
Residual analysis (RQ 1,2,3)	Analysis of the regression and linear mixed effects modeling efforts is followed by analysis of the residuals. This analysis is required to check that the assumptions of the modeling efforts have not been violated. The residuals are defined as the difference between the actual values and the fitted values. The analysis of residuals also provides means in which to locate possible outliers for inspection. Initial residual plots created were of the scatterplot (description above) variety. The residuals were plotted against each predictor. Any trends and patterns found identify possible model violation. The residuals were also plotted against the fitted values to test for heteroscedasticity. A final residual analysis method was used to verify that the distribution of the residuals is Gaussian. This step was accomplished by plotting the histogram of the residuals.

QQplot (RQ 1,2,3)	Another method used to validate the assumption of normality is the Quantile-Quantile Plot (QQPLOT). The QQplot computes the expected value for each data point based on the distribution. If the data follows the desired distribution then the points will generally fall on a straight line. If the points on a QQplot do not fall on the straight line then the assumed distribution must be questioned. For this research, the studentized residuals from the linear model are plotted against the theoretical quantiles. This will assess the distribution and validate our assumption of Gaussian distribution of residuals. To accomplish this, the qqPlot function in the car package of R (Fox and Weisberg 2011) was used.
Coefficient of Determination (RQ 1,2,3)	The Coefficient of Determination, commonly called "R-squared," is a measure of the variance captured by a particular model. It gives an assessment of how well actual data fits a model. The range of values for R-squared is from 0–1. A value of 1 represents a model that accounts for 100% of the data variance. The value is computed by first dividing the sum of squares of the residuals by the total sum of squares. Then subtracting this value from 1. This definition shows that the smaller the residual values, the less would be subtracted from one. This would leave a relatively high R-squared. This method was used to assess the model fit for multiple regression and for the mixed effects model. In the mixed effects model, R-squared is classified to be either from the fixed effects of the model or from the random factors.
Variance Inflation Factor (VIF) Analysis (RQ1,2,3)	Variance inflation factor analysis is used to test the modeling assumption that predictive variables must not be correlated. Analysis of VIF is used to test the independent variables when using multiple regression and with linear mixed effects modeling. The VIF tests how much inflation of the R-squared is present based on multicollinearity. A variance inflation factor of 1 means there is no inflation of the model's fit based on correlation of independent variables. Any VIF greater than 10 should be investigated further (Montgomery, Peck, and Vining 2012). This analysis was computed using the car package of the R statistical software (Fox and Weisberg 2011).
Significance Testing (RQ 1,2,3)	The significance of the model variables must be tested in order to have confidence in their usefulness. In this case, the slope of the trend line is tested. The hypothesis states that there is no trend between independent variables and the dependent variable. This is tested by assuming a hypothesis that the slopes of the model coefficients are zero. Then that is proven untrue. To prove the null hypothesis untrue, the T test and associated P value were chosen. This analysis was conducted using the R statistical analysis software program (R Development Core Team 2015).
Data Smoothing (RQ 4)	GPS data is inherently noisy because of accuracy problems already discussed. This noise comes from the random error of the

	measurements. For Research Question 4, a data-smoothing method was used to compensate for noisy data. It also simplified the large dataset. The method of smoothing based on bin means was used to accomplish the generalization of data. This straightforward method divides the dataset range into equally spaced segments called bins. All values of the dataset that fall within the range of each bin are used to compute the mean of that bin. The resulting mean is plotted to represent all data that falls within the bin range.
Curve Fitting (RQ4)	Curve fitting is a means by which an analytical expression is chosen to model data. In fact, linear modeling and mixed effects modeling mentioned above are forms of curve fitting. A generic form of this method is mentioned here. Curve fitting was used to arrive at coefficient values for the model in Research Question 4. This was the basis for the second order polynomial equation found to model navigation speed. In this case, the lm function in the basic stats package was used to approximate the polynomial coefficients (R Development Core Team 2015). Then manual tuning of the polynomial was used to settle on a fitted model which best approximated the theoretical underpinnings of speed vs slope.

Constraints and Limitations of the Methods

As discussed, the data collection had some significant constraints. Foremost, the researcher had limited control of the subjects since the training manager at USMA required that the testing not disrupt the normal training and routine. Similarly, the available measurement equipment for the observational study had limitations, which in turn had bearing on the results. Further, foundation data and orienteering maps collected from the GIS department at USMA had drawbacks. All three conditions had a significant impact on the findings of the research.

Constraints to Research Design

The major constraints for this research were the restrictions to total autonomy of research design. USMA cadet training could not be altered. While the summer training management office at USMA allowed for gathering of information from a large group of subjects, there were restrictions, which hampered flexibility of the study design. All elements of the navigational test were dictated by the training management office. Collection of data was allowed; however, altering of the sequencing, procedures, rules, and goals of the navigation exercise was prohibited. For example, affixing an activity monitor to a subject was allowed, but fitting them with an indirect calorimeter was considered too much of a mobility impediment.

Scheduling of the cadets was also dictated by the training manager, requiring that the researcher collect data on multiple subjects simultaneously. It was, therefore, impossible to observe the subjects while they navigated. In most cases, there were ~25 subjects navigating at any given time, which made it impossible to tell the exact activity of a navigator at a given time. It was impossible to tell if they were resting, looking at a map, walking, running, or talking with a friend. This forced several assumptions to be made during data analysis (discussed in the next chapter). Lack of observation made it difficult to assess the subject motivation; it is possible, although unlikely, that some subjects just didn't care about their navigational performance[6].

Another constraint was that subjects were not navigating alone. A second cadet was assigned by the training manager to travel with each subject, acting as a safety monitor. It is possible that the subject's behavior was altered by this companion. Perhaps the safety monitor hindered the movement speed or behavior of the subject, or assisted in the

---

[6] Since the subjects volunteered for the study, it is likely that they were motivated to score high on the navigational test but it is a constraint because motivation was not measured or observed.

navigation, despite having been instructed not to. Fundamentally, it is impossible to rule out interference during the data collection.

A major stipulation of the study was the inability to alter the load of the navigators. All subjects were directed to carry specific items, e.g., a camelback, canteen, ammunition, and other items. But the researcher was not able to request that additional items be carried. There was some variance in the load carried by the navigators since they were allowed to take items that were of their own preference.

The last significant constraint involved the study location. Due to many competing demands during summer training at West Point, the geographic terrain was mandated by the training manager. This gave the researcher no flexibility in designing the courses over which the subjects would navigate. The area chosen by the training manager was a relatively easy navigational experience. There were a number of roads that could be used when navigating. They were clearly marked on the map. There were also two significant water features in the middle of the training area, which could be used to easily reestablish a subject’s position while navigating.

The training manager also dictated that all eight companies of navigators use the same area. Since ~125 cadets were navigating during every iteration, paths were worn into the terrain, changing the terrain over the course of the navigational exercise. Paths were likely to be worn through the woods because of this usage. Terrain that began as vegetated might be classified as open woods by the last iteration. This would not be reflected in the data.

Constraints of the Measurement Devices

The measurement devices also came with constraints and limitations. The GNSS devices had limited accuracy due to systematic errors from satellite clocks, orbit errors, the Ionosphere, the Troposphere. Additionally, noise from the receiver and multipath error from the tree canopy produced multipath errors. Also, since the research was limited to lower-priced measurement devices, there was no ability for the devices to use the advantages of differential correction. A discussion of the literature reviewing GNSS error and positional accuracy can be found in the latter part of Chapter 2. It is likely that the GNSS produced locational errors near the magnitude of 4–6 meters (Rodríguez-Pérez et al. 2007). This amount of error affected our ability to determine which land-cover class the subjects were actually navigating through with absolute confidence. Figures 33, 34, and 35 visually describe some of the difficulties. Figure 33 shows a subset of the locations of the navigator as recorded by the GNSS (depicted by green dots). The dots represent the navigator’s position every 10 seconds. The green lines are simply drawn between the dots to depict navigator movement. Notice that there is a road running up and down (north and south) in the center of the figure. The dots generally follow the same direction as the road. It is likely that the error of the GNSS is shifting all positions of the navigator to the left (west) of the road. Most likely the navigator is traveling on the road but we can’t be sure.

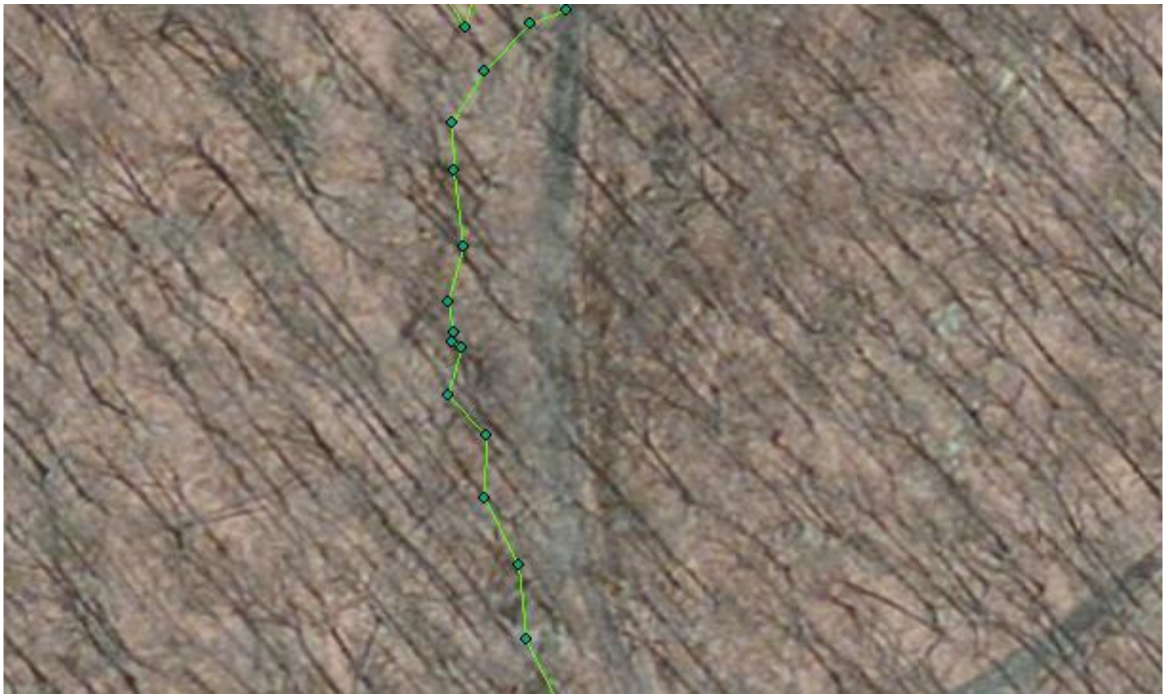

Figure 33: Subject tracks during navigation—unsure if they are walking in the woods or on the road.

Figure 34 shows classification of the land cover. The light blue corridor represents the road class; the light red represents the open woods class. Assumptions were made to account for GNSS errors near the roads. All land within 9 meters of the road centerline was classified as road. This assumption was made since navigators traveling near a road are most likely traveling on the road. Despite this assumption, Figure 34 shows an example of classification difficulty; here it is impossible to know if the navigator was actually on the road or not. The GNSS error could be systematically shifting points or the subject could be walking in the woods. To the contrary, Figure 35 shows a clear example of the same navigator walking down the middle of the road just 20 minutes after the examples shown in Figure 33 and Figure 34.

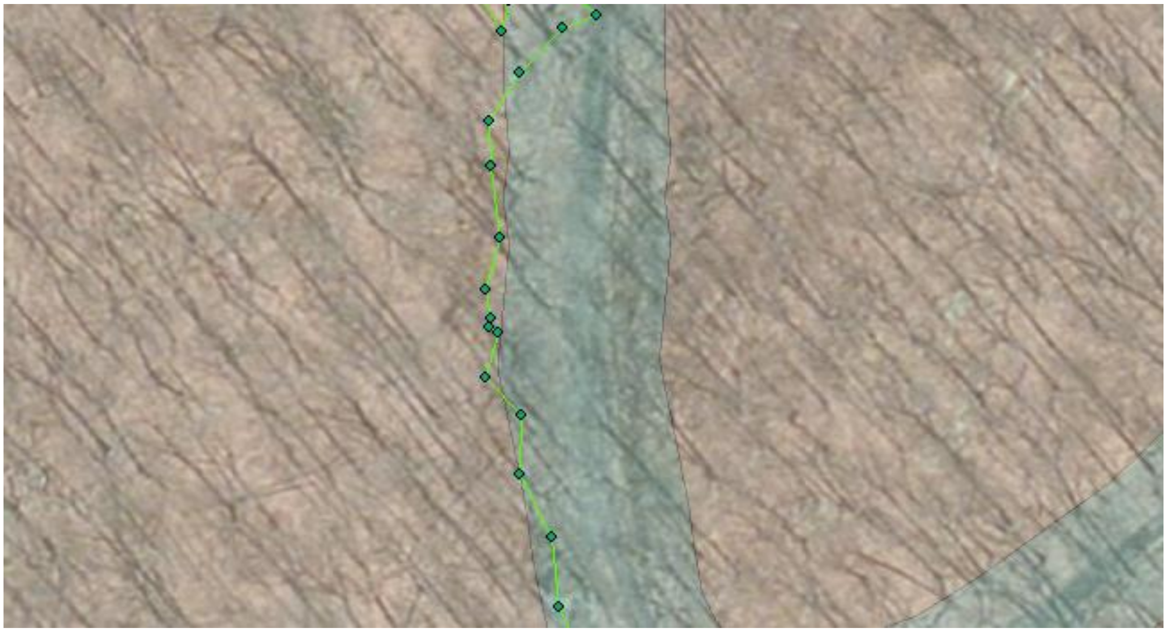

Figure 34: Subject tracks during navigation with land cover class overlay - unsure if they are walking in the woods or on the road

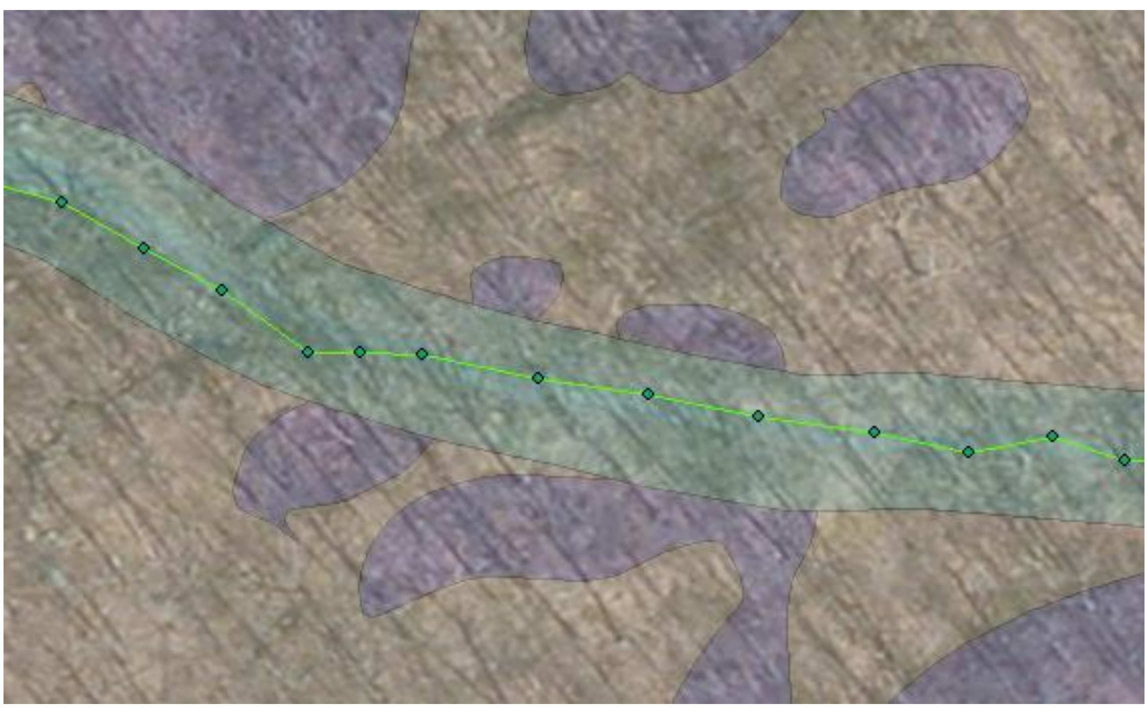

Figure 35: Subject tracks during navigation with land cover class overlay—most likely navigating on the road.

Another limitation of the study was the inability to gather an accurate DEM of the training area. Initially, there was an expectation to be able to extract elevation values from a DEM produced from aerial LIDAR. This would combat the inaccuracies produced by GNSS, especially in the elevation values. Unfortunately, the DEM that was received from USMA was filled with inaccuracies in elevation data. Often the tops of trees were part of the last return DEM, forcing the research to use GNSS elevation values to calculate slope values.

Figures 36 and 37 show the inaccuracies of the DEM. Figure 36 is a 1-meter resolution satellite imagery of a location in the navigation training area. The red dot on the left is shown in the middle of a tree canopy; the red dot on the right is 8 meters away. It is shown in an area where there was no tree canopy. Figure 37 shows the Digital Elevation Model of the same geographic area as that shown in Figure 36. In the DEM, light pixels represent higher elevations. The elevation from the DEM for the left dot is 291 meters above sea level. The elevation from the DEM for the dot on right is 272 meters. This 19-meter elevation difference is due to poor DEM creation. Due to these inaccuracies, GNSS elevations were used to calculate the slope between points.

Figure 36: Satellite Imagery of a tree and nearby point.

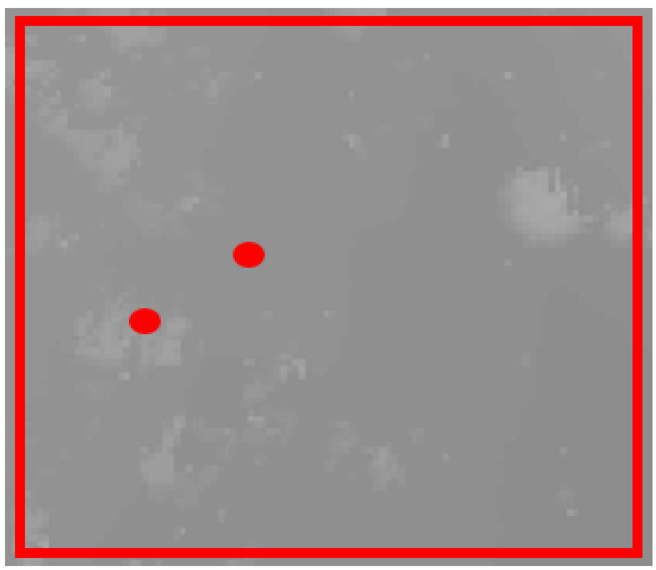

Figure 37: DEM of same location as the previous figure.

Yet another constraint of the research pertaining to the terrain was the inability to ground-truth the land-cover classification. Ideally, validation of the terrain classifications would be done by onsite inspection. Unfortunately, lack of time and manpower resources denied the ability to visually validate the land cover maps that were created from satellite imagery and orienteering maps. Some limited ground-truthing was generally executed during the study dates, but most of the classification of the terrain occurred after the researcher visited USMA.

A final constraint pertains to the understanding of our energy expenditure estimates. The Actigraph accelerometer is a measurement device that is commonly used in research studies. Researchers use this device because most of the behind-the-scenes algorithmic estimation of caloric expenditure is transparent and documented. Similarly, other movement-based variables—such as steps and sedentary analysis—also originate from published research. However, all of these processes rely on a common variable produced by the Actigraph GT3X and are computed by the Actilife 6 software. The raw data from the accelerometer is converted to "Actigraph counts" by the Actilife software. Unfortunately, the underlying process for determining Actigraph counts from the raw data is proprietary and undisclosed to the research community (Actigraph Software Department 2012; "What Is the Difference among the Energy Expenditure Algorithms? : ActiGraph Support" 2014; John and Freedson 2012). This constraint diminishes our ground level understanding of the human movement dynamics of navigation.

Summary of Methods Used

The methods described were implemented to solve the research questions stated in Chapter 1. An observational study involving 200 cadets from the USMA was conducted in West Point, NY. The study consisted of affixing wearable biosensors and GNSS devices to the subjects while they navigated through hilly, wooded terrain. The subjects were required to find 7 control points along a 3–5 km course.

Data from the study was downloaded and analyzed using Actigraph software and Qstarz software specific to the wearable devices. ArcGIS and the R statistics program were used for data analytics and visualization. Energy expenditure estimates were developed from the Crouter Refined 2-Regression Model within the Actilife 6 software. This algorithm is used to estimate METS per 10-second epoch (Crouter et al. 2010; Actigraph Software Department 2012). ArcGIS and the R statistics program were used to prepare, clean, and analyze the spatiotemporal dataset.

This chapter concludes by discussing particular constraints of the observational study design, the measuring devices, and the data. Constraints emplaced by the training management office at USMA limited the flexibility with which subjects could be controlled. The navigational test was directed and could not be altered. Higher quality and more precise measuring devices are available, but limited funding inhibited purchase.

Despite the constraints, data collection to answer the research questions was successful. Limitations of the research design were present but overcome. The Chapter 4 will describe in detail the data collected. The results and the conclusions are described in the final two chapters, demonstrating how these methods were adequate to successfully answer the posed research questions.

CHAPTER 4

DATA

The data collection, refinement, organization, and preparation were a key component of this project, which was designed to understand the energy consumption and speed at which navigators move from an origin to a destination. To model these variables required significantly large datasets. Chapter 3, Methods, described the steps taken to collect empirical data to answer the research questions. This chapter describes the data collected to support understanding navigation and the modeling process.

This chapter will describe the data and interpret the information compiled during the summer of 2015 at USMA. The defined settings of the wearable devices used will be discussed as well as the output of the data download of these devices. The computational methods used to organize and refine the large amounts of noisy data that were collected during the study are chronicled, specifying decisions made about data omission, curtailment, and outlier removal. Finally, it summarizes each dataset used to answer the four research questions. The intent of this chapter is to make the data reproducible and give future researchers a roadmap to understanding the applied data.

Data Collection Devise Settings

The data collection of the accelerometer and movement information was accomplished using an Actigraph GT3X device. Two days prior to use, this device was initialized and charged prior to use by the Actilife 6 software program. After the initialization, the device was constantly on. There was no on/off button. The device was programmed to start collecting data 2 minutes after initialization. This meant that there was always data on the device that was not associated with the observational study. This data was removed in subsequent steps as will be described later. The battery life and the storage capacity of the device allowed for approximately 25 days of data collection. Turning the device on early had no negative impact on the data collection. The settings for the GT3X are shown in Figure 38.

SETTING	VALUE
Start Time	Default
Stop Time	None
Device Time	Use Atomic Server Time
Sample Rate	30 Hz
LED Options	None
Wireless Options	Enable Wireless, Heart Rate
Record Option	Idle Sleep Mode Enabled
Subjects were instructed to wear on left hip and inspected before navigating	

Figure 38: Actigraph GT3X settings.

The Qstarz GNSS device collected the location data for the study. This device was also initialized two days prior to the observation of subjects. However, since this device had an on/off button, the Qstarz was turned off until the morning of the study. This device could only record about 4 hours of data on the settings used; the battery could last approximately 10 hours. Approximately 1 hour prior to collecting location data, the GNSS was turned on and left in an open outdoor area so that it could establish a satellite fix. The settings used for the Qstarz are shown in Figure 39.

SETTING	VALUE
GPS Log Setting	Hiking
Log Criteria	Every Second
1Hz mode	Default
Data Log Memory Mode	Overwrite
Date Mode	UTC Date Time
Fixed Mode	Valid
Navigation	Latitude, Longitude, Height, Speed, Heading
DOP	PDOP,HDOP,VDOP

Satellite Info	NSAT, SID, Elevation, Azimuth, SNR
Record Reason	RCR
Other	Distance

Figure 39: Qstarz BT-Q1300ST settings.

Data from the GT3X was downloaded from the device to the computer using the Actilife 6 software. The file was saved in a folder with an ID number and the subject's name (01 Lastname). The download options are listed in Figure 40.

SETTING	**VALUE**
Naming Convention	Custom by subject name
Clinical Report	No clinical report created
Add biometric and user information-	Yes
Create AGD file:	Yes
Epoch	1 second
# of Axis	3
Steps	Yes
Lux	Yes
Inclinometer	Yes
Low Frequency Ext	Yes
Heart Rate	Yes

Figure 40: Actigraph GT3X download settings.

The location data was downloaded using the Q Travel software from the Qstarz. This data was downloaded using the Raw Data Manager. The fields in Figure 41 were captured during the download:

VALUE
INDEX
RCR (Record method)
UTC DATE
UTC TIME
LOCAL DATE
LOCAL TIME
LATITUDE
LONGITUDE
HEIGHT
SPEED
HEADING
PDOP
HDOP
VDOP

NSAT(USED/VIEW)
SAT INFO (SID-ELE-AZI-SNR)
DISTANCE
*Descriptions found in APPENDIX C.

Figure 41: Qstarz downloaded fields.

Estimation of METS

The Actilife 6 software was used to estimate the Metabolic Equivalent of Tasks (METS). Estimated METS for each subject was based on a 10-second interval. Estimates were calculated using the Crouter Refined 2-Regression Model. These estimates were computed using the scoring tab in the Actilife 6 software. The following options were chosen in the software:

SETTING	VALUE
Energy Expenditure	No
METS	Yes, Crouter Adult (2010)
Cut Points and MVPA	No
Bouts	Yes
Sedentary Analysis	Yes
HREE	No
Exclude Non-Wear Times	No
Use Subject Log Diaries	No
Time filter	Yes: 05:00am -09:59 of study date

Figure 42: Actilife METS estimate settings.

After choosing to calculate the METS, the software prompts for reintegration of the data since it was collected at 1-second intervals. Ten-second interval data is needed for the computation. After agreeing to reintegrate, the METS data was exported by clicking the details button and choosing export all epochs. This process provided a comma separated values (CSV) file named "name1sec10sec AGD Details Epochs……." This file contained the fields listed in Figure 43.

FIELDS
DATE
EPOCH

AXIS1
AXIS2
AXIS3
VM
STEPS
HR
LUX
INCLINOMETER OFF
INCLINOMETER STANDING
INCLINOMETER SITTING
INCLINOMETER LYING
MET RATE
*Descriptions found in APPENDIX C.

Figure 43: Actilife 6 scoring output fields

Data Integration Steps

The first data integration step was combining the METS estimates with the location data. This phase was accomplished using an R script to join the tables. The script code can be found in APPENDIX B – Computer Programming Titled: R Script to join GNSS data and MET data. The output of this operation yields a CSV table that includes the fields listed in Figure 44.

FIELDS
DATETIME
DATE
EPOCH
AXIS1
AXIS2
AXIS3
VM
STEPS
HR
LUX
INCLINOMETER.OFF
INCLINOMETER.STANDING
INCLINOMETER.SITTING
INCLINOMETER.LYING
MET.RATE

INDEX
RCR
UTC.DATE
UTC.TIME
LOCAL.DATE
LOCAL.TIME
MS
VALID
LATITUDE
LONGITUDE
HEIGHT
SPEED
HEADING
PDOP
HDOP
VDOP
NSAT.USED.VIEW.
SAT.INFO..SID.ELE.AZI.SNR.
DISTANCE
*Descriptions found in APPENDIX C.

Figure 44: GNSS and Scoring data table fields.

Obviously some of these fields are not used in the data analysis but the information was maintained in the files because there was no real impact to the computational time or storage requirements associated with having additional data. This also allows future researchers to investigate related research questions that may need data fields not used in this analysis.

Step 2 for data integration was adding subject data to the data table. Information such as the subject's platoon, squad, Qstarz GNSS number, comments about the observation, sex, age, height, weight, loaded weight, and load were in a spreadsheet that was created during the data collection phase. An R script was used to accomplish this data organization task. The script can be found in APPENDIX B – Computer Programming Titled: R Script to Add

Administrative Data to GNSS/METs file. The output of these operations yields a CSV table, lastnameDataCombined, that has added the fields in Figure 45 to those listed previously.

FIELDS
MONITOR
ID
PLATOON
SQUAD
GPS_NUM
SEX
AGE
HEIGHT
WEIGHT
LOADEDWEIGHT
LOAD
COMMENTS
*Descriptions found in APPENDIX C.

Figure 45: Administrative data fields.

Step 3 of data integration was to convert the GNSS/METS data into a format that can be used for analysis and visualization in ArcGIS. The file created in the previous step was used with a Python Script to create a feature class from the GNSS/METS file, lastnameDataCombined. The script creates a feature class of points that represent the location of the subject (see APPENDIX B – Computer Programming Titled: Python Script to create a feature class of points from GNSS/METs data file)

Figure 46 shows a graphic of the created feature class. Each white point represents the location of the subject as he participated in the navigational event. There is a 10-second interval between each point. The attributes of each point are the same as those listed as data fields from Steps 1 and 2. The quantitative data associated with each row in the resulting feature class represents a measure of things that occur between the time of that point and the time associated with the next point. For example, Axis 1 counts from Row 1 in the feature class actually represents the number of Axis 1 counts between the time listed in Row 1 and the Time listed in Row 2. The exception to this statement is speed and distance. Those attributes represent the instantaneous speed and distance recorded from the GNSS device. Later the average speed and the distance between points were added.

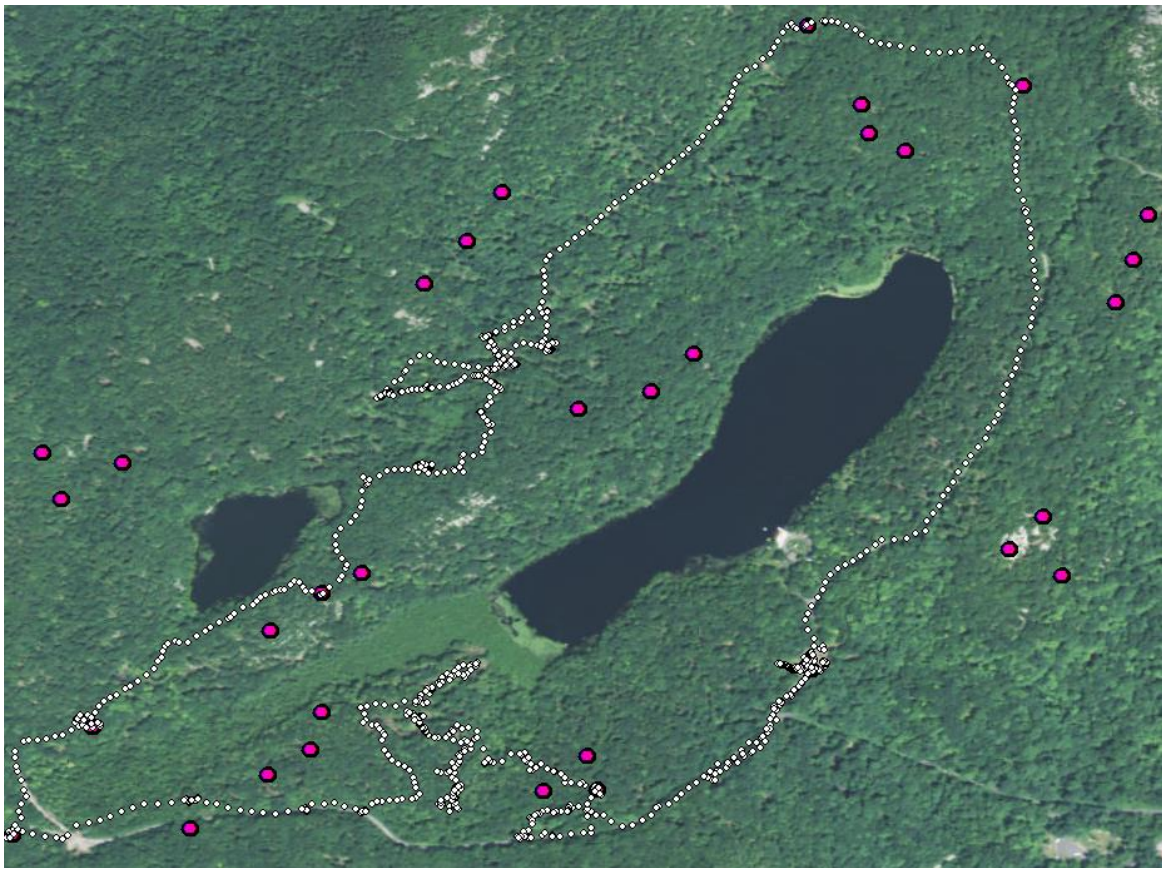

Figure 46: Point feature class created from GNSS/METS data.

Step 4 involved a spatial data preparation task accomplished with ArcGIS. The batch define projection tool was used to assign a datum (WGS84) to the spatial data. WGS 1984 is a standard datum used by GNSS devices and the U.S. Military. Coordinates were Geographic, i.e., raw latitude and longitude.

In Step 5, the duration, distance, speed and slope between points are computed and these values are added to the attribute table. These are computed using a Python script. The script can be found in APPENDIX B – Computer Programming Titled: Python Script to add duration, speed, distance and slope between points. This adds the fields listed in Figure 47 to the attribute table.

VALUE	**DESCRIPTION**
RASTERVALU	Elevation extracted from DEM at point
ZDIFF	Difference in elevation between points
DISTANCE_M_DATETIME	Distance between point and the next point (forward distance) in meters
DURATION_SEC_DATETIME	Duration between points in seconds
SPEED_MPS_DATETIME	Average Speed between points (forward looking) in m/s
COURSE_DEG_DATETIME	Direction of movement in degrees
SLOPE	Slope between points (forward looking) in percent

Figure 47: Fields added by converting points to tracks.

Step 6 involved turning the points into line segments. This is done since the data and analysis involves the behavior and biometrics of the subject between the points, not *at* the point. For example, the estimation of Calories expended was calculated for the 10-second interval between points, not as represented at a single point. The new line created in this step appropriately represents the data. Each line is a row in the data table. That data now coincide with the interval represented as a line. Calorie consumption that was representing the amount expended between two points is now the Calorie consumption along the line segment. The same can be said for the slope. This step was computed using another Python Script, and the script can be found in APPENDIX B – Computer Programming Titled: Python Script to convert points to lines. The same fields are associated with line segments as with the point file created in Step 5 above. Figure 48 shows the change of geometry.

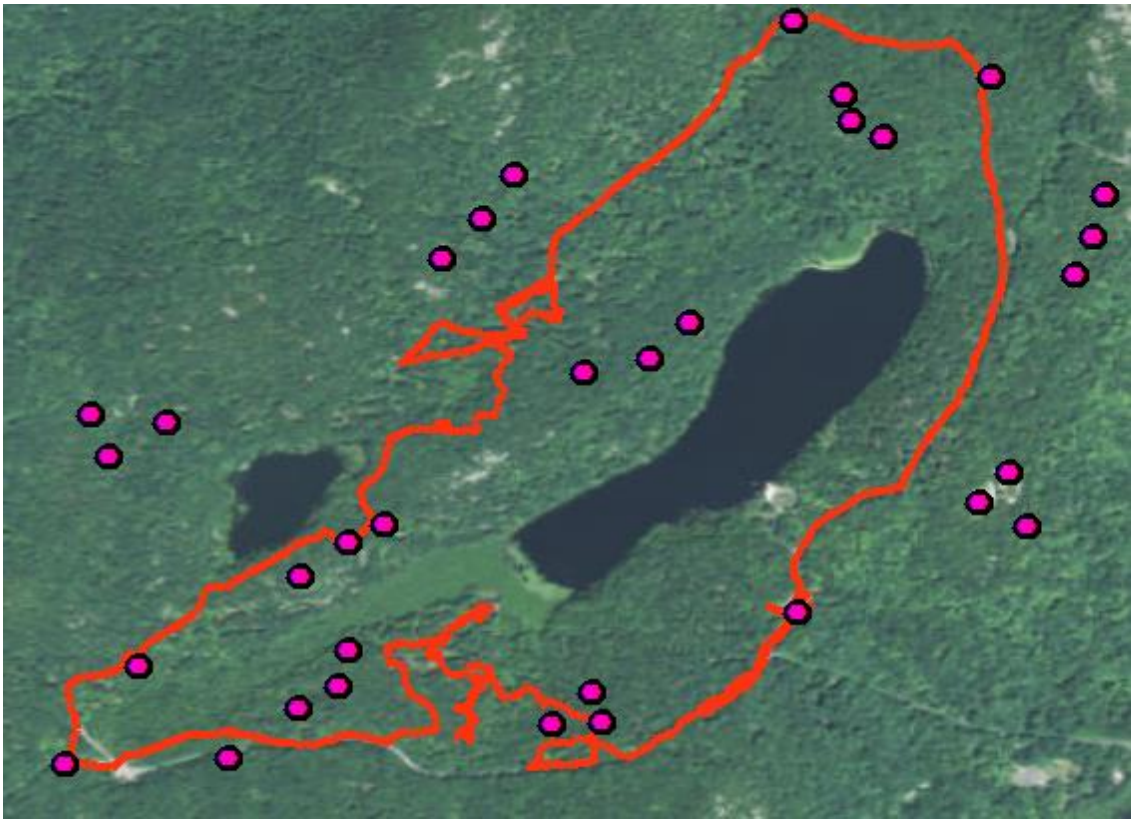

Figure 48: Line segments of subject's trajectory after conversion from points.

Step 7 entailed classifying the area over which the subjects moved. The land was classified into eight different land cover types as shown in Figure 49 and in the classification map, Figure 55. An orienteering map was used as the primary reference for this task. A second source of 1-meter satellite imagery was used to validate the classification effort. Finally, foundation data from the USMA Department of Public Works was used to verify road centerlines, water feature outlines, and vegetation type.

Class	**Description**
Roads	Paved and well maintained dirt roads.
Boulders or Rocky	Terrain that is predominately rocks. This terrain has very little soil or vegetation but the rocks provide an uneven surface to walk over. See Figure 50.
Water	Ponds or lakes
Light Vegetation	Forested areas with very light undergrowth such as small trees, shrubs, and branches. Mobility is hindered in these areas compared to open woods but not as much as in moderate vegetation. See Figure 51.
Moderate	Forested areas with some undergrowth such as small trees, shrubs, and branches. Mobility is

Vegetation	hindered in these areas compared to light vegetation but not as much as in heavy vegetation. See Figure 52.
Dense Vegetation	Forested areas with heavy undergrowth of trees, shrubs, and branches. Mobility is hindered in these areas compared to other classes of vegetation. See Figure 53.
Swamp	Areas of the forest which have wet soil either year-round or intermittently. Generally these areas also have undergrowth such as shrubs and small trees.
Open Woods	Terrain consisting of large trees with foliage but no undergrowth. The tree spacing in this area is at least 10 meters apart in most places. It allows subjects to walk along uneven ground without being hindered by shrubs and small trees.

Figure 49: Definition of land cover classes.

Figure 50: Depiction of rocky terrain.

Figure 51: Depiction of light vegetation.

Figure 52: Depiction of moderate vegetation.

Figure 53: Depiction of dense vegetation.

The orienteering map was obtained from the USMA Orienteering Team. The map was created in 1996, using satellite imagery and field surveys. The final map was validated by Mikell Platt and Bob Forbes. It was created at a 1:15000 scale with 5-meter contours, and is only available in hard copy. The map was scanned and georeferenced using a first order transformation of 12 control points. The control points were derived from a 1-meter ESRI world imagery basemap. The root mean square error of the transformation was 1.9m.

An assumption was made at this point regarding how to derive the roads on the classification map. A buffer of 9 meters from the road centerline created a polygon for the road class. This procedure was used to capture some of the GNSS error that occurred when subjects were walking on the road. It is assumed that subjects walking very near (<5meters) the road were most likely walking on the road. An expected GNSS error of 6–8 meters is expected in foliage as described in Chapters 2 and 3. Using a 9-meter buffer would help ensure that all subjects walking on the road were classified as such despite the GNSS error.

Figure 54 shows a subset of the orienteering map that was used. Figure 55 shows the derived classification map.

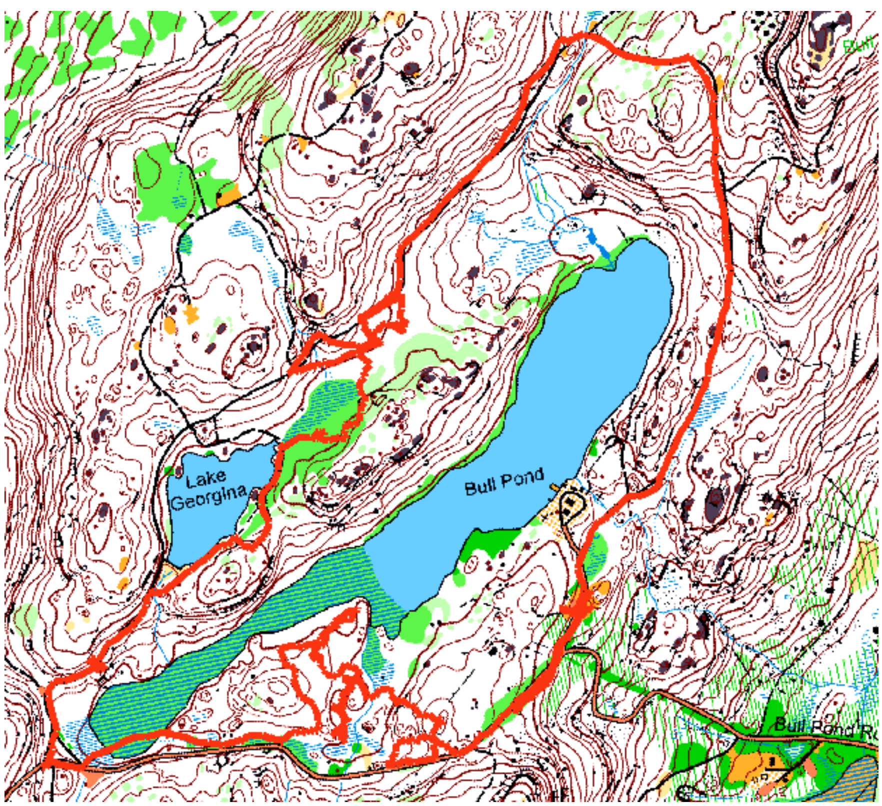

Figure 54: Orienteering map of navigation testing area.

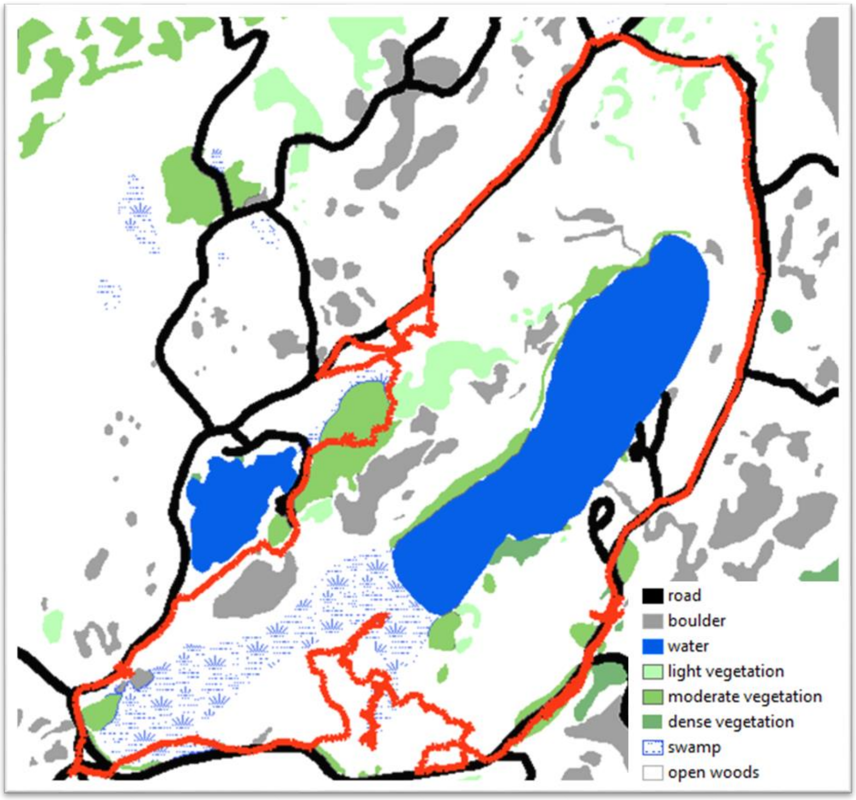

Figure 55: Classification map of navigation testing area.

Step 8 was to associate each line segment created in Step 6 with a land-cover class from the classification map. If the line segment crossed over one or more class boundaries, both class types were stored in the attribute field. The classes were stored in the feature class as a numerical ID. Figure 56 shows the IDs and the associated land cover types.

1	Roads
2	Boulders or Rocky
3	Water
4	Light Vegetation
5	Moderate Vegetation
6	Dense Vegetation
7	Swamp
8	Open Woods

Figure 56: Land cover codes.

This step was computed using a Python Script and a spatial join (see APPENDIX B – Computer Programming Titled: Python Script to add Class ID to trajectories). This operation added a field to the feature class that represents the class of the terrain over which the subject traveled at any given time.

Step 9 was to remove data in each of the files that was not related to the navigational task (e.g., lingering around the start point, making trips to the bathroom, returning to their sleeping areas) before beginning the navigational test. Removal of this data was essential to modeling actual navigation and was accomplished in several parts: (1) The erase tool in ArcGIS was used to remove all data within a 50-meter buffer of the start point. This removed all the unrelated movement before and after the navigational event. A few subjects navigated back through the start point during their test. This data was left in the dataset since it probably represented a part of navigational behavior. (2) Visualization techniques were used to look for trips to the bathroom and the bivouac area before and after the navigational exercise. This data was also cleaned from the dataset. (3) The feature classes were renamed to lastname_tracks for brevity.

In Step 10, average speed recorded from the GNSS was added to each line feature. This data was used for comparison to the GIS computed speed and for possible future research. The original GNSS data was captured at 1-second intervals, but it was down sampled to a 10-second temporal resolution. The down sampling was required for use with the Crouter Refined 2-Regression Model. The average GNSS reported speed for each 10-second temporal resolution was computed using an R script (found in APPENDIX B – Computer Programming Titled: R Script to Compute Average GNSS Speed). Then this information was joined to the line feature class using a Python script (found in APPENDIX B – Computer Programming Titled: Python Script to join AveGPS data.

In Step 11, run score and fitness scores for each subject were added as attribute fields by first adding new fields called “fitness_scr” and “run_scr.” Values were then added to the fields using the ArcGIS field calculator.

Fitness scores and run scores were taken from performance on the standard army fitness test. The subjects had all been tested by the training management office of USMA less than one month prior to the study. These scores were obtained from for use in this research.

The Army Physical Fitness Test (APFT) is a standard measure designed by the U.S. Army to measure a Soldier’s fitness. The APFT measures the upper and lower body muscular endurance of the subjects and indicates a Soldier’s ability to handle his or her body weight. The test scoring and standards are adjusted for age and physiological differences between genders. It consists of push-ups, sit-ups and a 2-mile run. Each event is scored on a scale between 0–100, with a maximum score of 300 points. This research used overall fitness score and the run score to model energy expenditure.

In Step 12 values were added to each line feature that represented the cumulative time and cumulative distance. Each record represented the time and distance from the time that the subject left the 50-meter buffer around the start point. The values were added to each subject’s trajectory feature class. New fields, called cumDist and cumTime, were added via the batch “add field” tool. Then the values were computed using the batch-field calculator. The field calculation was computed using the Python parser inside the field calculation tool. The script can be found in APPENDIX B – Computer Programming Titled: Python Script used in field calculation of cumulative distance and cumulative time.

These data-integration steps resulted in 200 line feature classes. Each feature class represented the trajectories that the subjects followed during the navigation test. Each line segment represented summarized navigational behavior, movement information, energy expenditure, individual characteristics of the subject, and terrain characteristics along the line. Most line segments represent the spatial movement covered over a 10-second period. Four segments represent a temporal resolution different than 10 seconds due to GNSS sampling errors. Two records had no duration, one record had a 20-second duration and one segment had a 70-second duration.

Finally, all the feature classes were merged into one final file using the Merge tool in ArcGIS. This was completed to have one final data matrix to use to analyze the navigation movement and biometric data. The attribute table was cleaned by deleting some unwanted fields and renaming attributes for clarity. The data matrix was exported as a text file from ArcGIS and then converted to a CSV file by Microsoft Excel. This data file contained attributes as described in fields 1–54 of APPENDIX C – Description of Data Matrix Fields. The data matrix had the following properties:

Number of Attributes:	54
Number of Records:	168645
Number of Subjects:	200
Approximate number of records per subject:	~840 *number of records per subject varied based on how quickly they finished the navigation test.
Each record represents a 10 second segment of navigation	

Figure 57: Properties of Data Matrix.

We have reviewed the steps and procedures used to convert the raw movement data, the administrative data, and GNSS trajectories into a format usable for modeling. The next section of this chapter will describe the cleaned data and measures taken to model energy expenditure and speed of navigation. The computation of additional variables needed for modeling will be discussed; the subject data will be described, as will the terrain data, energy expenditure, and speed data. Finally, it concludes by summarizing the final data matrix used to answer each specific research question.

Statistical analysis and calculation of new variables was accomplished in the R programing language. The first step in this process was to compute additional necessary variables from existing data, specifically, BMI, energy expenditure, energy expenditure per meter (EEnorm), and the slope of each line segment. Data was first read into R and certain fields were formatted to be used in the analysis. Finally, each mentioned variable was computed. The variables were computed using an R script (found in APPENDIX B – Computer Programming Titled: R Script used to compute additional variables).

General statistics of the resulting data set are listed below. These were derived using functions in the R programing language. The code for deriving these statistics can be found in APPENDIX B – Computer Programming Titled: R Script to compute general dataset statistics.

Summary of the Subjects

Number of Subjects: 200 Total, 162 Male, 38 Female

Age (Frequency): 17(16); 18(115); 19(47); 20(9); 21(4); 22(7); 23(2)

Weight: Range: 50.8–117.9 (kg)

Mean: 76.7 (kg)

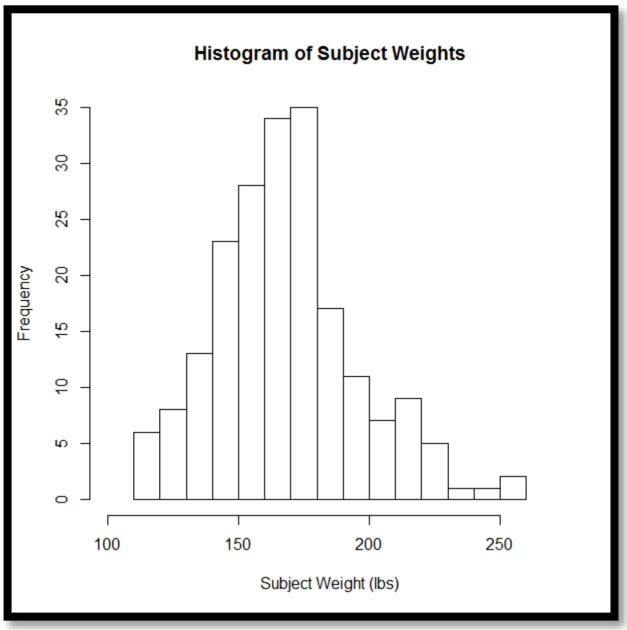

Figure 58: Subject weights.

Height: Range: 1.55–1.98 (m)

Mean: 1.75 (m)

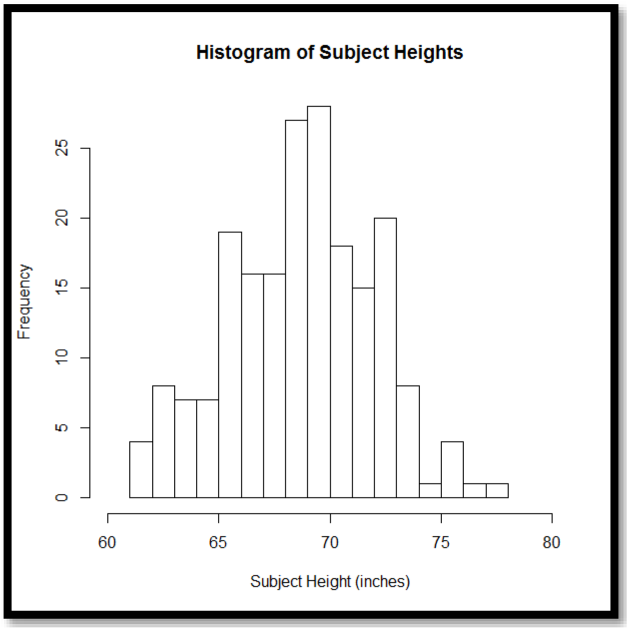

Figure 59: Subject heights.

BMI: Range: 18.46–34.23 (kg/m^2)

Mean: 24.81 (kg/m^2)

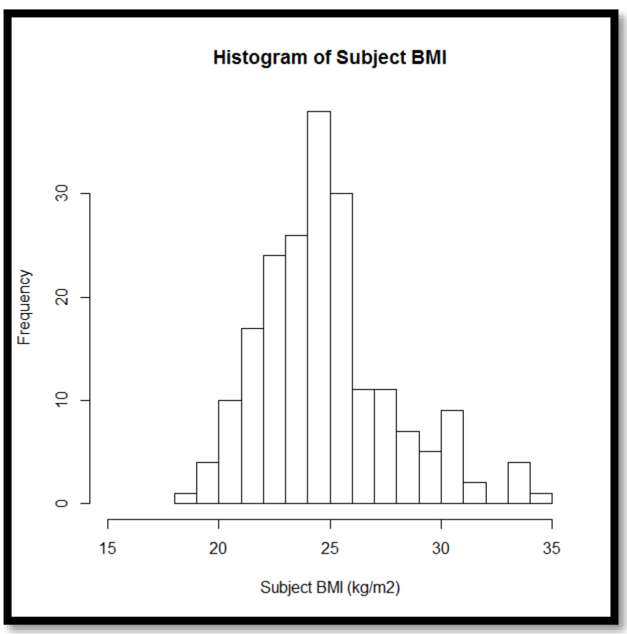

Figure 60: Subject BMI.

Load Carried: Range: 4.99–16.3 (kg)

Mean: 10.3 (kg)

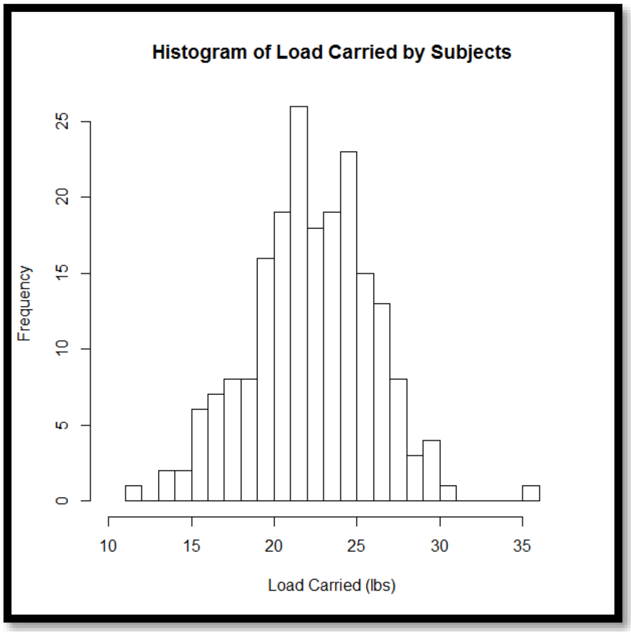

Figure 61: Load carried.

Fitness Score: Range: 61–300 (points)

Mean: 212 (points)

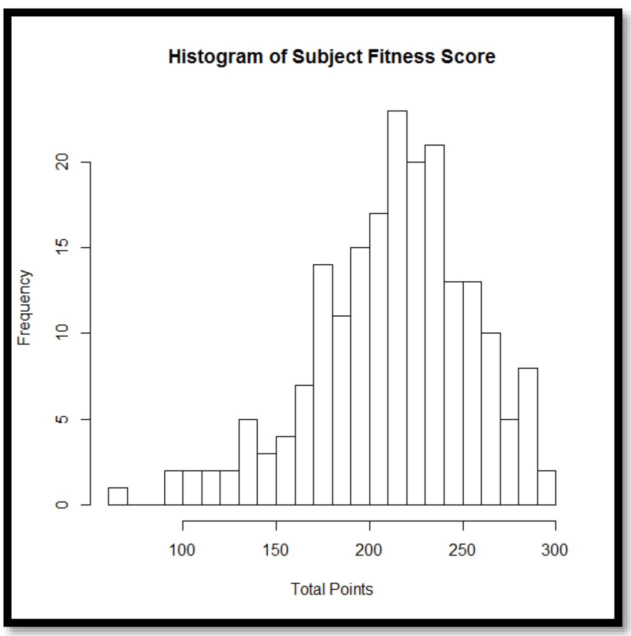

Figure 62: Subject fitness score.

Run Score: Range: 0–100 (points)

Mean: 75 (points)

Figure 63: Subject run score.

Statistics of the Terrain

Data was constrained to -50<Slope<50 since some segments had very little movement over the 10-second duration, which in turn created some extreme slope values. It is unlikely subjects would move over slopes of greater than ±50 percent. This reduces the total number of line segments used in the analysis from 168645 to 147343. The following two pages describe the slope and the land cover of the terrain.

Slope: Average Slope: 0.95%

*This makes sense since most subjects finished where they started.

Average Absolute value of slope: 14%

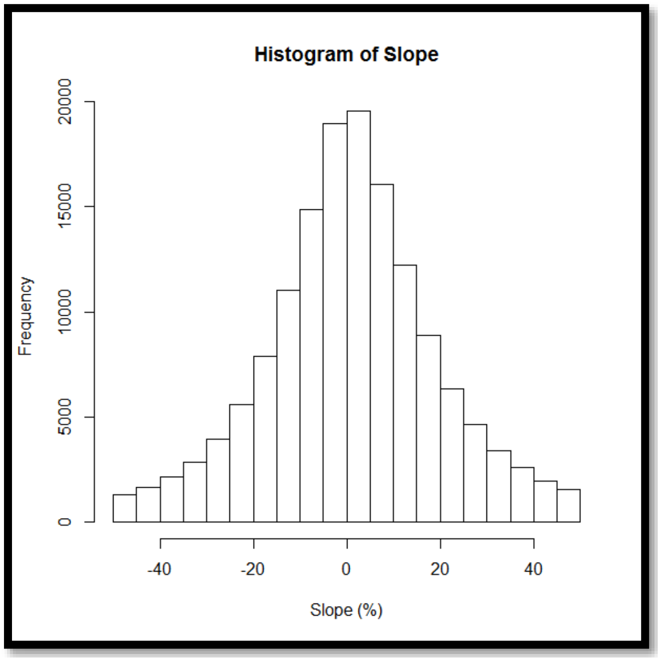

Figure 64: Slope traversed by subjects.

Number of line segments in each land cover class:

Land Cover Class	**# of Data Points**	**% of Data Points**
1 On Road	35348	24.00
2 Boulders	16380	11.00
3 Water	110	0.01
4 Light Vegetation	2672	1.80
5 Moderate Vegetation	2355	1.60
6 Heavy Vegetation	476	0.30
7 Swamp	1989	1.30
8 Open Woods	71763	49.00
Combination of Classes	16250	11.00

Figure 65: Summary of land cover passed through.

Statistics of Modeled Variables

This section describes the speed values and the energy expenditure data. Speed was computed by using a distance divided by time computation. The range, mean, and histogram of the speeds are used to describe the data. First, the speed of all line segments with slopes between ±50 percent is described. Then the dataset was curtailed by removing segments where there was no accelerometer movement or the speed was less than .14m/s. This curtailment is required to answer Research Questions 1–3. Rationale is given in the following section while describing the data.

Statistics of Speed

Range of speeds of segments: 0–4.4 m/s

Average speed of each segment: 0.74 m/s

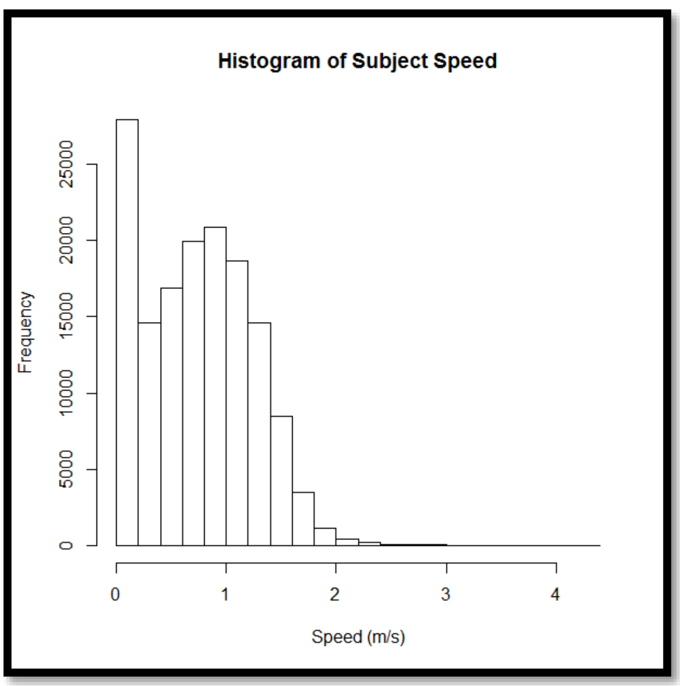

Figure 66: Histogram of subject speed.

Research Questions 1–3 are strictly concerned with modeling the energy expenditure of navigation while moving across terrain. The data where the subjects were simply standing is removed. In the dataset, this can be filtered by finding METS values =1 and when the speed is less than 0.14m/s.

Average speed with 0.14 m/s and lower taken out:

Range: 0.14–4.4 m/s

Mean: 1.03 m/s

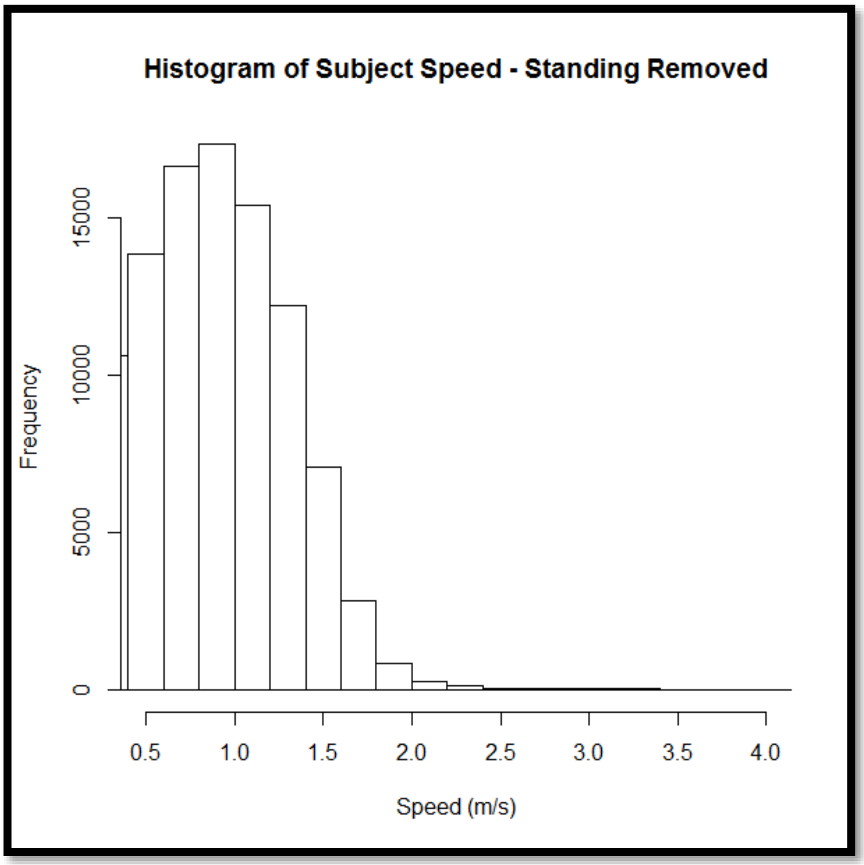

Figure 67: Subject speed with standing removed.

Energy Expenditure

The energy expended over certain terrain was modeled using the segments that are represented as rows in our data matrix. METS for each subject are derived by using the accelerometer counts and the Crouter refined 2 regression model as described in Chapter 3. Then the energy expenditure is computed based on the below equation (Humphrey 2006):

EE(Calories)=METS*Body Weight (kg)/57.

Since the subjects are carrying a modest load, the load is added to the body weight and used in the final equation for determining the energy expenditure:

EE(Calories)=METS*(Load+Body Weight (kg))/57.

Finally, the energy expenditure must be normalized over the distance of the segment for the statistical modeling. Calories/meter were finally derived by dividing the EE by the distance of the segment in meters. The general statistics of energy expenditure / meter are shown in Figure 68:

Range: 0.02–1.8 Calories/meter

Mean: 0.21 Calories/meter

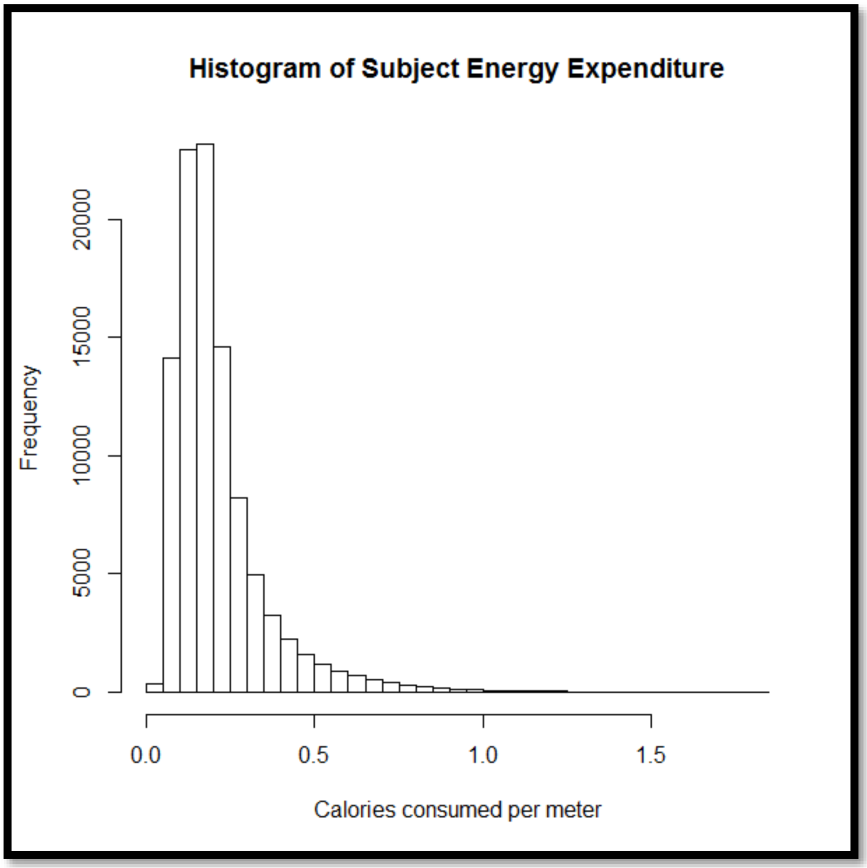

Figure 68: Subject energy expended per meter.

Overall Track Statistics

This section describes the overall statistics of routes taken by subjects, summarizing the length of time subjects navigated, the distance covered, and the terrain that was navigated through. Although overall statistics of the subjects was not used in the final modeling step, this review of the trajectories was used to validate the data, develop a general understanding of navigational behavior, and to find anomalies.

Distance of Navigation

Range of distance covered while navigating:	2358–7603 meters
Average distance covered while navigating:	5655 meters
Range of male average distance covered while navigating:	2358–7603 meters
Male average distance covered while navigating:	5682 meters
Range of female average distance covered while navigating:	3517–7468 meters
Female average distance covered while navigating:	5538 meters

Duration of Navigation

Range of times on the course:	88–185 min
Average time on the course:	141 min
Range of times on the course (Male):	88–185 min
Average time on the course (Male):	139 min
Range of times on the course (Female):	115–183 min
Average time on the course (Female):	147 min

Seven trajectories began at the start point but did not return to it, most likely due to a GNSS losing battery power or turning off. These seven trajectories were removed for our summary of the distance and duration of the routes, but were left in for the modeling effort since they would have significant influence on ranges and means of the overall trajectory summary but would not affect our modeling variables.

Summary of Data Matrix Used to Answer Research Questions 1–3 (EE)

Research Questions 1–3 involve energy expenditure by the subjects while they navigate. The modeling effort curtails the original data matrix described in in Figure 57. Line segments with slope greater than the absolute value of 50 percent were removed to ensure that noise from GPS error when subjects were not moving was minimized. It is unlikely that subjects navigated over such steep terrain. Additionally, work was done to remove line segments where standing occurred. Although standing is part of navigation, it is removed from the energy expenditure model so as to focus on the contribution of terrain and individual's physical attributes to Calorie consumption. All records without accelerometer movement were removed, as were records with speeds less than 0.14 m/s. It is very unlikely that someone walked less than 0.14 m/s along the length of a segment. This is shown by Tobler's hiking function that proves 0.14 m/s is lower than expected speeds between -50 and 50 percent slope (Tobler 1993). The finalized data matrix used to answer Research Questions 1, 2, and 3 contained attributes as described in APPENDIX C – Description of Data Matrix Fields. The data matrix had the following properties:

Number of Attributes:	54
Number of Records:	100348
Number of Subjects:	200
Approximate number of records per subject:	~500 *number of records per subject varied based on how quickly they finished the navigation test.
Each record represents a 10-second segment of navigation	

Figure 69: Properties of data matrix used to answer Research Questions 1, 2, and 3.

Summary of Data Matrix Used to Answer Research Question 4 (Speed)

Research Question 4 focused on the speed of navigation. This modeling task curtails the original data matrix described in Figure 57. For the same rationale as above, line segments with slope greater than ±50 percent were removed. However, data representing standing behavior remained in the data matrix. The finalized data matrix used to answer Research Question 4 contained attributes described in APPENDIX C – Description of Data Matrix Fields. The data matrix had the following properties:

Number of Attributes:	54
Number of Records:	130983
Number of Subjects:	200
Approximate number of records per subject:	~650 *number of records per subject varied based on how quickly they finished the navigation test.
Each record represents a 10-second segment of navigation	

Figure 70: Properties of data matrix used to answer Research Question 4.

The steps outlined in this chapter describe the processes applied to develop finalized datasets. The discussion described the procedures used when initializing the wearable devices used in the observational study; the download procedures were specified; computation techniques and scripts were provided to assist understanding data development, and summary statistics pertaining to the data variables were detailed. Further, all decisions made pertaining to the exclusion of data records have been recorded. This chapter concludes by outlining the finalized datasets, and aims to give structure to how the data was derived. Finally, it provides information for others interested in using the data in the future.

CHAPTER 5

RESULTS

The results of this research are summarized in this chapter, describing the outcomes of the methods outlined in Chapter 3. The chapter begins by presenting the results of Research Questions 1 and 2. It defines the process outputs when determining the contributing factors of EE when navigating as well as the contributing weights of BMI, fitness, sex, slope, land cover, and distance traveled. Further, it quantitatively assesses GRL's EE algorithm and concludes by comparing Tobler's hiking function and the speed at which subjects moved during the navigation test. This chapter provides a quantitative assessment of the process taken to answer each research question.

Results of Research Questions 1 and 2

Energy expenditure of navigation is affected by human physiological variables and the type of terrain over which a person travels. Energy expenditure modeling required collecting empirical EE estimates and independent variable data. Then, relationships between the variables and the EE data were used to develop model parameters.

Investigation of the energy expenditure estimates from the navigators yielded a non-Gaussian distribution. Review of Figure 68 clearly shows a positive skew. A box-cox transformation was applied to the energy expenditure data to make the distribution normal. Lambda from the Box Cox procedure is -0.22. A plot of the results of this procedural assessment are shown below in Figure 71. The Box Cox method recommends transforming the data with a log transform if the Lambda is -0.22. After applying this transformation, the distribution is more representative of a Gaussian distribution as can be seen in Figure 72.

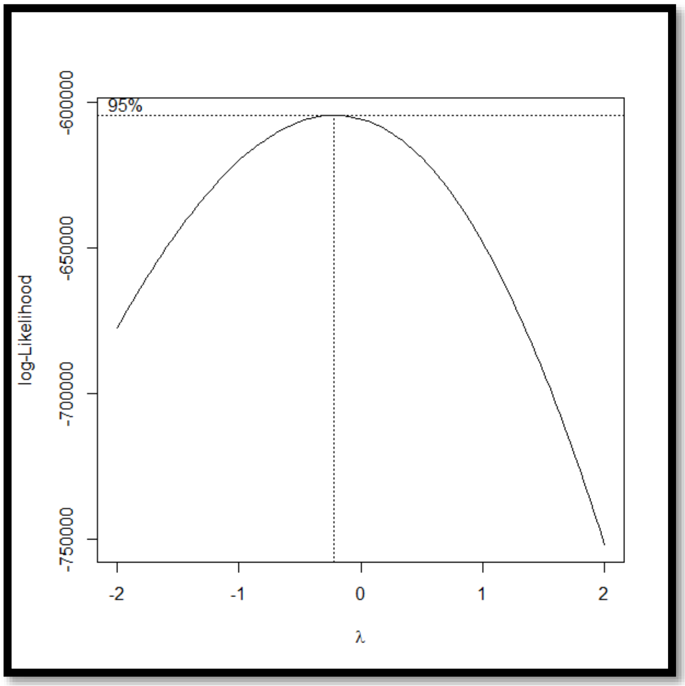

Figure 71: Box-Cox plot of energy expenditure values

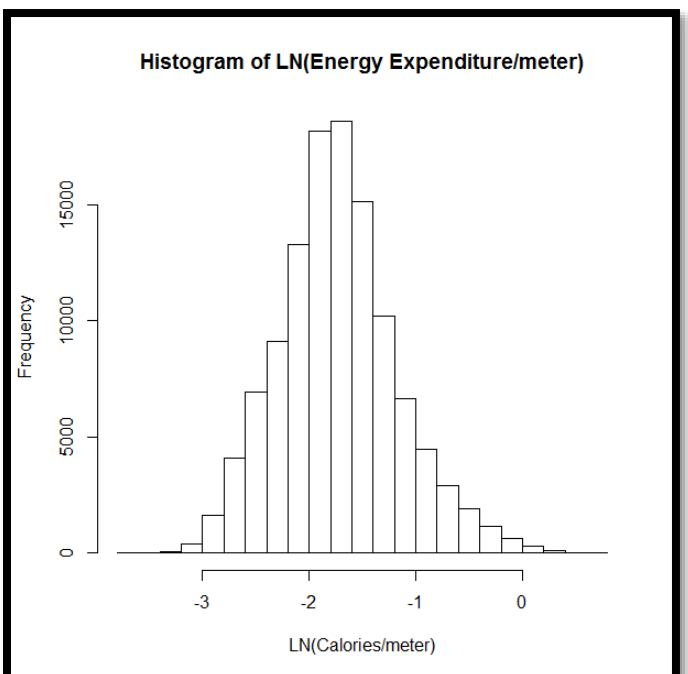

Figure 72: Transformed energy expenditure data

The variable of slope was also investigated. Linear modeling requires a linear relationship between the dependent variable and the independent variables. The relationship between the slope and the natural logarithm of the energy expenditure per meter is plotted in Figure 73, showing a non-linear relationship that approximates a quadratic polynomial with the vertex at -4 percent slope. This leads to squaring the slope term during model development in order to use it in the linear modeling effort.

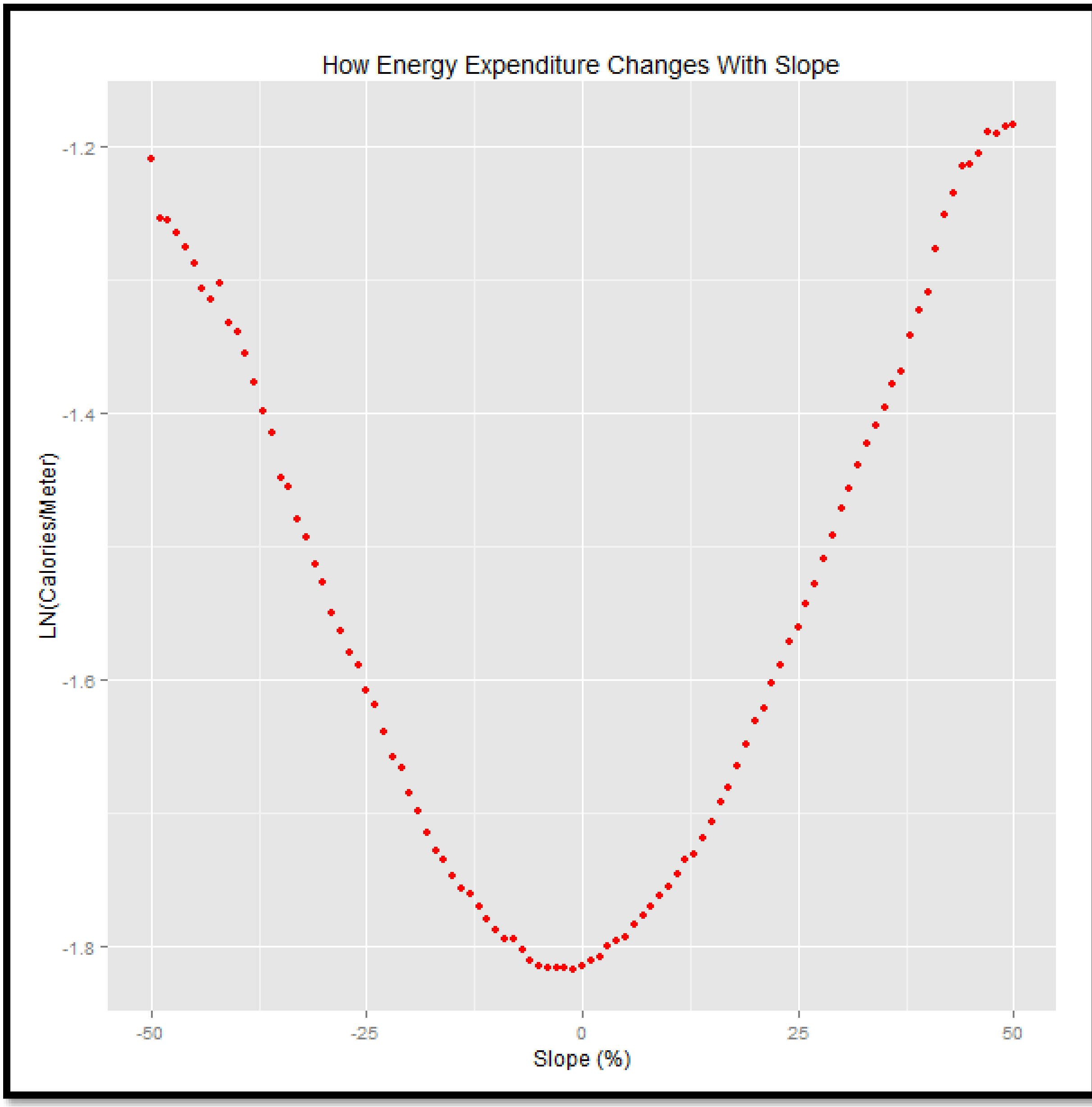

Figure 73: Nonlinear relationship between slope and energy expenditure.

Raw land cover data was also transformed for modeling. First, 0.1 percent of the data points were classified as water and were omitted from the model since subjects were instructed not to enter the lakes or ponds and it is unlikely that subjects swam or waded through water. Since there were so few points classified as water, it was impossible to obtain any significant results from modeling. Furthermore, even if subjects had waded through water, it is unlikely that the EE estimates produced by the Actigraph GT3X were accurate.

A second transformation was made in the land-cover data that involved class consolidation. All vegetation and swampy areas were consolidated into one class called “vegetation” for the following reasons: these classes had many fewer data points than other classes; there seemed to be little discernable difference in EE/meter of movement between these classes; and visual re-inspection of each class on satellite imagery provided little justification that there was measurable difference between these four classes.

Navigators mostly traversed areas other than vegetated or swampy areas during the navigational exercise. Swamp, light vegetation, moderate vegetation, and heavy vegetation represented only 5 percent of the data. This is significantly lower than other classes in the model.

These four classes had similar caloric expenditure. Figure 74 shows light, moderate, heavy, and swamp had similar caloric expenditure. The figure shows these four classes are clustered, and because there are few observations, it shows that the data is noisy. This is especially true at grades greater than ±15 percent slope. In fact, it shows that light vegetation requires slightly higher Calorie consumption than swamp, moderate vegetation, or heavy vegetation, which is counter-intuitive.

Satellite imagery was used to reinvestigate the classification. Swamp land appeared mostly dry. Many of the areas initially classified as swamp had areas with worn foot paths. Additionally, it was difficult to see much difference between the levels of light, moderate, and dense vegetation in the imagery. Similar to swampy areas, there were paths worn through the densely vegetated areas that were likely used by the subjects and would probably lead to similar EE.

Because these factors would suggest similarity in energy expenditure between light vegetation, moderate vegetation, heavy vegetation, and swamp they were grouped together to simplify the model. Figure 75 shows the energy expenditure of a consolidated vegetation class. It shows that moving over rocky areas require the most energy and that moving over vegetated areas is the next most energy intensive, while moving on the road is the easiest way to maneuver.

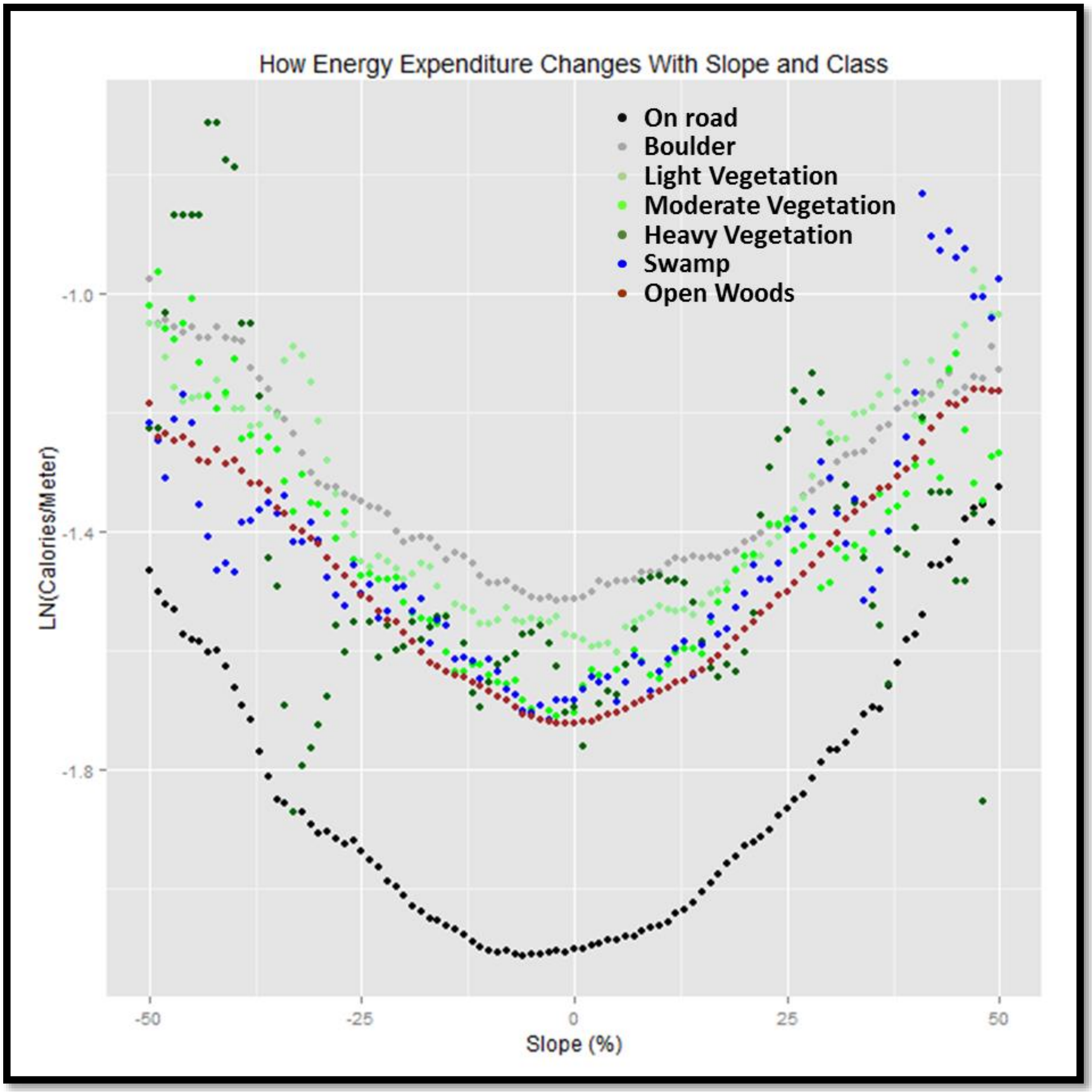

Figure 74: Energy Expenditure of different land cover classes.

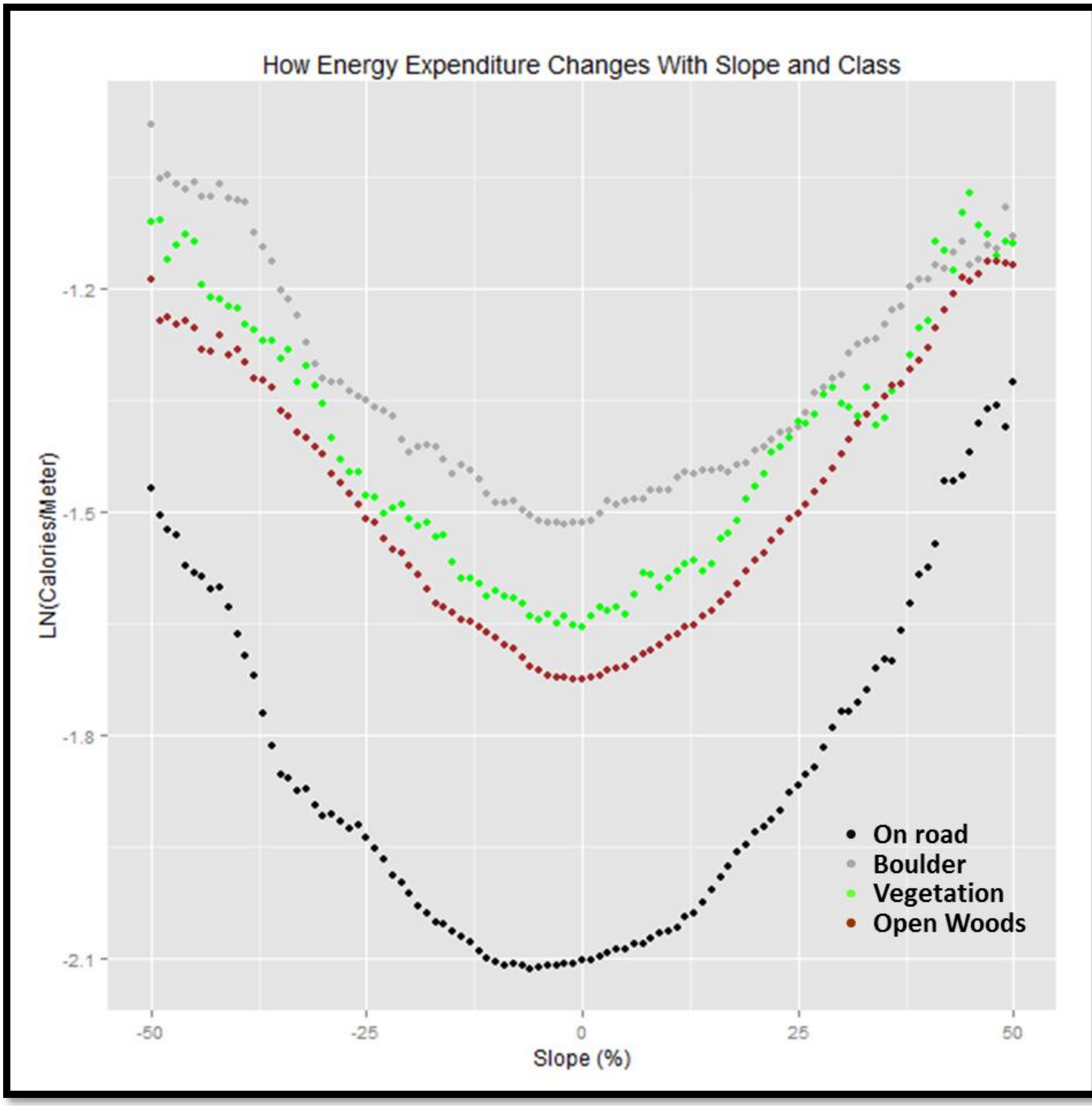

Figure 75: Energy Expenditure of different types of land cover.

The transformed variables were used in a multiple regression analysis to investigate how energy expenditure can be modeled by land cover, slope^2, distance from the origin, BMI, fitness, and sex. This technique provides evidence that there was a significant relationship between EE and the independent variables. The linear regression results are listed in Figure 76.

Model: LN (EE) ~ Land Cover + $Slope^2$ + Distance from Start + BMI + Sex + Fitness Score

Residual standard error: 0.487 on 100222 degrees of freedom

Model Coefficients and significance:

Variable	**Coefficient**	**P Value**	**Relative Importance**
Intercept	-2.83	$<2x10^{-16}$	
Land Cover			47%
Boulder	.505	$<2x10^{-16}$	
Vegetation	.437	$<2x10^{-16}$	
Open Forest	.347	$<2x10^{-16}$	
$Slope^2$	2.7x10-4	$<2x10^{-16}$	20%
Distance from Start	$-4.4x10^{-5}$	$<2x10^{-16}$	14%
Sex			5%
Male	.13	$<2x10^{-16}$	
BMI	$3.2x10^{-2}$	$<2x10^{-16}$	14%
Fitness Score	$-1.7x10^{-4}$	$2.14x10^{-4}$	<1%

Multiple R-squared: 0.25

AIC: 140095

Figure 76: Regression results with all independent variables included.

This method of analysis also provides insight into the significance of each independent variable. It can be used to identify variables that do not contribute to the model. Figure 76 shows that fitness score has a very low relative importance and requires further investigation.

The research investigated the variable “fitness score” for significance in several ways. First, the fitness metric was examined. The modeling effort considered replacing overall fitness with running fitness. The model described in Figure 76 uses the subject’s APFT score as a measure of fitness. The APFT metric includes a sum of scores from the three events which make up the APFT: pushups, sit-ups and a 2-mile run (fully detailed in the data chapter). It is reasonable to think that a running test may be more significant to a navigation model than one that included all three events. Results of replacing APFT score with a subject's run score are detailed in Figure 77. No significant improvement in modeling characteristics are present.

Model: LN (EE) ~ Land Cover + Slope^2 + Distance from Start + BMI + Sex + Run Score

Residual standard error: 0.487 on 100222 degrees of freedom

Model Coefficients and significance:

Variable	**Coefficient**	**P Value**	**Relative Importance**
Intercept	-2.83	$<2x10^{-16}$	
Land Cover			47%
Boulder	.505	$<2x10^{-16}$	
Vegetation	.437	$<2x10^{-16}$	
Open Forest	.347	$<2x10^{-16}$	
Slope^2	2.7x10-4	$<2x10^{-16}$	20%
Distance from Start	$-4.4x10^{-5}$	$<2x10^{-16}$	14%
Sex			5%
Male	.13	$<2x10^{-16}$	
BMI	$3.2x10^{-2}$	$<2x10^{-16}$	13%
Run Score	$-2.6x10^{-4}$	$2.13x10^{-4}$	<1%

Multiple R-squared:0.25

AIC: 140105

Figure 77: Multiple regression results after replacing fitness score with run score.

Multiple regression computation was conducted without the variable of "fitness score," providing an assessment of whether or not to include the fitness variable in the model. The model diagnostics of the multiple regression method without the fitness variable are shown in Figure 78. The results after removing the fitness variables indicate that run score or fitness score should be excluded from continued modeling efforts. The coefficients, R-squared, AIC and relative importance did not change significantly when the variables are omitted. Further discussion of this exclusion is provided in the conclusions.

Model: LN (EE) ~ Land Cover + Slope^2 + Distance from Start + BMI + Sex

Residual standard error: 0.487 on 100340 degrees of freedom

Model Coefficients and significance:

Variable	**Coefficient**	**P Value**	**Relative Importance**
Intercept	-2.87	$<2x10^{-16}$	
Land Cover			47%
Boulder	.505	$<2x10^{-16}$	
Vegetation	.438	$<2x10^{-16}$	
Open Forest	.348	$<2x10^{-16}$	
Slope^2	2.7x10-4	$<2x10^{-16}$	20%
Distance from Start	$-4.4x10^{-5}$	$<2x10^{-16}$	14%
Sex			5%
Male	.13	$<2x10^{-16}$	
BMI	$3.3x10^{-2}$	$<2x10^{-16}$	14%

Multiple R-squared: 0.25

AIC: 140287

Figure 78: Multiple regression results after removing fitness variable.

The Generalized Variance Inflation Factor (GVIF) is used to test model variables for multicollinearity. This factor quantifies how much the variance is inflated by multicollinearity in the model. Factors are derived for each of the independent variables. A factor near 1 shows nearly no correlation between that variable and any of the other independent variables. Figure 79 shows the GVIF of each independent variable. A GVIF greater than 10 is cause for concern (Kutner, Nachtsheim, and Neter 2004). All of the GVIFs computed for this model are near 1, thus proving nearly no correlation between independent variables.

Variable	GVIF
Land cover	1.1
Slope2	1.0
Distance from the start	1.1
Sex	1.07
BMI	1.06

Figure 79: Generalized Variance Inflation Factors (GVIF) of each coefficient.

Linear regression also requires that residuals are approximately normally distributed. Common methods of showing the distribution of data are histograms and box plots. Figure 80 shows a histogram of the studentized residuals. Figure 81 shows the mean of the studentized residuals by subject. These plots are used to validate the linear regression assumption that the residuals are normally distributed around zero.

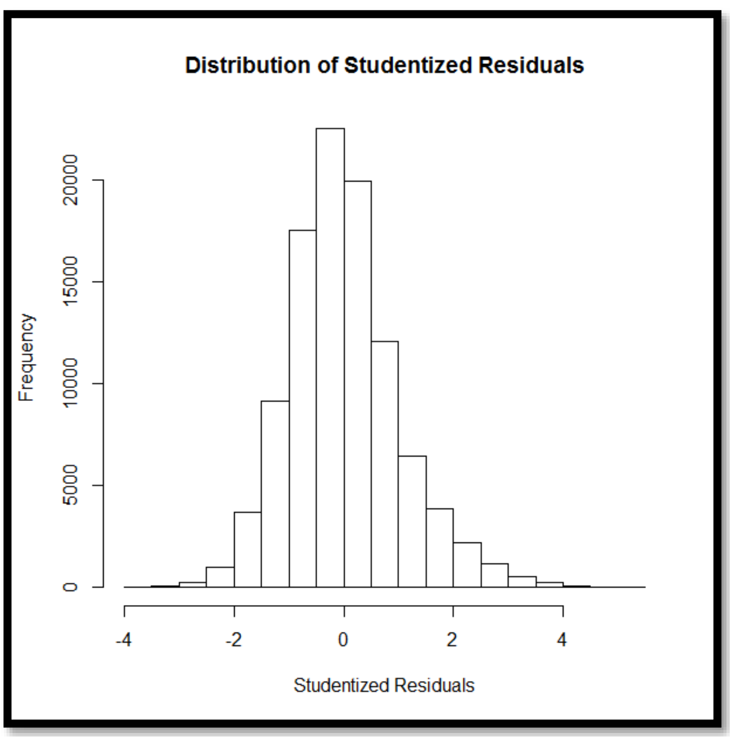

Figure 80: Studentized residuals.

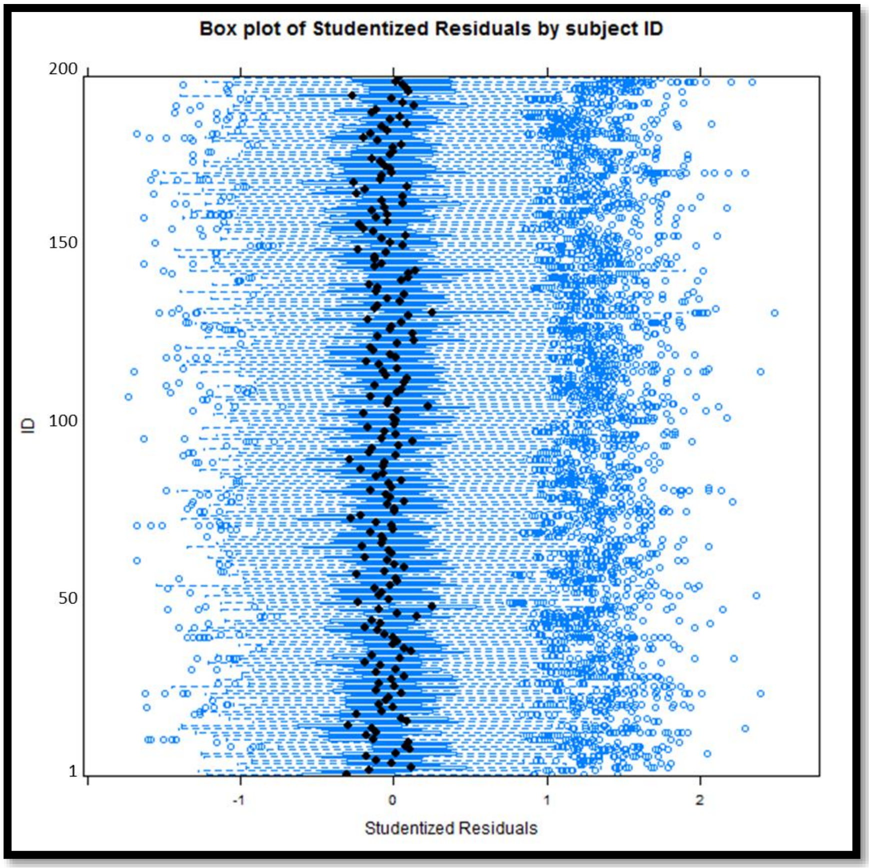

Figure 81: Box plot of studentized residuals grouped by subject ID.

One surprising result from the multiple regression procedure was obtaining a negative coefficient for Distance from Start. Initial expectations were to have a positive relationship between energy expenditure per meter and distance from the start point. The result displayed in Figure 78 led to further investigation of this variable. Figure 82 shows the relationship of how energy expenditure changes with distance from the origin. This plot was constructed by first creating equal interval bins based on the distance from the start point. Bins with a width of 100 meters were generated along the range of the data. Then data falling within each bin was generalized by computing the mean of the model's dependent variable. Further discussion of this variable and its relationship with energy expenditure is provided in Chapter 6.

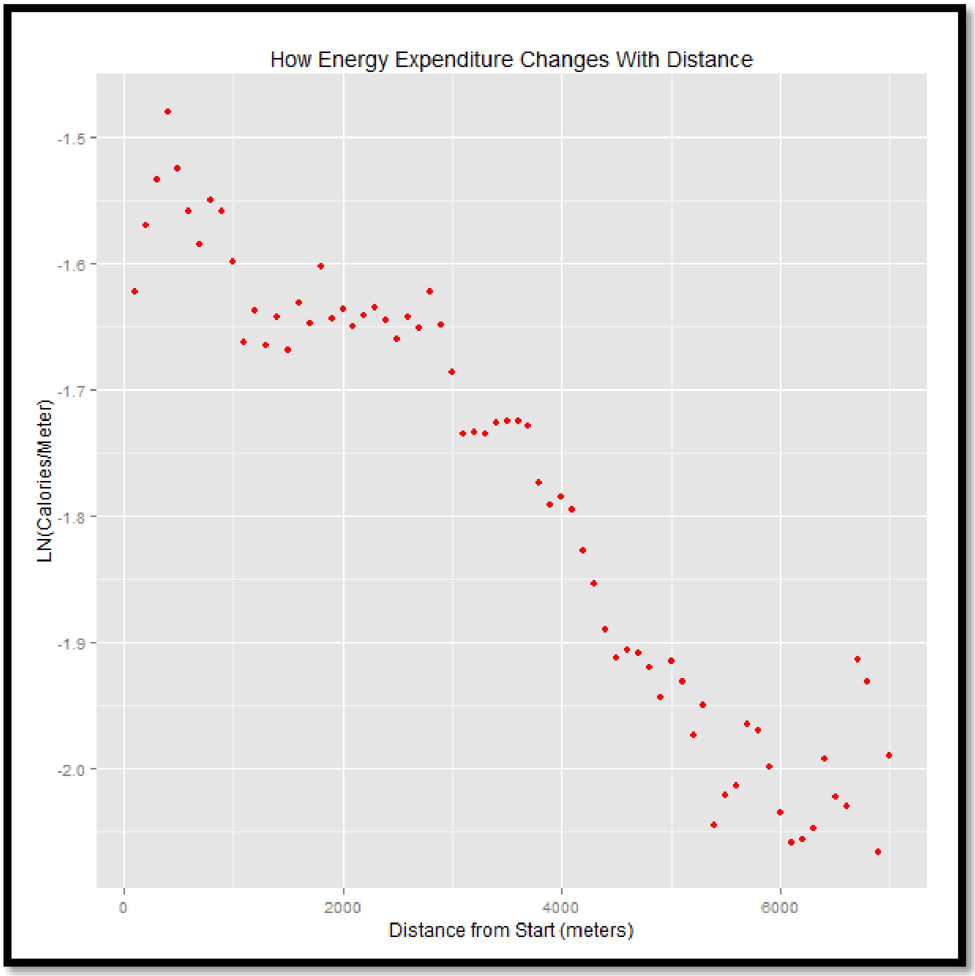

Figure 82: How energy expenditure changes with distance from start point.

Care must be taken in using multiple regression to analyze our model. Attention is required since the data violates the assumption that all instances are independent. The data used in this research is grouped on the basis of individual subjects and it follows that a statistical method that considers these issues must be used. Linear Mixed Effects Modeling can handle grouped data with varying slopes and intercepts. The results of this technique are described below.

Linear Mixed Effects Modeling uses computational regression techniques to account for the varying intercepts and slopes of the grouped data. This method was completed using the nlme package in R (Pinheiro et al. 2015). The varying intercepts and slopes of a model based on grouped data are considered the random effects of the model. Model coefficients of the random effects are provided for each data grouping. The method also provides an overall assessment of how much the random effects contribute to the modeled variables. Finally, it provides fixed effects model coefficients based on the entire sample. Figure 83 shows the results from the statistical analysis. A total of 30 percent of the variance is explained by the model. The fixed effects explained 23 percent of the variance and the random effects explained 7 percent of the variance.

Model: LN (EE) ~ Land Cover + Slope^2 + Distance from Start + BMI + Sex + Random Effects
Random effects: (1+ Land Cover + Slope^2 + Distance from Start|ID)

Residual standard error: 0.47

Model Coefficients and significance:

Variable	**Coefficient**	**P Value**
Intercept	-2.85	$<2x10^{-16}$
Land Cover		
Boulder	.45	$<2x10^{-16}$
Vegetation	.39	$<2x10^{-16}$
Open Forest	.31	$<2x10^{-16}$
Slope^2	2.7x10-4	$<2x10^{-16}$
Distance from Start	$-5.3x10^{-5}$	$<2x10^{-16}$
Sex		
Male	.11	$<2x10^{-16}$
BMI	$3.5x10^{-2}$	$<2x10^{-16}$

R Squared from fixed effects: 0.23
R Squared from fixed and random effects: 0.07

AIC: 134378

Figure 83: Linear mixed effects modeling results.

Similar to multiple regression, suitable Linear Mixed Effect models must have residuals that follow a normal distribution centered on zero. The validation of the studentized residual distribution is shown in Figure 84. This box plot shows that the distribution of the residuals follows the required distribution. Also, comparison of this plot with Figure 81 shows that including the random effects of the grouped data improves the bunching of the distributions around zero.

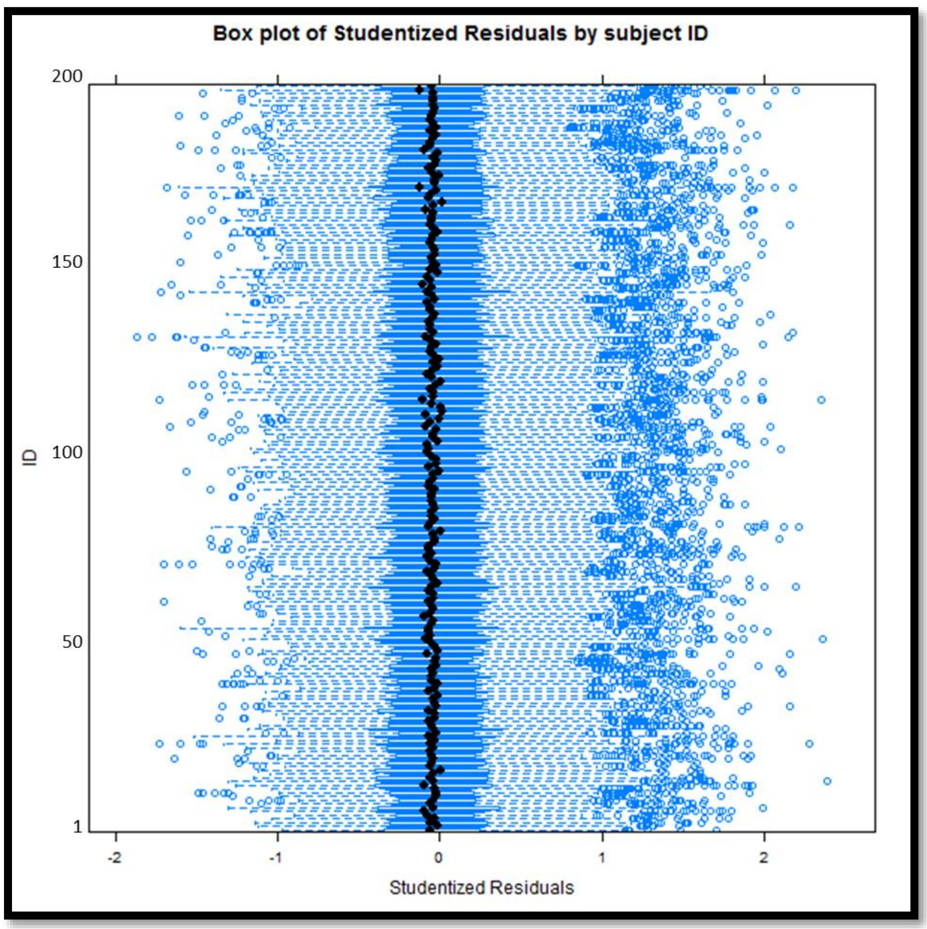

Figure 84: Box Plot of studentized residuals by group for the Linear Mixed Effects Model.

Results of Research Question 3

Question 3 assesses a currently used energy expenditure model. It uses EE estimates collected during experimentation to evaluate GRL's algorithm, which is currently being used by military planners to predict energy expenditure during navigation. This assessment serves as the first, field-based validation of this tool.

Each subject's trajectory was used to determine a trip-based energy expenditure value. This value was computed using the GRL algorithm to predict Calories consumed on each subject's route[7]. These values were then compared with empirically collected energy expenditure values from the Actigraph GT3X. The Pearson product-moment correlation coefficient (r) that measures the relationship between the predicted and actual values was determined. This coefficient was found to be 0.66. The mean absolute error (MAE) of the GRL model is 83 Calories. The mean percent error between estimated and actual value was 14 percent. Graphical depictions of the results are shown in Figure 85. This plot shows the model deviation from the actual values. If the modeled result exactly matched the predicted result it would fall on the superimposed line.

---

[7] The GRL algorithm uses terrain factors based on the type of land cover. These factors did not exactly match the classification. Terrain factors for road, path, light vegetation, heavy vegetation, and swamp are defined by the GRL algorithm. This research equated path movement with open-ground movement for calculations. Swamp and heavy vegetation areas were omitted for the same reasons discussed earlier in this chapter.

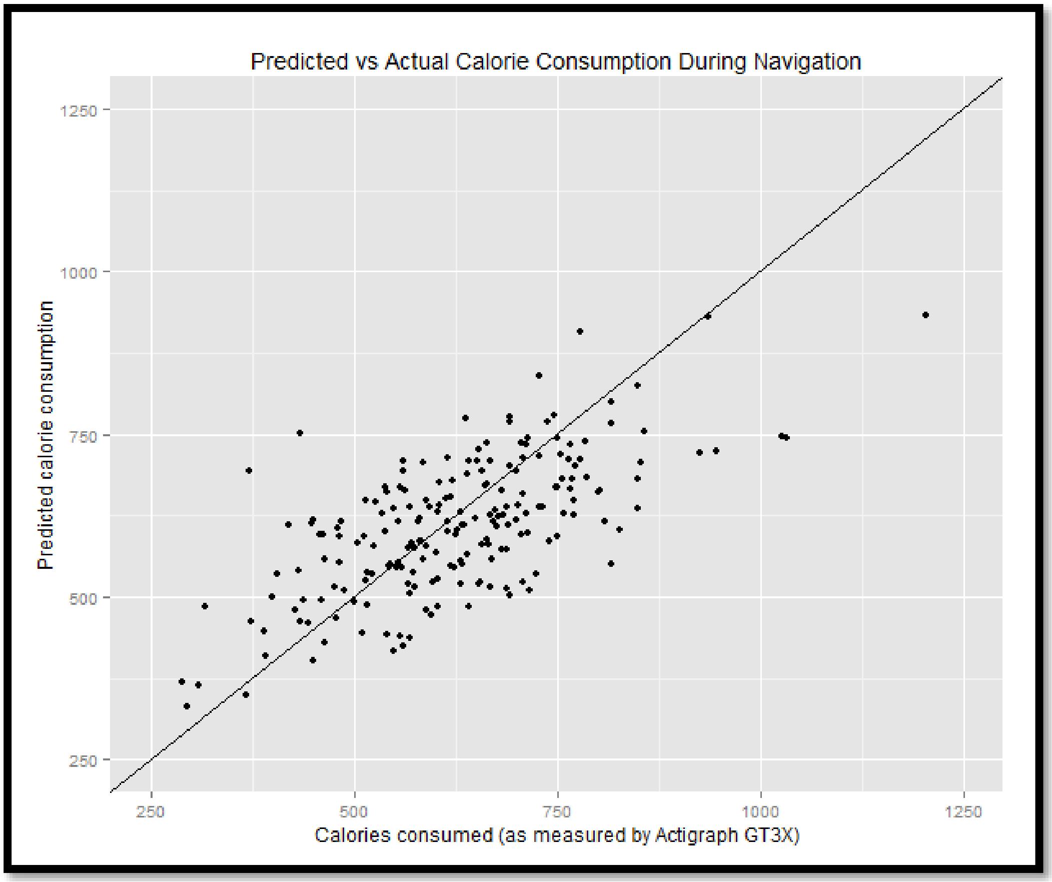

Figure 85: Predicted Calorie consumption vs actual Calorie consumption using GRL's model.

Results from the model created in this research are presented for comparison. Discussion of how the model was created can be found in the previous section. Pearson's r comparing predicted vs actual is 0.86. The MAE is 66 Calories and the average percent error is 11 percent. A plot showing the predicted values from the newly derived model is shown in Figure 86.

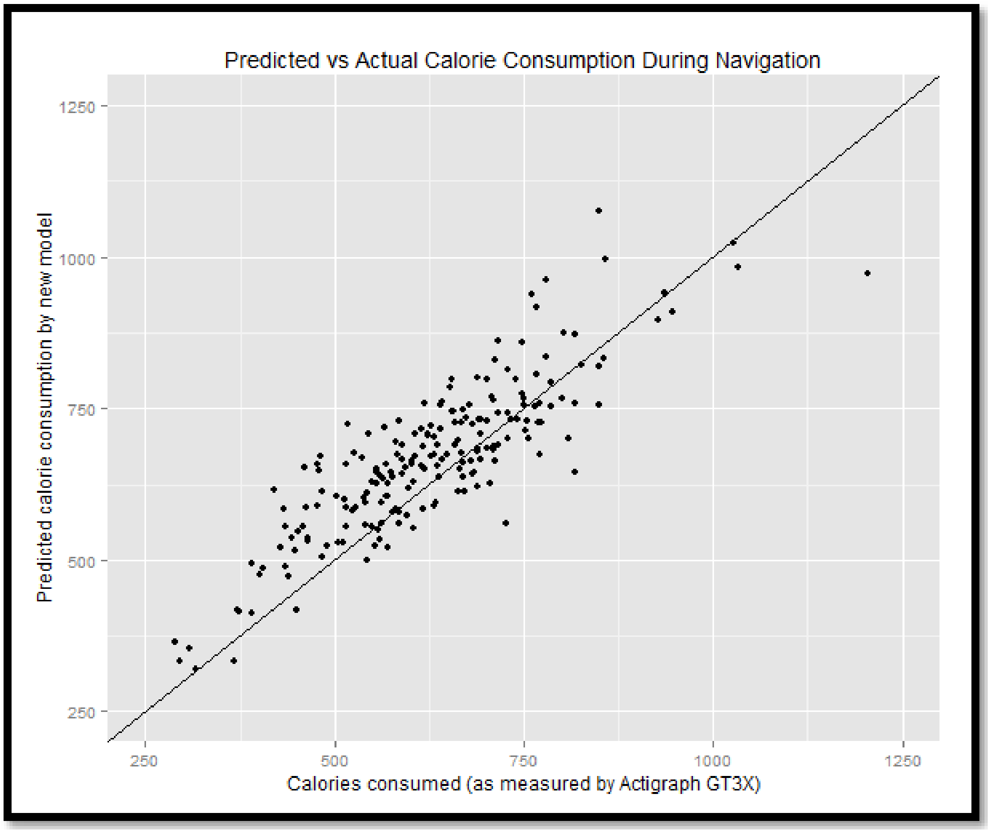

Figure 86: Predicted Calorie consumption vs actual Calorie consumption using the derived model.

Results of Research Question 4

Question 4 queries the speed at which humans conduct dismounted navigation. Understanding how long it will take a person to traverse a route is of great value for setting realistic scheduling goals and expectations. Estimation of human navigational speed is needed to successfully be at a specified location at a given time. This section describes the speeds at which the subjects navigated, quantitatively assessing the subjects' movement and developing a model to represent the speed of human navigation. The navigators' speed is shown at different slopes and a curve fitting effort is shown. Then results of the modeling effort are described.

Speeds of navigation were recorded every second for each navigator. For analytical purposes, data was categorized as either on- or off-road. These categories follow previous research in this domain completed by Tobler and Naismith (Tobler 1993; Naismith 1892). Initially, the data was down scaled to 10-second intervals. Speed over each interval was then related to the slope of the terrain traversed. The results are reported by generalizing bin means based on the data's slope.

Bins were created every 1 percent slope between ±50% and all trajectory segments with a slope value falling within a bin were grouped. Speeds of the grouped data were then averaged. Figure 87 shows the data collected while subjects navigated on roads. Figure 88 then shows the data from off-road movement. In both figures the blue represents the bin mean of the subject's speed. The red dots represent expected speeds of movement computed using Tobler's hiking function.

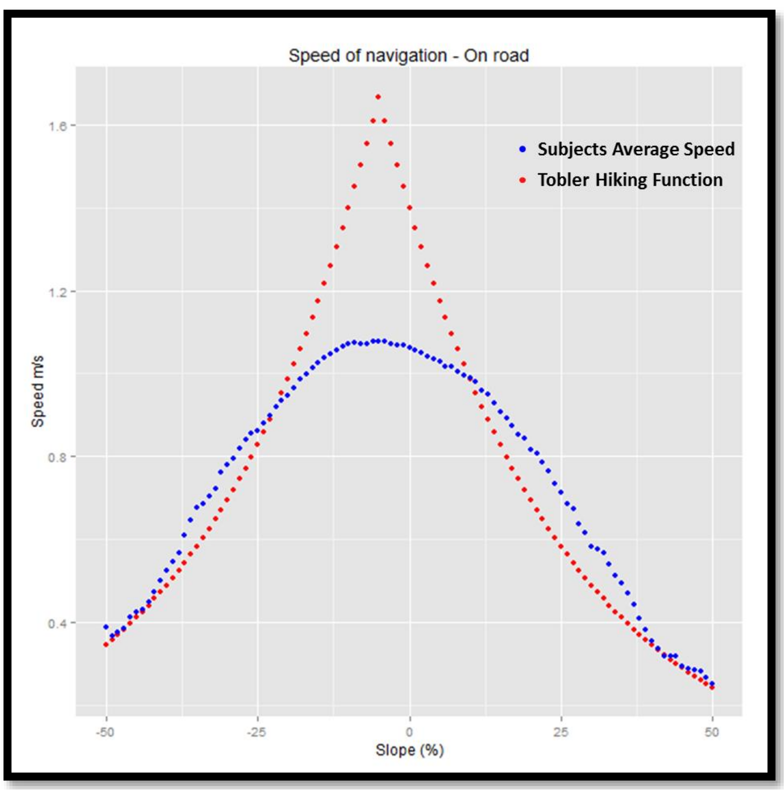

Figure 87: Speed of navigation at varying slopes—on road movement.

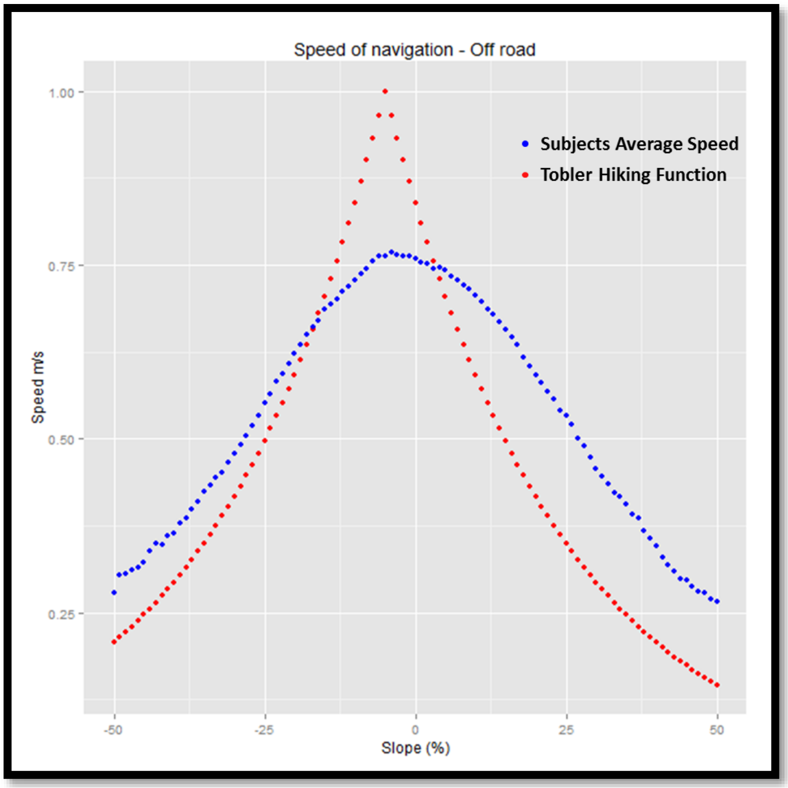

Figure 88: Speed of navigation at varying slopes—off road movement.

The graphical depiction of the data shows a clear relationship between navigation speed and slope, which can be modeled to quantitatively answer Question 4. The blue dots in the figure show a quadratic relationship between the two variables. Curve fitting was used to find a function that can best approximate the relationship. The R statistical package was used to estimate a curve to fit the data. The results of the curve fitting are shown in Figures 89, 90, 91, and 92.

These results show the model succeeds at representing navigator speed. However, the results only partially help to answer Question 4 and improve understanding of the dynamics of navigation. The following sections further discuss the model characteristics, the data and fully develop results to answer to the question.

The model can be quantitatively and visually assessed. Figure 89 shows the model diagnostics of our curve fitting effort for on-road movement. Modeled coefficients are listed in the figure. Model fit for slope values between ±30 percent were weighted heaviest since the majority of movement is accomplished at these grades (See Figure 64, histogram of slope values in Chapter 4, for specifics on the frequency of data by slope). Each coefficient reflected P Values near zero indicating a significant result. Figure 90 shows similar summary statistics for the off-road model.

The visual assessment of the model fit is shown in Figure 91 and 92. The curve is very good at approximating the data within the subset of slope range. The parabolic curve chosen to model on-road speed is slightly better at matching the extreme slope values than the one chosen for off-road speed. Both do an excellent job of matching the bin means between ±30 percent slopes (90th percentile of the data).

Model: Speed=1.07-.004Slope-.00045Slope^2 - On-Road Movement

Residual standard error: 0.048

Model Coefficients and significance:

Variable	**Coefficient**	**P Value**
Intercept	1.07	$<2x10^{-16}$
Slope	4x10-3	$<2x10^{-16}$
Slope^2	4.5x10-4	$<2x10^{-16}$

R Squared: 0.97

Figure 89: Quadratic model characteristics of on-road navigation speed.

Model: Speed=.765-.001Slope-.00035Slope² - Off-Road Movement Residual standard error: 0.04 Model Coefficients and significance:		
Variable	**Coefficient**	**P Value**
Intercept	.765	$<2x10^{-16}$
Slope	1x10-3	$<6x10^{-11}$
$Slope^2$	3.5x10-4	$<2x10^{-16}$
R Squared: 0.94		

Figure 90: Quadratic model characteristics of off-road navigation speed.

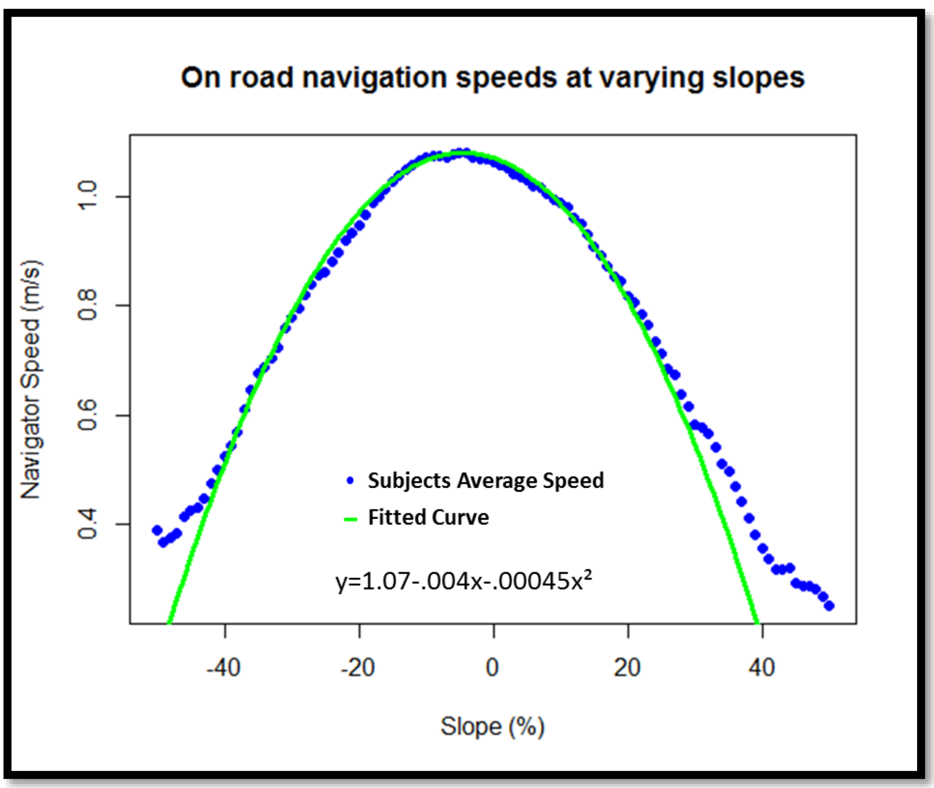

Figure 91: Graphical depiction of model representing on-road navigation speed.

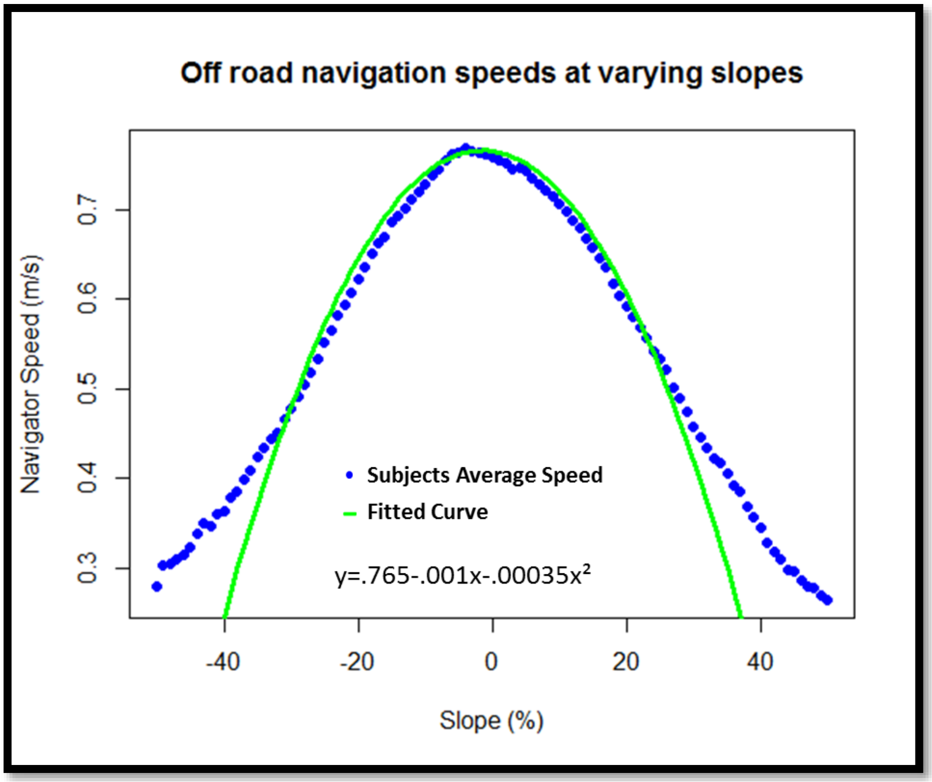

Figure 92: Graphical depiction of model representing off-road navigation speed.

The model of navigation speed derived by this research was assessed by comparing the estimated completion times of the subject using the new model with the actual completion times. The mean percent error between estimated and actual value was less than 10 percent. Graphical depictions of the results are shown in Figure 93. This plot shows the model deviation from the actual values. If the modeled result exactly matched the predicted result it would fall on the superimposed line.

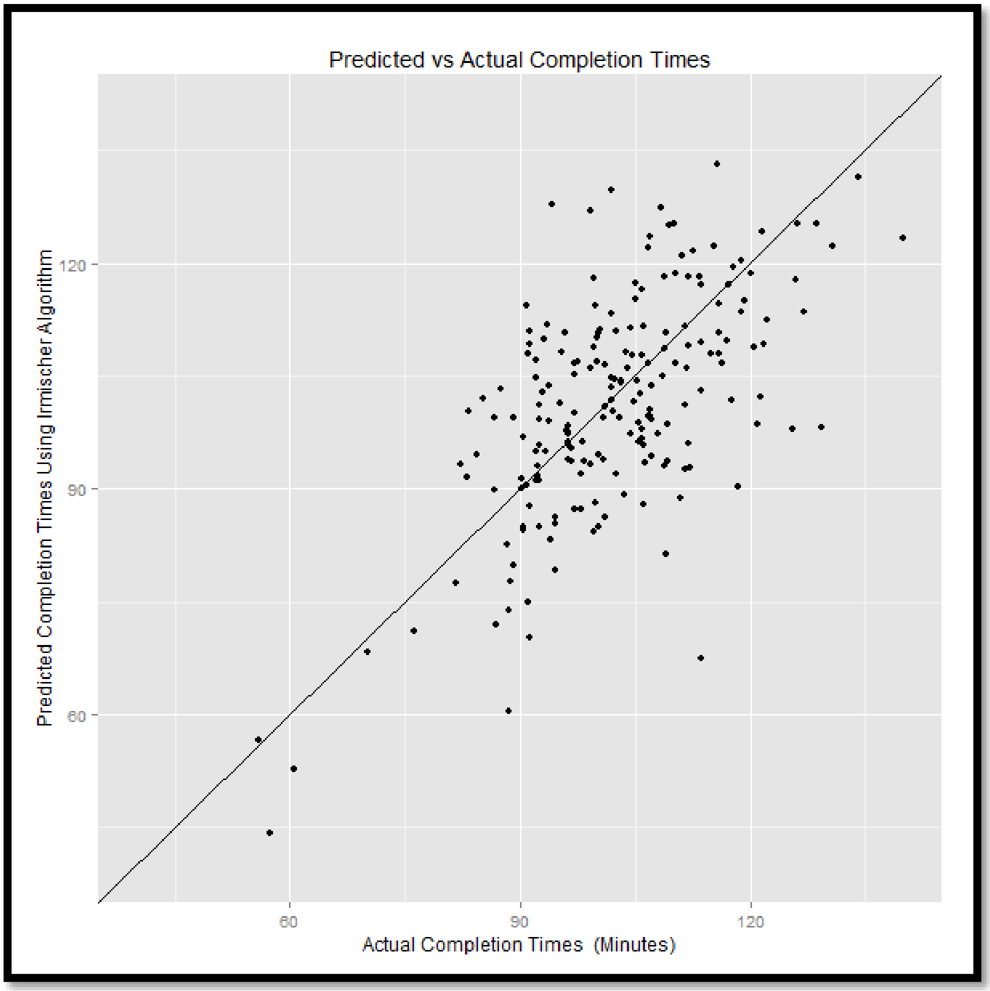

Figure 93: Comparison of model predicted completion times and actual completion times.

Research Question 4 posed in this dissertation asks "How does arduous wayfinding affect human movement speed when navigating in hilly wooded terrain?" Modeling speed of navigation partially answers this question. From Montello's definition of navigation, the cost of wayfinding during navigation can be found by subtracting navigational speed from locomotion speed (Montello 2005). Thus, a comparison of navigation speed against locomotion speed is required to fully answer this research question. Figures 94 and 95 compare the model of navigation speed against the speed of locomotion defined by Tobler's Hiking Function. The shaded region represents the difference between the two models, explaining the cognitive cost of wayfinding upon movement speed during navigation.

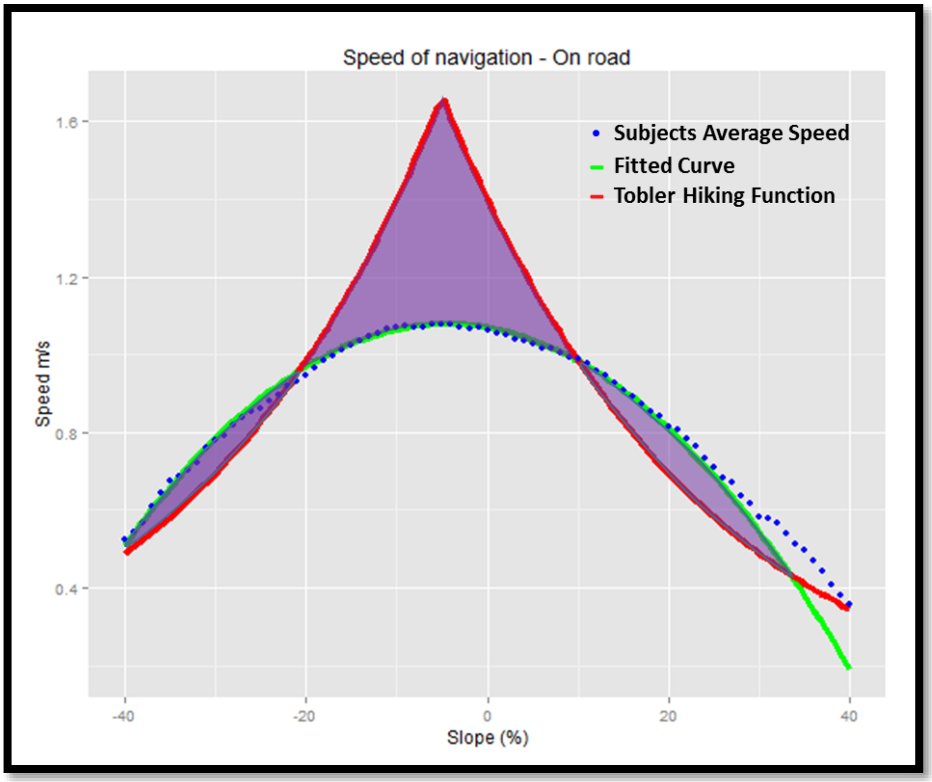

Figure 94: Speed of on-road navigation.

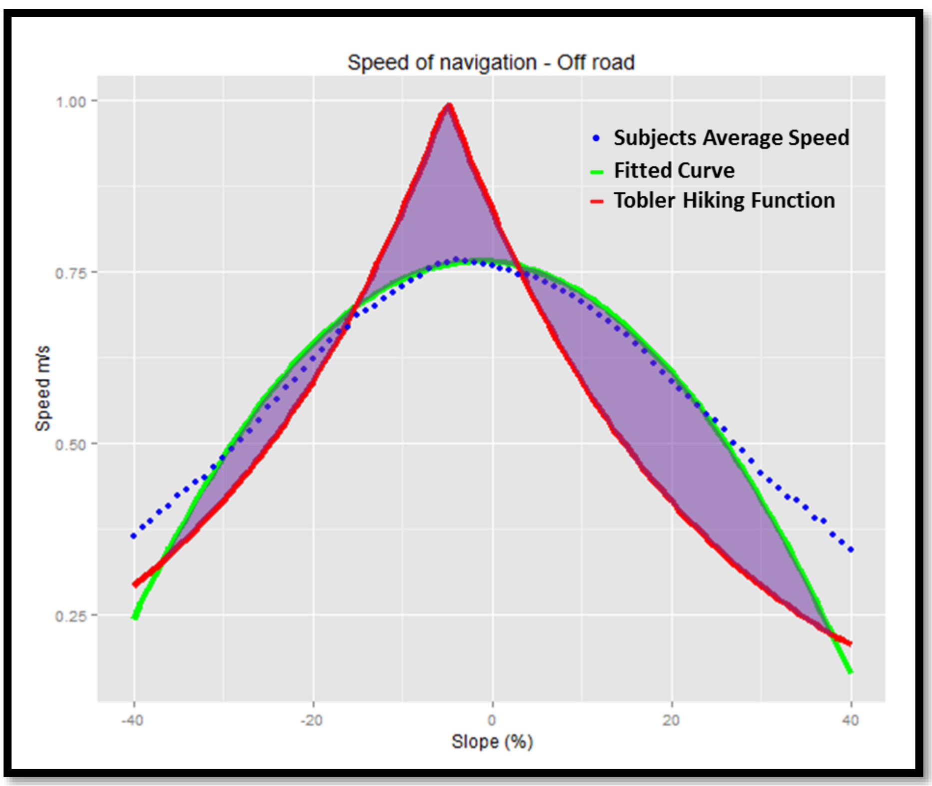

Figure 95: Speed of off-road navigation.

Summary of Results

This chapter has presented the results of the dissertation research. It has listed the products that were derived after applying the methods described in Chapter 3. First, data was collected throughout a series of observational studies. Then the data was organized and analyzed using a series of previously stated geostatistical procedures. These results help answer the four research questions initially posed.

Research Questions 1 and 2 involve modeling the energy expenditure of navigation. They specifically ask "What are the contributing weights of BMI, fitness, sex, slope, land cover, and distance traveled to energy expenditure during navigation?" The results are presented by defining coefficients for an energy expenditure model. The chapter also describes measures that assess the goodness of fit of the model (Figures 78 and 83).

The chapter also provides a quantitative assessment of GRL's energy expenditure algorithm. The outcomes of this assessment are found by comparing predicted values of the GRL model with the measured Calorie consumption of the subjects (Figure 85).

Finally, results are presented that answer the Question 4. Speed of navigation is modeled for on and off-road movement. The coefficient of determination (R-squared) for on and off-road models are 0.97 and 0.94, respectively. The diagnostics of the models are presented in Figures 89 and 90. These models are then compared to Tobler's hiking function. The difference found between the models defines the cognitive cost of wayfinding. These results are further discussed in the following chapter.

CHAPTER 6

CONCLUSIONS

Navigation is one of the fundamental tasks of our species. Similar to eating or breathing, navigation occurs repeatedly throughout the course of a day. This study has focused on the energy required to navigate and the speed at which individuals accomplish this essential undertaking.

Navigation occurs in many different environments. Whether it be indoors, on the urban street network of New York City, or in the wilderness, goal-directed movement requires human energy and is accomplished at some speed. Modeling these dynamics is essential to understanding this core function of humans.

This research focuses on woodland, wilderness navigation, executed in natural areas with forests, where few people live. Navigation is used by a variety of people for different purposes. Primitive tribes navigate to hunt and gather; hikers use navigation to visit unexplored areas or get away from society; search and rescue operators use it to find lost individuals; the military uses it during operations.

Discussion of Research Questions 1 and 2

Research Questions 1 and 2 investigate how human physiological characteristics and terrain differences affect energy expenditure during navigation. This research has proven that five of the six variables investigated are significant contributors of energy expenditure—land cover, slope, and distance traveled being the most significant contributors. BMI and sex were also significant variables. Analysis of data collected during this research has concluded that individual fitness does not contribute to energy expenditure during land navigation.

Land cover was found to be the most significant contributor to energy expenditure during navigation. The model diagnostics show this variable contributes 47 percent toward the model's power. This intuitively makes sense. If faced with a decision to bushwhack through dense bushes or climb a hill, most human beings would choose to climb the open hill.

The most energy-expensive ground to traverse was found to be rocky areas. It is 45 percent more energy consuming to walk over rocks instead of navigating on a road. It is 4 percent more intensive than traveling through vegetation and 10 percent more depleting than navigating through the open woods. In some respects this is an expected outcome. Bounding over rocks is extremely energy consuming. It requires significant movement, balance, and strength. However, it is somewhat surprising that energy expenditure over rocks is greater than energy expenditure through brush or swamps. Intuitively, some people may choose to walk over rocky areas instead of pushing through vegetation, especially if the vegetation is thick. This finding requires further discussion.

There are several possible explanations for the model concluding rocky areas as the most energy expensive. The most reasonable explanation is that movement over rocks requires the most muscular stimulation. Movement from rock to rock requires significant vertical and lateral movement and requires more movement than just horizontal movement in the direction of travel. This makes the rocky land cover class different from others tested. Also, rocky areas require more balance (i.e., muscular stimulation) than the other classes. It requires a separate contribution to total energy expenditure that is not as prevalent as when traveling over other terrain types.

It could be that subjects followed paths through vegetated areas. This delineation would not have been parsed during modeling. Only major roads were considered part of the road classification. Minor, intermittent paths through vegetated areas were not depicted on the orienteering maps used for classification. They were also difficult to find on 1-meter resolution satellite imagery, such that these paths were simply classified as vegetation. If it were the case that subjects were following paths through vegetated areas then energy expenditure per meter would be similar to moving through open woods. It would lower the model coefficient. This lowering could cause movement through vegetation to be incorrectly represented in the model.

Further, it was difficult to classify vegetation of varying levels. In the end, light, moderate, and heavy vegetation were consolidated with swampy areas into a single class called vegetation. The original classification shown in Figure 55 was called into question during data analysis. Figure 74 shows the original data plots of the energy expenditure of different classes. The data shown in Figure 74 for the four vegetation classes is concerning. It shows that energy expenditure measurements in different vegetation classes were erratic. Movement through swampy areas is shown to be less energy expensive than movement through light vegetation. It also shows little trend for moderate and heavy vegetation. In some cases, movement through heavy vegetation required less Calorie consumption than movement through light vegetation. This is a peculiar outcome. These results required reinspection of the classification and ultimately, consolidation of all vegetation classes. This generalization of the vegetation classification is a likely reason for the result indicating rocky travel to be the most energy consuming.

Slope was the second most powerful predictor of energy expenditure during navigation. Twenty percent of the model power can be attributed to slope variations when navigating. Moving uphill and downhill at extreme slopes requires more energy than navigating over gently sloped terrain. It is slightly more consuming to navigate uphill than downhill at corresponding slopes. Figure 73 shows the parabolic nature of the energy required to navigate at varying slopes. The least Calories are burned during navigation when moving slightly downhill. The figure shows the vertex of the upward opening parabolic curve is at a slope of -3 percent.

Subject BMI and distance from the start were equally important to the modeling effort. Both contributed 14 percent to the model. BMI is important to the model because it represents an assessment of the human physiology of the navigator. Stocky navigators generally burn more Calories than lanky individuals. Weight also contributes to this metric; it costs muscular energy to move weight. Hence, heavier individuals of the same height burn more Calories during locomotion.

The distance that a navigator has already traversed also contributes to the model. Interestingly, however, as a navigator moves away from the start point, his energy expenditure per meter decreases. This is shown in Figure 82. This is an intriguing result that was not anticipated; there are two possible explanations. These subjects are in their first navigational training event as cadets at USMA; it may be that they are learning efficient ways to move in the woods throughout the event.

A second explanation of this result involves the cognitive nature of modeling navigation. The most likely rationale of why energy expenditure per meter decreases throughout the event is interesting. It is probable that navigators make decisions that impact this variable. As the navigator gets further into the exercise, they choose routes that require less energy and take less energy-expensive risks. Initially they may choose to climb directly over hills rather than traversing level terrain around a hilltop or decide to fight their way over rocky terrain rather than taking a road. Yet as the event progresses, they intuitively make choices that would burn less energy per meter. They may decide to take a road although this may not be the most direct route. This finding has significant impacts to an overall model of navigational cognition, providing a valuable topic for future work.

The last significant variable in the model of energy expenditure during navigation is the individual's sex. Males expend 11 percent more energy per meter than females. This variable contributes the least to the overall model of energy expenditure of navigation—only 5 percent of the model's power.

The variable of subject fitness was found to be insignificant to modeling energy expenditure. The model was not improved by the consideration of the subject's Army Physical Fitness Score (APFT) or the subject's APFT run score. Results described in Figures 76, 77, and 78 show no model improvement using either of these variables. Most likely, this is due to the fact that the course is relatively short or that cadets represent a more fit segment of the overall population. Additionally, the design of the courses intentionally opted against creating a grueling event to test fitness. The engineering of the courses was done in a manner such that cadets could finish the course in 2.5 hours even walking at a slow pace. The straight-line distance of the routes required to finish the course is only 3 km. Most cadets walked 4–7 km to complete the exercise. Thus, the fitness of the cadets was not stressed during the event.

The other likely reason that the subject's fitness did not impact the model was that most cadets are in good shape. Fitness is an admissions consideration at USMA and applicants with poor physical fitness are not likely to be admitted. The variance in the subject's overall fitness is therefore limited. High overall fitness of the cadets also works conjointly with the first reason stated above; the navigational test simply did not stress the fitness of this group of subjects.

Research Questions 1 and 2 were asked in order to develop a model for energy expenditure during navigation. This research defined model coefficients, significance, and relative importance for each variable: land cover, slope, distance from start, BMI, sex, and subject fitness. Diagnostics of the model can be found in Figure 83. The answer to the research questions can be summarized by applying those diagnostics to the following equation:

$$\ln(EE)=Tc + 0.00027*G^2 - 0.053*D + S + 0.0035*B - 2.85$$

Where:

Variable	Description	Units	Notes
EE	Energy Expenditure per meter	Cal/meter	
Tc	Terrain Coefficient	N/A	0 if Road .31 if open forest .39 if vegetation .45 if rocky
G	Grade in percent (rise/run)*100	%	Applied -50% to 50% only
D	Distance from Start Point	km	Distance in km
S	Sex	N/A	0 if Female .11 if Male
B	Body Mass Index	kg/m^2	

Figure 96: Explanation of model variables.

This model captures 74 percent of the variance of total route energy expenditure for the 200 subjects tested. It predicts the actual Calories consumed during the navigational test with a Mean Average Error of 66 Calories. However, the power of this model is lowered when analyzing this equation over a segment-by-segment resolution. It accounts for only 23 percent of the energy expenditure variance when analyzing each segment of the trajectory separately. The low R squared at this resolution is likely due to limitations in the realm of land cover classification and the limitations of the data collection devices. Classifying the entire $9km^2$ area was challenging. Additionally, the caloric consumption information derived from the accelerometer and the locations/elevation from the GNSS provided noisy data for modeling. These limitations are explored further.

Effects of Poor Land Cover Classification

The principal factor limiting the modeling effort was the inability to accurately assess the land cover of each individual data segment. The model diagnostics show that land cover is the greatest contributor to an energy expenditure model. It follows that an accurate assessment of the terrain type over which a subject traverses is imperative.

Assumptions were made to help ensure the maximum number of segments were either correctly classified or not included in the analysis. If a segment was not completely contained within a single class it was excluded from the analysis. This procedure removed some uncertainty from the classification process. Additionally, to help account for GNSS errors when subjects were walking on roads, the roads were assumed to be slightly larger than depicted on the orienteering map and the imagery (as per Chapter 4).

Undoubtedly, classification errors affected the model fit even with these assumptions. Paths through the vegetation would provide difficulty for our model; a small group of clustered rocks in an otherwise open forest may not be identified correctly by our classification, and hence, misclassify the movement. A 4m x 4m natural spring, which may not have been captured in classification, would cause the ground in that area to be wet, which in turn, would make movement more difficult than in other areas. Fundamentally, classification required for a 10-second movement resolution is difficult and would probably require direct observation of each subject and ground-truth classification of the land cover. This is clearly beyond the scope of this research. Nevertheless, classification errors partially explain why the R squared for the model at the scale of a 10-second time resolution is low.

Effects of Error in the Measurement Devices

The energy expenditure estimates made from the Actigraph GT3X and the Crouter Refined 2 Regression Model also contribute to the low model fit at the segment-by-segment resolution. In laboratory conditions, estimation produced errors of ~25 percent (Crouter et al. 2010). It is likely that field-based use, such as this study, had even more error when estimating caloric consumption per meter. There are no known studies to reference that test this caloric estimation technique in such a rigorous environment and over a 6-km course.

Similarly, the location data that was used for the computation of distance traveled, speed, and elevation change had inherent error. GNSS devices are known to have associated measurement error. The reported x and y locations are likely to be off by 4–6 meters in dense vegetation. The elevation error is reported to be almost double that (Rodríguez-Pérez, Álvarez, and Sanz-Ablanedo 2007).

The horizontal error associated with the GNSS affects the analysis by introducing inaccuracy into several of the model variables. The distance between the end points of data segments is determined by the GNSS location readings. Further, identifying the type of land cover over which a subject is traveling depends on location data.

The distance between the data points is an important measurement that affects the modeling analysis. It is used to compute the dependent variable, the energy expenditure per meter. The Actigraph accelerometer provides energy expenditure values for each segment of the trajectory. These values are in Calories per segment. Those measurements require conversion into a per meter form. The distance between segment end points is used as the denominator for that computation.

The GNSS error is expected to affect distance measurement computations by approximately 10 percent. The mean segment length is 9 meters and the predicted error in the distance computation is about 0.9 meter. This estimate comes from reviewing the error budget of the GNSS. The receiver noise and multipath error are what primarily contributes to the distance errors. Although the entire GNSS horizontal error is 4–6 meters, most of this is systematic error, and does not contribute to distance computation error. The systematic error from satellite clocks, orbit errors, and the atmosphere shift each GNSS reading in the same manner. Thus, the distance measurement is not affected by this form of error. Even so, the 1-meter error from the receiver noise and multipath readings would significantly affect the model power and lower the goodness of fit.

The GNSS error also negatively impacts the land cover classification. Each trajectory segment is derived from the location data from the GNSS. Error from the Qstarz device could cause entire segments of data to be misclassified. While this applies to all data classification, it is especially true near roads where the classification area is narrow. Misclassification of data segments due to location error is a likely explanation for the low goodness of fit metrics of the model at micro resolutions. This is a contributing factor to why it is difficult to tune the model at the individual data segment resolution.

The errors associated with elevation values were an even greater problem than these horizontal errors. The slope of each route segment was computed directly from the GNSS elevation measurements. Initially, a high-resolution Digital Elevation Model (DEM) created from LiDAR was proposed to extract elevation values at recorded locations. This would avoid using Z values from the GNSS since it is known that Z value measurement by GNSS is less accurate than positional measurements (Brimicombe and Li 2009). Unfortunately, the available DEM was found to be laden with errors and elevation anomalies. The only practical method for obtaining elevation measurements associated with the location data was from the GNSS. The slope measurements used on the 10-second temporal resolution were a large source of error and one of the root causes for a low coefficient of determination at the segment level resolution.

Even with these limitations, this research has made significant findings. Generalization of the model provides optimism for our understanding of energy expenditure during navigation despite the low R squared observed at the segment level resolution. The horizontal and vertical GNSS error cause significant noise in the distance and the slope data. However, the substantial number of data records mitigates the noisy data when modeling over the entire trajectory. Generalization of the model over the entire trajectory can improve the model's performance and power. The correlation coefficient of the trajectory level results predicted by the EE equation compared to the measured EE is 0.86. It predicts total route Calorie consumption with an expected error of only 11 percent.

The conclusions that can be derived through inspection of Research Questions 1 and 2 are the following: (1) Land cover, slope, distance traveled, BMI, and sex are all significant factors in an energy expenditure model of navigation, with land cover being the most important consideration for determining the quantity of Calories consumed during navigation. (2) Limitations of the classification methods and the measurement equipment were major factors constraining model fit. (3) The model has trouble fitting to data at a micro resolution. The coefficient of determination of the energy expenditure equation is 0.23 when considering each individual segment as separate data. (4) However, the model is valuable for predicting route-based energy expenditure.

Discussion of Research Question 3

Research Question 3 investigates a U.S. Army energy expenditure model that had yet to be validated with empirical data while subjects were performing navigation tasks. This model was assessed by comparing the model estimates to actual energy expenditure measurements. First, the model was used to predict the total energy expenditure along each subject's trajectory. Then, this route-based energy expenditure value was compared to the measured Calorie consumption for the entire route. This methodology found that there was a correlation coefficient of 0.66 between the predicted and actual energy expended. The mean absolute error between the value predicted by the GRL algorithm and the empirical data was found to be 83 Calories. This value represents an average of 14 percent difference between the predicted Calorie consumption and the actual consumption during the 2.5-hour navigation course.

Several other findings were discovered while assessing this question. (1) There are some trends in the error. Figure 85 shows the difference between the predicted values and the actual energy expended. Each dot represents one of the 200 subjects. The GRL algorithm over predicts subjects that expended less than 500 Calories. Thirty-three subjects consumed less than 500 Calories. GRL’s algorithm over predicted these values in 31 of 33 instances. Conversely, for 34 of the 35 subjects who consumed more than 750 calories, the model proves to under predict. A potential area for future work is the investigation of these systematic differences between GRL’s algorithm and the empirical data.

Analytical assessment of the Irmischer EE algorithm shows the model developed in this research (outcome of RQ 1&2) outperforms the GRL algorithm when estimating Calorie consumption over a 4–8km route. The same methodology that was used to assess GRL’s algorithm was used to analyze the Irmischer model. The mean absolute error for the Irmischer model is 66 Calories. The average error is just 11percent of the actual Calorie consumption. This is a 3 percent improvement in Calorie estimation over the GRL model. Remarkably, the correlation coefficient between the predicted and empirical data is 0.86 for the Irmischer model. This explains 30 percent more variance than the GRL model. Figure 86 depicts this improved relationship. Similar to the GRL algorithm, routes with low energy expenditure are over predicted by the model. However, the Irmischer model estimates routes with higher energy consumption more evenly. There is no systematic error for values greater than 750 Calories.

There are several cautionary notes that must be appended to these conclusions. All the limitations pertaining to Research Questions 1 and 2 apply to Question 3. Measurement devices and classification introduce error into the measured and computed values. Speed and distance are required model variables and both have inherent error. This must be considered when making conclusions based on these results. The error inherent to the Actigraph GT3X Calorie consumption values is the most significant drawback in the assessment of the results for Question 3, which compares a model of energy expenditure to measured Calorie consumption that contains error. Reducing error in the measured Calorie consumption values will improve our findings and decrease uncertainty in our conclusions. Future work should consider using an indirect calorimeter for the empirical data collection to diminish this limitation.

Some data was omitted from analysis because of differences in land cover classes between the empirical data and GRL's terrain coefficients. GRL's algorithm is based on Pandolf's research. Only on-road, open woods, and vegetation were used in the assessment. GRL's algorithm has no terrain factor for rocky terrain, hence, segments that were classified as rocky were omitted from the analysis. Swampy areas were also omitted based on the uncertainty of the vegetation class. Finally, open woods were classified as paths in the GRL algorithm, as it has no "open woods" category. This research assumed that open woods navigation and path navigation would have similar energy expenditure.

The results describing the Irmischer model should be cautiously interpreted. The error reporting and the correlation values were computed using the same data that created the model. Therefore, cross validation of the model should be conducted to validate the results. The model should be tested against an independent dataset. This would describe how the model will generalize across the population. Unfortunately, collection of additional trajectories is impossible at this time. A list of future work includes testing the Irmischer model against new routes traversed by different subjects.

This assessment of GRL's algorithm provides the first validation study of the algorithm that uses subjects conducting navigation training in the field as the empirical data. The results indicate that there is room for model improvement if not replacement. These conclusions are of importance since this model is currently operationalized and in use as a planning tool. The algorithm produced from this research (Research Question 1 and 2) betters the Mean Absolute Error during prediction of Calorie consumption by 17 Calories per route. The new model explains 30 percent more variance. These findings are promising and use of the new model should be considered. However, the research is not without limitations. The results must be cross validated before hastily replacing the currently used algorithm. Better energy consumption measurements should be used for final assessment. Still, this research has identified potential areas of improvement that can drastically advance route planning for search and rescue, hikers, firefighters, and the military.

Discussion of Research Question 4

The final research question investigated the relationship between speed and slope during dismounted navigation. Movement speeds during navigation are a function of two fundamental processes: wayfinding and locomotion. The subjects used a map and a compass to maneuver their way across a wooded, hilly landscape from one control point to another. Wayfinding represents the constant cognitive processes involved with making decisions while navigating. It involves assessing one's current location by relating the map to the environment. Additionally, it considers time spent making routing decisions. Locomotion represents the physical muscular movement required for navigation. In this case it involves muscular contraction used to walk in the woods. Models of human locomotion while hiking in the woods exist but none of these models have addressed the cognitive aspect of difficult wayfinding. This research developed a model for human navigation that includes both wayfinding and locomotion. It then used previously defined locomotion equations to quantify the cognitive cost of navigation.

The navigation was first classified into on and off-road movement. Subsequently, movement rates of the 200 subjects were analyzed at varying slopes to create models of navigational speed. Separate models were created for on-road and off-road movement rates.

The model fitting process weighted slopes with the most data the heaviest. The subjects traversed terrain at level and moderate slopes more frequently than navigating extreme slopes. The majority of the movement occurred at slopes between ±30 percent. Ninety percent of the movement occurred at slopes within these bounds. Our modeling focused on this range.

Movement speed during on-road navigation was found to follow a parabolic relationship over the constrained ranges discussed above. Figure 91 shows the curve fitted to the data. Movement at a slight downhill slope (approximately 4.2 percent) was the fastest on roads. This is close to Tobler's estimate of -5 percent slope. The speed traveled at this grade during navigation is 1.08 m/s, which is 34 percent slower than Tobler's function that estimates the fastest on-road hiking speed to be 1.67 m/s.

The model for off-road navigation between ± 30 percent was also found to be parabolic. Figure 92 shows the navigation data and the resulting model, with the fastest off-road navigation speed being 0.77m/s at -4 percent slope. The parabolic model best fits the data with a vertex at -2 percent. Tobler estimates the maximum hiking speed for off-road movement to be 1 m/s.

The models of movement rates created in this research contribute to understanding the dynamics of navigation and the ability to predict and model the speed of navigation has widespread use. These models of navigation speeds can be used to help wilderness recreation aficionados plan how far they can travel in a day along specified routes. Archeologists can use these models to predict time-space computations of tribal travel. Back country search and rescue teams can use the equations to estimate ranges of lost persons. The military will undoubtedly benefit by using these models to plan missions that require overland navigation.

The equations created in this research have several peculiarities. Most human movement equations, like Tobler's, follow an exponential curve. The model proposed in this research used a quadratic equation to best fit the data. The use of the parabolic function is only useful between ± 30 percent slope. Grades greater than ±30 percent have a steep drop-off in speed that is not representative of human mobility. Further research and experimentation is needed to devise an equation that would represent movement at higher slopes. Inspection of the limited data available hints toward an exponential equation at higher slopes, but more data is needed to make any significant conclusions.

A second oddity in the data is apparent when comparing the subject's navigational speed to predictions made by Tobler's hiking function. It is expected that hiking speeds are equal to or faster than navigational speed. The hiking function primarily models locomotion and disregards most of the cognitive cost of wayfinding. The friction involved with cognition reduces movement speeds when navigating compared to hiking. Figure 87 shows the on-road navigation speeds in blue. Red are the predicted speeds estimated by Tobler's hiking function. This expectation is realized for areas of level ground. However, slopes less than -25 percent show that our subjects navigate faster than the prediction made by Tobler's hiking function. The same is true for slopes greater than 15 percent. This is unexpected and will be investigated during future research.

The cost of cognition attributed to wayfinding can be derived from the results of this research as well. Tobler's hiking function defines the speed of locomotion at differing slopes. The difference between Tobler's function and the model created in this research can be quantified. The 34 percent difference in maximum speed can be attributed to the cognition involved in wayfinding. Acts such as map reading, analyzing the terrain, decision making, assessing one's current position, and determining routes have costs associated with them. These costs are defined as the cost of cognition during navigation.

Figures 94 and 95 show the on and off-road differences between Tobler's hiking function and the model of navigation speed resulting from this research. The shaded region between the curves represents this difference. These figures show a significant difference between locomotion and navigation on gentle slopes. The cognitive cost of navigation is higher on level ground. Could it be that humans are more likely to engage in wayfinding activities at low grades? It is evident that there is less difference between hiking and the navigation model at higher grades. However, results must be considered inconclusive at this time since the navigational speeds exceed Tobler's predicted locomotion speeds. This finding is inconsistent with theoretical expectations and further research is required to definitively resolve these conclusions.

Nonetheless, investigation of this research question has successfully developed a model of navigation speed; at least when considering slopes within the range of ±30 percent slope. The model developed in this research predicts navigation times with better than 90% accuracy when route completion estimates are compared to actual data. Although the limitations intrinsic in the measurement devices and classification difficulties persist for this study, the vastness of the number of observations mitigates the limitations. Finally, this exploration of navigation, locomotion, and wayfinding has developed a methodology and framework to define the cognitive cost of navigation. This research has set the groundwork and structure to study the cost of cognition in the future.

Future Work

The work presented in this dissertation has generated as many new questions as it has answered. There are significant areas of future work that can be of great contribution to understanding the fundamentals of navigation. These will be discussed below.

Future work must address some of the limitations evident from this study. Improvements in the data used to model energy expenditure and speed are possible. For instance, better classification of the terrain, vis-à-vis on-ground observation of land cover, would improve the quality of the modeling effort. While the human capital required for this reconnaissance effort is large, the results would be worthwhile. Additionally, future collections of trajectory data could be gathered in future years with better GNSS devices. Devices with the capability to differentially correct location readings would improve the data quality and usefulness. Further, investment for indirect calorimeters would make a significant difference, eliminating almost all uncertainty inherent in the measured energy expenditure.

Improving the measurement devices would significantly improve the modeling effort and offers great potential for future work.

This research found that humans expend less energy per meter the further they move from the start point. I postulate that fatigue affects the cognitive processes. It seems likely that navigators choose slopes and land covers that minimize energy expenditure more often when stressed by fatigue. This area requires further investigation but holds potential for improving cognitive models of navigation.

This research made several general statements about the cognitive cost of navigation. It compared the navigational speeds of the subjects to the predicted speeds suggested by Tobler's hiking function. This work attempted to define the cost of wayfinding by comparing the cost of hiking to the cost of navigation. Comparing the cost of locomotion (hiking) to the cost of navigation provides a methodology to assess the cost of cognition. The costs investigated could be speed, energy expenditure, or any human dynamic. While this framework holds great promise, it requires additional experimentation.

One of the limitations of this study was the grouped nature of the navigators. It is possible that a subject's energy expenditure and speed of navigation was affected by the safety partner who was required to accompany each subject. Future studies might consider comparing these results to an exercise where subjects navigate alone through the woods, removing any uncertainty about the subject's speed and energy expenditure being influenced by partnered personnel.

Testing subjects in a different geographical area would add to the power of this study. The training area used for the navigational test had several significant terrain features that aided the subject's location-finding ability. For example, Figure 55 shows a prominent circular road network that follows a course similar to most of the navigational courses assigned to the subjects. This road could be consistently used by the subjects to reassess their location. Additionally, the study area contained two large, discernable, water features that could be used to regain orientation in the event of disorientation. Finding a study area that was void of these major terrain features would add rigor to the navigational exercise. The results obtained from observing subjects in more difficult terrain would certainly add breadth to the conclusions.

The use of the summer training environment at USMA has several advantages that should be highlighted in the context of future work. (1) The same subjects will be participating in another test this coming summer. It is possible to observe the same subjects in different conditions after a year of navigational training. Assessing the year-to-year differences in speed and energy expenditure would also be interesting. (2) The model of speed vs slope presented in this dissertation could not be generalized at slopes more extreme than ±30 percent. Future data collection should increase the number of observations at the more extreme slopes. These two areas of future work are relatively easy follow-up studies for the present work.

This research on the movement dynamics of navigation can also be combined with a sister study involving navigation proficiency and sense of direction. This research focused on the movement dynamics of navigation. In a parallel study involving the same subjects, data has been collected on the subjects' self-reported sense of direction and navigational proficiency. Combining these data holds great promise for the study of movement dynamics. Investigation of how speed and energy expenditure differs based on proficiency is exciting. Researching relationships between self-reported sense of direction and energy consumption decisions is possible. Measuring the correlations between sense of direction, proficiency and how often people stop to look at the map or their surroundings is entirely possible with the dataset available here.

Finally, there is a long list of additional research themes that can be accomplished by simply using the dataset created from this dissertation. Examining where people stop most often can lead to answering cognitive questions about decision making. Research of the characteristics of good and bad navigators is possible with the trajectories created here. Looking at navigational movement around control points can lead to identifying search patterns most common to human behavior. The opportunities for further analysis of the data amassed by this project are outstanding. They provide an excellent point of departure for improving our understanding of human land navigation.

Summary

The research presented here offers significant contributions to the fields of GIS and to the navigation modeling community. In the realm of GIS, this research has outlined methods for studying navigation using technology and GNSS devices. It has applied a big data solution to modeling navigation and studying human movement. Computer code and statistical tools have been developed to handle large quantities of GNSS-based tracking information, and it has developed methods for integrating biosensor and GNSS data for visualization and spatial analysis.

More specific contributions include development of a new algorithm to predict energy consumption of human beings while they navigate. The contributions of land cover, slope, distance from origin, BMI and sex have been defined. This research has tested a currently used algorithm and assessed its effectiveness. A model of navigation speed in hilly wooded terrain has been defined. Finally, this investigation has devised a framework for defining the cognitive cost of navigation.

These contributions provide a waypoint along a continued journey to understand the human dynamics of navigation. This dissertation is not a conclusion to humanity's intrigue with the complexity of navigation but rather a stepping stone to future discovery. Deeper investigation into speed and energy costs inherent to woodland navigation offers a rich topic of inquiry. Improving the scientific understanding of decision making in the wilderness remains compelling. The future work listed in this chapter provides a research agenda for a decade of investigation. This dissertation will undoubtedly propel scholarly acumen and investigation to new levels in the years to come.

REFERENCES

Actigraph. “Activity Monitors | Product Categories | ActiGraph.” Accessed December 30, 2014. http://www.actigraphcorp.com/product-category/activity-monitors/.

Actigraph. “What Is the Difference among the Energy Expenditure Algorithms? : ActiGraph Support.” Accessed October 22, 2014. https://help.theactigraph.com/entries/20744123-What-is-the-difference-among-the-Energy-Expenditure-Algorithms-.

Actigraph. “What Is the Difference among the MET Algorithms?” *ActiGraph Support*. Accessed January 2, 2015. https://help.theactigraph.com/entries/21383846-What-is-the-difference-among-the-MET-Algorithms-.

Actigraph Software Department. “ActiLife 6 Users Manual.” ActiGraph Software Department, April 3, 2012.

Ainslie, Philip N., Thomas Reilly, and Klass R. Westerterp. “Estimating Human Energy Expenditure.” *Sports Medicine* 33, no. 9 (August 1, 2003): 683–698. doi:10.2165/00007256-200333090-00004.

Aitken, Robert. “Wilderness Areas in Scotland.” Ph.D. Dissertation, University of Aberdeen (United Kingdom), 1977.

Allen, Gary L. “Spatial Abilities, Cognitive Maps, and Wayfinding.” *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, 1999, 46–80.

Andrienko, Gennady, Natalia Andrienko, Urska Demsar, Doris Dransch, Jason Dykes, Sara Irina Fabrikant, Mikael Jern, Menno-Jan Kraak, Heidrun Schumann, and Christian Tominski. “Space, Time and Visual Analytics.” *International Journal of Geographical Information Science* 24, no. 10 (2010): 1577–1600.

Andrienko, Natalia, and Gennady Andrienko. “Visual Analytics of Movement: An Overview of Methods, Tools and Procedures.” *Information Visualization* 12, no. 1 (January 1, 2013): 3–24. doi:10.1177/1473871612457601.

Berry, Joseph K. *Map Analysis: Understanding Spatial Patterns and Relationships*. GeoTec Media, 2007.

Box, G. E. P., and D. R. Cox. “An Analysis of Transformations.” *Journal of the Royal Statistical Society. Series B (Methodological)* 26, no. 2 (1964): 211–52.

Brandes, Mirko, Vincent T. Van Hees, Viviane HannöVer, and Soren Brage. “Estimating Energy Expenditure from Raw Accelerometry in Three Types of Locomotion:” *Medicine & Science in Sports & Exercise* 44, no. 11 (November 2012): 2235–2242. doi:10.1249/MSS.0b013e318260402b.

Brimicombe, Allan, and Chao Li. *Location-Based Services and Geo-Information Engineering*. Vol. 21. John Wiley & Sons, 2009.

Chen, Kong Y., and Ming Sun. “Improving Energy Expenditure Estimation by Using a Triaxial Accelerometer.” *Journal of Applied Physiology* 83, no. 6 (1997): 2112–22.

Clarke, Donald D., and Louis Sokoloff. “Circulation and Energy Metabolism of the Brain.” In *Basic Neurochemistry: Molecular, Cellular and Medical Aspects*, edited by George R. Siegel, R. Wayne Albers, Scott T. Brady, and Donald L. Price, 6th Edition., 6:637–669. Philadelphia, PA: Lippincott Raven, 1999.

Clarke, Keith C. *Analytical and Computer Cartography*. Prentice Hall PTR, 1995.

Clarke, Keith C. *Getting Started with Geographic Information Systems*. 5 edition. Upper Saddle River, N.J: Prentice Hall, 2010.

Crouter, Scott E., James R. Churilla, and David R. Bassett Jr. "Estimating Energy Expenditure Using Accelerometers." *European Journal of Applied Physiology* 98, no. 6 (2006): 601–612.

Crouter, Scott E., Kurt G. Clowers, and David R. Bassett Jr. "A Novel Method for Using Accelerometer Data to Predict Energy Expenditure." *Journal of Applied Physiology* 100, no. 4 (2006): 1324–1331.

Crouter, Scott E., Erin Kuffel, Jere D. Haas, Edward A. Frongillo, and David R. Bassett. "A Refined 2-Regression Model for the ActiGraph Accelerometer." *Medicine and Science in Sports and Exercise* 42, no. 5 (May 2010): 1029–1037. doi:10.1249/MSS.0b013e3181c37458.

Di Gregorio, Antonio, and Louisa JM Jansen. "Land Cover Classification System (LCCS): Classification Concepts and User Manual for Software Version 1.0," 2000.

Duggan, A., and M. F. Haisman. "Prediction of the Metabolic Cost of Walking with and without Loads." *Ergonomics* 35, no. 4 (April 1, 1992): 417–426. doi:10.1080/00140139208967822.

Egenhofer, Max J., and Reginald G. Golledge. *Spatial and Temporal Reasoning in Geographic Information Systems*. Oxford University Press, 1998.

Fisher, Peter. *Re-Presenting GIS*. John Wiley & Sons, 2005.

Fleury, Michelle, and Chantal Bard. "Effects of Different Types of Physical Activity on the Performance of Perceptual Tasks in Peripheral and Central Vision and Coincident Timing." *Ergonomics* 30, no. 6 (1987): 945–958.

Ford, L. R., and Delbert Ray Fulkerson. *Flows in Networks*. Vol. 1962. Princeton Princeton University Press, 1962. http://www.maa.org/publications/maa-reviews/flows-in-networks.

Freedson, Patty S., Edward Melanson, and John Sirard. "Calibration of the Computer Science and Applications, Inc. Accelerometer:" *Medicine & Science in Sports & Exercise* 30, no. 5 (May 1998): 777–81. doi:10.1097/00005768-199805000-00021.

Freedson P, and Miller K. "Objective Monitoring of Physical Activity Using Motion Sensors and Heart Rate." *Research Quarterly for Exercise and Sport* 71, no. 2 Suppl (June 2000): S21–29.

Fox, John, and Sanford Weisberg. *An {R} Companion to Applied Regression*. Second. Thousand Oaks, CA: Sage, 2011. http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Givoni, B., and R. F. Goldman. "Predicting Metabolic Energy Cost." *Journal of Applied Physiology* 30, no. 3 (March 1, 1971): 429–433.

Hagerstraand, Torsten. "What about People in Regional Science?" *Papers in Regional Science* 24, no. 1 (1970): 7–24. doi:10.1111/J.1435-5597.1970.TB01464.X.

Hall, Cameron, Arturo Figueroa, B. O. Fernhall, and Jill A. Kanaley. "Energy Expenditure of Walking and Running: Comparison with Prediction Equations." *Medicine and Science in Sports and Exercise* 36 (2004): 2128–2134.

Hegarty, Mary, Anthony E. Richardson, Daniel R. Montello, Kristin Lovelace, and Ilavanil Subbiah. "Development of a Self-Report Measure of Environmental Spatial Ability." *Intelligence* 30, no. 5 (2002): 425–447.

Humphrey, Reed. “Clinical Applications: The Exercise Caloric Challenge.” *ACSM’s Health & Fitness Journal* 10, no. 2 (2006): 40–41.
Jensen, John R. *Introductory Digital Image Processing 3rd Edition*. Upper Saddle River: Prentice Hall, 2005.
Jiang, Bin, and Xiaobai Yao. “Location-Based Services and GIS in Perspective.” *Computers, Environment and Urban Systems* 30, no. 6 (2006): 712–725.
John, Dinesh, and Patty Freedson. “ActiGraph and Actical Physical Activity Monitors: A Peek under the Hood.” *Medicine and Science in Sports and Exercise* 44, no. 1 Suppl 1 (January 2012): S86–89. doi:10.1249/MSS.0b013e3182399f5e.
Kennedy, Michael. *The Global Positioning System and ArcGIS, Third Edition*. CRC Press, 2009.
Keytel, L. R., J. H. Goedecke, T. D. Noakes, H. Hiiloskorpi, Raija Laukkanen, L. Van Der Merwe, and E. V. Lambert. “Prediction of Energy Expenditure from Heart Rate Monitoring during Submaximal Exercise.” *Journal of Sports Sciences* 23, no. 3 (2005): 289–297.
Kramer, Patricia Ann, and Adam D. Sylvester. “The Energetic Cost of Walking: A Comparison of Predictive Methods.” *PLoS ONE* 6, no. 6 (June 22, 2011): e21290. doi:10.1371/journal.pone.0021290.
Kuffel, Erin E., Scott E. Crouter, Jere D. Haas, Edward A. Frongillo, and David R. Bassett Jr. “Validity of Estimating Minute-by-Minute Energy Expenditure of Continuous Walking Bouts by Accelerometry.” *Int J Behav Nutr Phys Act* 8 (2011): 92.
Kwan, Mei-Po. “Space-Time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-Based Framework.” *Geographical Analysis* 30, no. 3 (1998): 191–216.
Kwan, Mei-Po. “Interactive Geovisualization of Activity-Travel Patterns Using Three-Dimensional Geographical Information Systems: A Methodological Exploration with a Large Data Set.” *Transportation Research Part C: Emerging Technologies* 8, no. 1 (2000): 185–203.
Langmuir, Eric. *Mountain Leadership: The Official Handbook of the Mountain Leadership Training Boards of Great Britain*. Edinburgh: Scottish Council of Physical Recreation, 1969.
Langmuir, Eric, Scottish Sports Council, and Mountain Leader Training Board. *Mountaincraft and Leadership: A Handbook for Mountaineers and Hillwalking Leaders in the British Isles*. Scottish Sports Council, 1984.
Langran, Gail. *Time in Geographic Information Systems*. CRC Press, 1992.
Lee, Jung-Min, Youngwon Kim, and Gregory J. Welk. “Validity of Consumer-Based Physical Activity Monitors:” *Medicine & Science in Sports & Exercise* 46, no. 9 (September 2014): 1840–48. doi:10.1249/MSS.0000000000000287.
Leyk, D., U. Rohde, W. Gorges, D. Ridder, M. Wunderlich, C. Dinklage, A. Sievert, T. Rüther, and D. Essfeld. “Physical Performance, Body Weight and BMI of Young Adults in Germany 2000–2004: Results of the Physical-Fitness-Test Study.” *International Journal of Sports Medicine* 27, no. 8 (August 2006): 642–647. doi:10.1055/s-2005-872907.
Lennie, Peter. “The Cost of Cortical Computation.” *Current Biology* 13, no. 6 (2003): 493–497.

Leonard, William R. “Laboratory and Field Methods for Measuring Human Energy Expenditure.” *American Journal of Human Biology: The Official Journal of the Human Biology Council* 24, no. 3 (June 2012): 372–84. doi:10.1002/ajhb.22260.

Levine, James A. “Measurement of Energy Expenditure.” *Public Health Nutrition* 8, no. 7a (October 2005): 1123–32. doi:10.1079/PHN2005800.

Li, Zhilin, Christopher Zhu, and Chris Gold. *Digital Terrain Modeling: Principles and Methodology*. CRC press, 2010.

McMinn, David, Rijan Acharya, David Rowe, Stuart Gray, and Julia Allan. “Measuring Activity Energy Expenditure: Accuracy of the GT3X and Actiheart Monitors.” *International Journal of Exercise Science* 6, no. 3 (July 16, 2013). http://digitalcommons.wku.edu/ijes/vol6/iss3/5.

McMurray, Robert G., and Kristin S. Ondrak. “Energy Expenditure of Athletes.” In *Sports Nutrition: Energy Metabolism and Exercise*, 127, 2008.

Miller, Harvey J. “Modelling Accessibility Using Space-Time Prism Concepts within Geographical Information Systems.” *International Journal of Geographical Information System* 5, no. 3 (1991): 287–301.

Montello, Daniel R. “Navigation.” In *The Cambridge Handbook of Visuospatial Thinking*, edited by P Shah and A Miyake, 257–94. Cambridge: Cambridge University Press, 2005.

Muehlenbein, Michael P., Jana L. Hirschtick, Julia Z. Bonner, and Ann M. Swartz. “Toward Quantifying the Usage Costs of Human Immunity: Altered Metabolic Rates and Hormone Levels during Acute Immune Activation in Men.” *American Journal of Human Biology* 22, no. 4 (July 1, 2010): 546–556. doi:10.1002/ajhb.21045.

National Research Council (U.S.), National Research Council (U.S.), National Research Council (U.S.), and National Research Council (U.S.). *New Research Directions for the National Geospatial-Intelligence Agency: Workshop Report*. Washington, D.C: The National Academies Press, 2010.

Pande, A., Y. Zeng, A. Das, P. Mohapatra, S. Miyamoto, E. Seto, E. K. Henricson, and J. J. Han. “Accurate Energy Expenditure Estimation Using Smartphone Sensors.” In *ACM Wireless Health 2013*, −, 2013.

Pandolf, K. B., B. Givoni, and R. F. Goldman. “Predicting Energy Expenditure with Loads While Standing or Walking Very Slowly.” *Journal of Applied Physiology*, 1977.

Pettee, K., Catrine Tudor-Locke, and Barbara E. Ainsworth. “The Measurement of Energy Expenditure and Physical Activity.” In *Sports Nutrition. Energy Metabolism and Exercise*, 159–189, 2008.

Peuquet, Donna 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): 441–461.

Pinheiro, Jose, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, and R Core Team. *{nlme}: Linear and Nonlinear Mixed Effects Models*, 2015. http://CRAN.R-project.org/package=nlme.

Potter, Adam W., William R. Santee, Cynthia M. Clements, Kelly A. Brooks, and Reed W. Hoyt. “Comparative Analysis of Metabolic Cost Equations: A Review.” *Journal of Sport and Human Performance* 1, no. 3 (September 18, 2013). doi:10.12922/9.

Pred, Allan. "The Choreography of Existence: Comments on Hägerstrand's Time-Geography and Its Usefulness." *Economic Geography* 53, no. 2 (April 1, 1977): 207–221. doi:10.2307/142726.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2015. http://www.R-project.org/.

Richards-Rissetto, Heather, and Kristin Landau. "Movement as a Means of Social (re)production: Using GIS to Measure Social Integration across Urban Landscapes." *Journal of Archaeological Science* 41 (January 2014): 365–375. doi:10.1016/j.jas.2013.08.006.

Robinson, Arthur Howard, Joel L. Morrison, Phillip C. Muehrcke, A. Jon Kimerling, and Stephen C. Guptill. *Elements of Cartography*. Vol. 6. Wiley New York, 1995.

Rodríguez-Pérez, J., M. Álvarez, and E. Sanz-Ablanedo. "Assessment of Low-Cost GPS Receiver Accuracy and Precision in Forest Environments." *Journal of Surveying Engineering* 133, no. 4 (2007): 159–167. doi:10.1061/(ASCE)0733-9453(2007)133:4(159).

Santee, William R., Laurie A. Blanchard, Karen L. Speckman, Julio A. Gonzalez, and Robert F. Wallace. *Load Carriage Model Development and Testing with Field Data*, June 2003.

Sauer, Carl O. "Foreword to Historical Geography." *Annals of the Association of American Geographers* 31, no. 1 (March 1, 1941): 1–24. doi:10.2307/2560961.

Shaw, Shih-Lung, and Hongbo Yu. "A GIS-Based Time-Geographic Approach of Studying Individual Activities and Interactions in a Hybrid Physical–virtual Space." *Journal of Transport Geography* 17, no. 2 (2009): 141–149.

Soule, Roger G., and Ralph F. Goldman. "Terrain Coefficients for Energy Cost Prediction." *J Appl Physiol* 32, no. 5 (1972): 706–708.

Swartz, Ann M., Scott J. Strath, David R. Bassett Jr, William L. O'Brien, George A. King, and Barbara E. Ainsworth. "Estimation of Energy Expenditure Using CSA Accelerometers at Hip and Wrist Sites." *Medicine and Science in Sports and Exercise* 32, no. 9 Suppl (2000): S450–456.

Tapia, Emmanuuel. "Using Machine Learning for Real-Time Activitiy Recognition and Estimation of Energy Expenditure." Massachusetts Institute of Technology, 2008.

Tharion, William J., Scott J. Montain, Catherine O'Brien, Ronald L. Shippee, and Jennifer L. Hoban. "Effects of Military Exercise Tasks and a Carbohydrate-Electrolyte Drink on Rifle Shooting Performance in Two Shooting Positions." *International Journal of Industrial Ergonomics* 19, no. 1 (January 1997): 31–39. doi:10.1016/0169-8141(95)00087-9.

Thomas, James J., and Kristin A. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society Press, 2005.

Thomas, James J., and Kristin A. Cook. "A Visual Analytics Agenda." *Computer Graphics and Applications, IEEE* 26, no. 1 (2006): 10–13.

Tobler, Waldo R. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46 (1970): 234–240.

Tobler, Waldo R. "Non-Isotropic Geographic Modeling." Santa Barbara, California: National Center for Geographic Information and Analysis, 1993.

"Trimble - GPS Tutorial - Why GPS?" Accessed October 15, 2014. http://www.trimble.com/gps_tutorial/.
Reynolds, Stephen C., and Richard L. Taylor. *Digital Terrain Data in Support of Land Combat Models*. DTIC Document, 1988.
Slocum, T. A., R. B. McMaster, F. C. Kessler, and H. H. Howard. *Thematic Cartography and Geographic Visualization. 2005*. Pearson, Prentice-Hall, n.d.
U.S. Army. *Army Physical Readiness Training*. Department of the Army, FM 7-22. Washington D.C.: Government Printing Office, 2012.
U.S. Army. *Foot Marches*. Department of the Army, FM 21-18. Washington D.C.: Government Printing Office, 1990.
U.S. Army. *Map Reading and Land Navigation*. Department of the Army, TC 3-25.26. Washington, D.C.: Government Printing Office, 2013.
U.S. Army. *Soldier's Manual of Common Tasks: Warrior Leader Skills Level 1*. STP 21-1-SMCT. Washington, D.C.: Government Printing Office, 2012.
Venables, W, and B Ripley. *Modern Applied Statistics with S*. 4th ed. New York, NY, USA: Springer, 2002. http://www.stats.ox.ac.uk/pub/MASS4.
Wade, Tasha, and Shelly Sommer. *A to Z GIS: An Illustrated Dictionary of Geographic Information Systems*. ESRI Press, 2006.
Welk, Gregory J., Jodee A. Schaben, and James R. Morrow Jr. "Reliability of Accelerometry-Based Activity Monitors: A Generalizability Study." *Medicine and Science in Sports and Exercise* 36, no. 9 (2004): 1637–1645.
Willems, Niels, Huub Van De Wetering, and Jarke J. Van Wijk. "Visualization of Vessel Movements." *Computer Graphics Forum* 28, no. 3 (2009): 959–966. doi:10.1111/j.1467-8659.2009.01440.x.
Willems, Niels, Roeland Scheepens, Huub van de Wetering, and Jarke J. van Wijk. "Visualization of Vessel Traffic." In *Situation Awareness with Systems of Systems*, edited by Piërre van de Laar, Jan Tretmans, and Michael Borth, 73–87. Springer New York, 2013.
Wilson, John P., and John C. Gallant. "Digital Terrain Analysis." In *Terrain Analysis: Principles and Applications*. John Wiley & Sons, 2000.
Wolinsky, Ira, and Judy Driskell. *Sports Nutrition: Energy Metabolism and Exercise*. Boca Raton, FL: CRC Press, 2008. http://www.crcpress.com/product/isbn/9780849379505.
Yuan, May, and Kathleen S. Horns by. *Computation and Visualization for Understanding Dynamics in Geographic Domains: A Research Agenda*. CRC Press, 2007.

APPENDIX A—Institutional Review Board (IRB) Documentation

IRB protocol application

DEPARTMENT OF THE ARMY
UNITED STATES MILITARY ACADEMY
West Point, New York 10996

REPLY TO
ATTENTION OF

MADN-G-ENE 6/17/15

MEMORANDUM THRU

Department Head or Activity Director,Geography and Environmental Engineering (MADN-G-ENE), West Point, NY 10996

FOR Chief of Clinical Investigation, Keller Army Community Hospital, West Point, NY 10996

SUBJECT: Application for the Use of Human Subjects in **Behavioral or Social Science Research**

*Required Fields

Select Only One
☒Request Expedited Review (Minimal risk and listed in one of the Expedited Review categories in the Federal Register)
☐Request Exemption (Minimal risk and listed in one of the six Exempt categories in 32 CFR 219.101(b))
☐Requires Full IRB Review

1. **Project Title*.** Space, Time and Energy in Dismounted Navigation

2. **Time Required for Completion*.**
   a. Expected Start Date: **04-15**
   b. Expected Completion Date: **11-16**

3. **Investigators.**
   a. **Principal Investigator**: Human Research Training Completed ☒ *(Training certificate enclosed.)

Rank/Title*: **CIV**	Full Name (First, M.I., Last, Suffix)*: **John Brockhaus**		Professional Credential (MD, CRNA, PhD): **PhD**
Academic Degree(s)*: **PhD**			
Job Title*: **GIS Program Director**			Status*: ☐ Military ☒ Civilian
Department/Division/Branch*: **GENE**		Major Activity Directorate*: DEAN	
Work Address*: Department of Geography and Environmental Engineering 745 Brewerton Road, 6th Floor West Point, NY 10996		Affiliation*: ☒ USMA Faculty or Staff ☐ Cadet ☐ DoD Employee (Not USMA) ☐ Contractor ☐ Other, specify:	
Work Phone*: 845.938.2063	Fax:	Mobile:	Email: John.Brockhaus@usma.edu

Version: 23 May 2011 Page 1 of 31

b. **Associate Investigator(s)**: *Include all individuals responsible for the design or conduct of the study. Enclose training certificates.*

Name (First, Middle, Last)/Degree	Department/ Activity, Division, or Affiliation	Role and Responsibilities *(i.e. recruit and consent volunteers, maintain study records, enter data into computer data base, etc.)*	Training Completed (check if training certificate is enclosed)
Irmischer, Ian J.	Geography / UCSB & USMA	Data Collection / Analysis	☒

c. *All investigators are aware of the requirement to inform participants, prior to the start of the study and throughout the research study, about:

- The risks and benefits of their participation in the research.
- The tasks that will be required of them to perform.
- Their right to stop participating in the research at anytime without penalty.

Yes ☒ No ☐ Explain why:

d. *The personal identity of participants may not be revealed in any form in the collected data or in the research report, unless specifically granted by the participant. Are all investigators aware of this right to privacy?

Yes ☒ No ☐

e. *Research documents are to be maintained by the research organization (i.e. department, MAD, etc.) for at least three years after the completion or termination of the study. Are all investigators aware of this policy?

Yes ☒ No ☐

3. **Location(s) of Study***. *List all facilities to be used, including all centers, laboratories, clinics, or outdoor facilities where the study is to be conducted. Include the investigator responsible for the conduct of research at each site. Provide the Federal-wide or DOD Assurance Number for each institution, other than USMA, engaged in the research. If an investigator is involved that is not operating under an approved Assurance, indicate whether a DoD Individual Investigator Agreement is included with the protocol.*

Facility and/or Location(s)	Investigator Responsible on Site	Federal-wide (FWA) or DoD Assurance Number
USMA Training Area - Land Navigation	Irmischer, Ian	
USMA GSL -, 5th Floor Washington Hall	Irmischer, Ian	

5. **Introduction.** *Describe the proposed research using language understandable to those IRB members whose expertise is non-scientific.*

a. **Objectives/Specific Aims/Research Questions*.** *Provide a description of the purpose and objectives of the study with detailed specific aims and/or research questions/hypothesis.*

The goal of this research is to model human energy expenditure (EE) of foot-based navigation. It will focus on military navigation which is referred to as dismounted navigation when the movement is conducted via walking. Soldiers in a combat environment perform less effectively if they are excessively fatigued from movement. Understanding energy expenditure of a proposed route is a vital element that should be considered when moving troops or supplies. Accurate assessment of energy expenditure is imperative to reducing operational risk. Currently no tools that fuse the physical terrain costs with human performance factors have been validated. This project will improve geospatial tools available to our military that focus on navigation. Improving geospatial-related technologies will support the United States Army Warfighter through enhanced knowledge of the environment, his/her capabilities, opportunities and risk.

The core contributions and the research products that will be developed are specifically focused on the dismounted routing implications of human energy expenditure. First, the research will define what are the relative weights assigned to individual characteristics that contribute to energy expenditure. Secondarily, the investigation will determine how significant environmental and terrain effects contribute to energy expenditure. Finally, the study will assess current energy expenditure models and provide improvements to existing military routing tools.

Research Questions and Hypotheses

1. What are the contributing weights of BMI, age, fitness and sex (individual characteristics) to an energy expenditure model of dismounted navigation?

Hypothesis: The above listed individual characteristics all contribute equally to the amount of energy a person expends during dismounted navigation.

2. What are the contributing weights of slope, land cover and distance traveled from origin (environmental/terrain variables) to energy expenditure while navigating?

Hypothesis: The above listed environmental and terrain variables contribute equally to the amount of energy a person expends during dismounted navigation.

3. How well do current energy expenditure models match the weights determined in questions 1 and 2 above?

Hypothesis: Current energy expenditure models can be improved by using the weights found by my research. This research will provide critical insight to the correct weighting of model parameters.

b. **Background***. *Describe any preliminary studies and findings that led to the development of the question. State what has been accomplished or published in the proposed area of study. Describe the way in which the project will relate to that which has been accomplished. Use references to support statements and list references in Enclosure 1.*

A fundamental task for human beings is to get from an origin to a destination. When the route to the destination is not easily discernable from the origin this task requires navigation. Navigation is defined as the coordinated and goal-directed movement through an environment (Montello 2005). It involves the physical act of moving and the cognitive aspects of deciding on and following a route. These two components of navigation are classified as locomotion and wayfinding.

Human locomotion during navigation can be accomplished in many ways such as in a vehicle, assisted by bicycle, in an airplane or by foot. Each of these forms of locomotion requires energy. This thesis will focus on unassisted human locomotion by walking. Locomotion via walking involves body movement due to muscular contraction. This muscular contraction requires energy from the human body. One of the primary goals of this research is to develop a model of energy expenditure (EE) due to navigation.

The second component, wayfinding, is a more cognitively centered aspect of navigation. Wayfinding involves planning movement, route selection, continued reassessment of location, and the constant decision making process to adjust your route. The use of the sensory processes and the brain require some energy to wayfind. Navigators continually update their location by recognizing landmarks and terrain features. Human beings primarily use vision to accomplish these tasks but also use hearing and the vestibular senses. However, these elements contribute to total energy expenditure of a human being in a manner insignificant in comparison to walking or muscular contraction (Clarke, Sokoloff, and others 1999; Lennie 2003). Still, a navigator uses muscular energy while wayfinding to assist in the sensory processes to look around, turn around or look behind themselves.

Metabolic Cost of Walking

Mathematical models that depict the metabolic cost of walking are at the heart of tools created to assist military officers with route selection. Current models of metabolic cost over terrain involve knowledge of the environment and individual characteristics such as gender, height, weight, BMI, fitness and load carried. Generally speaking, these equations are devised by affixing indirect calorimeters to subjects while they preform walking tasks over different types of terrain. Then regression analysis is used to fit a model that explains the variation in metabolic cost with a set of predictors. One of the most commonly used models of energy expenditure was devised by Givoni and Goldman in 1971 and refined by Pandolf in 1977 (Potter et al. 2013).

The Givoni and Goldman equation was derived by studying 26 subjects walking on treadmills at different speeds and grades. The research unsurprisingly found that metabolic rates increase with walking speed. Similarly, they proved that energy expenditure also rose with increasing grade in a near-linear fashion. Thirdly, the examination found that the metabolic cost of walking was linearly related to the summation of the person's body weight and the load (up to 30kg) with which they were burdened. The devised equation is listed below (Givoni and Goldman 1971). The proposed equation touted a correlation coefficient of .97.

$$MW = \eta\ (W+L) \bullet [2.3 + 0.32 \bullet (V-2.5)^{1.65} + G \bullet (0.2 + 0.07 \bullet (V - 2.5))]$$

Mw = metabolic cost of walking (in watts);
η = terrain factor (terrain for this equation was only considered as 1.0 as it accounted for treadmill surfaces only);

W = body mass (kilograms);
L = load mass (kilograms);
V = velocity or walk rate (kph); and
G = slope or grade (%)

The Pandolf equation was a refinement of the work done by Givoni and Goldman mentioned above (Pandolf, Givoni, and Goldman 1977). This equation is often used as a benchmark for validating other research relating to energy expenditure (Hall et al. 2004; Duggan and Haisman 1992; Kramer and Sylvester 2011, Potter et al. 2013). Pandolf used six subjects who walked for 15 minutes with backpacks. A second study was conducted with 10 subjects standing still with a loaded backpack to determine the energy expenditure of standing with different loads. The cost of different terrain factor coefficients were taken from a previous study at the same laboratory (Soule and Goldman 1972). A correction factor for downhill walking at a pace of 1.12 m/sec was added by Santee (Santee et al. 2003). Pandolf's original equation was devised with a correlation coefficient of .96 and the correction factor had an $r > .90$. Listed below is the finalized outcome (Pandolf, Givoni, and Goldman 1977; Santee et al. 2003):

1977 Equation
MW = 1.5 • W + 2.0 • (W + L) • (L / W)2 + ŋ • (W + L) • (1.5 • V2 + 0.35 • V • G)
Where:
Mw = metabolic cost of walking (or standing) (in watts)
W = body mass (kilograms)
L = load mass (kilograms)
ŋ = terrain factor
V = velocity or walk rate (m/s)
G = slope or grade (%)
The terrain factor categories are: 1.0 = black top road or treadmill; 1.1 = dirt road; 1.2 = light brush; 1.5 = heavy brush; 1.8 = swampy bog; 2.1 = loose sand; 2.5 = soft snow, 15 cm depth; 3.3 = soft snow 25 cm deep; 4.1 = soft snow, 35 cm depth (12).

2003 Correction
Mw = PE – CF   Where PE is the 1977 Pandolf Equation and CF is the correction factor listed below.
$CF = \eta \cdot [(G \cdot (W + L) \cdot V) / 3.5 - ((W + L) \cdot (G + 6)^2) / W) + (25V^2)]$
Where:
η = terrain factor
G = grade (%)
W = body wt (kg)
L = load wt (kg)
V = velocity (m/s)

Military Land Navigation is a basic skill required of every Soldier. It is taught at each form of individual basic training and during all US Army leadership schools. One of the major components of land navigation is route selection (U.S. Army 2013). Route selection is an element of every military operation involving dismounted troops. There are many geospatial tools available to assist Soldiers and

Marines during route selection (See ArcGIS Military Analyst among others). However, no validated tools address the variable of consumed energy. Energy consumption prediction and exhaustion avoidance is critical when conducting military operations and patrols (U.S. Army 1990). Similarly, no large, field based, navigation studies have validated the metabolic cost of walking equations mentioned above.

Human Energy Expenditure

Energy Expenditure in human beings can be defined as the amount of energy used over a given time. This is usually measured in kilocalories or Metabolic Equivalent of Task (MET). A kilocalorie in terms of physical activity and the energy stored in foods is the heat energy required to raise the temperature of one kilogram of water by one degree Celsius. A MET is defined as the rate of energy created per surface area of an average human while at rest. It is essentially a rating of how intense an activity is. If a task has a high MET then it requires more human energy expenditure than an activity with a lower MET. If an activity has a MET of 2 then it requires twice as much energy as resting (Wolinsky and Driskell 2008). Human beings require calories for a number of different reasons. Over 99% of energy used by the body is due to (www.fao.org):

- Basal metabolism - Functions that are essential for life, such as cell function
- Metabolic response to food - Energy for the ingestion and digestion of food
- Physical activity – Movement and other activities
- Growth - Energy needed to synthesize and support growing tissues
- Pregnancy - During pregnancy, extra energy is needed for the growing the fetus
- Lactation - The energy cost of lactation

There are a number of ways to measure human energy expenditure. The estimation of human energy expenditure over a period of time is known as calorimetry. Calorimetry is based on the principle that energy expended in a human can be calculated if the amount of heat transfer from the body over a given time is known. There are four main methods for determining a value for energy expended. Direct calorimetry measures the actual heat loss of an individual. Secondly, indirect calorimetry (IC) measures respiratory gases such as oxygen consumption or carbon dioxide production to estimate energy expenditure. A third method of measuring EE is using physical activity monitors[1]. These devices employ a number of different modeled variables such as body acceleration and heart rate to estimate the amount of energy expended. Lastly, a technique known as Doubly Labeled Water[1] (DLW) can be used to determine energy expenditure over long periods of time (2-21 days)[2](Pettee, Tudor-Locke, and Ainsworth 2008; McMurray and Ondrak 2008).

Direct Calorimetry calculates the energy used by the body directly from measuring the heat given off by a human (Wolinsky and Driskell 2008). This technique is most often conducted in a small sealed metabolic chamber equipped with specialized sensors to measure changes in the air temperature. Energy expended by the subject can be calculated very precisely based on the principles of heat transfer but applications are somewhat restricted due to the high cost of the equipment, confinement to a chamber and the complexity of the equipment.

Indirect Calorimetry is based on measuring the amount of oxidation in the body (Wolinsky and

[1] Sometimes the use of physical activity monitors and doubly labeled water can be considered forms of indirect calorimetry but most literature separate these techniques from indirect calorimetry.
[2] Doubly labeled water is considered the gold standard in energy expenditure determination but it will not be described in detail since it cannot be used for studies such as my research project that have a duration less than 48 hours.

Driskell 2008). It is important to understand that indirect and direct calorimetry do not measure the same energy. Indirect calorimetry estimates the energy expended based on consumed oxygen and produced carbon dioxide. Research has found that the amount of energy that the subject is using is proportional to the differences in the amount of O_2 and CO_2 inhaled and exhaled[3] (Leonard 2012). The quantity of O_2 and CO_2 inhaled and exhaled while breathing can be measured by a variety of bags/hoods, metabolic carts or portable facemasks. Examples are shown below in Figures 1 and 2 (www.cosmedusa.com):

Figure 1: COSMED Fitmate PRO Indirect Calorimetry device.

Figure 2: COSMED Fitmate GS Indirect Calorimetry device.

There are generally two types of indirect calorimetry. One uses a closed circuit system and the other uses an open circuit. The closed circuit system uses an airtight cylinder filled with oxygen. This system measures the amount of oxygen consumed over time to estimate expended energy. In this method, the person must breathe only through the mouthpiece connected to the oxygen supply. Closed circuit indirect calorimetry severely limits the mobility of the subject since they must continually be connected to the oxygen supply. A subject may use as much as 100 liters of oxygen over an hour long test (Levine 2005).

A second method of indirect calorimetry is an open circuit system. In this system subjects volume of oxygen consumption (VO_2) is measured by comparing a subjects inhaled air composition to the exhaled air composition. This comparison can reveal the amount of expired O_2 and CO_2. This method is measuring room air inhalation and therefore the subject need not be connected to an oxygen canister. There are several systems that can be used to compare the inhaled and exhaled air. These include a computerized cart, a bag or a portable device. Advantages of this IC system are that they can be lightweight and mobile and don't require oxygen tanks.

Physical activity monitors are wearable devices that collect biological and physiological information about an individual. The most common wearable physical activity monitors used to estimate energy expenditure are accelerometers and heartrate monitors. Accelerometers are biosensors that measure the acceleration of the body along either one, two or three axes. A predictive model can be used to estimate energy expended since acceleration is proportional to force applied (McMinn et al. 2013). Similarly, heart rate monitors have been shown to be capable of predicting energy expenditure after adjusting for age, gender and body mass (Keytel et al. 2005). The accelerometer is typically worn on a subject's hip or arm and heart rate monitors are usually worn around the chest. Several examples of typical research grade accelerometers and heart rate monitors are shown below in Figures 3,4, and 5 (Source: www.actigraphcorp.com):

[3] Occasionally, researchers include a urinary nitrogen measurement in their predictive model but most often it is excluded since it contributes less than 2% to total energy expenditure.

There are many factors that must be addressed when choosing a method for estimating energy expenditure. The differences among direct calorimetry, indirect calorimetry and accelerometers or heartrate monitors are significant. Tables 1, 2 and 3 describe some advantages and disadvantages of each type of device as summarized from Levine 2005; Wolinsky and Driskell 2008; McMurray and Ondrak 2008; Pettee, and Tudor-Locke, and Ainsworth 2008:

Direct Calorimeter

Major Characteristics	Advantages	Disadvantages	Typical Uses
-Chamber based measurement of heat lost by humans or animals	-Very Accurate	-Not Mobile -Expensive -Technically difficult to operate. -Cannot be used for short time scale. It takes 30 min to equilibrate before and after heat settings. -Difficult to maintain -Limited to exercises that can be evaluated inside a room or chamber.	-Lab based studies -Measure energy expenditure of sleeping, walking, ect. But not good for sports related tests -Metabolism experiments in animals and humans

Table 1: Characteristics, advantages, disadvantages and uses of direct calorimeters.

Indirect Calorimeter

Major Characteristics	Advantages	Disadvantages	Typical Uses
-Estimates heat production by measuring Oxygen consumption and CO2 production -Hoods, bags and facemasks -Oxidation can be measured with either an open circuit (most common) or closed circuit spirometers.	-Can be a portable device to use in field studies -Less expensive than direct calorimeter and almost as accurate -Good for measuring metabolic rate over short periods of time	-Less accurate than Direct Calorimeters -Much more expensive than Accelerometers -More limiting in the amount of unconstrained physical activity that can be tested.	-Exercise physiology experiments involving energy expenditure -Sports related studies involving fitness and metabolism

Table 2: Characteristics, advantages, disadvantages and uses of indirect calorimeters

Accelerometers or Heart Rate Monitors

Major Characteristics	Advantages	Disadvantages	Typical Uses
-Small inexpensive biosensors -Attach to hip, wrist and chest	-Very mobile -Technologically easy to understand -Durable -Quick to set up -Lowest cost	-Least accurate of all three types	-Physical Activity estimation of free living activities -Personal health monitoring -Research studies where funding makes Indirect -Calorimeters impossible -Research studies over long periods of time -Studies where hindering subjects with an indirect calorimeter would negatively affect research design

Table 3. Characteristics, advantages, disadvantages and uses of accelerometers and heart rate monitors

This study will primarily focus on accelerometers to measure energy expenditure due to their mobility, unobtrusive nature, cost and simplicity. The most accurate energy expenditure estimates come from indirect or direct calorimetry but the high cost, intricacy, invasiveness and human resources required make these devices infeasible for this study. Accelerometers can easily be transported to a field collection site for use. They are easily affixed to subjects that comprise the research sample and do not restrict the movement of the subjects. Additionally, the cost of accelerometers is significantly lower than indirect calorimetry. Research grade accelerometers cost less than $300 per unit which is minimal in comparison to indirect calorimetry devices which cost in the ballpark of $30,000 (Actigraph 2014). Finally, accelerometers are relatively simple to calibrate, set up and use for physical activity data collection (Wolinsky and Driskell 2008).

Accelerometers have been shown to predict energy expenditure to greater than 90 percent accuracy while walking (Crouter et al. 2010; Kuffel et al. 2011; Brandes et al. 2012). They are a common method for assessing energy expenditure in field based research (Welk, Schaben, and Morrow Jr 2004). Technological advances in micro-computing and data storage have led to an increased variety of research and consumer grade accelerometers. Some of the more common devices on the market are Actiheart, Actical, Actigraph, Body Media Fit, Fitbit Zip, Jaw Bone Up and Nike Fuel Band. Most smartphones are now equipped with an accelerometer which has given rise to a new form of energy expenditure estimation device (Pande et al. 2013). This research will use primarily Actigraph accelerometers as data collection devices. The Actigraph accelerometers are research grade devices that are commonly used as instruments in scientific studies to estimate energy expenditure (Lee, Kim, and Welk 2014). They also use open source algorithms to estimate energy expenditure which is necessary for our scientific understanding. Finally, these devices are paired with a software package which allows researchers to easily download raw acceleration data and analyze results.

Most attempts to estimate energy expenditure from accelerometers use regression to develop models of energy expenditure from accelerometer count information (Crouter, Clowers, and Bassett 2006; Crouter et al. 2010; Tapia, 2008, Freedson, Melanson, and Sirard 1998, (Actigraph 2015b). There have been models devised using linear regression (Freedson, Melanson, and Sirard 1998; Swartz et al. 2000) and non-linear regression (Chen and Sun 1997; Crouter, Clowers, and Bassett 2006; Crouter et al. 2010). The equations are developed by relating the accelerometer counts over a given time with energy expenditure information collected with a closed circuit IC device. The most accurate equations are particular to a specific activity such as walking, running, or mopping the floor (Crouter, Churilla, and Bassett Jr 2006).

A novel model of energy expenditure for persons walking was developed in 2006 by Crouter and his research team (Crouter, Clowers, and Bassett 2006). His previous research tested the validity of published energy expenditure models created from accelerometers. This study found that a single regression equation for estimating energy expenditure from accelerometer counts tended to overestimate walking and running (Crouter, Churilla, and Bassett Jr 2006).

Crouter subsequently created a new algorithm to estimate EE by first determining if the activity was indeed walking/running or if it was some other type of activity. He accomplished this by analyzing the variation of counts every 10 second epoch[4] for each whole minute. If there was little variation in counts/epoch then the whole minute was considered walking/running. If there was significant variation then the minute was classified as another activity. Two regression equations were created; one for walking/running and another for all other activities. This was an improvement because there is significant difference between EE/count when walking/running compared to other activities. The new algorithm was more accurate than all others he had tested in his previous study.

[4] An epoch represents the amount of time (i.e., seconds) over which movement data (i.e., activity counts) are measured.

In 2010, Crouter and his collogues further refined this energy expenditure model to better classify activities based on superior temporal analysis. The refined method examines each 10-second epoch in comparison to all arrangements of the neighboring five 10-second epochs. This is in contrast to the 2006 model which analyzed each 10 second epoch versus other epochs in each whole minute. The method essentially created a sliding temporal range for each epoch opposed to a minute by minute assessment. Again, they used the variation of counts/epoch to determine if the activity is walking/running or something else. Lastly, if the counts/epoch are less than 8 then the human is considered at rest during that epoch and by definition given a MET=1 for that time period.

This work has led to a state of the art algorithm to estimate energy expenditure listed below (Crouter et al. 2010):

1. If the counts˙10 sec−1 are > 8

a. CV of the counts per 10 sec are ≤ 10, then energy expenditure (METs) = 2.294275 * (exp(0.00084679 * ActiGraph counts˙10 sec−1)) (R^2 = 0.739; SEE = 0.250),

b. CV of the counts per 10 sec are > 10, then energy expenditure (METs) = 0.749395 + (0.716431 * (Ln(ActiGraph counts˙10 sec−1))) − (0.179874 * (Ln(ActiGraph counts˙10 sec−1))2) + (0.033173 * (Ln(ActiGraph counts˙10 sec−1))3) (R^2 = 0.840; SEE = 0.863)

2. If the counts˙10 sec−1 are ≤ 8, energy expenditure = 1.0 MET

Where:

$$\text{CV} = \frac{\textit{Standard Deviation of Data}}{\textit{Mean of the Data}}$$

The 2010 Crouter refined 2-regression model has been validated in reputable research studies (Kuffel, 2011). Additionally, this algorithm is built into the Actilife 6 data analysis software program that can be used to analyze accelerometer data (Actigraph Software Department 2012). Finally, it is built specifically for the two activities that occur during navigation: walking and rest.

Heart rate monitors are also used as energy expenditure estimation devices. Research completed by Keytel and his colleagues remarkably demonstrated a regression model that can predict energy expenditure from heart rate that explains 73 percent of the variance (Keytel et al. 2005). This data will be collected to compare results against accelerometer based estimates.

c. **Assessment of Benefits.** *(See Section 3-2, USMA Reg. 70-25 for more information.)*

(1) **Novelty of the Scientific Question***. *Describe, in detail, the novelty of the scientific question. How does this project differ from previous studies? Note: Repeat of a completed research project is not research because it does not develop or contribute to new knowledge.*

The goal of this research is to model human energy expenditure (EE) of foot-based navigation. It will focus on military navigation which is referred to as dismounted navigation when the movement is conducted via walking. Soldiers in a combat environment perform less effectively if they are excessively fatigued from movement. Understanding energy expenditure of a route is a vital element that should be considered when executing movement of troops or supplies. Accurate assessment of energy expenditure is imperative to reducing operational risk. Currently no tools that fuse the physical terrain costs with human performance factors have been validated. This project will improve geospatial tools available to our military that focus on navigation. Improving geospatial-related technologies will support the United States Army Warfighter through enhanced knowledge of the environment, his/her capabilities, opportunities and risk.

The core contributions and the research products that will be developed are specifically focused on the dismounted routing implications of human energy expenditure. First, the research will define what are the relative weights assigned to individual characteristics that contribute to energy expenditure. Secondarily, the investigation will determine how significant environmental and terrain effects contribute to energy expenditure. Finally, the study will assess current energy expenditure models and provide improvements to existing military routing tools.

There have been no studies of dismounted land navigation that have collected energy expenditure data, GPS trajectories and individual characteristics. This study will be the first of its kind to use high resolution geospatial data, individual attributes and energy expenditure in order to model dismounted human navigation over varied terrain.

(2) **Potential Benefits***. *Describe real and potential benefits of the results of the research to subjects, a specific community, or society. Why is this project important enough to expose human subjects to any amount of risk? In evaluating benefits, the IRB considers only those benefits that may result from the research (as distinguished from benefits that subjects would receive even if not participating in the research). Ensure that the benefits are not overstated.*

Military Land Navigation is a basic skill required of every Soldier and Marine. It is taught at each form of individual basic training and during all US Army leadership schools. One of the major components of land navigation is route selection (U.S. Army 2013). Route selection is an element of every military operation involving dismounted troops. There are many geospatial tools available to assist Soldiers and Marines during route selection (See ArcGIS Military Analyst among others). However, no validated tools address the variable of consumed energy. Energy consumption prediction and exhaustion avoidance is critical when conducting military operations and patrols (U.S. Army 1990). The omission of tools that provide routing assistance based on potential exhaustion leaves a gap in current US military capability.

This research will collect energy expenditure estimates of Soldiers while they navigate on foot over hilly, wooded terrain. The estimates will provide an opportunity to model the contribution of both individual and terrain factors to energy expenditure during dismounted navigation. Knowledge of the influences contributing to energy expenditure will provide a basis for assessing and improving current GIS routing tools. Improving routing algorithms and geospatial tools for the mission planner will be a great improvement to current capabilities. It will increase operational safety and military success. Finally, it will provide better tools to access information about how the environment and Soldier capabilities affect missions.

NOTE: Payment and/or other compensation for participation (i.e. extra credit, gifts, loans, use of products or facilities, etc.) are not considered to be benefits and must be addressed in a separate section.

(3) **Intent to Benefit**. *If volunteers cannot give their own consent to participate in an experimental study, and 10 USC 980 applies, a clear intent to benefit each volunteer must be described. If not applicable, leave blank.*

d. **Major Safety Concerns***: *Briefly highlight major safety concerns for human subjects, such as potentially severe emotional distress or physical injuries. Provide a more complete description in paragraph 9.b. If none, leave blank.*

There are minimal risks to cadets during the experimentation.

The collection of cadet individual attribute data such as height, weight, sex and fitness score are administrative data that is already being collected and databased at USMA. All cadets had to submit admissions data with height, weight, sex, fitness score so it is unlikely these attributes will impact the emotional distress level of individuals. Height and weight will collected at the time of the experimentation to calibrate activity monitors.

The collection of GPS location data: This task will be completed by cadets regardless of this research. Safety mechanisms are in place by the Chain of Command of the Cadet Summer Training to mitigate risk to Cadets. This includes having trained military officers and medical personnel onsite at all times.

Collection of acceleration data and heart rate data: The wearing of an accelerometer and heart rate monitor have minimal safety risks. The cadets movement and decision making will not be impacted by the accelerometer and heart rate monitor. All activities will be completed by cadets regardless of this research activity. The accelerometer is worn on the hip and weighs 19 grams. Heart rate monitor us worn under the clothing and weighs less than 10 ounces. Safety mechanisms are in place by the Chain of Command of the Cadet Summer Training to mitigate risk to Cadets. This includes having trained military officers and medical personnel onsite at all times. It is possible that minor chaffing occurs due to the monitor belts but the risk is minimal.

6. **Research Plan.** *Outline expected accomplishments in enough detail to show a clear course of action. Include technological validity of procedures and chronological steps to be taken.*

a. **Research Design*.** *Describe the systematic investigation and how its design will develop or contribute to generalizable knowledge.*

- *Define the study variables and describe how they will be operationally measured.*
- *Describe the methods that will be used to obtain a sample of volunteers from the accessible population (i.e., convenience, snowball, simple random, stratified random, systematic random, cluster (area) random, multi-stage, etc.).*
- *If applicable, describe the subject to group assignment process (e.g., randomization, block randomization, stratified randomization, age-matched controls, alternating group, or other procedures).*
- *If applicable, explain the specific actions to accomplish the group assignment (e.g., computer assignment, use of table of random numbers).*

The research will be comprised of several experiments to gather information about how much human energy is expended while cadets at the United States Military Academy (USMA) navigate. The experiments will collect movement, location and time data using wearable devices. The devices will be affixed to cadets while they participate in a navigational test that is part of their summer military training. The movement, location and time data will be collected over varying terrain conditions including trees, swamps, roads, trails, brush, hills and cliffs. Finally, I will obtain individual characteristic data from the United States Military Academy including: age, fitness level, and sex. Height and weight will collected at the time of the experimentation to calibrate activity monitors.

The military training area where the navigational testing will take place falls with the following bounding box (Figure 6):

Upper Left MGRS Coordinate: 18TWL7573578270

Lower Right MGRS Coordinate: 18TWL7848976111

Figure 6: United States Military Academy Land Navigation Testing Site

I have obtained the following terrain data for the training area where the navigational testing will occur:

1. High resolution (< 1 Meter) aerial imagery of the test site mentioned above.
2. High resolution (<1 meter) digital elevation model of the testing site above.
3. LiDAR point clouds of the test site.
4. Vector data of paths and Land Cover of the test site.

Description of the Population and the Sample:

Experimentation will consist of asking cadets from the United States Military Academy to voluntarily wear accelerometers while they are tested on their ability to navigate through the woods in West Point, NY. Convenience sampling of volunteers will be used to obtain subjects. The population is junior military officers between the ages of 17-25. This is representative of the segment of the military most directly involved with route decision making for operations. The sample can be described as:

- Approximately 240 cadets at USMA who are participating in land navigation testing during Cadet Basic Training (CBT) in July and August, 2015.
- Consisting of both genders at an approximate ratio of 80% male and 20% female which is similar to the population referenced above.

- Cadets will be wearing military field uniforms with a pistol belt and camelback hydration system
- Consisting of cadets with varying heights, weights, and fitness levels
- Age of a cadet is between 17 and 25.
- Cadets will be using a map and compass as navigational aids.
- Each subject will be carrying a Samsung Galaxy S4 mobile phone with GPS device. The phone is not used by the cadets except for emergencies. The per-second trajectory information of each cadet is recorded by the phone and downloaded to a hard drive at the end of the experiment.

Experimental Description:

Approximately 1000 cadets participate in Cadet Basic Training each summer. They are divided into eight subgroups called companies. This leaves roughly 125 cadets in each company. Each company executes the land navigation testing on a different day in July and August. This will allow the investigator to collect data on 8 different days and maximize use of the limited number of accelerometers that will be purchased with funding provided from the Army Geospatial Labs. Approximately 30 accelerometers and heart rate monitors will be purchased as data collection devices for this experiment. We anticipate roughly 20% equipment malfunction leading us to a sample size of nearly two hundred persons. The investigator will obtain individual differences data such as age, gender, and fitness score from empirical data collection and the USMA admissions office. Height and weight will collected at the time of the experimentation to calibrate activity monitors.

The experiment will collect individual movement data using the Actigraph GT3X (Figure 7) or a comparable device (Actigraph 2014). This sensor determines acceleration values using a triaxial Microelector-Mechanical-System accelerometer. The biosensor is capable of detecting static (from force of gravity) and dynamic accelerations (from movement) in three directions. The accelerometer detects change in acceleration through measuring variations in the sensor's electric charge storage. The device allows for analysis of raw acceleration data or pre-processed activity counts such as steps. The Actigraph GT3X has 256 MB of onboard memory to record and store data in the time scale of days. This is sufficient for our experimentation. The GT3X can be worn around the wrist or waist (John and Freedson 2012) based on experimental design preferences. The data from the accelerometers will be used to estimate each cadet's expended energy every 10 seconds. The investigator will use the Crouter Refined 2-Regression Model within the Actilife 6 software to estimate METs per epoch (Crouter et al. 2010; Actigraph Software Department 2012).

Figure 7: Actigraph GT3X with belt and Heartrate monitor

Figure 8: Samsung Galaxy S4 with GPS device.

The subject's trajectory will be recorded by the GPS sensor inherent to the Samsung Galaxy S4 phone (Figure 8). The Cadet Summer Training (CST) navigation testing environment requires every cadet to carry the mobile phone in a pocket on their left shoulder. The spatio-temporal GPS tracks will be collected every second while the subjects are navigating. This space-time information collected via GPS and the energy expenditure estimates from the Actigraph GT3X can be directly associated via the timestamp since both use Coordinated Universal Time (UTC). Slope and land cover can be associated with this data by comparing the GPS spatial information, the high resolution digital elevation model and the USMA land cover vector layers.

Timeline

This research project will be divided into four execution phases. The first phase will consist of knowledge review and preparation for experimentation. Phase two is focused on experimentation at the United States Military Academy in West Point, NY. During the third phase, the investigator will organize and analyze the collected data. The final phase of the research plan centers on publishing the research findings.

Phase one of the research is ongoing. It involves a thorough review of previous work in the fields of navigation, space-time GIS and energy expenditure. Additionally, the investigator will use phase one to further solidify an existing relationship with the navigation community at the USMA summer training program. This work will finalized prior to Cadet Basic Training in July and August of 2015.

Phase Two of the research focuses on data collection. The data will be collected during several experiments in West Point, NY. Experimentation will be executed during the Cadet Basic Training at the United States Military Academy in July and August of 2015. The investigator will use wearable devices to extract spatiotemporal data from subjects participating in navigation training as described above. This data will be used to analyze and assess the geostatistical validity of current algorithms. Most importantly, the data will help determine contributing weights of individual and terrain variables for an energy expenditure model.

Phase Three of this research plan is focused on data analysis and model development. This will occur during late summer and fall of 2015. The data analysis and model development will stochastically compare the results of current EE algorithms to the EE calculated from indirect calorimetry. Secondly, it will apply parametric and non-parametric tests to a large database of individual and terrain variables. This will allow the investigator to find the contribution of each independent variable on EE in order to create a robust EE model.

Phase Four of the research includes writing and publishing the research findings. The investigator will finalize the results and findings of the research effort during the winter of 2015-16. The research will translate into publishable journal articles to inform the military and the field of GIS. This phase will also be used to document new findings and theory.

b. **Target Population***. *Describe the target population (to whom the study findings will be generalized)*

Junior Military Officers between the ages of 17 and 25.

c. **Accessible Population ***. *Describe the nature, approximate number and pertinent demographic characteristics (sex, race, age range, etc.) of the accessible population at the study site (population from which the sample will be recruited/drawn). The accessible population should be closely representative of the target population. The scientific review must estimate the probability of meeting the enrollment goals from this population and ensure selection is equitable.*

(1) Approximate Number Accessible*: 1000 *Note: This is not the sample size.*

(2) Age Range*: 17 to 25

(3) Other Pertinent Demographic Characteristics (Gender, Race, Profession, Geographic location, etc.)

~20 % women, 80% men, 70 % white, 12% African Americans, 10% Hispanics, 8% other -

d. **Sample Size***. *The protocol must specify the approximate number of volunteers that will be enrolled. If the protocol involves multiple sites, the number enrolled at each site should be stated. The minimum number of subjects should be used that yield statistically valid results.*

(1) Number of Subjects*: 240

If more than one site:

Location	Number of Subjects

(2) Rationale*: *(An appropriate sample size justification (for example, power analysis) must be included to ensure that the sample size is appropriate to meet the objectives of the study.)*

A sample size of 240 is desired. The research requires this relatively large sample because it needs to control for large variability in its measures. For one, subjects are free to choose a route, and these routes may differ greatly from each other, in terms of lengths, numbers of turns, terrain, and so on. Also, individual navigators differ substantially from each other in several of the measures being recorded. Thus, we need a large sample size to be able statistically to evaluate different potential confounds and contributions to error. Also, a goal of

the research is to evaluate how its conclusions generalize to different terrains and individuals. Again, more data are thus required to systematically evaluate this generalization. It is also anticipated that we will lose some data, either because of equipment malfunction or subjects opting out of certain procedures. Finally, the navigation training course involves over one thousand participants, and the research needs to take advantage of the potential for stronger research conclusions engendered by a larger sample. Our limited equipment keeps the number down to about 240.

e. **Inclusion/Exclusion Criteria**. *Selection of subjects must be equitable. This means that there is equity both in assuming the burdens and in receiving the benefits of human subject research. Include justification of any age, race, ethnicity, gender or other constraints or limitations required for the study.*

(1) **Inclusion Criteria.** *If certain characteristics are required of subjects to be included in the study, clearly state them and provide justification. Criteria must be related to the study objectives. Example: Only personnel who spent their adolescence in a rural area of North America.*

N/A

(2) **Exclusion Criteria**. *If certain characteristics of potential subjects do not support the study objectives, clearly state them and provide justification. This includes professions, ethnicity, or other group characteristics that should be excluded. Example: Personnel who are diagnosed with post-traumatic stress disorder will be excluded from the study.*

N/A

f. **Data Collection Methods**. *Copies of data collection forms and/or instruments administered (tests, surveys, etc) must be provided. Data collection forms should be adequate and accurate according to the data collection plan described. Whenever possible, identifiers should be removed from data collection forms. Critical measurements used as endpoints should be identified. (See Appendix E, USMA Reg. 70-25 for more information.)*

(1) **Method to obtain information about the living individual** (Check all that apply.):*

(a) ☐ **Interaction** *(Example: Surveys, Interviews, E-mail correspondence, etc.)* or ☒ **Intervention** *(i.e. blood draw, electrodes, wearing or use of equipment, designed exposure to environmental conditions, etc.)*

Procedures. *Describe the research interaction or intervention that the volunteer will experience. Provide sufficient detail in chronological order for a person uninvolved in the research to understand what the volunteer will experience and when it will occur. Provide a schedule of study evaluations and follow-up procedures. Enclose all case report forms, data collection forms, questionnaires, rating scales, and interview guides, etc. that will be used in the study.*

Individuals will be asked to volunteer in a research study to better our understanding of human navigation and in order to create new tools that can be helpful to junior military officers when they plan dismounted operations. They will be asked if they are willing to wear an accelerometer and a Polar heart rate monitor throughout the duration of their Cadet Basic Training (CBT) day 3 land navigation testing. The accelerometer is a small, 2 ounce box worn around the wrist with a strap, while the heart rate monitor is easily worn around the chest with another strap (see www.theactigraph.com for photos). These are designed to be small and unnoticeable, and to be worn comfortably without disrupting daily activity. The chest transmitter is comfortably worn against the bare skin with an elastic chest strap. Information from the unit (minute by minute heart rate, accelerometer counts, and number of steps taken) can be downloaded onto a computer. " Volunteers will be requested on Day 2 of the navigation training. Weight, height and age data will be taken from the subjects on this day if they agree to participate. On day 3 of the training the subjects will be asked to put on the accelerometer and heart-rate monitor prior to the land navigational test. Subjects will return the monitors when they finish the land navigation test. The test will last approximately 3 hours.

(b) ☒ **Secondary Data** *Include the source and describe any identifiers being collected.*

Source of the Accessible Data

Check all that apply	**Source**	**Name and Location of Source** *(Organization/Author, Web Site, Public Record)*

☒	System of Records *(includes USMA)*	
☐	Health Care Covered Entity *(Examples: a medical records database or a physician who provides health care services and conducts clinical trials)*	
☐	Prior Research. *Enclose original consent document authorizing use of data for other research purposes for which it was not originally intended or enclose a proposed consent document to gain informed consent.*	
☐	Publicly Available	
☐	Other	

Type of Data Accessed

Describe the existing data that the investigator will access to create a data set for the study.

Data Accessed		Data Set Created	**Identifiers** *(List personal identifiers to be collected or attach a data collection sheet that will be used to record the data)*
☒	Identifiable *Describe how this data will be protected in paragraph 9.d.*	☐ De-identified Data Set ☐ Limited Data Set *(Complete a "Data Use Agreement")*	
☐	Non-Identifiable. *(Aggregate (tabular) data when the data are not at the individual level of analysis and information about individuals cannot be recovered from the tabulations or Individualized data sets that do not contain direct or indirect identifiers).*		

(2) **Volunteer Identification.** *If unique identifiers or a specific code system will be used to identify volunteers, describe this process. Otherwise, state how the data will be de-identified or kept anonymous.*

Each participant will complete a consent form and receive a participant ID number. Information linking the participant and the ID number will be stored and locked in a file cabinet. All anonymized data will be stored on a password-controlled computer. Only members of this research group will have access to the data. The information will be stored for ten years.

NOTE: Requirements for reporting sensitive information to military, state, or local authorities should be addressed in the protocol and the consent form. Examples of sensitive information that may require reporting include illegal residency, child or spouse abuse, or participation in other illegal activities. These requirements will vary from state to state. Investigators should consult an IRB within a particular location for assistance with state/local requirements.

g. **Statistical/Data Analysis Plan***. *Describe the data analysis plan. The data analysis plan should be consistent with the study objectives.*

Data Analysis Techniques

Data will be analyzed using a temporal resolution of 10 seconds. A large data matrix will be created in which each row will represent movement over a 10 second epoch. Each row will have the following fields:

Subject ID

Subject Height

Subject Weight

Subject BMI
Subject Age
Subject Fitness
Subject Sex
Beginning Time
End Time
Beginning Location
End Location
Distance traveled
Slope
Land Cover
Energy Expenditure / METs
Accelerometer Counts
Heart rate

The individual characteristics and experimentation results will provide a rich set of explanatory variables that can be used to model EE. I will apply parametric and non-parametric tests to this large database of geo-temporal data to modeling the relationship between variables. These results will be used to provide suggested improvements to existing routing tool kits.

h. **Dissemination Plan*.** *Describe how the results of the study are intended to be presented to develop or contribute to generalizable knowledge. How will the results be disseminated to enrolled participants, professionals within a field of study, or society in general?* (Check all that apply)

☒ Presentation: *Describe any intended seminars, conferences, or groups for which the results will be presented.*

AutoCarto, 2016

☒ Publication: *Indicate the intended publication(s) (journal, book, etc.)*

IJGIS

Approximate Year of Publication: 2017

7. **Description of Devices/Equipment**. If the protocol uses devices (monitoring devices, testing equipment, etc.) to conduct the research, provide the following information *(Otherwise, leave blank)*:

- Description of device *(Provide complete names and composition of all supplements or devices.)*:

Actigraph WGT3X-BT Monitor with Polar heart rate monitor.

- Source *(manufacturer, issuer, dealer, retail store, etc.)*:

ActiGraph
49 East Chase Street
Pensacola, FL 32502
850.332.7900

- Intended Use *(Explain how the device will be used. If the device will be used in a way that it was not originally intended, identify the reasons for the altered use. Explain potential risks associated with using the device/equipment in paragraph 9.b):*

The WGT3x-BT will be worn on the hip of the cadets as shown in pictures above as prescribed by the manufacturer. No altered use will be used. The heart rate monitor will be worn around the chest prescribed by the manufacturer.

8. **Conflict of Interest Disclosure***. *Investigators and key study staff must disclose any real or apparent conflicts of interest (financial or other). Financial conflicts of interest must be disclosed using the Financial Disclosure Form for Investigators available on the USMA HRPP website. Information regarding conflicts of interest should be disclosed to volunteers in the consent form. All protocols that support development of a drug, device, or other intellectual property require completion of a conflict of interest declaration by all investigators on the protocol. (See Section 3-6, USMA Reg. 70-25 for more information.)*

(Check one)*

☒ Investigators and their immediate family members DO NOT have significant conflicts of interest in this study.

☐ One or more investigators or immediate family members of the investigators have a significant financial conflict of interest in this study. A Financial Disclosure Form is enclosed.

☐ One or more investigators have a significant non-financial conflict of interest in this study *(Disclose any significant non-financial conflicts of interest and the measures taken to mitigate their impact)*:

9. **Risk Assessment**. *Estimate the likelihood and severity of potential harm, discomfort, or inconvenience to human subjects during and after the conduct of research and identify how the likelihood or severity can and will be minimized. (See Section 3-2, USMA Reg. 70-25 for more information.)*

a. Vulnerable Populations (*Check all that apply*) *Use of these populations should be avoided unless they are required to meet the study objectives. Additional safeguards are required to protect these populations. Address the risks and safeguards in subparagraphs b and c of this paragraph and specifically describe how they will be recruited in paragraph 9 "Recruitment Process".*

☒ Cadets (*See Appendix E, USMA REG 70-25 for special procedures that apply when using cadets*)

☐ Pregnant/Nursing Women or Fetuses (*See 45 CFR 46, Subpart B for more information*)

☐ Prisoners (*See 45 CFR 46, Subpart C for more information.* ***The involvement of prisoners of war as human subjects of research is prohibited [DoDD 3216.02, para. 4.4.2]***)

☐ Children (*See 45 CFR 46, Subpart D for more information*)

☐ Military Members (Soldiers, Airmen, Marines, or Sailors)

☐ Employees of the principal or associate investigators

☐ Students enrolled in a class taught by a principal or associate investigator

☐ Cognitively-impaired persons

☐ Illiterate persons

☐ Non-English-Speaking Persons

☐ Other potentially vulnerable participants: Who and why?

b. **Foreseeable Risks***. *Clearly identify any and all foreseeable risks that a volunteer may be subjected to as a result of participation in the protocol. Consider physical, psychological, political, legal, social, and economic risks. If the risks are unknown, state so. Describe the risk in terms of an event of harm (ex. Invasion of privacy caused by a lost or stolen laptop). Do not attempt to evaluate the level of risk (i.e. minimal risk).*

The collection of cadet individual attribute data such as height, weight, sex and fitness score are administrative data that is already being collected and databased at USMA. All cadets had to submit admissions data with height, weight, sex, fitness score so it is unlikely these attributes will impact the emotional distress level of individuals. Height and weight will collected at the time of the experimentation to calibrate activity monitors.

The collection of GPS location data: This task will be completed by cadets regardless of this research activity. Safety mechanisms are in place by the Chain of Command of the Cadet Summer Training to mitigate risk to Cadets. This includes having trained military officers and medical personnel onsite at all times.

Collection of acceleration data and heart rate data: The wearing of an accelerometer and heart rate monitor have some safety risks. The cadets movement and decision making will not be impacted by the accelerometer and heart rate monitor. All activities will be completed by cadets regardless of this research activity. The accelerometer is worn on the hip and weighs 19 grams. Heart rate monitor us worn under the clothing and weighs less than 10 ounces. Safety mechanisms are in place by the Chain of Command of the Cadet Summer Training to mitigate risk to Cadets. This includes having trained military officers and medical personnel onsite at all times. It is possible that minor chaffing occurs due to the monitor belts but the harm to individuals is unlikely.

c. **Risk Management and Emergency Response.**

- Measures to be taken to minimize and/or eliminate risks to volunteers and study personnel or to manage unpreventable risks.* *Describe all safety measures in place to mitigate risk (e.g., heart rate monitoring, observation periods, and special procedures to avoid disclosure of potentially damaging information)*:

Safety mechanisms are in place by the Chain of Command of the Cadet Summer Training to mitigate risk to Cadets. This includes having trained military officers and medical personnel onsite at all times. Additionally, each cadet will carry an active mobile phone to make emergency calls. Finally, the mobile phone has an active GPS chip that is used to transmit each cadet's location to a base station tracking display. This allows the chain of command to actively monitor each cadet's movement and activity.

- Planned responses. *Describe, in detail, any pre-planned responses, such as pre-determined alert values and other safeguards. If none, then leave blank. (The study site must have adequate personnel and equipment to respond to expected adverse events and the nearest medical treatment facility should be identified in the emergency response plan.)*:

Safety mechanisms are in place by the Chain of Command of the Cadet Summer Training to mitigate risk to Cadets. These individuals will have total control of response to emergency. The closest hospital is Keller ACH.

- Provision of Care for Research-Related Injuries. *If a volunteer may require emergency care or treatment for an adverse event, discuss the overall plan for provision of care for research related injuries to include who will be responsible for the cost of such care. If none, then leave blank. (For example, if a study sponsor or institution has committed to providing care for research related injury at no cost to non-Military volunteers, this provision should be explained in the protocol.)*:

- Special precautions to be taken by the volunteers before, during, and after the study *(e.g., medication washout periods, dietary restrictions, hydration, fasting, pregnancy prevention, etc.). Describe any special precautions. If none, then leave blank.*

- Special care. *Describe any special care (e.g., wound dressing assistance, transportation due to side effects of research intervention impairing ability to drive) or equipment (e.g., thermometers, telemedicine equipment) needed for volunteers enrolled in the study. If none, then leave blank.*

d. **Privacy and Confidentiality**. *(See Section 3-3, USMA Reg. 70-25 for more information.)*

(1) Protection measures*. *Explain measures taken to protect the privacy of research volunteers and maintain confidentiality of private information, particularly identifiable private information.*

Each participant will complete a consent form and receive a participant ID number. Information linking the participant and the ID number will be stored and locked in a file cabinet. All anonymized data will be stored on a password-controlled computer with encryption. Only members of this research group will have access to the data. The information will be stored for ten years.

(2) Access to Records, Data, and Specimens*.

- Personnel: *State the personnel that will have access to research records, data, and specimens. Include whether these individuals will have full access or any restrictions that may apply.*

Irmischer, Ian J.	Full Access

(3) Disposition of Data*. *Describe where data (both electronic and hard copy) will be stored, who will keep the data, how the data will be stored, and the length of time data will be stored. If the data will be destroyed, explain the method of destruction. Personally identifiable information on portable electronic media (laptops, flash/thumb drives, etc) must be encrypted.*

LTC Irmischer will store all data in locked file cabinet in his office at UCSB until he relocates to USMA for assignment. At that time it will be locked in his office at USMA. The digital data will be held on LTC Irmischer's computer which is password protected and the data will be encrypted.

10. **Recruitment Process***. *Address who will identify potential volunteers, who will recruit them, and what methods will be used to recruit them. Investigators must ensure potential subjects are not exposed to undue influence or coercion. (See Section 3-4, USMA Reg. 70-25 for more information.)*

Investigator will ask for volunteers from each company while they are participating in Day 2 of the land navigation training. During this day, the cadets will be participating in a classroom instructor led training class. A voluntary consent statement will be read before asking for volunteers.

a. **Incentives**. *If volunteers will be compensated for participation in the research study, a detailed description of the compensation plan must be included in the protocol. Otherwise, leave blank. Ensure that the compensation plan is fair and does not provide undue influence. If the study requires multiple visits, a plan for pro-rating payments in the event of volunteer withdrawal should be considered. Incentives may include money, extra credit in a course, or gifts.*

b. **Recruitment and Advertisement Materials***. *Recruitment materials should not be coercive or offer undue inducements and should accurately reflect the research. An ombudsman should be considered for use with particularly vulnerable populations. See USMA Reg 70-25 for additional information specific to recruitment of military personnel and additional guidance on recruitment materials.*

Are recruitment or advertisement materials being used to recruit human subjects?

☒ No. ☐ Yes. Enclose copies with this protocol for review.

c. **Volunteer Screening Procedures**. *Please note that some screening procedures may require a separate consent or a two-stage consent process. Informed consent must be obtained prior to initiation of any procedures for collection of individually identifiable private information for the purposes of determining eligibility. List and describe any evaluations (e.g., laboratory procedures, personal history, records review, or physical/psychological examination) that are required to determine eligibility/suitability for study participation and the diagnostic criteria for entry.*

11. **Informed Consent Process**. *Specifically describe the plan for obtaining initial informed consent from study volunteers. (See Section 4-3, USMA Reg. 70-25 for more information.)*

a. **Process***. Informed Consent will be obtained in accordance with 32 CFR 219.116 (a) and (b). A consent procedure that excludes or alters some or all of the elements of informed consent IAW 32 CFR 219 (c) or (d) ☐ **is** / ☒ **is not** requested.

If requested, describe and justify the alteration or exclusion:

b. **Responsible Investigator(s) and Relationship to Subjects***: *Identify who is responsible for explaining the study, answering questions and obtaining informed consent. Explain any other relationship this person may have with a subject other than for research purposes.*

LTC Irmischer will be explaining the study, answering questions and obtaining informed consent. He has no relationships with any potential subject.

c. **Timing and Setting***: *Include information regarding the timing(s) and setting(s) of the informed consent process (include locations and anyone, other than the responsible investigator(s), who will be present).*

The informed consent will be given in the GSL on the 5th floor of Washington hall at USMA. It will be given during a classroom instruction that each cadet will attend.

d. **Subject's Consideration of Participation***: *Describe, in sufficient detail, the amount of time and privacy that will be provided to potential volunteers for decision-making and whether or not the potential volunteer will be allowed to discuss the study with anyone before making a decision.*

Subjects will have 24 hours to discuss the study with anyone of their choice before making a decision whether to participate. They may opt out of participation at any time during the navigation training and they may contact the PI after the navigation training to withdraw from the study.

e. Documentation of Initial Consent*: (Check one)

☒ A consent document will be used IAW 32 CFR 219.117(a) and (b) and is enclosed for review.

☐ Request a waiver from documentation of informed consent IAW 32 CFR 219.117(c).

Reason for Request:

f. Ongoing Consent: *Address the need for obtaining ongoing consent or for re-assessing capacity over the course of a long term study and describe any relevant procedures to assure continued consent. If this does not apply, leave blank.*

g. Special Situations. *Certain situations require special procedures for obtaining informed consent, depending on the sample population being studied. Check the boxes that apply to this study and address relevant issues in the space provided.*

☐ Altered Mental Capacity *(e.g., altered capacity due to administration of any mind-altering substances such as tranquilizers, conscious sedation, or anesthesia, brain injury, stress life situations, or volunteer age)*: *Address issues relevant to the mental capacity of the potential volunteer.*

☐ Legally Authorized Representatives: *If volunteers will be included in the study that cannot give their own consent, describe the plan for the consent of the individual's legally authorized representative to be obtained prior to the volunteer's participation.*

☐ Illiterate Volunteers: *If illiterate volunteers are anticipated, explain how the consent process will be enhanced to accommodate them.*

☐ Multilingual Study: *If it is anticipated that some or all volunteers will not speak the primary language of the host country, all documentation provided to volunteers (consent form, information sheets, etc.) should be translated with a copy provided to the IRB for review. Describe the plan for obtaining consent and ensuring that volunteers' questions can be addressed during the consent process and throughout the study.*

Address issues relevant to obtaining informed consent for the situation(s) checked above. If not applicable, leave blank.

12. **Withdrawal from the Protocol**.

a. Volunteers may discontinue participation in the research at any time without penalty or loss of benefits to which the volunteer is otherwise entitled.

b. Withdrawal Procedure. *If appropriate, describe the procedure in place to support an orderly end of the volunteer's participation. (e.g., exit exam or follow-up safety visits outside of the context of the research study, information regarding prorated payment for partial participation, etc.).*

Volunteers will be given a contact card to contact LTC Irmischer to withdraw from research. It will contain his email address and mailing address.

c. Consequences of Withdrawal. *Describe any consequences of a volunteer's decision to withdraw from the study early that may negatively affect their privacy, confidentiality, health, or wellbeing.*

None

d. Criteria for Mandatory Withdrawal. *Describe any anticipated circumstances under which the volunteer's participation may be terminated by the investigator or others (e.g., non-compliance, safety issues, loss of funding, deteriorating health, etc.).*

13. **Modifications to the Protocol**. Any major modification to the protocol, consent form, or other materials (questionnaires, data collection sheets, recruitment flyers, etc.) previously approved by the IRB or a termination of the protocol before completion will be submitted to the IRB as an amendment for review and approval prior to implementation. Major modifications include, but are not limited to, a change in PI, addition of a research site, changes in study design, and addition or widening of a study population. Amendments for minor changes, such as the correction of typos on a consent form, may be submitted with the continuing review report to the IRB for acceptance.

14. **Protocol Deviations**. Any deviation to the protocol that may have an effect on the safety or rights of the volunteers or the integrity of the study will be promptly reported to the HPA.

15. **Reporting of Serious Adverse Events and Unanticipated Problems**. Reporting will be conducted in accordance with USMA Reg 70-25.

- HPA – Call 845-938-7385 or DSN 688-0761
- EDO – Call (845) 938-7385 or DSN 688-7385
- IRB Administrator – Call (845) 938-4821 or DSN 688-4821

If the IRB assigns a medical/safety monitor, the monitor will review all unanticipated problems involving risk to volunteers or others, serious adverse events, and all volunteer deaths associated with the protocol and provide an unbiased written report of the event to the IRB. At a minimum, the monitor should comment on the outcomes of the event or problem and in the case of a serious adverse event or death comment on the relationship to participation in the study. The monitor should also indicate whether he/she concurs with the details of the report provided by the study investigator. Reports for events determined by either the investigator or monitor to be possibly or definitely related to participation and reports of events resulting in death will be promptly forwarded to the AHRPO."

16. **Continuing Review and Final Report**. The continuing review and/or final reports, as applicable, will be submitted to the IRB as scheduled or as soon as these documents become available. *See the IRB SOP for guidance on required reports. See Section 5-5, USMA REG 70-25 for information about Continuing Reviews.*

17. **Principal Investigator Assurance**.

a. I understand that I am ultimately responsible for the conduct of the research.

b. I understand my responsibilities as a principal investigator as stated in USMA Reg 70-25 and I will comply with all regulatory requirements.

c. I will abide by the decision of the IRB and conduct the research in accordance with this protocol and any additional guidance provided by the IRB.

John Brockhaus
GIS Program Director
Department of Geography and Environmental Engineering

Encls
1. Bibliography/Literature Review
2. Consent Form or Partially Completed DA Form 5303
3. Principal Investigator's Curriculum Vitae
4. Principal Investigator's Training Certificate
5. Associate Investigator's Curriculum Vitae (required for each associate investigator, if any)
6. Associate Investigator's Curriculum Vitae (required for each associate investigator, if any)

Enclosure 1. Bibliography/Literature Review. *(List references used in preparing the protocol or developing the study question.)*

1. 32 CFR 219, *Protection of Human Subjects*
2. USMA REG 70-25, *Human Research Protection Program*
3. Actigraph. "Activity Monitors | Product Categories | ActiGraph." Accessed December 30, 2014. http://www.actigraphcorp.com/product-category/activity-monitors/.
4. Actigraph. "What Is the Difference among the MET Algorithms?" *ActiGraph Support*. Accessed January 2, 2015. https://help.theactigraph.com/entries/21383846-What-is-the-difference-among-the-MET-Algorithms-.
5. Actigraph Software Department. "ActiLife 6 Users Manual." ActiGraph Software Department, April 3, 2012.
6. Ainslie, Philip N., Thomas Reilly, and Klass R. Westerterp. "Estimating Human Energy Expenditure." *Sports Medicine* 33, no. 9 (August 1, 2003): 683–98. doi:10.2165/00007256-200333090-00004.
7. Aitken, Robert. "Wilderness Areas in Scotland." Ph.D. Dissertation, University of Aberdeen (United Kingdom), 1977.
8. Allen, Gary L. "Spatial Abilities, Cognitive Maps, and Wayfinding." *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, 1999, 46–80.
9. Andrienko, Gennady, Natalia Andrienko, Urska Demsar, Doris Dransch, Jason Dykes, Sara Irina Fabrikant, Mikael Jern, Menno-Jan Kraak, Heidrun Schumann, and Christian Tominski. "Space, Time and Visual Analytics." *International Journal of Geographical Information Science* 24, no. 10 (2010): 1577–1600.
10. Andrienko, Natalia, and Gennady Andrienko. "Visual Analytics of Movement: An Overview of Methods, Tools and Procedures." *Information Visualization* 12, no. 1 (January 1, 2013): 3–24. doi:10.1177/1473871612457601.
11. Berry, Joseph K. *Map Analysis: Understanding Spatial Patterns and Relationships*. GeoTec Media, 2007.
12. Brandes, Mirko, Vincent T. Van Hees, Viviane HannöVer, and Soren Brage. "Estimating Energy Expenditure from Raw Accelerometry in Three Types of Locomotion:" *Medicine & Science in Sports & Exercise* 44, no. 11 (November 2012): 2235–42. doi:10.1249/MSS.0b013e318260402b.
13. Brimicombe, Allan, and Chao Li. *Location-Based Services and Geo-Information Engineering*. Vol. 21. John Wiley & Sons, 2009.
14. Chen, Kong Y., and Ming Sun. "Improving Energy Expenditure Estimation by Using a Triaxial Accelerometer." *Journal of Applied Physiology* 83, no. 6 (1997): 2112–22.
15. Clarke, Donald D., and Louis Sokoloff. "Circulation and Energy Metabolism of the Brain." In *Basic Neurochemistry: Molecular, Cellular and Medical Aspects*, edited by George R. Siegel, R. Wayne Albers, Scott T. Brady, and Donald L. Price, 6th Edition., 6:637–69. Philadelphia, PA: Lippincott Raven, 1999.
16. Clarke, Keith C. *Analytical and Computer Cartography*. Prentice Hall PTR, 1995.
17. Clarke, Keith C. *Getting Started with Geographic Information Systems*. 5 edition. Upper Saddle River, N.J: Prentice Hall, 2010.
18. Crouter, Scott E., James R. Churilla, and David R. Bassett Jr. "Estimating Energy Expenditure Using Accelerometers." *European Journal of Applied Physiology* 98, no. 6 (2006): 601–12.
19. Crouter, Scott E., Kurt G. Clowers, and David R. Bassett Jr. "A Novel Method for Using Accelerometer Data to Predict Energy Expenditure." *Journal of Applied Physiology* 100, no. 4 (2006): 1324–31.
20. Crouter, Scott E., Erin Kuffel, Jere D. Haas, Edward A. Frongillo, and David R. Bassett. "A Refined 2-Regression Model for the ActiGraph Accelerometer." *Medicine and Science in Sports and Exercise* 42, no. 5 (May 2010): 1029–37. doi:10.1249/MSS.0b013e3181c37458.
21. Di Gregorio, Antonio, and Louisa JM Jansen. "Land Cover Classification System (LCCS): Classification Concepts and User Manual for Software Version 1.0," 2000.
22. Duggan, A., and M. F. Haisman. "Prediction of the Metabolic Cost of Walking with and without Loads." *Ergonomics* 35, no. 4 (April 1, 1992): 417–26. doi:10.1080/00140139208967822.

23. Egenhofer, Max J., and Reginald G. Golledge. *Spatial and Temporal Reasoning in Geographic Information Systems*. Oxford University Press, 1998.
24. Fisher, Peter. *Re-Presenting GIS*. John Wiley & Sons, 2005.
25. Fleury, Michelle, and Chantal Bard. "Effects of Different Types of Physical Activity on the Performance of Perceptual Tasks in Peripheral and Central Vision and Coincident Timing." *Ergonomics* 30, no. 6 (1987): 945–58.
26. Ford, L. R., and Delbert Ray Fulkerson. *Flows in Networks*. Vol. 1962. Princeton Princeton University Press, 1962. http://www.maa.org/publications/maa-reviews/flows-in-networks.
27. Freedson, Patty S., Edward Melanson, and John Sirard. "Calibration of the Computer Science and Applications, Inc. Accelerometer:" *Medicine & Science in Sports & Exercise* 30, no. 5 (May 1998): 777–81. doi:10.1097/00005768-199805000-00021.
28. Freedson P, and Miller K. "Objective Monitoring of Physical Activity Using Motion Sensors and Heart Rate." *Research Quarterly for Exercise and Sport* 71, no. 2 Suppl (June 2000): S21–29.
29. Givoni, B., and R. F. Goldman. "Predicting Metabolic Energy Cost." *Journal of Applied Physiology* 30, no. 3 (March 1, 1971): 429–33.
30. Hagerstraand, Torsten. "What about People in Regional Science?" *Papers in Regional Science* 24, no. 1 (1970): 7–24. doi:10.1111/J.1435-5597.1970.TB01464.X.
31. Hall, Cameron, Arturo Figueroa, B. O. Fernhall, and Jill A. Kanaley. "Energy Expenditure of Walking and Running: Comparison with Prediction Equations." *Medicine and Science in Sports and Exercise* 36 (2004): 2128–34.
32. Hegarty, Mary, Anthony E. Richardson, Daniel R. Montello, Kristin Lovelace, and Ilavanil Subbiah. "Development of a Self-Report Measure of Environmental Spatial Ability." *Intelligence* 30, no. 5 (2002): 425–47.
33. Jensen, John R. *Introductory Digital Image Processing 3rd Edition*. Upper Saddle River: Prentice Hall, 2005.
34. Jiang, Bin, and Xiaobai Yao. "Location-Based Services and GIS in Perspective." *Computers, Environment and Urban Systems* 30, no. 6 (2006): 712–25.
35. John, Dinesh, and Patty Freedson. "ActiGraph and Actical Physical Activity Monitors: A Peek under the Hood." *Medicine and Science in Sports and Exercise* 44, no. 1 Suppl 1 (January 2012): S86–89. doi:10.1249/MSS.0b013e3182399f5e.
36. Keytel, L. R., J. H. Goedecke, T. D. Noakes, H. Hiiloskorpi, Raija Laukkanen, L. Van Der Merwe, and E. V. Lambert. "Prediction of Energy Expenditure from Heart Rate Monitoring during Submaximal Exercise." *Journal of Sports Sciences* 23, no. 3 (2005): 289–97.
37. Kramer, Patricia Ann, and Adam D. Sylvester. "The Energetic Cost of Walking: A Comparison of Predictive Methods." *PLoS ONE* 6, no. 6 (June 22, 2011): e21290. doi:10.1371/journal.pone.0021290.
38.
39. Kuffel, Erin E., Scott E. Crouter, Jere D. Haas, Edward A. Frongillo, and David R. Bassett Jr. "Validity of Estimating Minute-by-Minute Energy Expenditure of Continuous Walking Bouts by Accelerometry." *Int J Behav Nutr Phys Act* 8 (2011): 92.
40. Kwan, Mei-Po. "Space-Time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-Based Framework." *Geographical Analysis* 30, no. 3 (1998): 191–216.
41. Kwan, Mei-Po. "Interactive Geovisualization of Activity-Travel Patterns Using Three-Dimensional Geographical Information Systems: A Methodological Exploration with a Large Data Set." *Transportation Research Part C: Emerging Technologies* 8, no. 1 (2000): 185–203.
42. Langmuir, Eric. *Mountain Leadership: The Official Handbook of the Mountain Leadership Training Boards of Great Britain*. Edinburgh: Scottish Council Of Physical Recreation, 1969.
43. Langmuir, Eric, Scottish Sports Council, and Mountain Leader Training Board. *Mountaincraft and Leadership: A Handbook for Mountaineers and Hillwalking Leaders in the British Isles*. Scottish Sports Council, 1984.
44. Langran, Gail. *Time in Geographic Information Systems*. CRC Press, 1992.

45. Lee, Jung-Min, Youngwon Kim, and Gregory J. Welk. "Validity of Consumer-Based Physical Activity Monitors:" *Medicine & Science in Sports & Exercise* 46, no. 9 (September 2014): 1840–48. doi:10.1249/MSS.0000000000000287.

46. Leyk, D., U. Rohde, W. Gorges, D. Ridder, M. Wunderlich, C. Dinklage, A. Sievert, T. Rüther, and D. Essfeld. "Physical Performance, Body Weight and BMI of Young Adults in Germany 2000 - 2004: Results of the Physical-Fitness-Test Study." *International Journal of Sports Medicine* 27, no. 8 (August 2006): 642–47. doi:10.1055/s-2005-872907.

47. Lennie, Peter. "The Cost of Cortical Computation." *Current Biology* 13, no. 6 (2003): 493–97.

48. Leonard, William R. "Laboratory and Field Methods for Measuring Human Energy Expenditure." *American Journal of Human Biology: The Official Journal of the Human Biology Council* 24, no. 3 (June 2012): 372–84. doi:10.1002/ajhb.22260.

49. Levine, James A. "Measurement of Energy Expenditure." *Public Health Nutrition* 8, no. 7a (October 2005): 1123–32. doi:10.1079/PHN2005800.

50. Li, Zhilin, Christopher Zhu, and Chris Gold. *Digital Terrain Modeling: Principles and Methodology*. CRC press, 2010.

51. McMinn, David, Rijan Acharya, David Rowe, Stuart Gray, and Julia Allan. "Measuring Activity Energy Expenditure: Accuracy of the GT3X and Actiheart Monitors." *International Journal of Exercise Science* 6, no. 3 (July 16, 2013). http://digitalcommons.wku.edu/ijes/vol6/iss3/5.

52. McMurray, Robert G., and Kristin S. Ondrak. "Energy Expenditure of Athletes." In *Sports Nutrition: Energy Metabolism and Exercise*, 127, 2008.

53. Miller, Harvey J. "Modelling Accessibility Using Space-Time Prism Concepts within Geographical Information Systems." *International Journal of Geographical Information System* 5, no. 3 (1991): 287–301.

54. Montello, Daniel R. "Navigation." In *The Cambridge Handbook of Visuospatial Thinking*, edited by P Shah and A Miyake, 257–94. Cambridge: Cambridge University Press, 2005.

55. Muehlenbein, Michael P., Jana L. Hirschtick, Julia Z. Bonner, and Ann M. Swartz. "Toward Quantifying the Usage Costs of Human Immunity: Altered Metabolic Rates and Hormone Levels during Acute Immune Activation in Men." *American Journal of Human Biology* 22, no. 4 (July 1, 2010): 546–56. doi:10.1002/ajhb.21045.

56. National Research Council (U.S.), National Research Council (U.S.), National Research Council (U.S.), and National Research Council (U.S.). *New Research Directions for the National Geospatial-Intelligence Agency: Workshop Report.* Washington, D.C: The National Academies Press, 2010.

57. Pande, A., Y. Zeng, A. Das, P. Mohapatra, S. Miyamoto, E. Seto, E. K. Henricson, and J. J. Han. "Accurate Energy Expenditure Estimation Using Smartphone Sensors." In *ACM Wireless Health 2013*, –, 2013.

58. Pandolf, K. B., B. Givoni, and R. F. Goldman. "Predicting Energy Expenditure with Loads While Standing or Walking Very Slowly." *Journal of Applied Physiology*, 1977.

59. Pettee, K., Catrine Tudor-Locke, and Barbara E. Ainsworth. "The Measurement of Energy Expenditure and Physical Activity." In *Sports Nutrition. Energy Metabolism and Exercise*, 159–89, 2008.

60. Peuquet, Donna 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.

61. Potter, Adam W., William R. Santee, Cynthia M. Clements, Kelly A. Brooks, and Reed W. Hoyt. "Comparative Analysis of Metabolic Cost Equations: A Review." *Journal of Sport and Human Performance* 1, no. 3 (September 18, 2013). doi:10.12922/9.

62. Pred, Allan. "The Choreography of Existence: Comments on Hägerstrand's Time-Geography and Its Usefulness." *Economic Geography* 53, no. 2 (April 1, 1977): 207–21. doi:10.2307/142726.

63. Richards-Rissetto, Heather, and Kristin Landau. "Movement as a Means of Social (re)production: Using GIS to Measure Social Integration across Urban Landscapes." *Journal of Archaeological Science* 41 (January 2014): 365–75. doi:10.1016/j.jas.2013.08.006.

64. Robinson, Arthur Howard, Joel L. Morrison, Phillip C. Muehrcke, A. Jon Kimerling, and Stephen C. Guptill. *Elements of Cartography*. Vol. 6. Wiley New York, 1995.

65. Rodríguez-Pérez, J., M. Álvarez, and E. Sanz-Ablanedo. "Assessment of Low-Cost GPS Receiver Accuracy and Precision in Forest Environments." *Journal of Surveying Engineering* 133, no. 4 (2007): 159–67. doi:10.1061/(ASCE)0733-9453(2007)133:4(159).

66. Santee, William R., Laurie A. Blanchard, Karen L. Speckman, Julio A. Gonzalez, and Robert F. Wallace. *Load Carriage Model Development and Testing with Field Data*, June 2003.

67. Sauer, Carl O. "Foreword to Historical Geography." *Annals of the Association of American Geographers* 31, no. 1 (March 1, 1941): 1–24. doi:10.2307/2560961.

68. Shaw, Shih-Lung, and Hongbo Yu. "A GIS-Based Time-Geographic Approach of Studying Individual Activities and Interactions in a Hybrid Physical–virtual Space." *Journal of Transport Geography* 17, no. 2 (2009): 141–49.

69. Soule, Roger G., and Ralph F. Goldman. "Terrain Coefficients for Energy Cost Prediction." *J Appl Physiol* 32, no. 5 (1972): 706–8.

70. Swartz, Ann M., Scott J. Strath, David R. Bassett Jr, William L. O'Brien, George A. King, and Barbara E. Ainsworth. "Estimation of Energy Expenditure Using CSA Accelerometers at Hip and Wrist Sites." *Medicine and Science in Sports and Exercise* 32, no. 9 Suppl (2000): S450–56.

71. Tapia, Emmanuuel. "Using Machine Learning for Real-Time Activitiy Recognition and Estimation of Energy Expenditure." Massachusetts Institute of Technology, 2008.

72. Tharion, William J., Scott J. Montain, Catherine O'Brien, Ronald L. Shippee, and Jennifer L. Hoban. "Effects of Military Exercise Tasks and a Carbohydrate-Electrolyte Drink on Rifle Shooting Performance in Two Shooting Positions." *International Journal of Industrial Ergonomics* 19, no. 1 (January 1997): 31–39. doi:10.1016/0169-8141(95)00087-9.

73. Thomas, James J., and Kristin A. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society Press, 2005.

74. Thomas, James J., and Kristin A. Cook. "A Visual Analytics Agenda." *Computer Graphics and Applications, IEEE* 26, no. 1 (2006): 10–13.

75. Tobler, Waldo R. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46 (1970): 234–40.

76. Tobler, Waldo R. "Non-Isotropic Geographic Modeling." Santa Barbara, California: National Center for Geographic Information and Analysis, 1993.

77. "Trimble - GPS Tutorial - Why GPS?" Accessed October 15, 2014. http://www.trimble.com/gps_tutorial/.

78. Reynolds, Stephen C., and Richard L. Taylor. *Digital Terrain Data in Support of Land Combat Models*. DTIC Document, 1988.

79. Slocum, T. A., R. B. McMaster, F. C. Kessler, and H. H. Howard. *Thematic Cartography and Geographic Visualization. 2005*. Pearson, Prentice-Hall, n.d.

80. U.S. Army. *Foot Marches*. Department of the Army, FM 21-18. Washington D.C.: Government Printing Office, 1990.

81. U.S. Army. *Map Reading and Land Navigation*. Department of the Army, TC 3-25.26. Washington, D.C.: Government Printing Office, 2013.

82. U.S. Army. *Soldier's Manual of Common Tasks: Warrior Leader Skills Level 1*. STP 21-1-SMCT. Washington, D.C.: Government Printing Office, 2012.

83. Welk, Gregory J., Jodee A. Schaben, and James R. Morrow Jr. "Reliability of Accelerometry-Based Activity Monitors: A Generalizability Study." *Medicine and Science in Sports and Exercise* 36, no. 9 (2004): 1637–45.

84. Willems, Niels, Huub Van De Wetering, and Jarke J. Van Wijk. "Visualization of Vessel Movements." *Computer Graphics Forum* 28, no. 3 (2009): 959–66. doi:10.1111/j.1467-8659.2009.01440.x.

85. Willems, Niels, Roeland Scheepens, Huub van de Wetering, and Jarke J. van Wijk. "Visualization of Vessel Traffic." In *Situation Awareness with Systems of Systems*, edited by Piërre van de Laar, Jan Tretmans, and Michael Borth, 73–87. Springer New York, 2013.

86. Wilson, John P., and John C. Gallant. "Digital Terrain Analysis." In *Terrain Analysis: Principles and Applications*. John Wiley & Sons, 2000.

87. Wolinsky, Ira, and Judy Driskell. *Sports Nutrition: Energy Metabolosm and Exercise*. Boca Raton, FL: CRC Press, 2008. http://www.crcpress.com/product/isbn/9780849379505.

88. Yuan, May, and Kathleen S. Hornsby. *Computation and Visualization for Understanding Dynamics in Geographic Domains: A Research Agenda*. CRC Press, 2007.

IRB Consent Form

KELLER ARMY COMMUNITY HOSPITAL
IRB
APPROVED:
JULY 9, 2015
APPROVAL EXPIRES:
JULY 8, 2016

UNITED STATES MILITARY ACADEMY
CONSENT TO PARTICIPATE IN RESEARCH

Study Title: **Space, Time and Energy in Dismounted Navigation**

You are asked to participate in a research study conducted at the United States Military Academy Land Navigation Training Area by Dr. John Brockhaus. Your participation in this study is voluntary. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

PURPOSE OF THE STUDY

The goal of this dissertation research is to model human energy expenditure (EE) of foot-based navigation. It will focus on military navigation which is referred to as dismounted navigation when the movement is conducted via walking. Soldiers in a combat environment perform less effectively if they are excessively fatigued from movement. Understanding energy expenditure of a route is a vital element that should be considered when moving troops or supplies. Accurate assessment of energy expenditure is imperative to reducing operational risk. Currently no tools that fuse the physical terrain costs with human performance factors have been validated. This project will improve geospatial tools available to our military that focus on navigation. Improving geospatial-related technologies will support the United States Army Warfighter through enhanced knowledge of the environment, his/her capabilities, opportunities and risk.

The core contributions and the research products that will be developed are specifically focused on the dismounted routing implications of human energy expenditure. First, the research will define what are the relative weights assigned to individual characteristics that contribute to energy expenditure. Secondarily, the investigation will determine how significant environmental and terrain effects contribute to energy expenditure. Finally, the study will assess current energy expenditure models and provide improvements to existing military routing tools.

EXPECTED DURATION OF PARTICIPATION

You will be expected to spend 10-15 minutes during day 2 of land navigation understanding the study and providing height weight measurements. Then you will be expected to wear an accelerometer and a heart rate monitor for approximately 2 hours the next day (day 3) while you conduct land navigation training.

PROCEDURES

If you volunteer to participate in this study, we will ask you to do the following things:

On day 2 of land navigation training you will be asked to have your height and weight measured in order for the accelerometer to be calibrated. On day 3 of land navigation training you will be wearing a small accelerometer on your waist that is the size of a

Version 1.0, dated 8 July 2015

small pager. You will also be wearing a heart rate monitor. You will wear these devices while you navigate through the woods and will return the accelerometer and the heart rate monitor devices to the investigator when you have completed your land navigation course.

POTENTIAL RISKS AND DISCOMFORTS
The accelerometer and the heart rate monitor may cause slight chaffing during the event. The devices are designed to minimize chaffing by using smooth flexible plastic as the surface that touches the body. It is not likely that the chaffing will be greater than wearing a belt or undergarments.

ANTICIPATED BENEFITS
This research will collect energy expenditure estimates of Soldiers while they navigate on foot over hilly, wooded terrain. The estimates will provide an opportunity to model the contribution of both individual and terrain factors to energy expenditure during dismounted navigation. Knowledge of the influences contributing to energy expenditure will provide a basis for assessing and improving current GIS routing tools. Improving routing algorithms and geospatial tools for the mission planner will be a great improvement to current capabilities. It will increase operational safety and military success. Finally, it will provide better tools to access information about how the environment and Soldier capabilities affect missions.

MEDICAL CARE FOR RESEARCH RELATED INJURY
Should you be injured as a result of your participation in this study, you will be given medical care for that injury at no cost to you. Medical care is limited to the care normally allowed for Department of Defense health care beneficiaries (patients eligible for care at military hospitals and clinics).

NUMBER OF PEOPLE THAT WILL TAKE PART IN THIS STUDY
Approximately 240 cadets participating in land navigation testing during Cadet Basic Training (CBT) in July and August 2015 will be enrolled in this study.

CONFIDENTIALITY
The principal investigator will keep your research records. These records may be looked at by staff from the Keller Army Community Hospital Institutional Review Board (IRB), the Army Clinical Investigation Regulatory Office (CIRO), and other government agencies as part of their duties. These duties include making sure that the research participants are protected. Confidentiality of your records will be protected to the extent possible under existing regulations and laws but cannot be guaranteed. Complete confidentiality cannot be promised, particularly for military personnel, because information bearing on your health may be required to be reported to appropriate medical or command authorities. Your name will not appear in any published paper or presentation related to this study.

COMPENSATION FOR PARTICIPATION

You will not receive any payment for being in this study.

PARTICIPATION AND WITHDRAWAL BY YOU
Your participation in this research is voluntary. If you choose not to participate, that will not affect your relationship with investigators, the United States Military Academy or your right to health care or other benefits or services to which you are otherwise entitled. If you decide to participate, you are free to withdraw your consent and discontinue participation at any time without prejudice.

WITHDRAWAL OF PARTICIPATION BY THE INVESTIGATOR
The investigator may withdraw you from participating in this research if circumstances arise which warrant doing so. The investigator will make the decision and let you know if it is not possible for you to continue. The decision may be made either to protect your health and safety, or because it is part of the research.

NEW FINDINGS
During the course of the study, you will be informed of any significant new findings (either good or bad), such as changes in the risks or benefits resulting from participation in the research or new alternatives to participation, that might cause you to change your mind about continuing in the study. If new information is provided to you, your consent to continue participating in this study will be re-obtained.

POINTS OF CONTACT
In the event of a research related injury or if you experience an adverse reaction, immediately contact the following:

Dr. John Brockhaus, 845.938.2063, bj9296@usma.edu
LTC Ian Irmischer, 808-312-7047, ian.j.irmischer.mil@mail.mil

If you have specific questions about the conduct of the research, please contact one of the investigators listed below.

LTC Ian Irmischer, 808-312-7047, ian.j.irmischer.mil@mail.mil

If you have any questions about your rights as a volunteer in the research, please feel free to contact the USMA Human Protections Administrator at 845-938-7370.

RIGHTS OF RESEARCH SUBJECTS
You may withdraw your consent at any time and discontinue participation without penalty. You are not waiving any legal claims, rights or remedies because of your participation in this research study.

SIGNATURE OF RESEARCH SUBJECT
I have read the information provided above. I have been given an opportunity to ask questions and all of my questions have been answered to my satisfaction. I have been given a copy of this form.

____
Name of Subject

____________________________________ ____________
Signature of Subject Date

____
Address

SIGNATURE OF WITNESS
My signature as witness certifies that the subject signed this consent form in my presence as his/her voluntary act and deed.

____
Name of Witness

____________________________________ __________________________
Signature of Witness Date (same as subject's)

Version 1.0, dated 8 July 2015

UCSB Human Subjects Approval Letter

UNIVERSITY OF CALIFORNIA

BERKELEY • DAVIS • IRVINE • LOS ANGELES • MERCED • RIVERSIDE • SAN DIEGO • SAN FRANCISCO • SANTA BARBARA • SANTA CRUZ

SANTA BARBARA

Office of Research
Human Subjects Committee
Santa Barbara, CA 93106-2050

Web: http://www.research.ucsb.edu

12/21/2015

VERIFICATION OF ACTION BY THE UCSB HUMAN SUBJECTS COMMITTEE

RE: HUMAN SUBJECTS PROTOCOL SUBMISSION ID 15-0313

TYPE: New

TITLE(S):
SPACE, TIME AND ENERGY IN DISMOUNTED NAVIGATION

PROJECT #: 20

INVESTIGATOR(S):
Keith Clarke
Ian Irmischer

The UCSB Human Subjects Committee (HSC) conducts review of human subjects research under FederalWide Assurance #FWA00006361 and in compliance with applicable regulations as described therein.

The above identified research project has undergone review and approval by the USCB HSC and may commence on 7/9/2015. It was approved by Expedited review.

Date of expiration of HSC approval: 7/8/2016

RENEWAL OF PROTOCOLS

If you wish to continue your research beyond the above expiration date, your protocol must be renewed before it expires. To ensure that your research can continue uninterrupted, the following schedule should be followed:

Full Board Review: Submit 5 weeks before expiration date.
Expedited Review: Submit 3 weeks before expiration date.
Exemption Review: Does not expire. Resubmit only if changes are made.

All research must cease under this protocol on its expiration date unless you have received notice of renewal from the HSC.

AMENDMENTS/MODIFICATIONS/CHANGES:

Any change in the design, conduct, or key personnel of this research must be approved by the HSC prior to implementation.

UNANTICIPATED PROBLEMS/ADVERSE EVENTS:

If any study subject experiences an unanticipated problem involving risk to subjects or others, and/or a serious adverse event, the HSC must be informed promptly. An e-mail or phone call must be received within 7 days of reporting to the Investigator(s). Further reporting requirements will be determined by the HSC at that time.

If you have any questions about the information provided above, please contact the, Human Subjects Committee Coordinator at:

805-893-3807
(805) 893-2611 (fax)
hsc@research.ucsb.edu

For more details on this protocol, go to the ORahs website: https://orahs.research.ucsb.edu/

For more information about human subjects research, go to http://www.research.ucsb.edu/compliance/human-subjects/.

USMA IRB Approval Letter

DEPARTMENT OF THE ARMY
U.S. ARMY MEDICAL DEPARTMENT ACTIVITY
WEST POINT, NEW YORK 10996-1197

REPLY TO
ATTENTION OF:

MCUD 10 July 2015

MEMORANDUM FOR John Brockhaus, PhD, GIS Program Director, Department of Geography and Environmental Engineering, United States Military Academy, West Point, New York 10996

SUBJECT: Notification of Initial IRB approval for KACH Protocol 15-020, IRBNet# 413717-1, Space, Time and Energy in Dismounted Navigation, PI: John Brockhaus, PhD

1. Expedited review of the study protocol (version dated 17 June 2015), informed consent (version dated 8 July 2015), and the Data Collection Form was conducted and approved by the IRB Chair on 9 July 2015 in accordance with 32 CFR 219.110 (b) (1), Category 4.

2. The KACH IRB was informed of the expedited review approval at the convened IRB meeting of 9 July 2015.

3. There are no outstanding human subjects' protection issues to be resolved. The study has been assessed as Minimal Risk.

4. You may implement/start your study with the issuance of this KACH IRB approval letter dated 10 July 2015.

5. The study is approved for enrollment of up to 240 subjects.

6. The IRB Chair approved the initiation of this Minimal Risk study for a one-year period effective 9 July 2015 – 8 July 2016. The IRB stamped approved informed consent (version dated 8 July 2015) should be used when consenting subjects.

7. In accordance with 32 CFR 219.109(e), the Principal Investigator must submit a continuing review report for this protocol to the KACH IRB. A continuing review report with a copy of the current protocol and informed consent must be submitted by 8 June 2016 to ensure approval on or before 8 July 2016.

8. The Principal Investigator is responsible for fulfilling reporting requirements to the KACH IRB.

JEFFREY A. LAWSON
COL, MC
Chair, KACH IRB

IRB Scientific Review Approval

APPENDIX A-1

DEPARTMENT OF THE ARMY
U.S ARMY MEDICAL DEPARTMENT ACTIVITY
REPLY TO ATTENTION OF
West Point, New York 10996-1197

Date 31 March 2015

RE: SCIENTIFIC REVIEW OF RESEARCH PROTOCOL – **Space, Time and Energy in Dismounted Navigation**, PI – Dr. John Brockhaus

1. **Description:** Navigation on foot through the wilderness is used by hikers, search and rescue, firefighters, the military, and many others. The human dynamics of foot-based navigation are critical to understanding individual capabilities, opportunities, requirements, and risk. Models of energy expenditure for dismounted navigation are not sufficient to assist users during route planning and logistical estimation. Improving our understanding of calorie consumption, during dismounted land navigation, will provide benefit across humanity. Hikers will be able to estimate caloric needs based on trail choice. Search and rescue will improve load planning and routing. Societal uses of an energy expenditure model for on foot navigation are infinite.

One of the specific focuses of this research is military land navigation. Military land navigation is a basic skill required of every Soldier and Marine. It is taught at each form of individual basic training, and during all US Army leadership schools. One of the major components of land navigation is route selection (U.S. Army 2013). Route selection is an element of every military operation involving dismounted troops. There are many geospatial tools available to assist Soldiers and Marines during route selection. However, no validated tools address the variable of consumed energy. Energy consumption prediction, and exhaustion avoidance, is critical when planning military operations and patrols. The omission of tools that provide route planning assistance based on potential exhaustion, leaves a gap in current US military capability.

2. **Standard Review Criteria:**

Significance: This research will collect energy expenditure estimates of Soldiers, while they navigate on foot over hilly, wooded terrain. The estimates will provide an opportunity to model the contribution of both individual, and terrain factors, to energy expenditure during dismounted navigation. Knowledge of the influences contributing to energy expenditure, will provide a basis for assessing and improving current GIS routing tools. Improving routing algorithms, and geospatial tools for the mission planner, will be a great improvement to current capabilities. It will increase operational safety, and military success. Finally, it will provide better tools to access information about how the environment, and Soldier capabilities affect missions.

Approach: This research will collect energy expenditure estimates of Soldiers while they navigate on foot over hilly, wooded terrain. The estimates will provide an opportunity to model the contribution of both individual and terrain factors to energy expenditure during dismounted navigation. Knowledge of the influences contributing to energy expenditure will provide a basis for assessing and improving current GIS routing tools. The research will be comprised of several experiments to gather information about how much human energy is expended while cadets at the United States Military Academy navigate. The experiments will collect movement, location, and time data using wearable devices. The devices will be affixed to cadets while they participate in a navigational test that is part of their summer military training. The movement, location, and time data will be collected over varying terrain conditions including trees, swamps, roads, trails, brush, hills and cliffs. Finally, I will obtain individual characteristic data from the United States Military Academy including: height, weight, fitness level, and sex.

Investigator: Dr. John Brockhaus

Environment: United States Military Academy Military Reservation

3. **Overall Evaluation:**

Synopsis: The purpose of this research is to model human energy expenditure of foot-based navigation. First, the research will define what are the relative weights assigned to individual characteristics that contribute to energy expenditure. Secondarily, the investigation will determine how significant environmental and terrain effects contribute to energy expenditure. Finally, the study will assess current energy expenditure models and provide improvements to existing military route planning tools.

Strengths: Availability of human subjects, the sample size, and study site provide a unique set of circumstances for the study of dismounted land navigation. Conditions such as these have not been available to researchers in previous studies of this type.

Weaknesses: None

Impact: This research will significantly add to the understanding of the impact of terrain on physical exertion during dismounted land navigation.

Military Relevancy: This research will significantly add to the understanding of the impact of terrain and environmental factors on physical exertion during military land navigation training.

Literature Review: The researcher, MAJ Ian Irmischer, has completed an exhaustive and current review of the scientific literature pertaining to land navigation.

Human Use Issues: Human subjects are required for the completion of this research. However, monitoring of human subjects will be conducted in conjunction with existing land navigation training exercises. Each human subject will be outfitted with minimally invasive equipment, a wrist mounted Actigraph.

Specific recommendations for improvement: None

4. ***Make one of the following 5 recommendations for the protocol:***

Approve without modification

CONFLICT OF INTEREST

There is no financial or professional interest or personal circumstance that will impair my ability to provide an objective review. I understand the confidential nature of the protocol and agree to destroy or return all review-related materials and to not discuss these materials or review proceedings with any individual except the Scientific Review Committee Chairperson.

COL Wiley Thompson, Head
Dept. of Geography & Environmental Engineering
United States Military Academy, West Point NY

Date: 31 March 2015

____
Signature

Department of Defense Individual Investigator Agreement

US Military Academy at West Point

Human Research Protection Program

DoD INDIVIDUAL INVESTIGATOR AGREEMENT

Part 1
AGREEMENT INFORMATION

This Department of the Defense (DoD) Individual Investigator Agreement describes the responsibilities of the individual who is engaged in human subject research and is not an employee of the assured institution, but is associated with the assured institution for the purpose of conducting research. This Agreement also describes the responsibilities of the assured institution. This Agreement, when signed, becomes part of the institution's Federal Assurance for the Protection of Human Research Subjects (e.g., DoD Assurance or Department of Health and Human Services (DHHS) Federal Wide Assurance (FWA)).

A. Name of Investigator: Ian J. Irmischer

B. Institution with the Assurance:

Name: USA MEDDAC – West Point
DoD Assurance Number: DOD A 2015
DHHS FWA Number [if applicable]:

C. Scope:

This Agreement applies to all research performed by this Investigator and supported by the Institution with the Assurance, unless specified below.

Limitation of Scope (if applicable):

D. Effective Date:

This Agreement is effective as of the date signed by the DoD Component Designated Official and expires on the date listed in Part 4, paragraph D.

NOTE: This Agreement is applicable only to the research listed above and does not apply to other research in which the investigator may be involved.

1

Part 2
INVESTIGATOR RESPONSIBILITIES

As the Investigator named above, I:

A. Have reviewed: a) *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*; b) the U.S. Department of Defense (DoD) regulations for the protection of human subjects at 32 Code of Federal Regulations, Part 219 (32 CFR 219) and DoD Directive 3216.02; c) the Assurance of the institution referenced above; d) the DoD Component policies identified in Part 3 of the DoD Assurance (if applicable); and e) the relevant institutional policies and procedures for the protection of human subjects.

B. Understand and accept the responsibility to comply with the standards and requirements stipulated in the above documents and to protect the rights and welfare of human subjects involved in research conducted under this Agreement.

C. Will comply with all other applicable federal, DoD, international, state, and local laws, regulations, and policies that provide protections for human subjects participating in research conducted under this Agreement.

D. Will complete any education and training required by the Institution and the IRB prior to initiating research covered under this Agreement (attach documentation).

E. Will abide by all determinations of the Institutional Review Board(s) (IRB) designated under the Institution's Assurance and will accept the final authority and decisions of the IRB, including but not limited to directives to terminate my participation in designated research activities.

F. Will not enroll subjects or start research activities under this Agreement prior to its review and approval by the IRB and the Institution.

G. Will comply with requirements from the IRB when responsible for enrolling subjects, to include obtaining, documenting, and maintaining records of informed consent for each such subject or each subject's legally authorized representative as required under DoD regulations at 32 CFR 219.

H. Acknowledge and agree to cooperate with the IRB for initial and continuing review, report for the research referenced above, and provide all information requested by the IRB or Institution in a timely fashion.

I. Will seek prior IRB review and approval for all proposed changes in the research except where necessary to eliminate apparent immediate hazards to subjects or others.

J. Will report immediately to the IRB a) unanticipated problems involving risks to subjects or others and b) serious or continuing non-compliance

K. Will comply with recordkeeping requirements for research protocols referenced above.

L. Will make all other notifications as specified by the IRB and the Institution.

M. Acknowledge my primary responsibility for safeguarding the rights and welfare of each research subject, and that the subject's rights and welfare will take precedence over the goals and requirements of the research.

Part 3
ASSURED INSTITUTION'S RESPONSIBILITIES

This Institution will apply the terms of its assurance to the Investigator and the research as specified in the scope of this Agreement, Part 1.

Part 4
AGREEMENT BETWEEN AN INVESTIGATOR AND AN ASSURED INSTITUTION

The Investigator or an official of the assured Institution may unilaterally terminate this agreement upon written notification to other signatories.

A. Investigator:

I understand my responsibilities as described in this Agreement and the policies referenced in Part 2A above. I acknowledge and accept my responsibility for protecting the rights and welfare of human research subjects and for complying with all applicable provisions of the Institution's Assurance.

Signature: Date: 14April2015
Name: Ian J. Irmischer (LTC, US Army)
Title: PhD Student (ACS Fellowship)
Telephone number: 808-312-7047
FAX number: N/A
Email address:ian.irmischer@gmail.com

Mailing Address:7448 San Carpino Drive, Goleta Ca 93117

B. Acknowledgement by Investigator's Institution (or DoD Supervisor if DoD Employee):

I am aware that Ian Irmischer is entering into this agreement.

Signature: Michael Witherell
Date: 4/14/2015

Name: Michael Witherell, PhD
Title: Vice Chancellor for Research
Telephone number: (805) 893-8270
FAX number: (805) 893-2611
Email address: witherell@research.ucsb.edu

Mailing Address: 3227 Cheadle Hall
Santa Barbara, CA 93106-2050

C. Institutional Official of the Assured Institution:

Acting in an authorized capacity on behalf of this Institution and with an understanding of the Institution's responsibilities under the Institution's Assurance, I will provide oversight of the Investigator and the research conducted under this Agreement.

Signature:
Date:

Name: BG Timothy Trainor
Institutional Title: Dean of the Academic Board/Institutional Official
Telephone Number: 845-938-2000
FAX Number: 845-938-2202
Email Address: Timothy.Trainor@usma.edu

Mailing Address:
US Military Academy at West Point
646 Swift Road
Office of the Dean
West Point, NY 10996

IRB Authorization Agreement from UCSB

Print Form

Institutional Review Board (IRB)/Independent Ethics Committee (IEC) Authorization Agreement

Name of Institution or Organization Providing IRB Review (Institution/Organization A):
United States Military Academy

IRB Registration #: DOD A 2015 Federalwide Assurance (FWA) #, if any: ______

Name of Institution Relying on the Designated IRB (Institution/Organization B):
University of California, Santa Barbara (UCSB) FWA #: FWA00006361

The Officials signing below agree that University of California, Santa Barbara (UCSB) may rely on the designated IRB for review and continuing oversight of its human subjects research described below: (*check one*)

☐ This agreement applies to all human subjects research covered by Institution B's FWA.

☒ This agreement is limited to the following specific protocol(s):

Name of Research Project: Space, Time and Energy in Dismounted Navigation

Name of Principal Investigator (Institute A): Dr. John Brockhaus

Name of Principal Investigator (Institute B): Ian Irmischer and Professor Keith Clarke (Faculty Advisor)

Sponsor or Funding Agency: Army Geospatial Lab Award Number, if any: ______

☐ Other (*describe*) ______

The review performed by the designated IRB will meet the human subject protection requirements of Institution B's OHRP-approved FWA. The IRB at Institution/Organization A will follow written procedures for reporting its findings and actions to appropriate officials at Institution B. Relevant minutes of IRB meetings will be made available to Institution B upon request. Institution B remains responsible for ensuring compliance with the IRB's determinations and with the Terms of its OHRP-approved FWA. This document must be kept on file by both parties and provided to OHRP upon request.

Signature of Signatory Official (Institution/Organization A):

Timothy E. Trainor Date: 16 April 2015

Print Full Name: BG Timothy Trainor Institutional Title: Dean of the Academic Board

Signature of Signatory Official (Institution/Organization B):

Michael Witherell Date: 4/27/2015

Print Full Name: Michael Witherell Institutional Title: Vice Chancellor for Research

IRB Data Collection Form

DATA COLLECTION FORM

DATA COLLECTED	SUBJECT 1	SUBJECT 2 ECT
Date		
Last Name		
First Name		
Middle Name		
COMPANY		
PLATOON		
SQUAD		
HEIGHT		
WEIGHT		
Activity Monitor #		
GPS #		

APPENDIX B – Computer Programming

Python Script to synchronize subject's navigation time

```
import arcpy
import tkFileDialog
import datetime
dirname=tkFileDialog.askdirectory(initialdir="C:/Users/ian/Dropbox/West Point Data 2015",title='Please
select a directory')
arcpy.env.workspace=dirname
arcpy.env.overwriteOutput=True
allFC=arcpy.ListFeatureClasses(wild_card="*tracks") # create a list of all feature classes


codeblock='''
total = -10
def cumsum(inc):
   global total
   total+=inc
   return total'''

#d = datetime.datetime.strptime("2016-03-29 06:38:00","%Y-%m-%d %H:%M:%S")

exp = '''
def add_date():
  return datetime.datetime.strptime("2016-03-29 06:38:00","%Y-%m-%d %H:%M:%S")'''


for files in allFC:
  arcpy.AddField_management(files, "cumtime3", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED",
"")
  arcpy.AddField_management(files, "futureStart", "date", "", "", "", "", "NULLABLE", "NON_REQUIRED",
"")
  arcpy.AddField_management(files, "futureTime", "date", "", "", "", "", "NULLABLE", "NON_REQUIRED",
"")
  arcpy.CalculateField_management(files, "cumtime3", "cumsum(!DURATION!)", "PYTHON_9.3", codeblock)
  arcpy.CalculateField_management(files, "futureStart","add_date()", "PYTHON_9.3",exp)
  arcpy.CalculateField_management(files,
"futureTime","datetime.datetime.strptime(!futureStart!,'%m/%d/%Y %H:%M:%S %p') +
datetime.timedelta(seconds=!cumtime3!)", "PYTHON_9.3")
```

R Script to join GNSS data and MET data

```
## Code to Join the GPS location data and the METs estimates from the accelerometer data.
## Code to look through folders and join the _q file with the “lastname1sec10sec AGD Details Epochs
        2015-08-27_01-57-48”
## This combines the lat/long with the METS. The output file is saved in the folder lastnameJoined.
## Caution - "Only folders with the GPS data and MET data can be in the folder. Other folders in the
        parent directory will throw an error."

#Select the parent folder that this code will crawl through to find all GPS files and MET scoring data
dirList <- list.dirs(choose.dir(caption = "Select folder"),full.names=TRUE, recursive = FALSE)

# Loop through each folder in the parent folder to find the GPS (_q) file and METS ("details") file
for (i in 1:length(dirList)){

 #Crawls through the above selected folder to find your GPS Files with"_q" pattern
 corGPSfilename <- list.files(dirList[i],pattern="*_q",ignore.case = TRUE)

 #Crawl through the above selected folder to find your MET Files with "Details" pattern
 METdatafilename <- list.files(dirList[i],pattern="Details")
```

```
 # Assign the directory to the file name
 corGPSfilename=paste(dirList[i],"/",corGPSfilename,sep="")
 METdatafilename=paste(dirList[i],"/",METdatafilename,sep="")

 #extract the subject name from the file name
 directory <- dirList[i]
 subjectFolder <- basename(directory)
 subjectName <- substring(subjectFolder,4)

 # Makes sure that the subjectName tested is lower case.
 subjectName <- tolower(subjectName)

 #Read the files into R
 corGPS <- read.csv(corGPSfilename,header=TRUE,sep=",")
 METdata <- read.csv(METdatafilename,skip=1,header=TRUE,sep=",")

 # Do some minor adjustments to the data tables

 # Start with MET data. Get date time information into a usable and joinable format
 a <- as.character(METdata$epoch)
 b <- as.character(METdata$date)
 c <- paste(b,a)
 METdata$datetime<- format(strptime(c,"%m/%d/%Y %H:%M:%S %p" ),"%Y/%m/%d %H:%M:%S")

 #Now to alter the corGPS data a little to get date time information into a usable and joinable format
 a <- gsub(" ", "", corGPS$LOCAL.DATE, fixed = TRUE)
 b <- gsub(" ", "", corGPS$LOCAL.TIME, fixed = TRUE)
 corGPS$datetime <- paste(a,b)

 # Change Longitude to negative to represent West of prime meridian.
 corGPS$LONGITUDE <- corGPS$LONGITUDE*-1

 # Now we join the data and write it to a table
 combinedData <- merge(x=METdata,y=corGPS,by.x="datetime",by.y="datetime")
 z <- combinedData [complete.cases(combinedData [,23]),]
 joinedTableName=paste(directory,"/",subjectName,"Joined.csv",sep="")
 write.table(format(z, digits=10),joinedTableName,sep=",",row.names=FALSE)
}
```

R Script to Add Administrative Data to GNSS/METs file

```
## This Script takes administrative data such as Last Name, First Name, Height, and Weight from a
spreadsheet and adds that to files that have GNSS data and METs information.
## Caution: Spreadsheet Last Name and Last Name on the data file must be exactly the same
## Caution: Problems arise if there are duplicate last names

# Choose the parent directory that contains the GNSS/METs data files
# and Select the GNSS/METs files that have a pattern ="joined"
fileList <- list.files(choose.dir( caption = "Select folder"),full.names=TRUE, pattern="joined",
recursive=TRUE, include.dirs=TRUE,ignore.case=TRUE)

# Select the spreadsheet from which admin data will be extracted.
# Must be a CSV file. Reads into R as a dataframe called subjectData
subjectSpreadsheet <- choose.files(default = "C:\\Users\\ian\\Dropbox\\West Point Data 2015\\A
Company\\*.*", caption = "Select spreadsheet", multi = TRUE)
subjectData <- read.csv(subjectSpreadsheet,header=TRUE,sep=",")

# convert everything admin spreadsheet dataframe to a lowercase
subjectData <- as.data.frame(sapply(subjectData,tolower))

#loop through the list to copy data from spreadsheet to each GNSS/METs file in fileList
for (i in 1:length(fileList)){
 directory <- dirname(fileList[i])
 subjectFolder <- basename(directory)
```

```
 subjectName <- substring(subjectFolder,4)
 subjectName <- tolower(subjectName)

 #Create a dataframe from each GNSS/METs file each time it loops through the fileList
 newDataTable<- read.csv(fileList[i],header=TRUE,sep=",")

 #Create a data frame called a from information from the admin spreadsheet....
 # if the GNSS/METs subject name equals the last name in the spreadsheet
 a <-subjectData[subjectData$Last_Name==subjectName,]

 #Assign the data in the new dataframe a to the GNSS/METs file
 newDataTable$Monitor <- a[1,1]
 newDataTable$Last_name <- a[1,2]
 newDataTable$First_name <- a[1,8]
 newDataTable$ssn <- a[1,9]
 newDataTable$Platoon <- a[1,3]
 newDataTable$Squad <- a[1,4]
 newDataTable$GPS_num <-a[1,5]
 newDataTable$sex <- a[1,10]
 newDataTable$age <- a[1,11]
 newDataTable$height <- a[1,13]
 newDataTable$weight <- a[1,14]
 newDataTable$loadedWeight <- a[1,7]
 newDataTable$load <- a[1,15]
 newDataTable$comments <- a[1,6]

 #write the data to a csv file
 newFileName <- paste(directory,"/",subjectName,"DataCombined.csv",sep="")
 write.table(format(newDataTable, digits=10),newFileName,sep=",",row.names=FALSE)
}
```

Python Script to create a feature class of points from GNSS/METs data file

```
## This script creates a feature class of points for each GNSS/METs file
## Each point represents the location of the navigator with a 10 second interval between points.

import arcpy
import Tkinter, tkFileDialog
import glob
import os
from fnmatch import fnmatch
root = Tkinter.Tk()
# Define the parent directory where you will search for the GNSS/METs files via GUI
dirname = tkFileDialog.askdirectory(parent=root,initialdir="D:/Test_for_writing_data_chapter/B
Company",title='Please select a directory')

#Define the pattern of the filename to search
pattern="*DataCombined*"
allFileNames=[]

#Search through the parent folder for files
for path, subdirs, files in os.walk(dirname):
  for name in files:
    if fnmatch(name, pattern):
      #Generate python list of all filenames called allFileNames
      fileNames=os.path.join(path, name)
      allFileNames.append(fileNames)
# Create a feature class using the arcpy library and the makeXYEventLayer_management tool. Save as the
subjects last name.
i=0
for files in allFileNames:
  i=i+1
  layerName="Layer"+str(i)
  arcpy.MakeXYEventLayer_management(files, "Longitude", "Latitude", layerName, "", "")
```

```
  baseName=os.path.basename(files)
  subjectName=baseName[:-16]
  geoDatabasePath="D:\\Test_for_writing_data_chapter\\B
Company\\B_Company_2015LandNavSubjectTracks.gdb\\"
  featureClassName=geoDatabasePath+subjectName

  ARCPY.COPYFEATURES_MANAGEMENT(LAYERNAME,FEATURECLASSNAME, "", "0", "0", "0")
```

Python Script to add duration, speed, distance and slope between points

```
## Script to determine values between points of Average Speed, Distance, Duration, and Slope from
extracted DEM.
## Uses a feature class of points and a DEM
## Output is a point feature class called lastname_W_Z saved in the geodatabase where the original
point file originated.

import arcpy
import tkFileDialog
arcpy.CheckOutExtension("spatial")
arcpy.CheckOutExtension("tracking")
# Define the geodatabase where you will search for the point files via GUI
dirname=tkFileDialog.askdirectory(initialdir="D:/Test_for_writing_data_chapter/B
Company/B_Company_2015LandNavSubjectTracks.gdb",title='Please select a geodatabase to search')
arcpy.env.workspace=dirname
arcpy.env.overwriteOutput=True
# create a list of all feature classes
allFC=arcpy.ListFeatureClasses()
##Select DEM to use as the elevation base layer
DEMfileName="C:\\Users\\ian\\GIS Research\\Tobler and Naismith Assessment IJGIS\\tobler and naismith
assessment IJGIS.gdb\\dem_50cm_a2_WestPoint"
for files in allFC:
  # Process: Add Field (2)
  arcpy.AddField_management(files, "zDiff", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
  # Process: Extract Values from DEM to Points
  OutputFileName=files+"_W_Z"
  arcpy.gp.ExtractValuesToPoints_sa(files, DEMfileName, OutputFileName, "INTERPOLATE", "VALUE_ONLY")
  valueList = []
  #Calculate difference in elevation by subtracting Z value of current row from Z value from next row.
  with arcpy.da.SearchCursor(OutputFileName, ["RASTERVALU"]) as cursor:
    for row in cursor:
      valueList.append(row[0])
  del cursor
  x = 1
  with arcpy.da.UpdateCursor(OutputFileName, ["RASTERVALU", "zDiff"]) as cursor:
    for row in cursor:
      try:
        value = valueList[x]
      except:
        value = valueList[-1]
      row[1] = row[0] - value
      x += 1
      cursor.updateRow(row)
  del cursor
  arcpy.TrackIntervalsToFeature_ta(OutputFileName, "datetime", "", "CURRENT_AND_NEXT_FEATURE", "",
"01033-English_(United_States)", "AM", "PM", "METERS", "Distance_M_datetime", "SECONDS",
"Duration_SEC_datetime", "METERS_PER_SECOND", "Speed_MPS_datetime", "DEGREES", "Course_DEG_datetime")
  # Process: Add Field
  arcpy.AddField_management(OutputFileName, "Slope", "FLOAT", "", "", "", "", "NULLABLE",
"NON_REQUIRED", "")
  # Process: Calculate Field - Slope
  arcpy.CalculateField_management(OutputFileName, "Slope", "(( !zDiff!)/ !Distance_M_datetime!)*(-
100)", "PYTHON_9.3", "")
```

Python Script to convert points to lines

```
## This script creates lines from a track points.
## Previous attributes had been associated to points that represented behavior and data between
points. This information is joined back to the file after the points are converted to a line.
## Output of this code is a line feature class saved to the same geodatabase that the point feature
class has originated.
## Output name is lastname_W_Z_line_join

import tkFileDialog
import arcpy
from FeatureclassConversion import outFeatureClass
arcpy.CheckOutExtension("tracking")
_# GUI to have user choose Geodatabase to load point featureclasses
dirname = tkFileDialog.askdirectory()
arcpy.env.workspace=dirname
arcpy.env.qualifiedFieldNames = False
arcpy.env.overwriteOutput = True
# create a list of all feature classes. In this case, all the point files contain a _Z so we search
the geodatabase by this wildcard.
allFC=arcpy.ListFeatureClasses(wild_card="*_Z*")
for files in allFC:
  #Create a file name containing the full path
  fullFileName=str(dirname+"\\"+files)
  #Create a name for the output file
  out_feature_class=dirname+"/"+files+"_line"
  #Convert points to line
  arcpy.TrackIntervalsToLine_ta(files, out_feature_class, time_field='datetime', time_field_format =
"MM/dd/yyyy HH:mm:ss",distance_field_units = "METERS",
                 distance_field_name = "DISTANCE",duration_field_units = "SECONDS",duration_field_name
= "DURATION",speed_field_units = "METERS_PER_SECOND"
                 ,speed_field_name = "SPEED",course_field_units = "DEGREES",course_field_name =
"HEADING")
  # The above function strips attributes previously associated to the points.
  # Rejoin the Attributes to the line file
  arcpy.MakeFeatureLayer_management (out_feature_class, "track_intervals")
  arcpy.AddJoin_management("track_intervals", "Start_Time", files, "datetime")
  # define the name of the output file and save
  out_feature_class=out_feature_class+"_join"
  # Copy the layer to a new permanent feature class
  arcpy.CopyFeatures_management("track_intervals", out_feature_class)
```

Python Script to add Class ID to trajectories

```
## This script uses a spatial join to assign a land cover class to each line segment in a subjects
trajectory
## Output file is called lastname_W_Z_line_join_class
import tkFileDialog
import arcpy
arcpy.CheckOutExtension("tracking")
#GUI to have user choose Geodatabase where trajectories are stored.
dirname = tkFileDialog.askdirectory(title='Please select a geodatabase where trajectories are stored')
arcpy.env.workspace=dirname
arcpy.env.qualifiedFieldNames = False
arcpy.env.overwriteOutput = True
#Select all trajectory line files. In our case they have the word "join" in the filename
allFC=arcpy.ListFeatureClasses(wild_card="*join*")


# Create the spatial join to get the class ID in each record of the attribute table
for files in allFC:
```

```
  fullFileName=str(dirname+"\\"+files)
  #Define the full path of the classification layer
  join_features="C:/Users/ian/Dropbox/West Point Data 2015/Classification Layers/classes_final.shp"
  #Define the output feature class name
  output_feature_class=fullFileName+"_class"
  #The below code is required to add multiple values to the attribute table if the line crosses a
class boundary.
  #If this code is not used, the spatial join function will only one of the class IDs to the attribute
table.
  fieldmappings = arcpy.FieldMappings()
  fieldmappings.addTable(fullFileName)
  fieldmappings.addTable(join_features)
  class_ID_index=fieldmappings.findFieldMapIndex("class_ID")
  fieldmap=fieldmappings.getFieldMap(class_ID_index)
  field=fieldmap.outputField
  field.type="Text"
  field.length=10
  fieldmap.outputField=field
  fieldmap.mergeRule="Join"
  fieldmap.joinDelimiter=","
  fieldmappings.replaceFieldMap(class_ID_index, fieldmap)
  arcpy.SpatialJoin_analysis(fullFileName, join_features, output_feature_class,
field_mapping=fieldmappings)
```

R Script to Compute Average GNSS Speed

```
##Script to add average speed to GNSS data.
##It uses forward 10 speeds for the average since the accelerometer data also represents quantities
     that occur over that interval.
##The codescrolls through folders and find _q files since this was how the GNSS data was saved.
##Output is a csv file called lastnameAveGPS.csv
##This file will later be joined to the feature classes.

library(zoo)
#Select the folder that this code will crawl through to find all GNSS data
dirList <- list.dirs(choose.dir(caption = "Select folder"),full.names=TRUE, recursive = FALSE)

for (i in 1:length(dirList)){
 #Crawls through the above selected folder to find your GPS Files with"_q" pattern
 corGPSfilename <- list.files(dirList[i],pattern="*_q",ignore.case = TRUE)
 corGPSfilename=paste(dirList[i],"/",corGPSfilename,sep="")
 directory <- dirList[i]
 subjectFolder <- basename(directory)
 subjectName <- substring(subjectFolder,4)
 # Makes sure that the subjectName tested is lower case.
 subjectName <- tolower(subjectName)
 corGPS <- read.csv(corGPSfilename,header=TRUE,sep=",")
 #Now to alter the corGPS data a little so we can join later based on time.
 a <- gsub(" ", "", corGPS$LOCAL.DATE, fixed = TRUE)
 b <- gsub(" ", "", corGPS$LOCAL.TIME, fixed = TRUE)
 corGPS$navTime <- paste(a,b)
 corGPS$LONGITUDE <- corGPS$LONGITUDE*-1
 corGPS$SPEED=as.character(corGPS$SPEED)
 #Units are saved as part of the text string. Remove the km/hr from the speed.
 corGPS$Speed_GPS<-substring(corGPS$SPEED,1,nchar(corGPS$SPEED)-7)
 #Convert it to a number
 corGPS$Speed_GPS<-as.numeric(corGPS$Speed_GPS)
 # Convert from Km/hr to m/s
 corGPS$Speed_GPS<-corGPS$Speed_GPS*0.277778
 corGPS$ave10=rollmean(corGPS$Speed_GPS,10,na.pad=TRUE,align="left")
 #Also compute GNSS distance over the 10 seconds for potential later use.
 corGPS$DISTANCE=as.character(corGPS$DISTANCE)
 corGPS$DISTANCE<-substring(corGPS$DISTANCE,1,nchar(corGPS$DISTANCE)-2)
 corGPS$DISTANCE<-as.numeric(corGPS$DISTANCE)
 corGPS$DISTANCE10=rollmean(corGPS$DISTANCE,10,na.pad=TRUE,align="left")*10
 #Save table as a CSV
```

```
 joinedTableName=paste(directory,"/",subjectName,"AveGPS.csv",sep="")
 write.table(format(corGPS, digits=10),joinedTableName,sep=",",row.names=FALSE)
}
```

Python Script to join AveGPS data

```
#Script to join AveGPS speed to feature class.
#Feature class name is lastname_tracks.
import tkFileDialog
import os
import arcpy
arcpy.CheckOutExtension("tracking")
arcpy.env.qualifiedFieldNames = False
arcpy.env.overwriteOutput = True
#Convert CSV to Table
#Ensure all your Ave GPS files are in a folder that contains all the Ave GPS csv files and only the Ave GPS csv files
#Ensure all the names of the Ave GPS files are lastnameAveGPS
tabledirname = tkFileDialog.askdirectory(initialdir=r'D:\28OCTprocessing\H Company\AveGPS') #GUI to have user choose folder
geoname=tkFileDialog.askdirectory(initialdir=r'D:\28OCTprocessing\H_Company_2015LandNavSubjectTracks.gdb')
allFiles=os.listdir(tabledirname)
# To create the table called nameAveGPS. Must use this step because the join does not work with a CSV.
for files in allFiles:
  tablename=tabledirname+"/"+files
  #to remove the .csv from the file name
  name=files[:-4]
  arcpy.TableToTable_conversion(tablename,geoname,name)
# To join the AveGPS file to the tracks File
for files in allFiles:
  name=files[:-4]
  #to remove the aveGPS from the file name we use the name[:-6]
  infeatures=geoname+"/"+name[:-6]+"_tracks"
  joinTable=geoname+"/"+name
  arcpy.JoinField_management(infeatures,"Start_Time",joinTable,"navTime")
```

Python Script used in field calculation of cumulative distance and cumulative time

Calculate Cumulative Distance

Pre-Logic Script Code:

```
totalDistance = 0
def accumulateDistance(inDist):
  global totalDistance
  totalDistance += inDist
  return totalDistance
```

Field Calculation box: cumDist=

```
accumulateDistance(!DISTANCE!)
```

Calculate Cumulative Time

Pre-Logic Script Code:

```
totalTime = 0
def accumulateDistance(inTime):
  global totalTime
  totalDistance += inTime
  return totalTime
```

Field Calculation box: cumTime=

```
accumulateTime(!DURATION!)
```

R Script used to compute additional variables

** Note – Additional Libraries shown here for use in later scripting efforts

```
library(foreign)
library (ggplot2)
library (rpart)
library (lme4)
library(MASS)
library(relaimpo)
library (car)
require(lattice)
require(nlme)
require(MuMIn)

data <- read.csv(file="C:\\Users\\ian\\Dropbox\\West Point Data
2015\\All_tracks_merged_20Nov_CSV.csv", na.strings = "NA", header = TRUE, sep = ",", fill = TRUE)

data$Platoon=as.numeric(data$Platoon)
data$Squad=as.numeric(data$Squad)
data$age=as.numeric(data$age)
data$hr=as.numeric(data$hr)
data$loadedWeight=as.numeric(data$loadedWeight)
data$load=as.numeric(data$load)
data$fitness_sc=as.numeric(data$fitness_sc)
data$run_scr=as.numeric(data$run_scr)
data$class_ID <- as.factor(data$class_ID)#Make Class_ID into a factor so that it can be treated as a
catagorical variable
data$sex=as.factor(data$sex)
data$Start_Time=as.POSIXct(data$Start_Time, format = "%m/%d/%Y %H:%M:%S")
data$End_Time=as.POSIXct(data$End_Time, format = "%m/%d/%Y %H:%M:%S")


# Caclulate and enter BMI into the dataframe.
data$BMI <- 703*(data$weight/((data$height)^2))

#calculate a new column called EE for Expended Energy.
data$EE<- (data$MET_rate*(data$loadedWeight/2.20462)*(data$DURATION/60)/57)

#calculate a new column called EEnorm for Expended Energy. This is just EE/Distance or Calories/meter.
data$EEnorm<-data$EE/data$DISTANCE
ifelse (data$sex=="m",data$EEhr<-(data$DURATION/60)*((-55.0969 + 0.6309*data$hr +
0.1988*0.453592*data$loadedWeig + 0.2017*data$age)/4.184),data$EEhr<-(data$DURATION/60)*((-20.4022 +
0.4472*data$hr + 0.1263*0.453592*data$loadedWeig + 0.074*data$age)/4.184))
data$EEhrnorm= data$EEhr/data$DISTANCE

#Compute the slope of each segment using the GPS elevation values
data$Z_GPS=as.character(data$Z_GPS)
data$Z_GPS=gsub(" M","",data$Z_GPS)
data$Z_GPS=as.numeric(data$Z_GPS)

for (i in 1:length(data$Z_GPS)){
 y=i+1
```

```
 ifelse((data$Start_Time[y]-data$Start_Time[i]==10), data$zGPSdiff[i]<-data$Z_GPS[y]-
data$Z_GPS[i],data$zGPSdiff[i]<-NA)
}
data$Slope_ZGPS=(data$zGPSdiff/data$DISTANCE)*100
```

R Script to compute general dataset statistics and results

** Note – Additional Libraries shown here for use in later scripting efforts

```
library(foreign)
library (ggplot2)
library (rpart)
library (lme4)
library(MASS)
library(relaimpo)
library (car)
require(lattice)
require(nlme)
require(MuMIn)
library(plyr)

data <- read.csv(file="C:\\Users\\ian\\Dropbox\\West Point Data
2015\\All_tracks_merged_20Nov_CSV.csv", na.strings = "NA", header = TRUE, sep = ",", fill = TRUE)

data$Platoon=as.numeric(data$Platoon)
data$Squad=as.numeric(data$Squad)
data$age=as.numeric(data$age)
data$hr=as.numeric(data$hr)
data$loadedWeight=as.numeric(data$loadedWeight)
data$load=as.numeric(data$load)
data$fitness_sc=as.numeric(data$fitness_sc)
data$run_scr=as.numeric(data$run_scr)
data$class_ID <- as.factor(data$class_ID)#Make Class_ID into a factor so that it can be treated as a
catagorical variable
data$sex=as.factor(data$sex)
data$Start_Time=as.POSIXct(data$Start_Time, format = "%m/%d/%Y %H:%M:%S")
data$End_Time=as.POSIXct(data$End_Time, format = "%m/%d/%Y %H:%M:%S")
data<- transform(data, ID = as.numeric(interaction(Last_name, First_name, drop=TRUE)))


# Caclulate and enter BMI into the dataframe.
data$BMI <- 703*(data$weight/((data$height)^2))

#calculate a new column called EE for Expended Energy.
data$EE<- (data$MET_rate*(data$loadedWeight/2.20462)*(data$DURATION/60)/57)

#calculate a new column called EEnorm for Expended Energy. This is just EE/Distance or Calories/meter.
data$EEnorm<-data$EE/data$DISTANCE

#Compute the slope of each segment using the GPS elevation values
data$Z_GPS=as.character(data$Z_GPS)
data$Z_GPS=gsub(" M","",data$Z_GPS)
data$Z_GPS=as.numeric(data$Z_GPS)
for (i in 1:length(data$Z_GPS)){
 y=i+1
 ifelse((data$Start_Time[y]-data$Start_Time[i]==10), data$zGPSdiff[i]<-data$Z_GPS[y]-
data$Z_GPS[i],data$zGPSdiff[i]<-NA)
}
data$Slope_ZGPS=(data$zGPSdiff/data$DISTANCE)*100

#Statistics of the subjects
subjects=ddply(data, .(ID), head, n = 1)

#To find the frequency of different ages:
table (subjects$age)

#For General statistics of Subjects
summary(subjects)
```

```
#Histograms for subject data
hist(subjects$weight,20, xlab="Subject Weight (lbs)",main="Histogram of Subject Weights",
xlim=c(100,270))
hist(subjects$height,20, xlab="Subject Height (inches)",main="Histogram of Subject Heights",
xlim=c(60,80))
hist(subjects$BMI,20, xlab="Subject BMI (kg/m2)",main="Histogram of Subject BMI", xlim=c(15,35))
hist(subjects$load,20, xlab="Load Carried (lbs)",main="Histogram of Load Carried by Subjects",
xlim=c(10,37))
hist(subjects$fitness_scr,20, xlab="Total Points",main="Histogram of Subject Fitness Score",
xlim=c(60,300))
hist(subjects$run_scr,20, xlab="Total Points",main="Histogram of Subject Run Score", xlim=c(0,100))

#General Statistics of Terrain
slopeSub=subset(data, data$Slope_ZGPS>-50&data$Slope_ZGPS<50)
slopeSub$ABSslope=abs(slopeSub$Slope_ZGPS)
mean(slopeSub$Slope_ZGPS,na.rm=TRUE)
mean(slopeSub$ABSslope)
hist(slopeSub$Slope_ZGPS,20, xlab="Slope (%)",main="Histogram of Slope")
table(slopeSub$class_ID)

#General Statistics of the study variables
mean(slopeSub$SPEED)
summary(slopeSub$SPEED)
hist(slopeSub$SPEED,20, xlab="Speed (m/s)",main="Histogram of Subject Speed")

METSlopeSub<- subset(slopeSub, slopeSub$MET_rate>1 & slopeSub$SPEED>.5)
summary(METSlopeSub$SPEED)
mean(METSlopeSub$SPEED)
hist(METSlopeSub$SPEED,20, xlab="Speed (m/s)",main="Histogram of Subject Speed - Standing Removed",
xlim=c(.5,4))

mean(METSlopeSub$EEnorm)
hist(METSlopeSub$EEnorm,20, xlab="Calories consumed per meter",main="Histogram of Subject Energy
Expenditure")
summary(METSlopeSub$EEnorm)


# Code for Chapter 5 Results

# Make sure we have the correct data
slopeSub=subset(data, data$Slope_ZGPS>-50&data$Slope_ZGPS<50)

#Remove areas that cross over land cover classes.
slopeSub=subset(slopeSub,slopeSub$class_ID=="1"|slopeSub$class_ID=="2"|slopeSub$class_ID=="3"|slopeSub
$class_ID=="4"|slopeSub$class_ID=="5"|slopeSub$class_ID=="6"|slopeSub$class_ID=="7"|slopeSub$class_ID=
="8")

#Remove areas where standing is occuring
METSlopeSub<- subset(slopeSub, slopeSub$MET_rate>1 & slopeSub$SPEED>.14)

#Research Questions 1 and 2 Go back to original data because

# Use the below to determine the best transform for EENorm which does not have a normal distribution
out <- boxcox(lm(METSlopeSub$EEnorm~1))
range(out$x[out$y > max(out$y)-qchisq(0.95,1)/2])


#Take Ln of EEnorm to get a normal distribution

METSlopeSub$LogEEnorm=log(METSlopeSub$EEnorm)
hist(METSlopeSub$LogEEnorm,30,xlab="LN(Calories/meter)",main="Histogram of LN(Energy
Expenditure/meter)")


#Now must investigate EE vs slope.  It is not linear so we must transform.

############################
#Must average the bins to generalize the noisy data
############################
```

```
########################
#Filtering Technique
########################

## Creat bins at different Slopes
lBins <- seq(-50.5,49.5,1)
rBins <- seq(-49.5,50.5,1)
midBins <- seq(-50,50,1)

binStats <- data.frame(ID=numeric(0),leftBinValue=numeric(0),
                       rightBinValue=numeric(0),
                       binMidPoint=numeric(0),
                       aveLogEEnorm=numeric(0))#creates a new data frame to save the statistics of
each bin

## For loop to subset data based on bins
for(i in 1:length(lBins)){

  currentBin <- subset(METSlopeSub,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsAll=binStats

########################
#Now to plot results
########################


#Plot the Log of EE vs Slope
p=ggplot(binStatsAll,aes(x))
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm), col="red")
p=p+ggtitle("How Energy Expenditure Changes With Slope") +labs(x="Slope (%)",y="LN(Calories/Meter)")
p


##############################
#Now analyze all classes to see if we can generalize the classes
##############################
METSlopeSubC1=subset(METSlopeSub,METSlopeSub$class_ID==1)
## For loop to subset data based on bins
for(i in 1:length(lBins)){

  currentBin <- subset(METSlopeSubC1,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsC1=binStats

####################################################################

METSlopeSubC2=subset(METSlopeSub,METSlopeSub$class_ID==2)
## For loop to subset data based on bins
for(i in 1:length(lBins)){
```

```
  currentBin <- subset(METSlopeSubC2,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsC2=binStats

######################################################################

METSlopeSubC4=subset(METSlopeSub,METSlopeSub$class_ID==4)
## For loop to subset data based on bins
for(i in 1:length(lBins)){

  currentBin <- subset(METSlopeSubC4,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsC4=binStats

#################################################################
METSlopeSubC5=subset(METSlopeSub,METSlopeSub$class_ID==5)
## For loop to subset data based on bins
for(i in 1:length(lBins)){

  currentBin <- subset(METSlopeSubC5,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsC5=binStats

##################################################################

METSlopeSubC6=subset(METSlopeSub,METSlopeSub$class_ID==6)
## For loop to subset data based on bins
for(i in 1:length(lBins)){

  currentBin <- subset(METSlopeSubC6,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsC6=binStats

###################################################################

METSlopeSubC7=subset(METSlopeSub,METSlopeSub$class_ID==7)
## For loop to subset data based on bins
```

```
for(i in 1:length(lBins)){

  currentBin <- subset(METSlopeSubC7,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsC7=binStats

#################################################################

METSlopeSubC8=subset(METSlopeSub,METSlopeSub$class_ID==8)
## For loop to subset data based on bins
for(i in 1:length(lBins)){

  currentBin <- subset(METSlopeSubC8,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsC8=binStats

###################################################################
METSlopeSubC47=subset(METSlopeSub,METSlopeSub$class_ID==4|METSlopeSub$class_ID==5|METSlopeSub$class_ID
==6|METSlopeSub$class_ID==7)
## For loop to subset data based on bins
for(i in 1:length(lBins)){

  currentBin <- subset(METSlopeSubC47,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsC47=binStats

########################################################################
combo=data.frame(binStatsAll,binStatsC1,binStatsC2,binStatsC4,binStatsC5,binStatsC6,binStatsC7,binStat
sC8,binStatsC47)

#Plot all the classes

p=ggplot(combo,aes(x))
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.1), col="black")# On road
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.2), col="dark grey")# Boulders
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.3), col="light green")# Light Vegetation
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.4), col="green")# Moderate Vegetation
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.5), col="dark green")# Heavy Vegetation
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.6), col="blue")# Swamp
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.7), col="brown")# Open Woods
p=p+ggtitle("How Energy Expenditure Changes With Slope and Class") +labs(x="Slope
(%)",y="LN(Calories/Meter)")
p

# Plot the vegetation and the swamp together with other classes
```

```
p=ggplot(combo,aes(x))
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.1), col="black")# On road
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.2), col="dark grey")# Boulders
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.8), col="green")# Vegetation
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm.7), col="brown")# Open Woods
p=p+ggtitle("How Energy Expenditure Changes With Slope and Class") +labs(x="Slope
(%)",y="LN(Calories/Meter)")
p

###########################
#Multiple regression
##########################

#Remove water and consolidate the vegetation
METSlopeSub=subset(METSlopeSub,METSlopeSub$class_ID!="3")
METSlopeSub$class_ID2=METSlopeSub$class_ID
METSlopeSub$class_ID2[METSlopeSub$class_ID=="5"]<-"4"
METSlopeSub$class_ID2[METSlopeSub$class_ID=="6"]<-"4"
METSlopeSub$class_ID2[METSlopeSub$class_ID=="7"]<-"4"

M1=lm(METSlopeSub$LogEEnorm~METSlopeSub$SlopeSq+METSlopeSub$class_ID2 + METSlopeSub$cumDist_1 +
METSlopeSub$sex+METSlopeSub$BMI+METSlopeSub$fitness_scr)
summary (M1)
calc.relimp(M1, type = c("lmg"),rela = TRUE)
AIC(M1)
vif(M1)

M2=lm(METSlopeSub$LogEEnorm~METSlopeSub$SlopeSq+METSlopeSub$class_ID2 + METSlopeSub$cumDist_1 +
METSlopeSub$sex+METSlopeSub$BMI+METSlopeSub$run_scr)
summary (M2)
calc.relimp(M2, type = c("lmg"),rela = TRUE)
AIC(M2)
vif(M2)

M3=lm(METSlopeSub$LogEEnorm~METSlopeSub$SlopeSq+METSlopeSub$class_ID2 + METSlopeSub$cumDist_1 +
METSlopeSub$sex+METSlopeSub$BMI)
summary (M3)
calc.relimp(M3, type = c("lmg"),rela = TRUE)
AIC(M3)
sresid <- studres(M3)
hist(sresid, 30, main="Distribution of Studentized Residuals",xlab = "Studentized Residuals")
bwplot(ID~resid(M3),data=METSlopeSub, xlab="Studentized Residuals",main="Box plot of Studentized
Residuals by subject ID")
vif(M3)

#####################################
#Mixed Effects Modeling Effort
#####################################

M1.lme=lme(LogEEnorm~SlopeSq+class_ID2+cumDist_1+BMI+sex, data =
METSlopeSub,random=~1+SlopeSq+class_ID2+cumDist_1|ID)
bwplot(ID~resid(M1.lme),METSlopeSub,xlab="Studentized Residuals",main="Box plot of Studentized
Residuals by subject ID")
summary(M1.lme)
r.squaredGLMM(M1.lme)
```

```
#####################################################################
#Assessment of cumulative distance in the model.
#####################################################################

a=METSlopeSub # all
#a=subset(METSlopeSub,METSlopeSub$Slope_ZGPS>-5&METSlopeSub$Slope_ZGPS<0) # We could also look at just
one specific class
lBins <- seq(50,6950,100)
rBins <- seq(150,7050,100)
midBins <- seq(100,7000,100)

binStats <- data.frame(ID=numeric(0),leftBinValue=numeric(0),
                       rightBinValue=numeric(0),
                       binMidPoint=numeric(0),
                       aveLogEEnorm=numeric(0))#creates a new data frame to save the statistics of
each bin

## For loop to subset data based on bins
for(i in 1:length(lBins)){

  #currentBin <- subset(METSlopeSub,METSlopeSub$cumDist_1>(lBins[i]) &
METSlopeSub$cumDist_1<(rBins[i]))#subsets all rows that fall within a given bin
  currentBin <- subset(a,a$cumDist_1>(lBins[i]) & a$cumDist_1<(rBins[i]))#subsets all rows that fall
within a given bin
  binStats[i,1] <- i
  binStats[i,2] <- lBins[i]
  binStats[i,3] <- rBins[i]
  binStats[i,4] <- midBins[i]
  binStats[i,5]= mean(currentBin$LogEEnorm, na.rm = TRUE)

}
binStatsDist=binStats

#Plot the Log of EE vs Cumulative Distance
p=ggplot(binStatsDist,aes(x))
p =p+geom_point(aes(x=binMidPoint,y=aveLogEEnorm), col="red")
p=p+ggtitle("How Energy Expenditure Changes With Distance") +labs(x="Distance from Start
(meters)",y="LN(Calories/Meter)")
p

####################################################################################
#To Answer Research Question 3
####################################################################################

######################
#Compute EE from Pandoff
######################

# 1977 Equation
# MW = 1.5 • W + 2.0 • (W + L) • (L / W)2 + ŋ • (W + L) • (1.5 • V2 + 0.35 • V • G)
# Where:
#   Mw = metabolic cost of walking (or standing) (in watts) - 1 kCal (A food Calorie is actually 1000
calories, small c, or 1 kCal) = 4.18 kJ
# 1 watt=1 joule/sec
# 1 joule = 0.00023900573614 food calories

# W = body mass (kilograms)
# L = load mass (kilograms)
# ŋ = terrain factor
# V = velocity or walk rate (m/s)
# G = slope or grade (%)
# The terrain factor categories are: 1.0 = black top road or treadmill; 1.1 = dirt road; 1.2 = light
brush; 1.5 = heavy brush; 1.8 = swampy bog; 2.1 = loose sand; 2.5 = soft snow, 15 cm depth; 3.3 = soft
snow 25 cm deep; 4.1 = soft snow, 35 cm depth (12).
```

```
#
# 2003 Correction
# Mw = PE – CF  Where PE is the 1977 Pandolf Equation and CF is the correction factor listed below.
# CF = η • [(G • (W + L) • V) / 3.5 - ((W + L) • (G + 6)^2) / W) + (25V2)]
# Where:
#   t = terrain factor
# G = grade (%)
# W = body wt (kg)
# L = load wt (kg)
# V = velocity (m/s)
rq3df=METSlopeSub
rq3df$comments=NULL #remove these columns b/c they don't have much to add and have NA values
rq3df$ssn=NULL
rq3df$fitness_sc=NULL
rq3df$fitness_scr=NULL
rq3df$run_scr=NULL
sum(is.na(rq3df))#verify no NA values

rq3df$tfL[rq3df$class_ID==1]<-1.0
rq3df$tfL[rq3df$class_ID==2]<-NA#Rocky
rq3df$tfL[rq3df$class_ID==3]<-NA #Water
rq3df$tfL[rq3df$class_ID==4]<-1.2#light
rq3df$tfL[rq3df$class_ID==5]<-NA#moderate
rq3df$tfL[rq3df$class_ID==6]<-NA#heavy
rq3df$tfL[rq3df$class_ID==7]<-NA#swamp
rq3df$tfL[rq3df$class_ID==8]<-1.1

W=rq3df$weight*0.453592 #in KG
L=rq3df$load*0.453592 # in KG
V=rq3df$SPEED # in m/s
G=rq3df$Slope_ZGPS # in %
t=rq3df$tfL

rq3df$PE= 1.5*W + 2.0*(W+L)*(L/W)^2 + t*(W+L)*(1.5*V^2 + 0.35*V*G)
rq3df$CF = t*((G*(W+L)*V)/3.5)-(((W+L)*((G+6)^2))/W)+(25*(V^2))
rq3df$MW=(rq3df$PE-rq3df$CF)*rq3df$DURATION*0.00023900573614 #Calories / segment

##########################################################
#Compute EE from Irmischer Coefficents
#########################################################

B0=-2.85
B1=.00027
B3=-.000053
B4=.11
B5=.035

rq3df$tfii[rq3df$class_ID==1]<-0#onroad
rq3df$tfii[rq3df$class_ID==2]<-.45#Rocky
rq3df$tfii[rq3df$class_ID==3]<-NA #Water
rq3df$tfii[rq3df$class_ID==4]<-.39#light
rq3df$tfii[rq3df$class_ID==5]<-.39#moderate
rq3df$tfii[rq3df$class_ID==6]<-.39#heavy
rq3df$tfii[rq3df$class_ID==7]<-.39#swamp
rq3df$tfii[rq3df$class_ID==8]<-.31#open offroad

rq3df$sexCode[rq3df$sex=="m"]<-1
rq3df$sexCode[rq3df$sex=="f"]<-0


rq3df$LNEEii<-B0+B1*rq3df$SlopeSq+rq3df$tfii+B3*rq3df$cumDist_1+B4*rq3df$sexCode+B5*rq3df$BMI

rq3df$EEii=rq3df$DISTANCE*exp(rq3df$LNEEii)
rq3df=rq3df1
#rq3df=subset(rq3df,rq3df$sexCode==1)# Put in to look at different subsets of the data

###########################################################
#Compute summaries of EE
###########################################################
```

```
colSums(is.na(rq3df))
rq3results<-data.frame(ID=numeric(0),
                       EE_GT3X=numeric(0),
                        EE_LAREDO=numeric(0),
                       EE_ii=numeric(0))

for (i in 1:200){
  a=subset(rq3df,rq3df$ID==i) # focus on just one subject at a time.
  a=a[complete.cases(a),]# Remove all rows with NA. These are rows where Pandolf and classification do
not match.
  ID=mean(a$ID)
  b=sum(a$EE,na.rm = TRUE)
  c=sum(a$MW, na.rm = TRUE)
  d=sum(a$EEii)
  e=(sum(is.na(a$EE)))
  f=(sum(is.na(a$MW)))
  g=(sum(is.na(a$EEii)))
  h=length(a$EE)
  j=length(subset(data$EE,data$ID==i))
  k=(sum(a$DISTANCE))
  aa=subset(data,data$ID==i)
  l=max(aa$cumDist_1)
  cat (ID, b,c,d,e,f,g,h,j,k,l,"\n")
  rq3results[i,1]=ID
  rq3results[i,2]=b
  rq3results[i,3]=c
  rq3results[i,4]=d
}

#rq3results=subset(rq3results,rq3results$ID>0)#Put in if you want to analyzed subsets of the results
cor(rq3results$EE_GT3X,rq3results$EE_LAREDO)
cor(rq3results$EE_GT3X,rq3results$EE_ii)


# Calcualte MAE and percent error for LAREDO Algorithm

rq3results$ErrorL=(rq3results$EE_GT3X-rq3results$EE_LAREDO)
rq3results$absErrorL=abs(rq3results$EE_GT3X-rq3results$EE_LAREDO)
rq3results$percentErrorL=(rq3results$absErrorL/rq3results$EE_GT3X)*100
paeL=mean(rq3results$percentErrorL)
maeL=mean(rq3results$absErrorL)
paeL
maeL

# Calculate MAE and percent error for II Model
rq3results$Errorii=(rq3results$EE_GT3X-rq3results$EE_ii)
rq3results$absErrorii=abs(rq3results$EE_GT3X-rq3results$EE_ii)
rq3results$percentErrorii=(rq3results$absErrorii/rq3results$EE_GT3X)*100
paeii=mean(rq3results$percentErrorii)
maeii=mean(rq3results$absErrorii)
paeii
maeii


############################################
#Plot the estimated vs the actual values
############################################


p=qplot(EE_GT3X,EE_LAREDO,data=rq3results)
p=p+geom_abline()
p=p+ggtitle("Predicted vs Actual Calorie Consumption During Navigation") +labs(y="Predicted calorie
consumption by LAREDO algorithm",x="Calories consumed (as measured by Actigraph GT3X)")
p=p+xlim(250,1250)+ylim(250,1250)
p

p=qplot(EE_GT3X,EE_ii,data=rq3results)
p=p+geom_abline()
```

```
p=p+ggtitle("Predicted vs Actual Calorie Consumption During Navigation") +labs(y="Predicted calorie
consumption by new model",x="Calories consumed (as measured by Actigraph GT3X)")
p=p+xlim(250,1250)+ylim(250,1250)
p



#####################################
#Research Question 4 Results
#####################################

#Create on road and offroad dataframes
slopeSub=subset(data, data$Slope_ZGPS>-50&data$Slope_ZGPS<50)
slopeSubC1=subset(slopeSub,slopeSub$class_ID==1)
slopeSubC2_8=subset(slopeSub,!grepl("1", slopeSub$class_ID))

## Creat bins at different Slopes
lBins <- seq(-50.5,49.5,1)
rBins <- seq(-49.5,50.5,1)
midBins <- seq(-50,50,1)

######################################
#Compute bin averages
######################################

####################
#On Road
####################

binStats <- data.frame(ID=numeric(0),
                       binMidPoint=numeric(0),
                       aveSpeed=numeric(0),
                       binSize=numeric(0))

## For loop to subset data based on bins
for(i in 1:length(lBins)){


  # Change the below line to match what subset you want.
  currentBin <- subset(slopeSubC1,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all rows
that fall within a given bin

  binStats[i,1] <- i
  binStats[i,2] <- midBins[i]
  binStats[i,3]= mean(currentBin$SPEED)
  binStats[i,4]=nrow(currentBin)
}
binStats$Tobler=.277778*6*exp(-3.5*abs(binStats$binMidPoint*.01+.05))
binStats$TobleroOffroad=.277778*.6*6*exp(-3.5*abs(binStats$binMidPoint*.01+.05))#multiply above by .6
as described in Tobler
binStatsC1=binStats

###############################
# Must do it all again to create another data frame with data from Classes 2-8
###############################

binStats <- data.frame(ID=numeric(0),
                       binMidPoint=numeric(0),
                       aveSpeed=numeric(0),
                       binSize=numeric(0))

## For loop to subset data based on bins
for(i in 1:length(lBins)){


  # Change the below line to match what subset you want.
  currentBin <- subset(slopeSubC2_8,Slope_ZGPS>(lBins[i]-3) & Slope_ZGPS <(rBins[i]+3))#subsets all
rows that fall within a given bin
```

```
  binStats[i,1] <- i
  binStats[i,2] <- midBins[i]
  binStats[i,3]= mean(currentBin$SPEED)
  binStats[i,4]=nrow(currentBin)
}
binStatsC2_8=binStats

combo=data.frame(binStatsC1,binStatsC2_8)# Combine all data into one dataframe

p=ggplot(combo,aes(x))
p =p+geom_point(aes(x=binMidPoint,y=Tobler), col="red")
p= p+geom_point(aes(x=binMidPoint,y=aveSpeed), col="blue")
p=p+ggtitle("Speed of navigation - On road") + labs(x="Slope (%)",y="Speed m/s")
p

p=ggplot(combo,aes(x))
p =p+geom_point(aes(x=binMidPoint,y=ToblerOffroad), col="red")
p= p+geom_point(aes(x=binMidPoint,y=aveSpeed.1), col="blue")
p=p+ggtitle("Speed of navigation - Off road") + labs(x="Slope (%)",y="Speed m/s")
p

###################################################
#Curve fitting Effort
###################################################

x=binStatsC1$binMidPoint
y=binStatsC1$aveSpeed
xx <- seq(-50,50, length=101)

# Fit On road curve
fit1=lm(y-1.07~x + I(x^2))#I fix the apex to be 1.07 which is very similiar to my max value
fit1$coefficients[1]=1.07
fit1$coefficients[2]=-.004
fit1$coefficients[3]=-.00045
plot (x,y,pch=19, col="blue", main="On road navigation speeds at varying slopes",ylab="Navigator Speed
(m/s)", xlab="Slope (%)")
lines(xx, predict(fit1, data.frame(x=xx)), col="green",lwd=3.5)
summary (fit1)

#############################
#Offroad
#############################

u=binStatsC2_8$binMidPoint
v=binStatsC2_8$aveSpeed
z=binStatsC2_8$count

plot (u,v,pch=19, col="blue", main="Off road navigation speeds at varying slopes",ylab="Navigator
Speed (m/s)", xlab="Slope (%)")

fit2=lm(v-.76~u + I(u^2),weights = z)
#force coefficents to 0.765000    -0.00100    -0.00035
fit2$coefficients[1]=.765
fit2$coefficients[2]=-.001
fit2$coefficients[3]=-.00035
lines(xx, predict(fit2, data.frame(x=xx)), col="green", lwd=3.5)
summary (fit2)

quantile(slopeSub$Slope_ZGPS, probs = c(0.05, 0.95))

###########################################################
#Plot Tobler vs II On road
###########################################################

combo2=subset(combo,combo$binMidPoint>=-40&combo$binMidPoint<=40)
f=ggplot(data.frame(x=c(-40,40)), aes(x))
f=f+stat_function(fun=function(x)-.00045*x^2-.004*x+ 1.07,colour="green", size =1.5)
f=f+stat_function(fun=function(x)(.277778*6*exp(-3.5*abs(x*.01+.05))),colour="red", size=1.5)
```

```
f= f+geom_point(data=combo2,mapping=aes(x=binMidPoint,y=aveSpeed), col="blue")
f=f+ggtitle("Speed of navigation - On road") + labs(x="Slope (%)",y="Speed m/s")
f


###########################################################
#Plot  Tobler vs II off road
###########################################################

f=ggplot(data.frame(x=c(-40,40)), aes(x))
f=f+stat_function(fun=function(x)-.00035*x^2-.001*x+.765,colour="green",size=1.5)
f=f+stat_function(fun=function(x).6*(.277778*6*exp(-3.5*abs(x*.01+.05))),colour="red",size=1.5)
f= f+geom_point(data=combo2,mapping=aes(x=binMidPoint,y=aveSpeed.1), col="blue")
f=f+ggtitle("Speed of navigation - Off road") + labs(x="Slope (%)",y="Speed m/s")
f



######################################################
#Test significance that veg vs boulder vs onroad vs open woods are different
######################################################

#combine all veg and swamp and call it Class 4.
METSlopeSubCveg=METSlopeSubC47
METSlopeSubCveg$class_ID=4
#combine all the classes
METSlopeSub1=rbind(METSlopeSubC1,METSlopeSubC2,METSlopeSubCveg,METSlopeSubC8)
boxplot(METSlopeSub1$LogEEnorm~METSlopeSub1$class_ID)
veg.aov=aov(METSlopeSub1$LogEEnorm~METSlopeSub1$class_ID)
summary(veg.aov)
TukeyHSD(veg.aov)

#######################################################
#Comparative values of different types of hike vs navigating
#######################################################

# 7 km hike at zero slope with on road Tobler vs Mine - Time difference
#Tobler Onroad=6*EXP(-3.5*ABS(Slope/100+0.05)) - Slope in Percent
#Tobler equation at 0 slope = 6exp(-3.5*(.05))=5.0367km/hr
Tobler=7/5.0367*60 #How many min
Irmischer=7000/1.07/60 #How many min
Tobler
Irmischer
print(Irmischer/Tobler)
# 7 km hike at zero slope with off road Tobler vs Mine - Time difference
#Tobler Offroad=(3/5)*6*EXP(-3.5*ABS(Slope/100+0.05)) - Slope in Percent
ToblerOff=7/(5.0367*(3/5))*60
IrmischerOff=(7000/.765)/60
ToblerOff
IrmischerOff
print(IrmischerOff/ToblerOff)

# 7 km hike at -5% slope with on road Tobler vs Mine - Time difference
ToblerOn=7/(6*exp(-3.5*abs(-5/100+0.05)))*60
IrmischerOn=7000*(1.07-(0.004*(-5))-(0.00045*(-5^2)))/60
ToblerOn
IrmischerOn
print(IrmischerOn/ToblerOn)
## 7 km hike at -5% slope with off road Tobler vs Mine - Time difference
ToblerOff=(5/3)*7/(6*exp(-3.5*abs(-5/100+0.05)))*60
IrmischerOn=7000*(.765-(0.001*(-5))-(0.00035*(-5^2)))/60
ToblerOff
IrmischerOff
print(IrmischerOff/ToblerOff)


#########################################################
##Determine total time of navigation vs predicted by Tobler and Irmischer
#########################################################
```

```
slopeSub30=subset(data, data$Slope_ZGPS>-30&data$Slope_ZGPS<30)
slopeSub30C1=subset(slopeSub30,slopeSub30$class_ID==1)
slopeSub30C2_8=subset(slopeSub30,!grepl("1", slopeSub30$class_ID))
slopeSub30C1$toblerSpeed=(6*exp(-3.5*abs((slopeSub30C1$Slope_ZGPS/100)+0.05)))
slopeSub30C1$toblerTime=(3.6*slopeSub30C1$DISTANCE)/(6*exp(-
3.5*abs((slopeSub30C1$Slope_ZGPS/100)+0.05)))

slopeSub30C2_8$toblerSpeed=(3/5)*(6*exp(-3.5*abs((slopeSub30C2_8$Slope_ZGPS/100)+0.05)))
slopeSub30C2_8$toblerTime=(3.6*slopeSub30C2_8$DISTANCE)/((3/5)*(6*exp(-
3.5*abs((slopeSub30C2_8$Slope_ZGPS/100)+0.05))))


slopeSub30C1$ianSpeed=1.07-.004*slopeSub30C1$Slope_ZGPS-
.00045*(slopeSub30C1$Slope_ZGPS*slopeSub30C1$Slope_ZGPS)
slopeSub30C1$ianTime=slopeSub30C1$DISTANCE/slopeSub30C1$ianSpeed

slopeSub30C2_8$ianSpeed=.765-.001*slopeSub30C2_8$Slope_ZGPS-
.00035*(slopeSub30C2_8$Slope_ZGPS*slopeSub30C2_8$Slope_ZGPS)
slopeSub30C2_8$ianTime=slopeSub30C2_8$DISTANCE/slopeSub30C2_8$ianSpeed

routeTime=aggregate(DURATION~ ID, slopeSub30C1, sum)
routeTime$toblerC1=aggregate(toblerTime~ ID, slopeSub30C1, sum)
routeTime$actualC1=aggregate(DURATION~ ID, slopeSub30C1, sum)
routeTime$toblerC2_8=aggregate(toblerTime~ ID, slopeSub30C2_8, sum)
routeTime$actualC2_8=aggregate(DURATION~ ID, slopeSub30C2_8, sum)

routeTime$ianC1=aggregate(ianTime~ ID, slopeSub30C1, sum)
routeTime$ianC2_8=aggregate(ianTime~ ID, slopeSub30C2_8, sum)

routeTime$actual=routeTime$actualC1$DURATION+routeTime$actualC2_8$DURATION
routeTime$tobler=routeTime$toblerC1$toblerTime+routeTime$toblerC2_8$toblerTime
routeTime$ian=routeTime$ianC1$ianTime+routeTime$ianC2_8$ianTime
routeTime$diffTI=routeTime$tobler-routeTime$ian
routeTime$diffAI=routeTime$actual-routeTime$ian
routeTime$diffAT=routeTime$actual-routeTime$tobler

summary(routeTime$diffAI)
summary(routeTime$diffAT)
summary(routeTime$diffTI)

hist(routeTime$diffAI)
hist(routeTime$diffAT)
hist(routeTime$diffTI)

maeIan=mae(routeTime$actual,routeTime$ian)#mean absolute error
maeTobler=mae(routeTime$actual,routeTime$tobler)
rsmeIan=rmse(routeTime$actual,routeTime$ian)
rsmeTobler=rmse(routeTime$actual,routeTime$tobler)

mapeI = mean(abs((routeTime$actual - routeTime$ian)/routeTime$actual))#Mean absolute percentage error

mapeT = mean(abs((routeTime$actual - routeTime$tobler)/routeTime$actual))
cor(routeTime$actual, routeTime$ian)


maeIan
maeTobler
rsmeIan
rsmeTobler

mapeI

mapeT

routeTime$ianMin=routeTime$ian/60
routeTime$actualMin=routeTime$actual/60
routeTime$toblerMin=routeTime$tobler/60
```

```
############################################
#Plot the estimated vs the actual values
############################################

p=qplot(actualMin,ianMin,data=routeTime)
p=p+geom_abline()
p=p+ggtitle("Predicted vs Actual Completion Times") +labs(y="Predicted Completion Times Using
Irmischer Algorithm",x="Actual Completion Times  (Minutes)")
p=p+xlim(40,140)+ylim(40,140)
p

p=qplot(actualMin,toblerMin,data=routeTime)
p=p+geom_abline()
p=p+ggtitle("Predicted vs Actual Completion Times") +labs(y="Predicted Completion Times Using Tobler
Algorithm",x="Actual Completion Times  (Minutes)")
p=p+xlim(40,140)+ylim(40,140)
p
```

APPENDIX C – Description of Data Matrix Fields

OBJECTID	A unique identifier of each data record.
DISTANCE	The distance a subject covered over the duration of the segment. Usually this is over 10 seconds. Units are meters.
DURATION	The duration of the segment. Usually this is over 10 seconds. Units are seconds.
SPEED	This is a speed calculated in ArcGIS by taking the Distance/Duration when using the track interval to feature. Units are meters/second.
HEADING	Heading person is looking. This is determined by the ArcGIS desktop tool "track intervals to feature".
START_TIME	DateTime when the subject was at the beginning of the 10 second segment.
END_TIME	DateTime when the subject was at the beginning of the segment. Usually 10 seconds after the START_TIME.
CLASS_ID	Class of terrain through which the segment passes
AXIS1	Number of counts from the accelerometer on the vertical axis. Determined from proprietary formula by the Actilife software based on raw accelerometer data.
AXIS2	Number of counts from the accelerometer on the horizontal axis. Determined from proprietary formula by the Actilife software based on raw accelerometer data.
AXIS3	Number of counts from the accelerometer on the perpendicular (out from the body) axis. Determined from proprietary formula by the Actilife software based on raw accelerometer data.
VM	Vector Magnitude of the number of counts from the 3 axis of the accelerometer.
STEPS	Estimate of the number of steps taken by a subject. Calculation by the Actilife6 using only the axis 1 counts and no individual data.
HR	Heart-rate in Beats/Min from a polar HR monitor working in conjunction with the Actigraph GT3X.
INCLINOMETER_OFF	Determines how many seconds the device inclinometer was turned off during a segment. The devices were on over 99% of the time.
INCLINOMETER_STANDING	Determines how many seconds a person was standing. The number represents how many sec of the 10 second epoch a person was standing.
INCLINOMETER_SITTING	Determines how many seconds a person was sitting.

	The number represents how many sec of the 10 second epoch a person was sitting.
INCLINOMETER_LYING	Determines how many seconds a person was standing. The number represents how many sec of the 10 second epoch a person was sitting.
MET_RATE	Metabolic Rate as determined by the Actilife software.
LATITUDE	WGS 84 Latitude of the Subject at the beginning of the segment
LONGITUDE	WGS 84 Longitude of the Subject at the beginning of the segment
Z_GPS	Elevation in Meters from the Qstarz GNSS at the beginning of the segment.
MONITOR	Actigraph GT3X monitor number worn by the subject.
PLATOON	Subject's platoon during USMA Cadet Basic Training.
SQUAD	Subject's squad during USMA Cadet Basic Training.
GPS_NUM	Number of the backup GNSS worn by the subject
SEX	Sex of the subject
AGE	The GPS number of the beacon they carried. Not the Qstarz GPS number that I used for my data collection.
HEIGHT	Height of the subject
WEIGHT	Weight of the subject wearing cargo pants, socks and a brown tshirt.
LOADEDWEIGHT	Weight of the subject with all gear worn during navigation event. Weight measured just before they started the navigation event.
LOAD	Loaded Weight - Weight
COMMENTS	Any comments made about the observation.
ZDIFF	Computed difference in elevation from the end of the segment to the beginning of the segment. Positive is uphill movement. Elevation values extracted from the LiDAR DEM.
RASTERVALU	Elevation values extracted from the LiDAR DEM. Represents the elevation at the beginning of each segment.
SLOPE	Computed % slope from ZDIFF. 100*ZDIFF/DISTANCE
FITNESS_SCR	Fitness Score of the cadet from their Army Physical Fitness Test. Range is 0-300.
RUN_SCR	Run Score of the cadet from the 2 mile run portion of the Army Physical Fitness Test. Range is 0-100.

CUMDIST_1	Distance the subject has walked in meters since the start to the end of the segment.
CUMTIME	Duration in seconds the subject has walked in meters since the start to the end of the segment.
COMPANY	The company the subject was in during USMA cadet basic training.
VALID	Status of the Qstarz GNSS reading
PDOP	(Positional Dilution Of Precision) ; Position accuracy; 3Dcoordinates
HDOP	(Horizontal Dilution Of Precision); horizontal accuracy; 2Dcoordinates
VDOP	(Vertical Dilution Of Precision); vertical accuracy; height
NSAT_USED_VIEW_	Number of Satellite (in Used, in View)
SAT_INFO__SID_ELE_AZI_SNR_	Information about the satellites used
NAVTIME	Same as start time
SPEED_GPS_1	Instantaneous GNSS Speed at the beginning of the segment
AVE10_1	Average Speed as computed by GNSS over the segment. Not used for dissertation.
ID	Numeric ID of the subject. Used to remove names from the dataset
BMI	Body Mass Index as computed using Height and Weight
EE	Energy Expended (Calories) per segment. Computed using METS and loaded weight.
EENORM	Energy Expended (Calories) per meter. Computed using EE and Distance of the segment.