DETECTING AND ASSESSING AGGREGATE HUMAN PEDESTRIAN MIGRATION PATTERNS USING HIGH RESOLUTION MULTISPECTRAL IMAGERY AND GEOGRAPHIC INFORMATION SYSTEMS

THESIS

Presented to the Graduate Council of
Texas State University–San Marcos
in Partial Fulfillment
of the Requirements

for the Degree

Master of SCIENCE

by

Warren D. Grannis, B.A.

San Marcos, Texas
May 2012

DETECTING AND ASSESSING AGGREGATE HUMAN PEDESTRIAN MIGRATION PATTERNS USING HIGH RESOLUTION MULTISPECTRAL IMAGERY AND GEOGRAPHIC INFORMATION SYSTEMS

Committee Members Approved:

Dr. Nathan A. Currit, Chair

Dr. Yongmei Lu

Dr. Jennifer Jensen

Approved:

J. Michael Willoughby
Dean of the Graduate College

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

ACKNOWLEDGMENTS

I would like to acknowledge the many contributions that made this thesis possible: Dr. Nate Currit, for his unwavering patience as I worked through numerous obstacles and challenges with this thesis. Many thanks to my thesis committee members Drs. Lu and Jensen, for their infectious enthusiasm about geography, and opening my eyes to the unlimited opportunities of learning more about our world through scientific inquiry.

I am grateful for the assistance of the employees of the U.S. Customs and Border Protection Agency, and the volunteers and staff of the Coalición de Derechos Humanos humanitarian group of Arizona in helping me get the data I needed for the thesis.

Lastly but most importantly, I would like to express my unending love and eternal gratitude for my wife and children, for their patience, love, and never-ending support as I pursued my schooling. Without their sacrifices, none of this would be possible. Thank you from the bottom of my heart.

This manuscript was submitted on April 17, 2012.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

ABSTRACT

DETECTING AND ASSESSING AGGREGATE HUMAN PEDESTRIAN MIGRATION PATTERNS USING HIGH RESOLUTION MULTISPECTRAL IMAGERY AND GEOGRAPHIC INFORMATION SYSTEMS

by

Warren Douglas Grannis, B.A.

Texas State University-San Marcos

May 2012

SUPERVISING PROFESSOR: NATE CURRIT

This project seeks to use a combination of high resolution multispectral imagery analysis and geographic analysis to predict pedestrian migration patterns, especially those intended to evade detection. Methods of analyzing multispectral imagery will be explored, and combine those methods with geographic analysis of known migrant locations and likely routes, thereby demonstrating that a correlation can be established between geographic factors and migration patterns.

I. INTRODUCTION

While illegal immigration has been a point of socioeconomic discussion and tension, as well as a major topic in political discourse for decades in North America, it has not been until 1993 with the then-newly established Clinton administration that the American federal government has taken a more aggressive approach at curbing said illegal immigration (Cornelius 2001). From the beginning of the 20th century, migrant numbers from Mexico fluctuated in time with the economic circumstances of the United States, until “an era of institutionalized migration for Mexican labor” (Harner 1995) between 1942 and 1964 known as the ‘bracero program’ solidified the demand for migrant labor in large numbers, when the demand exceeded the legal supply (Harner 1995). From the mid-1960s to the present day, illegal immigration from Mexico into the United States has continued in response to this demand. The 1980s and 1990s saw a series of laws passed with the intent to curb the numbers of illegal immigrants entering the United States. In the ensuing years since federal border security funding increased in 1993, increasingly more aggressive effort to reduce the numbers of illegal immigrants entering the United States have changed migration patterns, shifting once-common crossings into California and Texas, into less-stringently patrolled regions of Arizona (Cornelius 2001).

As a result of increased funding and staffing, border security enforcement has increased all along the Mexico/United States border in all but the harshest terrains and environments, where border enforcement via foot and vehicle patrols simply is not feasible (Harner 1995). In response to the increased border security elsewhere, much of the pedestrian illegal immigration routes have shifted to a desert region of the Mexico/United States border located in Arizona that is largely inhospitable (Harner 1995). This particular section of the border is formed by the southernmost boundaries of Pima and Santa Cruz counties (figure 1), located in an eastern section of the Sonoran desert known as Altar valley.

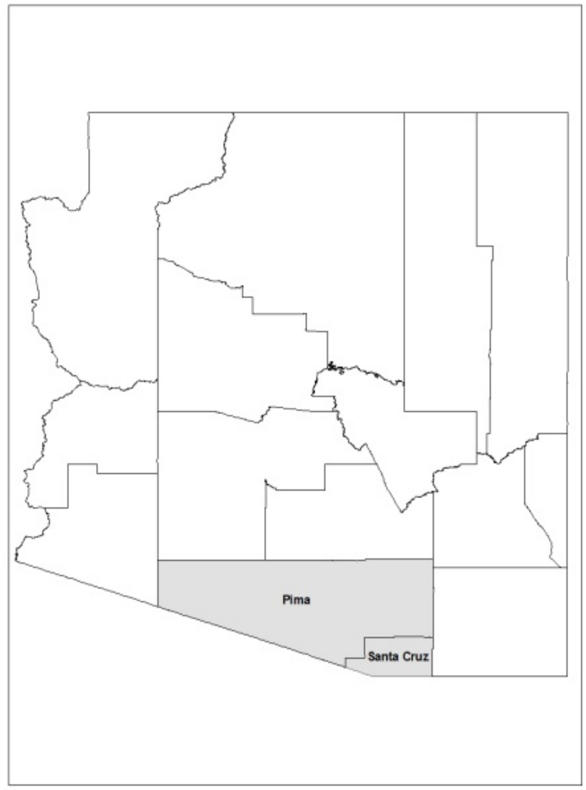

Figure 1—Study Area

Southern Arizona, including Pima and Santa Cruz counties, is part of the Sonoran desert, one of the largest and hottest deserts in North America. The terrain of the desert in the southern Arizona region consists largely of low mountain ranges and desert valleys, with elevations as low as 30.5 meters above sea level, and peaking at 1,255 meters above sea level on Baldy Peak in the Sierrita Mountains. Temperatures have been known to reach above 51.6^{o} C during the summers, with the potential to drop as low as 23.8^{o} C at night during the summer months (Desert Research Institute 2011). During winter months, daily temperature averages have the potential to reach 21.1^{o} C during the daytime, and fall to freezing temperatures at night. Rainfall averages for the region are around 7.6 to 10.1cm a year (Desert Research Institute 2011).

According to the 2010 U.S. Census, the two largest cities in Arizona are Phoenix with a population of 1,445,632, and Tucson with a population of 520,116 (U.S. Census 2010). These two cities are also the two primary initial destinations of pedestrian illegal immigrants crossing the border through southern Arizona (Coalición de Derechos Humanos 2011).

Illegal migration routes provide a unique challenge to the field of geography, and geographic information science, in particular, because they introduce a unique human component that is not often introduced in other types of geographic analyses: pedestrian behavior. How does one detect and analyze human pedestrian behavior that is affected by local geographic factors, especially when that behavior is intended to evade detection?

Most research of human pedestrian behavior deals with the urban, built environment (Greenwald 2001, Handy 2006). Factors such as crowds, obstacles, social relationships, evacuation and egress have all been discussed at length, and seemingly have little application to a “raw” and harsh environment that seems to have little potential for development. Despite these challenges in research, there are aspects of the research that prove useful to understanding how illegal pedestrian migrants choose their routes to their target destinations. As geographic information science seeks to understand identifiable, quantifiable, data, the 'hidden' nature of illegal migration routes requires creative methods for extracting potential information that may not be easily found or identified, and various research provides insights for methods of identifying them.

While border enforcement presents one challenge to the route decision-making process for those who embark on illegal pedestrian transborder crossings, those crossing routes face the potential to be additionally complicated by a plethora of environmental and geographic factors, such as weather, terrain, vegetation, water (or lack thereof), and others. For example, triple-digit summer weather may compel migrants to seek routes that provide access to shade, water sources, and concealment from law enforcement as they wait out the hottest parts of the day. Heavy rains may cause flash flooding or heavy water flows that create dangers for migrants who would otherwise swim or ford water crossings, forcing them to seek out other routes. Steep or otherwise rugged terrain may limit migrants' route options. Local vegetation or the lack thereof can affect migrants' decision-making process when considering concealment and fuel sources for food, warmth and shade.

The question I pose is: Can transborder routes be identified using GIS related techniques, including analysis of relevant high resolution multispectral imagery and recovered migrant remains data, in conjunction with an understanding of pedestrian behavior? By utilizing various analysis techniques, I seek to demonstrate that a correlation can be established between geographic factors and migration patterns. That is, I intend to identify migrant travel routes based on geographic factors associated with discrete measurement locations along an existing, but unknown, pedestrian route. This question will require analysis in three different areas. First, what are the spatiotemporal patterns of the locations of recovered remains of illegal immigrants? Second, based on known geographic factors influencing pedestrian behavior, what routes are illegal immigrants likely taking? Third, can those projected routes be verified by using high resolution multispectral imagery?

II. CONCEPTUAL FRAMEWORK

While literature exists that discusses at length various techniques for trail detection, and other literature discusses the impacts of geography on migration behavior, there has not been any literature made available to the public regarding techniques to identify potential pedestrian migration routes by combining trail detection via multispectral analysis and analyses of geographic influences on pedestrian behavior. While there are various articles and papers published touching on techniques to detect pedestrian trails, and papers and articles discussing geographic influences on the behaviors of immigrants, no research as far as I am aware has sought to combine the approaches for route detection.

Pedestrian Behavior in Selecting Routes

i Bort et al. (2010) assemble a comprehensive assessment as to how pedestrians behave when making route choices in walkable networks. While the setting for their research is urban Paris, a section of their assembled literature discusses at length a rubric of sorts for the parameters that come into play.

The authors note that a pedestrian's route, regardless of the settings and circumstances, will incorporate a tactical process that seeks to create a continuous, if not non-stop, path. Pedestrians will have a fixed starting point and a given destination point, and a potential network that includes alternative possibilities for arriving at the destination. These alternative possibilities can serve multiple purposes: to accomplish an intermediate objective, to avoid an unforeseen obstacle, or some other possibility, but ultimately all alternative possibilities exist to comply with "basic human natures: maximisation of utility." (i Bort et al. 2010) In the end, the various route possibilities serve to satisfy the decision-making process that best minimizes walking costs of the pedestrian's network. i Bort et al. (2010) take specifically note that of all possible factors influencing route decision, "the steepness of slopes accounts for the most important quality factor" in network accessibility. This key point, addressing the critical importance of a network's steepness of slopes, is a readily quantifiable attribute I can measure in my project. The authors go on to state that slopes of up to 25% (22.5^{o}) can be surmounted by healthy pedestrians, but would pose a significant challenge to those who are mobility-impaired or otherwise disabled. This provides me with a readily identifiable parameter that I can incorporate into my analysis. i Bort et al. (2010) identify a number of other factors that influence pedestrian behavior, but those factors are largely limited to the built environment, and therefore do not apply to this research.

Geographic Factors Influencing Migration Patterns

McIntyre and Weeks (2002) provide useful information that helps properly frame the project. Their paper addresses the various types of environmental impacts illegal immigrants may have as they attempt to cross the border. For one, the authors point out that wood is a common fuel source for migrant groups throughout the world. The authors also indicate most illegal immigrant deaths on the US/Mexico border are attributed to dehydration, hyperthermia, drowning, and motor vehicle accidents. With the exception of vehicular accidents, these factors have environmental or geographic components that can potentially be factored into this analysis, namely vegetation and natural fuel sources (also referred to as “duff”), proximity to water sources, water flow rates, and inclement weather. For example, if the region of focus is arid or semi-arid, permanent water sources such as rivers and lakes become more likely to have a significant influence on migrant routes in order to avoid dehydration. In addition, arid or semi-arid regions could present greater risks of death if they are large enough that they present a significant potential for death to pedestrians from exposure, and thus avoided if less risky routes present themselves. If the region of focus has enough water or moisture to support significant areas of woody plants and tall shrubs or trees, those wooded areas may become more important geographically to migrants for their potential fuel, shade, and cover from detection by law enforcement patrols. If particular areas are especially prone to flash flooding due to inclement weather, they could alter potential migrant routes due to the dangers they present.

McIntyre and Weeks (2002) mention that environmental impacts from transborder crossings are similar to environmental impacts associated with recreational forest use. Such impacts would be those which result from hiking and camping, which are essentially identical behaviors to those employed by pedestrians crossing the international border illicitly. While this is a valid point, the implication the authors make is that since the activities and end results are identical, the behaviors of transborder crossers will be essentially identical to those of recreational campers. This is erroneous largely due to the fact that recreational campers do not necessarily attempt to conceal their whereabouts or movements from other observers, whereas pedestrians illegally crossing the border have a motivation to conceal themselves as much as possible. The end impacts are the same (loss of duff, creation of new trails, soil disturbances and erosion, etc.) but the underlying differences in behavior between recreational campers and pedestrian migrants engaging in illegal transborder crossing are important. Migrants on the other hand, are not necessarily concerned with preservation of the current environment, but will take whatever evasive actions they need to in order to prevent being observed by the authorities. The point still stands, however, that there will be similar environmental impacts on the part of illegal border crossers in terms of hiking, camping, the burning of duff, and trails to water sources.

Rossmo et al. (2008) discuss methods of detecting crossing patterns. First, the authors raise the topic of viewsheds, and how they relate to the selection of border crossings. This rectifies the implied assumption made by McIntyre and Weeks (2002), as Rossmo et al. (2008) suggest that migrants are likely to select a location that is not easily surveilled by patrolling law enforcement officials. This reinforces my belief that efforts to evade surveillance will alter behavior patterns. Unfortunately, Rossmo et al. (2008) focus primarily on crossing locations as a single stationary point, and fail to consider the entire route migrants might take once they have actually crossed. What’s more, the authors make no suggestion of utilizing remote sensing in identifying potential routes between crossing and detention, despite having obtained data that identified where individuals were detained.

Smith, Bond, and Townsley's (2009) research provides an understanding of the decision-making process of individuals as it pertains to their geographic selection of a crossing point and route (an important factor in illegal immigration, as demonstrated by Rossmo et al. [2008]). Smith et al. (2009) make it a point as they discuss "journey to crime," to recognize that until the recent development of environmental criminology, focus on the geographic aspects of crime tended towards the final 'when and where.' With little or no consideration to the factors that influenced the end result, crimes were largely investigated as a single static moment in time, isolated from the time and space leading up to, and following the singular event. "Environmental criminology" (Smith et al. 2009) has greatly benefitted from the inclusion of journey-to-crime as a factor in the occurrence of crime, as it forces an assessment of the associated behaviors and decisions leading up to the event, and a significant number of cases have been solved because of that inclusion.

Another aspect touched on by Smith et al. (2009) is the discussion of crime pattern theory, which "postulates that individuals have certain psychologically intimate or familiar locations…called nodes or anchor points. Certain routes between nodes are usually preferred over others." This point furthers i Bort et al.'s (2010) assertion that pedestrians have a fixed starting point and target destination. This in turn strengthens the premise that there is purpose and intent behind an individual's choice of where they cross the border, and where they are attempting to arrive. Smith et al. (2009) point out that environmental criminology expands from an investigation of the singular event as a static, isolated event, to an explanation of crime events "in terms of an individual's interaction with their physical setting" (Smith et al. 2009). They go further by stressing that it is crucial to assess what the most appropriate criteria are for analysis. To put it succinctly, "issues of aggregation need to be resolved before proper conclusions can be arrived at." In other words, what factors correlate best with the end results?

While Smith et al. (2009) raise a number of excellent points in the discussion of journey-to-crime as related to the motivations and mindsets for the selection of nodes and routes, what is problematic in regards to this project is that Smith et al. (2009) describe the journey as for the commission of crime, not that the journey is a commission of crime. The reason this approach will not work for this project is that Smith et al.'s (2009) assertion that journey-to-crime is largely based on preference, convenience, or other factors. Therefore, by their reasoning, journey-to-crime has the potential to vary and take any number of possible routes, as long as the perpetrator still arrives at the intended destination, because the route is not the most critical factor, but simply an influencing factor. In the case of this project, the assumption is that since the route is better described as journey-as-crime as opposed to journey-to-crime, the route will lack emphasis on preference, convenience, or other factors that can be controlled by the pedestrian. Instead, I anticipate that due to the harshness of the environment in the study area, more emphasis in the pedestrian's decision-making process for determining routes will be placed on maximization of efficiency and utilization of resources.

To summarize the literature, several key factors for the project have been identified: first, slope of the environment's terrain is a primary consideration in a pedestrian's behavior when selecting routes. Second, 1m resolution multispectral imagery is sufficiently adequate for PCA and NDVI analysis for trail detection. Finally, the act itself of pedestrian illegal immigration, by nature being a crime, helps set out the determination for what routes the pedestrian migrant will take, due to the need to remain unobserved by law enforcement during the journey.

The primary gaps, and by extension, shortcomings in the literature revolve mainly around pedestrian behavior. Nearly all research on pedestrian behavior is set in the urban environment, while the project's setting is an arid desert environment. In regards to the discussion of slope, insufficient discussion is placed on the scale of the analysis necessary for valid, quantifiable results: at what resolution will the analysis of slope help predict route selection? What is the minimum resolution necessary for an accurate estimation? Key assumptions are made about high resolution multispectral imagery, in that it is readily available for the area studied, and of sufficient quality to be properly analyzed in such a way as to get useable results. Finally, can all the techniques discussed in this variety of literature work together in such a way as to provide an adequate means of predicting pedestrian routes?

Trail Identification Using Remotely Sensed Imagery

Kaiser et al. (2004) most directly address the usage of techniques for analyzing multispectral imagery in detecting illicit transborder routes. Their article lends its greatest value in a broad discussion of a number of different methods and analysis techniques. Kaiser et al. (2004) point out the value of deriving trail maps from imagery by mentioning the fact that field surveys of trails consume inordinate amounts of time and money, and are relatively incomplete. Unfortunately, this is an assertion that seems to contradict itself. Kaiser et al. (2004) postulate that due to the time, expense and incompleteness of field surveys of trails, processing and analyzing remotely sensed imagery would logically be a superior process. They reassert this position in their findings in regards to accuracy, stating “that most errors in visual based mapping of small trails are associated with omission rather than commission errors." Indeed, as they illustrate various processed images, those images that have undergone analysis (specifically on PCA and NDVI analyses) and subjected to trail extraction show a significantly larger number of trails appearing. But when the authors illustrate the trail lines created from field surveys, they delineate between “obvious” and “suspected” trail features, despite the fact that the trail map is much sparser than maps that have been created based on PCA and NDVI analyses. It is implied that an ideal map identifying potential trails is one created from on-site field surveys: given unlimited time and resources, the field survey map created represents the best of human capabilities.

This contradiction raises the argument—does image processing "detect" non-existent trails? Or is image processing such a superior process that field surveys are more than just prohibitively expensive and time-consuming, they are actually inferior to what technology has to offer? Given that this paper is among some of the earliest published research on this topic, it would have been wise on the authors' part to either give more clarification on the point of whether or not image processing is a feasible substitute for field work. Barring that, it would have been beneficial if the authors had better described exactly how thorough the field survey map was, in terms of time and effort. In either case, Kaiser et al. (2004) render the contradiction a moot point by suggesting that rather than leave image processing a fully automated process, errors in processing can be significantly reduced by creating a hybrid process—an automated feature extraction combined with a human, visual analysis of false color imagery.

Cao et al. (2007) provide a more current discussion of the same topic. The paper differs from Kaiser et al. (2004) in that while Kaiser et al. (2004) focus on trail extraction, Cao et al. (2007) focus on the potential to revise and update extracted features based on changes in trail usage. That is, Cao et al. (2007) expand on the idea of digitally identifying and extracting potential trails, by adding the concept of analyzing images of the same region from different dates, and identifying differences in potential trails. The end idea is that by analyzing different images of the same region (covering different dates), researchers can more readily identify changes in trail usage, without the need to redo field work for the entire area in question. In the discussion of challenges related to accurate trail map revision, Cao et al. (2007) point out that ultimately, automatic trail extraction and updating creates a greater challenge than automatic road extraction, “due to the fact that trails have less geometric and radiometric consistency than roads, and trails are often broken by overhanging vegetation and shadows."

Witztum and Stow (2002), despite not having a focus on transborder trails, provide a relevant discussion by discussing human impacts on the environment via newly established trails. This has bearing on properly understanding the nature of transborder trails, especially those that are being newly established. Witztum and Stow (2002) establish the usefulness of high resolution (1m) imagery in trail detection, and examine the limits of this kind of imagery in detecting trail features. The authors’ main stipulation on the usage of high resolution imagery is that radiometric and geometric corrections need to be made to datasets to ensure accuracy of any analyses. Without these corrections, false results may occur and render the analysis inaccurate. Witztum and Stow (2002) note, as have Cao et al. (2007) and others in similar literature, that there is comparatively little published research on trail detection. This appears to be due to the complexities introduced by the natural landscape, primarily a lack of geometric predictability and spectral variations among vegetation.

Coulter and Stow (2008) further explore multispectral imagery analysis, again touching on methods for detecting human impacts on vegetation. Coulter and Stow’s usage of 1m multispectral imagery helps establish a baseline for what would be considered a sufficient resolution for this type of assessment. Earlier papers, especially in the light of a lack of published research, did not inspire much confidence in the sufficiency of 1m resolution in identifying pedestrian trails, despite the reference of “high resolution.” Coulter and Stow (2008) also reinforce the idea that this sort of analysis is both cost-effective in terms of time and money, and is sufficiently acceptable in terms of accuracy when detecting new trails.

Another key factor Coulter and Stow (2008) introduce is the idea of generating change products covering a significant period of time. This concept, generating an output that shows changes in trails, ties in with research Cao et al. (2007) discuss in terms of identifying changes in pedestrian trail patterns. This suggests the feasibility of using imagery from similar seasons but different years, to detect changes in recent transborder migration trends. For example, if historical data suggesting particular seasonal patterns in the springtime follow a particular route over several years, and recently obtained imagery shows a consistent, predictable, deviation from that historical norm, researchers could then assume that transborder migration trends have responded to some particular influence (changes in terrain, law enforcement patrolling routes, environmental factors such as unseasonal drought or rain, etc.) and respond to those changes accordingly.

While Coulter and Stow (2008) do not discuss trail detection specifically, the techniques they employ in their research are very similar to techniques cited earlier, strengthening the argument that such techniques are both feasible and acceptable for route detection. This is especially significant, given that studies in this particular type of data extraction frequently comment on the unique challenges trails present in identification.

By utilizing a GIS analysis of slope and other factors representing pedestrian cognition and selection of routes as they relate to pedestrian migrant data and border crossings, I hope to predict potential pedestrian migrant routes intended to evade detection by border enforcement in southern Arizona, and verify those predictions via PCA and NDVI analysis.

III. METHODOLOGY

Data

Interviews

Interviews occurred between March and May of 2011with staff and volunteers from the Coalición de Derechos Humanos humanitarian group of Arizona. Eight interviews were conducted via email, and a number of informal "chats" were conducted via social chat messenger programs. The interviews were conducted to gain an understanding behind the social constructs and behavior of pedestrian migrants crossing the international border from Mexico into southern Arizona.

From these interviews, the following information was obtained: pedestrian migrants crossing the border rely on the services of 'coyotes,' guides who are familiar with the region being traversed, and employed specifically for the purpose of assisting groups in crossing. The 'coyotes' work for the Mexican drug cartels, and actively work to prevent routes from being discovered. Additionally, 'coyotes' take routes that avoid rugged terrain and hills whenever possible, but also seek to avoid detection by law enforcement by avoiding known observation points. Lastly, pedestrian migrants entering southern Arizona from Mexico almost always have one of two target destinations in mind: Phoenix or Tucson. The reason for this is because simply entering into the United States is not the ultimate goal or destination: oftentimes there is another destination somewhere in the United States where the immigrant has pre-established contacts, and Phoenix and Tucson serve as hubs for continuing on to the final destination. Nothing about the journey is accidental: the funds paid to the ‘coyote’ are not just for getting to Phoenix or Tucson, they are for getting in touch with individuals called ‘reiteros,’ who arrange transportation to the final destination.

Recovered Migrant Remains Database

During interviews with Coalición de Derechos Humanos, information was given regarding the Arizona Recovered Human Remains Project, a program by Derechos Humanos to record the numbers and locations of migrant remains recovered by authorities. The data as they provide it for each set of remains consists of the name, gender, age, country of origin, date the remains were discovered, location the remains were discovered, and cause of death. Each field is filled out as comprehensively as can be, but often detail is lacking and the names, genders, ages, and causes of death are left blank (Table 1).

Table 1—Sampling of raw data provided by Derechos Humanos

Nmbr	Name	Gender	Age	Country	Date Disc'd	Location Discovered	Cause of Death
131	Unknown	F	Unknown	Unknown	5/29/2005	5.5mi S of Hickwan	Exposure
132	Unknown	M	54	Mexico	5/29/2005	Elephant Head Rd W. of I-19	Exposure
133	Unknown	M	Unknown	Unknown	5/30/2005	N 31 59.389 W 112 078	Exposure
134	Unknown	M	Unknown	Unknown	6/1/2005	Desert Area 1.5mi S. of Via Montana Vista	Exposure
135	Unknown	M	Unknown	Unknown	6/2/2005	Arivaca and Amado Rd	Exposure
136	Unknown	M	Unknown	Unknown	6/2/2005	Arivaca and Amado Rd	Exposure

As explained by i Bort et al. (2010), pedestrians who embark on a route will have a definite starting point, an intended destination, and a route with potential alternative detours. Whether pedestrian migrants successfully arrive at the intended destination, are apprehended by law enforcement en route, or pass away due to circumstances related to their attempted migration, the general route will still remain unchanged.

The Arizona Recovered Human Remains Project data was digitized and converted into a georeferenced database of spatiotemporal points for analysis. Because the Project's data covered the period from 2000 to the present day, and has over 2,000 entries, it was necessary to eliminate some data to make for a more meaningful analysis. The initial data pool was limited to entries recorded in fiscal years 2005 and 2007, to coincide with NAIP high resolution multispectral image captures available from USGS covering the southern Arizona region. Next, any entries that did not have latitude and longitude entries for where the remains were discovered were eliminated from the two data sets. This was the result of needing as accurate a location as possible. Next, remaining entries were checked to see if they had valid latitude and longitude coordinates. If there were obvious errors (coordinates that placed a recovery location in an unfeasible location, such as in the ocean, clearly outside the study area, for example), they were checked for any potential data entry flaws (misplaced decimals, extra digits, etc.) before being eliminated if the errors could not be reconciled. Fiscal year 2005 had 30 useable samples and fiscal year 2007 had 100 useable samples, for a total of 130 samples (Figure 2). From these samples, the only attribute data retained for the recovered migrant remains was of discovery location coordinates.

GIS data

GIS data used for the project consisted of, in addition to the recovered migrant remains data, elevation data obtained from the USGS in the form of digital elevation maps. The area covered by the elevation data consisted of the southern Arizona region, including the extent of the spatiotemporal points of the recovered migrant remains data, and the cities of Phoenix and Tucson. The data has a 3m spatial resolution, and was utilized to create a slope raster to be used in least cost path calculations (Figure 2).

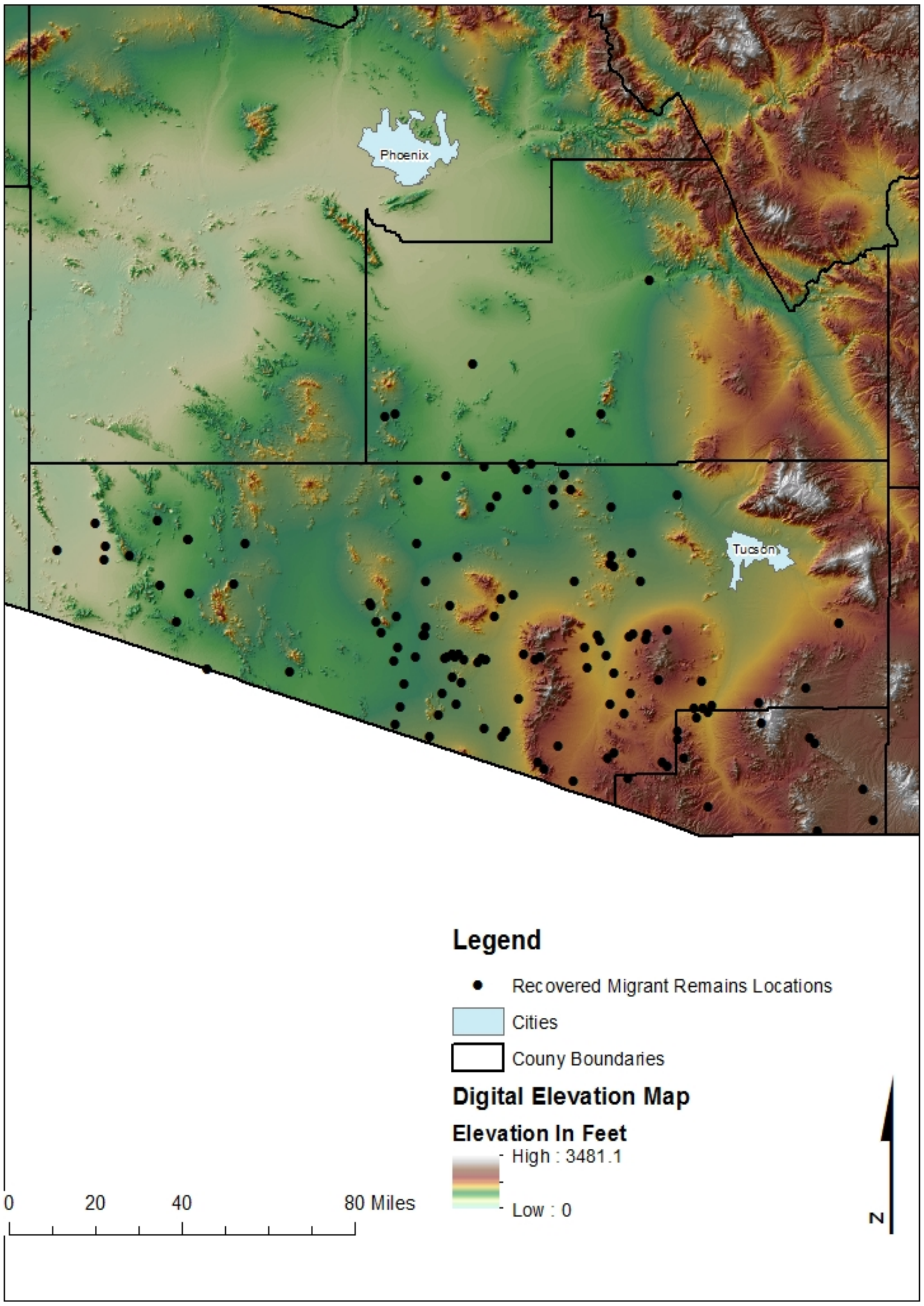

Figure 2—Illustration of recovered migrant remains from the years 2005 and 2007

Elevation data extended from the Arizona-Mexico border to north of Phoenix and spanned the entire east-west distance of Arizona. The objective was to cover as broad an area as possible to allow for a comprehensive least cost path analysis that would not be unnecessarily limited and potentially give a false projection of potential trails.

The other GIS data utilized consisted of the political boundaries of Phoenix and Tucson, and the international border separating Arizona and Mexico. These were also used in the least cost path analysis.

Satellite Imagery

Satellite imagery, like the elevation data, was obtained from the USGS Seamless server (USGS 2011). The remotely sensed imagery consists of 2007 NAIP multispectral 1m high resolution imagery consisting of four spectral bands: red, green, blue, and near-infrared. Due to the size of individual images and the constraints of processing and analysis, it was unfeasible to download NAIP imagery at the same extent of geographic coverage as the elevation data. Instead, imagery tiles were downloaded for areas that had been identified as having higher-than-average concentrations of recovered migrant remains locations. This would allow for one image to be processed and yet allow for an analysis of multiple potential trail locations near recovered migrant remains locations, rather than a single image for a single location.

In addition to using concentrations of recovered migrant remains locations, imagery was downloaded for regions that were identified as containing potential routes for multiple recovered migrant remains locations, but not actually in the proximity of those locations. This was done to allow for analysis verification of projected routes from the border to recovered migrant remains locations.

Methods

The goals of the methods are (1) to determine the spatial pattern of recovered remains, (2) build a least cost path model to determine likely immigrant routes, (3) identify actual migrant paths, and (4) assess the accuracy of the least cost path model by comparing its resultant routes to actual paths followed by migrants. The methods used in this thesis are Nearest Neighbor Analysis, Least Cost Path Analysis, and image analysis consisting of Normalized Difference Vegetation Index and Principal Components Analysis, and a manual verification of results.

GIS Analysis

Nearest Neighbor Analysis

Nearest Neighbor analysis is a type of spatial pattern analysis. The technique calculates a statistical index based on the average distance from each feature to its nearest neighboring feature. In ArcGIS (ESRI, Redlands, CA), usage of this tool provides five values: the expected mean distance between features, the observed mean distance between features, a nearest neighbor index, a z-score, and a p-value. These results measure statistical significance which will indicate whether or not to reject the null hypothesis, which in the case of Average Nearest Neighbor analysis, is that features are randomly distributed.

The nearest neighbor index is expressed as the ratio of the observed mean distance to the expected mean distance. The expected mean distance is the average distance between features in a hypothetical random distribution. A value of less than 1 in the index indicates that a pattern exhibits clustering. If the value is greater than 1, the pattern tends toward dispersion. It is important to accurately define the extent of the study area for this analysis.

Least Cost Path Analysis

A Least Cost Path analysis (LCPA) is intended to calculate a path from a source to a destination based on a travel cost from source location to destination location. The analysis will produce an output raster that records the path or paths that lead from selected locations to the closest source cell defined by a cost surface. The process is based on Dijkstra's Algorithm, which finds all possible paths between a source vertex and a destination vertex, and assigns a value to each path as it is encountered in the process. After all possible least-cost paths have been assigned values; the optimal path is identified and selected.

Image Analysis

Normalized Difference Vegetative Index

Analyzing multispectral imagery using the Normalized Difference Vegetative Index is well-suited for this project because a key advantage of this process is that it cuts back on many forms of radiometric noise present in multispectral imagery, and allows for a better identification, in this case, of obscured features such as pedestrian trails. The algorithm that defines NDVI is as thus:

$$NDVI = \frac{\mathrm{P}nir - \mathrm{P}red}{\mathrm{P}nir + \mathrm{P}red}$$

where P*nir* is near-infrared radiant flux and P*red* is red radiant flux in a multispectral image, and takes advantage of the inverse relationship between red and near-infrared reflectance associated with healthy vegetation.

Principal Components Analysis

Principal Components Analysis is a technique that projects a remotely sensed dataset into a new set of uncorrelated variables that represents the information in the original dataset. Like NDVI, PCA also has a benefit of reducing extraneous radiometric noise, allowing for a better detection of obscured features. Each variable or component is derived from the original dataset to account for the maximum proportion of each variance.

Least Cost Path Model Verification

The analysis of LCPA projections are verified through a count of actual versus expected trails found in processed images. Route projections represented by LCPA results are broken into portions represented by individual routes and merged, aggregate routes. Among the route portions, a random sampling of route projections are selected and given quarter mile buffers from the route center line, for a total of a half-mile buffer, to account for minor delineations and variations in the terrain which might affect pedestrian travels, such as small hills, rock outcroppings, gullies caused by erosion, etc.

Within route portions, estimates are given for the number of expected trails to be found in the processed image, and each image is manually inspected to identify potential trails.

IV. FINDINGS

Nearest Neighbor Analysis

The recovered migrant remains data underwent Nearest Neighbor (NN) analysis to determine whether the locations of discovered remains were randomly dispersed, uniformly distributed or clustered. If the results indicated that the recovered migrant remains were not randomly distributed, this would indicate that many migrants follow (and die) along the same routes, an indication of pedestrian behavioral patterns. The sample set of recovered migrant remains for 2005 after undergoing NN analysis was determined to be non-random, and clustered (Table 2).

Table 2—Average Nearest Neighbor Summary for 2005 sample of recovered migrant remains data. Given the z-score of -2.67, there is less than 1% likelihood that this clustered pattern could be the result of random chance.

Observed Mean Distance:	0.168117
Expected Mean Distance:	0.223264
Nearest Neighbor Ratio:	0.752997
z-score:	-2.673055
p-value:	0.007516

The sample set of recovered migrant remains for 2007 after undergoing NN analysis was determined to be non-random and clustered (Table 3).

Table 3—Average Nearest Neighbor Summary for 2007 sample set of recovered migrant remains data. Given the z-score of -4.99, there is less than 1% likelihood that this clustered pattern could be the result of random chance.

Observed Mean Distance:	0.079808
Expected Mean Distance:	0.106733
Nearest Neighbor Ratio:	0.747737
z-score:	-4.992023
p-value:	0.000001

Aggregated, the remains recovered in 2005 and 2007 were determined to be non-random, and clustered (Table 4).

Table 4—Average Nearest Neighbor Summary for aggregated recovered remains data covering 2005 and 2007. Given the z-score of -6.73, there is less than 1% likelihood that this clustered pattern could be the result of random chance.

Observed Mean Distance:	0.062797
Expected Mean Distance:	0.090213
Nearest Neighbor Ratio:	0.696097
z-score:	-6.730058
p-value:	0.000000

Because all three combinations of the data sets (2005 alone, 2007 alone, and 2005 and 2007 aggregated) returned results of non-random clustered patterns, this was a good indication of data reflecting pattern-based pedestrian behavior for this project.

Least Cost Path Analysis

Variables used in the Least Cost Path Analysis were slope, the southern borders of Pima and Santa Cruz counties (which form part of the international border between Mexico and the United States), points representing the locations of recovered migrant remains, and the city boundaries of Phoenix and Tucson. The slope raster was created from DEMs obtained from the USGS Seamless Server.

The permanent water bodies data were eliminated from consideration due to a lack of any permanent water bodies between the international border and RMR locations. All permanent water bodies were well north of RMR locations, and therefore unable to be a factor in the decision-making process of pedestrians for selecting their routes (Figure 3). The vegetation rasters were eliminated from consideration as well, due to the homogeneity of vegetation classification. Since there was no significant variance in vegetation types, it would have minimal influence in GIS analysis when trying to determine potential pedestrian routes.

Transportation networks were eliminated from consideration due to a unique problem they presented for analysis, which could not be reconciled: given that at some point from the time a pedestrian crosses the international border to arriving at his or her intended destination, the pedestrian would have to approach and cross a transportation corridor. However, despite the necessity of approaching and crossing routes on the transportation network, for the most part pedestrian migrants are also avoiding transportation networks to avoid potential detection from border enforcement who may be patrolling the roadways. An adequate technique or process could not be devised that would represent the contradictory nature of both avoiding and approaching transportation networks (Figure 4).

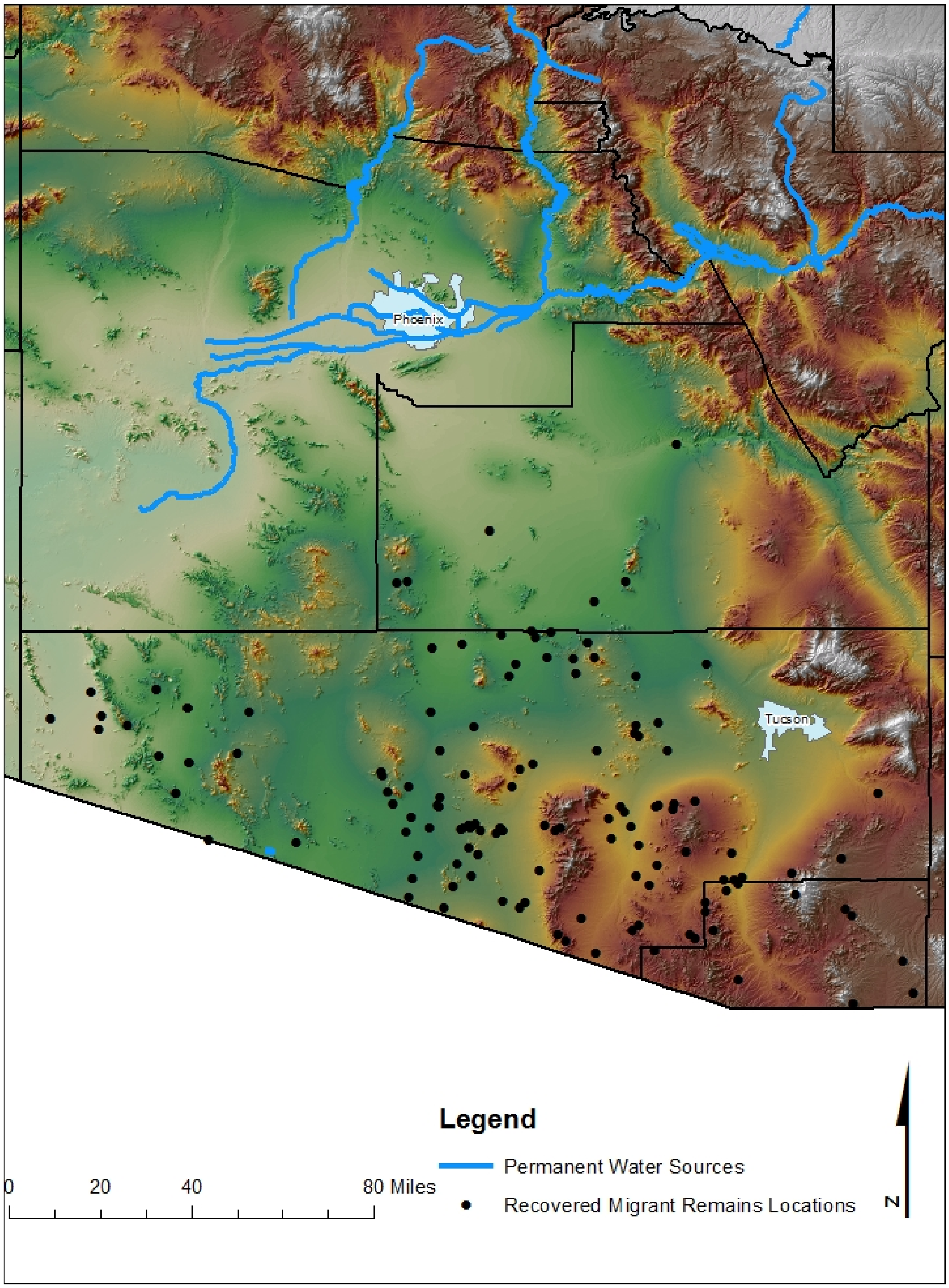

Figure 3—Permanent water sources in relation to recovered migrant remains locations

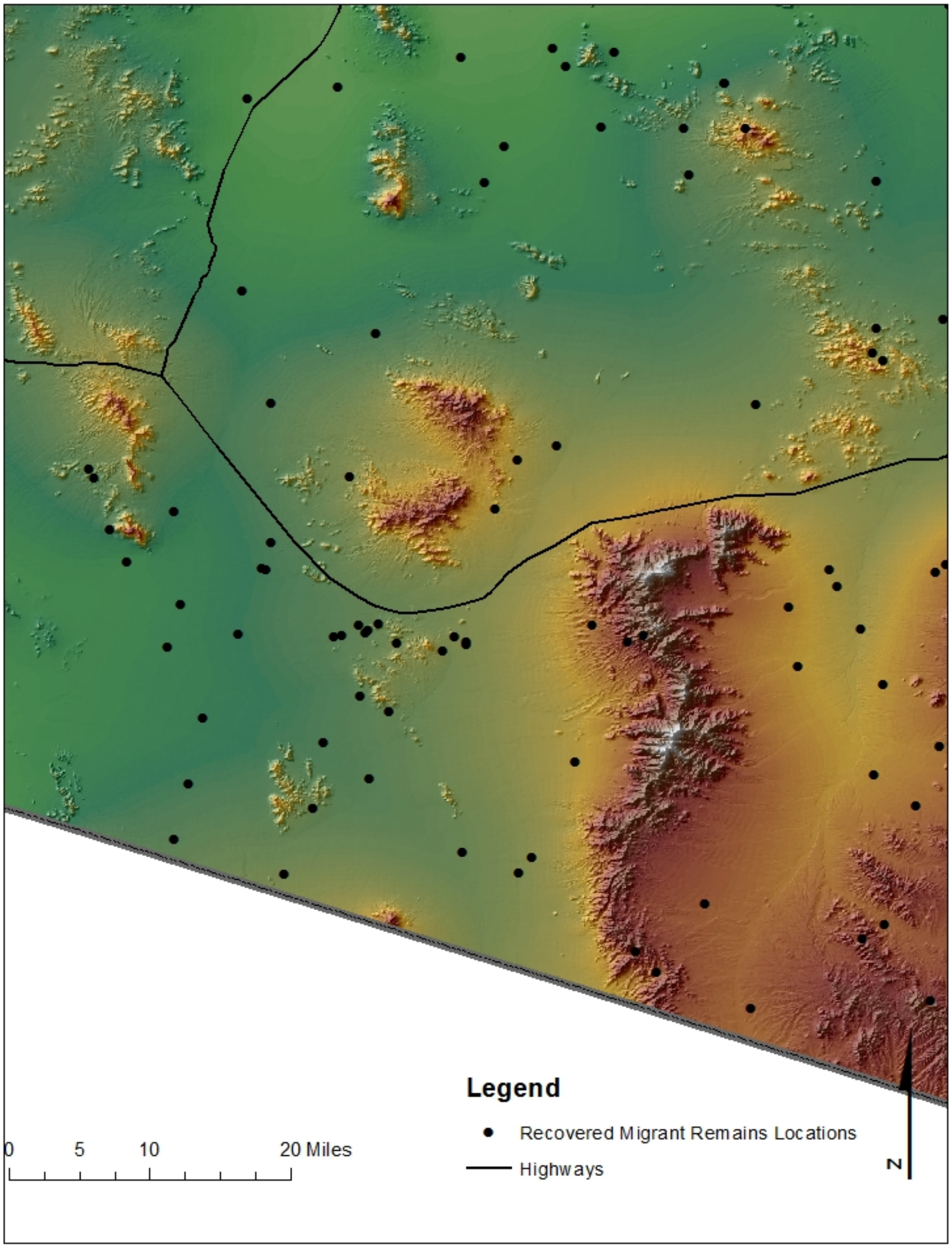

Figure 4—Illustration of relationship between highways and recovered migrant remains. Some pedestrian migrants, based on the locations of discovered remains, clearly have approached road networks, and then traveled away from them.

Political boundaries (other than the noted exemptions of the international border and city boundaries of Tucson and Phoenix) such as state, county, and city borders were determined to be largely irrelevant to the analysis. While they were used to assist in visual and spatial comprehension when reviewing data, they were not utilized in actual analysis.

Two versions of a final cumulative cost surface raster were created: one that utilized straight line distance cost from city boundaries and slope as cost weights in relation to remains locations and the border, and one that utilized slope only. These two different cumulative cost rasters created very different results for projected pedestrian routes (Figures 5, 6 and 7). I chose to use the results from the cumulative cost surface raster utilizing only slope, as the projected trails gave the most sensible results in terms of routes, as trails did not travel away from the destination cities.

While it seemed logical at the time to include city boundaries as a weight for projecting pedestrian trails, some resulting trails did not make sense—pedestrians appeared to be headed for, and indeed, arrived in Tucson, but then veered away from the city and headed for Phoenix. Those results directly contradict expert opinion that Tucson is indeed a destination city, and there would be no need to head away once arriving there.

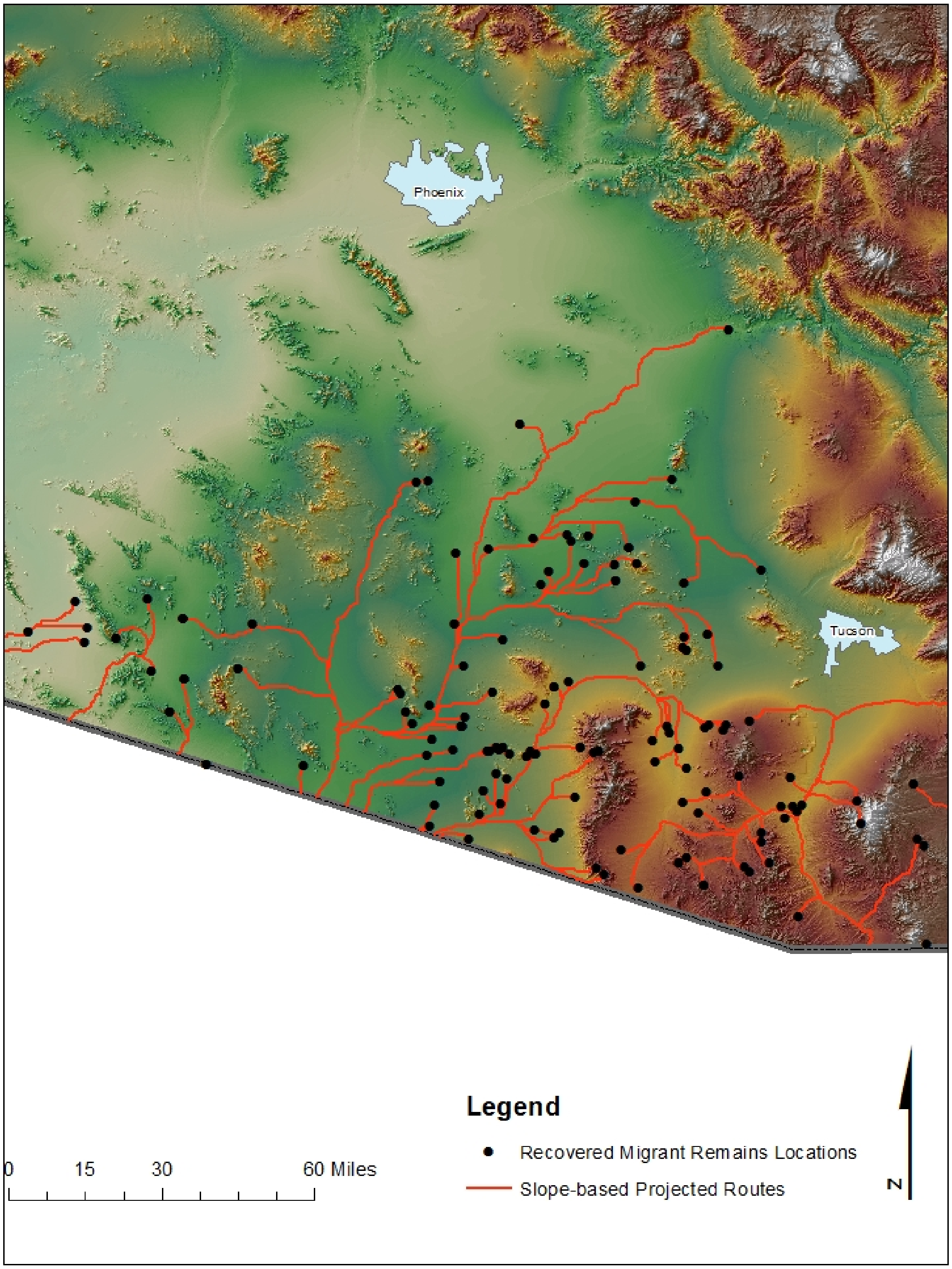

Figure 5—Slope-based routes. This map illustrates projected pedestrian routes utilizing only slope, as opposed to including weights based on the cities of Phoenix and Tucson.

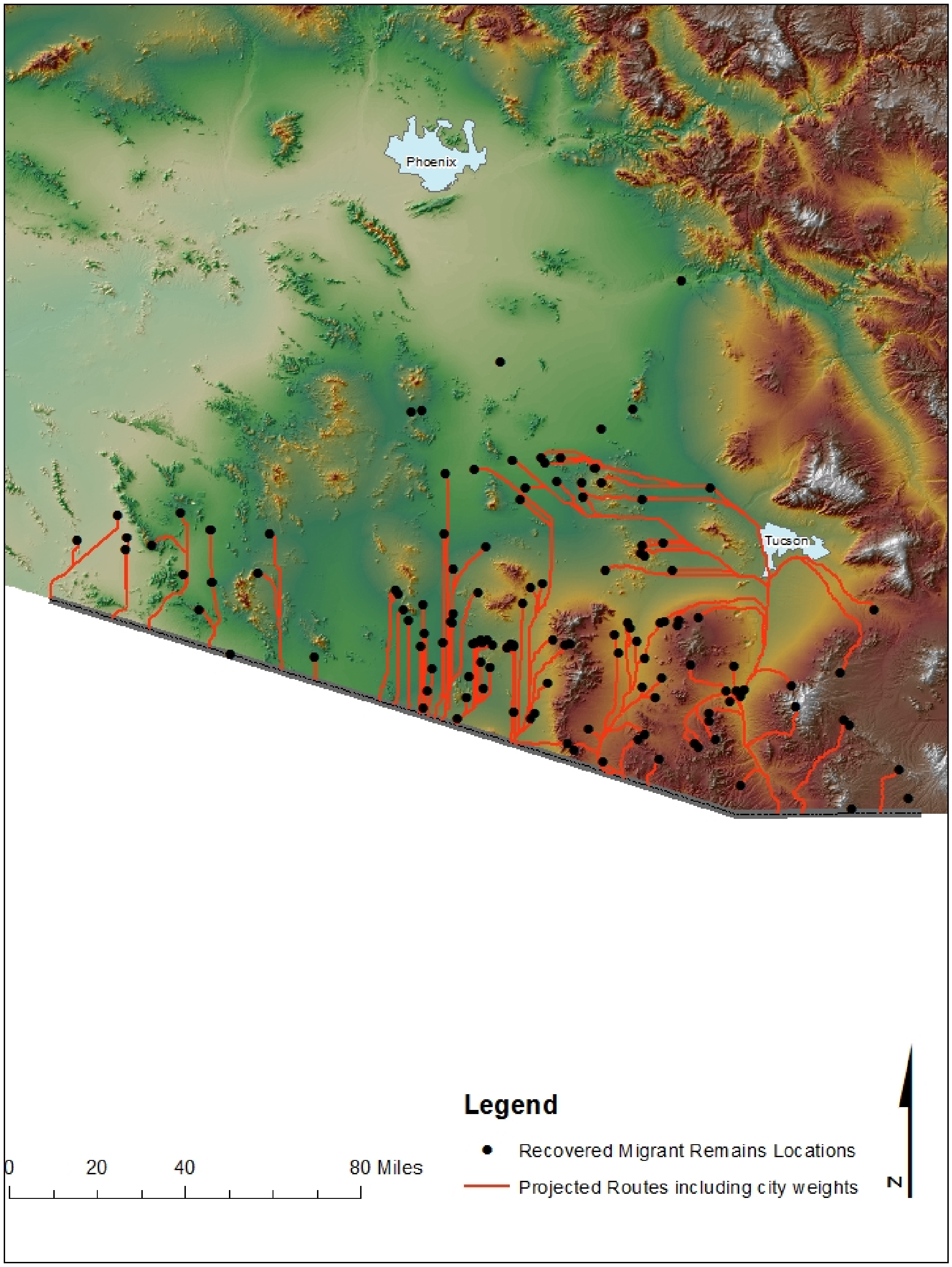

Figure 6—Routes with cost distance weight from city included. This map illustrates projected pedestrian routes factoring in weights from both slope and cities.

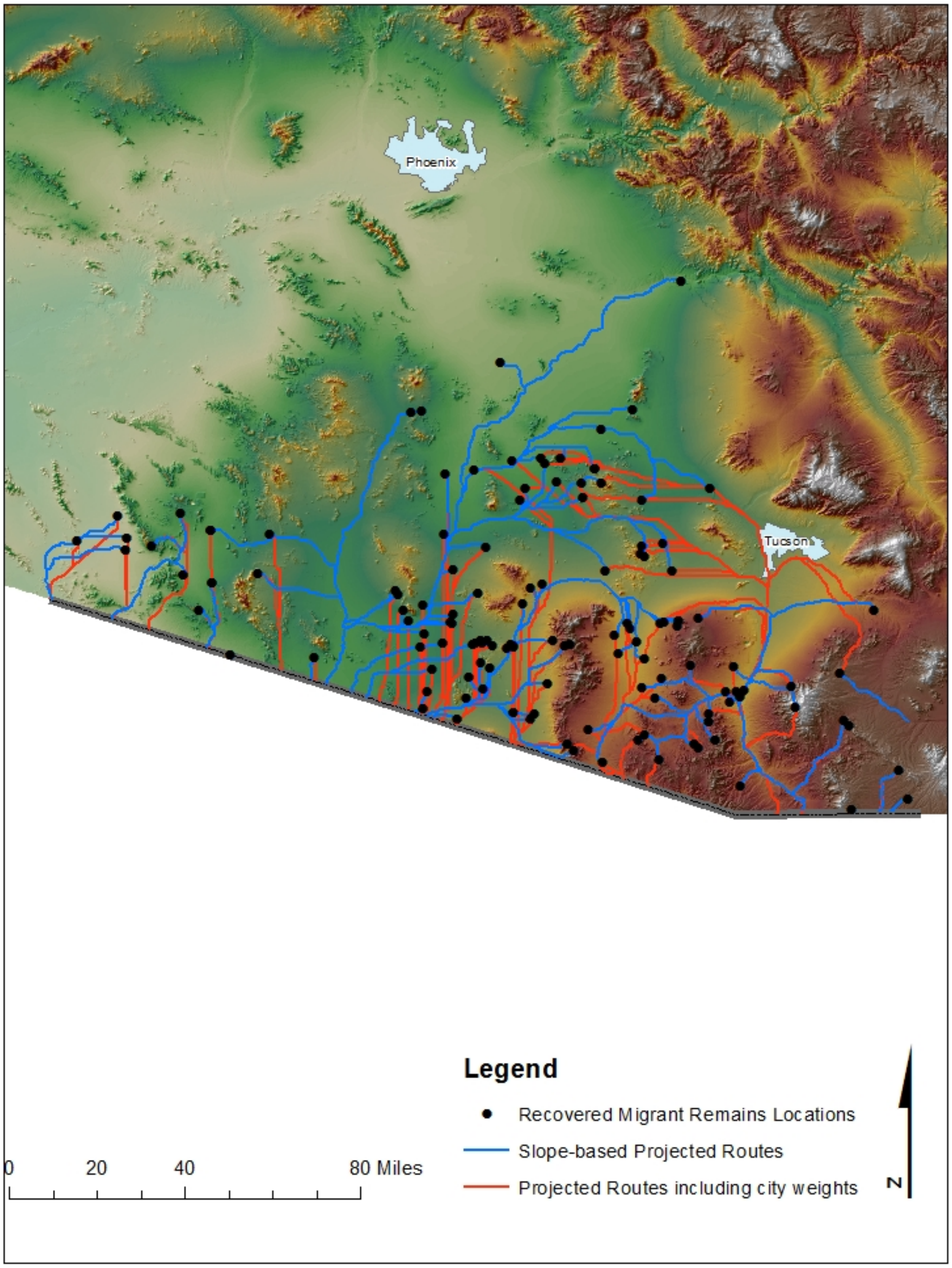

Figure 7—Comparative map illustrating slope-only and city-included projected routes. Note the differences in results illustrated in figures 5 and 6.

Image Analysis

PCA and NDVI analysis was utilized to facilitate manual trail identification by sharpening the differences between trails and surrounding land features. Unaltered multispectral imagery, while useable for identifying obvious transportation routes such as paved roadways, presented difficulty in allowing for ready identification of potential pedestrian trails. Little difference in advantages was seen in initial multispectral analysis between NDVI and PCA methods in terms of trail identification (Figure 8). In addition, it was discovered during the data acquisition phase of the project, that NAIP imagery from the USGS Seamless Server for 2005 in Arizona was only available in true-color composition (RGB spectral bands). While this still could have been utilized for trail identification, I elected to use only 2007 data for all imagery analysis for several reasons: first, the 2007 imagery included near-infrared spectral band data; and second, of the two sample sets, 2005 had the smaller group.

Usage of NDVI-analyzed imagery allowed for easier identification trails by sharpening differences between vegetative and non-vegetative surfaces, and reducing radiometric noise in the images. With this increased clarity in the images, features that previously could not be identified with certainty as trails took on a more definitive and distinct trail-type appearance: rather than seeming to be potentially natural, they stood out as definitely not being a “normal” part of the surrounding environment.

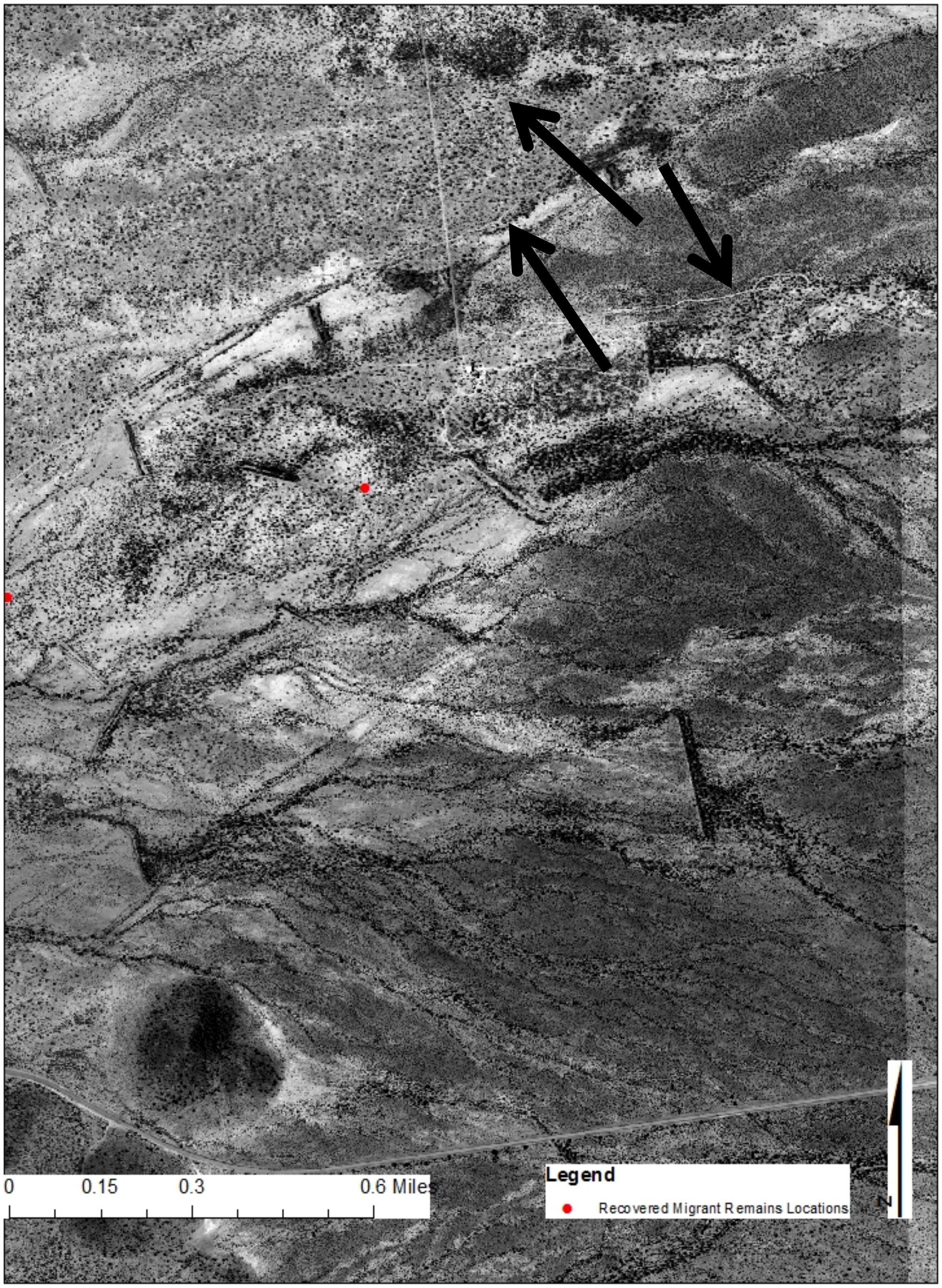

Figure 8—Map showing multispectral imagery after undergoing NDVI analysis. While previously visible trails still are visible, there are additional trails (identified with arrows) now identifiable.

Verification Analysis

Verification analysis was carried out by utilizing both data that resulted from GIS analysis and from multispectral imagery analysis. Thirty points were randomly selected for verification (Table 5).

Table 5—Average Nearest Neighbor Summary for set of verification samples. Given the z-score of .01, there is 98% likelihood that this pattern could be the result of random chance**.**

Expected Mean Distance:	.325167
Observed Mean Distance:	.257183
Nearest Neighbor Ratio:	.790926
z-score:	.013299
p-value:	.9894

30 spatiotemporal points selected for verification analysis in areas where projected trails were thought to exist. Each point was located within the projected routes' buffer zones (quarter-mile radius), of which each buffer zone containing a verification point selected to be inspected for potential trails.

Verification of projected routes led to some interesting results. All buffer zones containing verification points had at least one trail within a quarter mile of projected routes, and generally following the projection, if not always within a quarter mile (Figure 9). The number of observed potential trails did not always

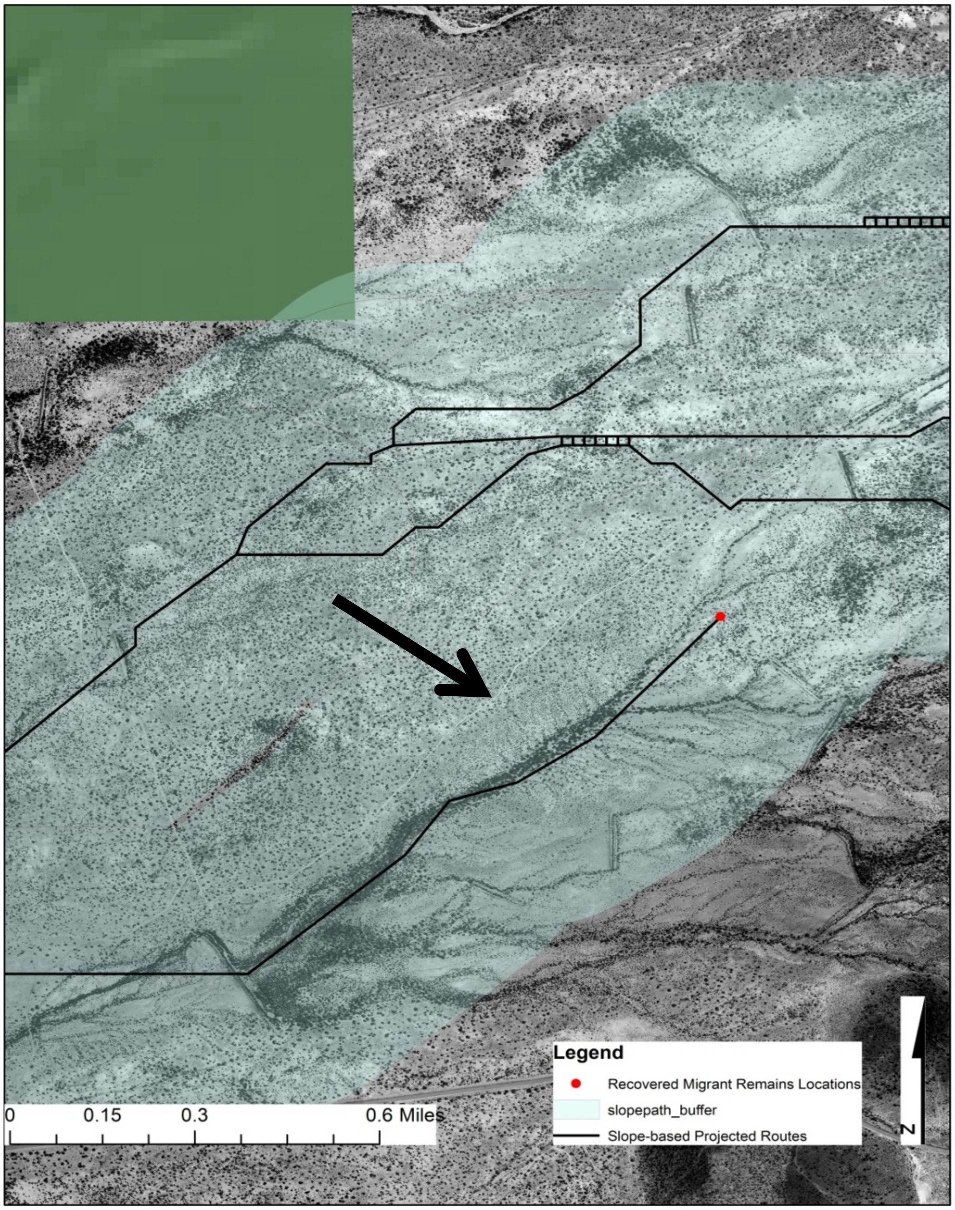

Figure 9—Map illustrating typical image results for sample image verification sites. Note a potential trail (noted by arrow) running roughly parallel between two main projected routes, within the projected routes' buffer zone (Image is portion of buffer containing site 11, with 2 expected, and 2 observed).

match the expected number of trails. In addition, potential trails were observed that fell outside of the projection buffer, but were not counted as my verification counts were limited solely to determining how many potential trails could be identified within the buffers.

This seeming discrepancy has a reasonable explanation: the route projections are based on the results of 130 sampled reference points from two different years. These 130 sampled reference points came from a potential pool of over 2,000 recovered migrant remains, of which are just a small sampling of the millions of pedestrian migrants who have crossed the international border from Mexico into Arizona over the last few decades and survived. The projected routes based on the sample set of recovered migrant remains do not preclude the possibility of other potential pedestrian migrant routes that might be projected from other sample sets.

Another possibility that exists is that some potential trails that were observed might not be the result of pedestrian migrants, but rather of animals or mundane human behavior such as camping and hiking. A final possibility for the discrepancies in observance of potential trails is the potential of normal erosion or washout from periodic flash flooding. Without the ability to do on-site verification, it is beyond my ability to explain with certainty what I observed in areas where expectations did not meet projections of potential route observations.

Table 6—Results of image verification sample sites. Contains a contrast of expected versus observed potential routes.

Sample Site	Expected	Observed	Difference
1	1	1	0
2	1	1	0
3	3	3	0
4	1	2	+1
5	2	1	-1
6	1	1	0
7	2	2	0
8	1	1	0
9	1	1	0
10	2	1	-1
11	2	2	0
12	3	1	-2
13	1	1	0
14	2	2	0
15	1	1	0
16	1	2	+1
17	3	3	0
18	2	2	0
19	1	1	0
20	1	1	0
21	3	3	0
22	1	1	0
23	1	1	0
24	2	2	0
25	1	1	0
26	1	2	+1
27	2	2	0
28	2	2	0
29	1	1	0
30	1	1	0

V. CONCLUSIONS

My objective for this project was to identify whether I could detect potential transborder routes using GIS related techniques, including analysis of relevant high resolution multispectral imagery and recovered migrant remains data, in conjunction with an understanding of pedestrian behavior. I sought to accomplish my objective by answering these three questions: first, what are the spatiotemporal patterns of the locations of recovered remains of illegal immigrants? Second, based on known geographic factors influencing pedestrian behavior, what routes are illegal immigrants likely taking? Third, can those projected routes be verified by using high resolution multispectral imagery?

Based on my findings, I was able to draw conclusions for each question. First, my analysis of spatiotemporal patterns of recovered migrant remains shows there is a pattern of clustering among recovered migrant remains, indicating a lack of randomness where pedestrian migrants travel. Second, based on the knowledge of what geographic factors influence pedestrian behavior, I am able to project potential routes taken by pedestrian migrants utilizing data indicative of where recovered migrant remains were located. Finally, those projected routes could be verified using high resolution multispectral imagery.

As a result of the outcomes, I conclude that potential transborder routes for pedestrian migrants can indeed be identified using a combination of GIS techniques and analysis of geographic data and factors, in conjunction with relevant high resolution multispectral imagery, to interpret data that gives an indication of spatiotemporal points of pedestrian locations at a point of time.

These findings in this project tell me a number of things about behavioral geography: with adequate understanding of geographic factors as they relate to human behavior, projections can be made about where people might go or where they came from based on a known static point. As behavioral geography relates to illegal immigration, these findings give an indication that it is possible to predict routes with relative accuracy.

The results indicate that while 'coyotes' may take active measures to physically conceal their routes and migration patterns from authorities, their natural human behaviors are still reflected on a scale that cannot be hidden by minor alterations in route selection: they ultimately are still trying to get from point A to point B as is economically feasible, and that pedestrian economy cannot be entirely concealed.

There are many opportunities for further research and exploration of this topic. Assuming one were to know the starting point of a migrant's route, as well as the interception point (due to medical assistance, detainment by authorities, etc.) could a route be predicted with even greater accuracy? What about the incorporation of aspect into the calculation—for example, a west-facing slope, with a potential for north-south travel, creating a surface that has no functional slope? How might the model be affected in other geographic regions and environments?

This research lends itself to other possible applications: one might use various geographic data to predict the potential paths a lost hiker might have taken based on their last known location, or it could be used as a planning tool in creating recreational trails for hikers and bikers, that support environmental conservation and preservation.

References

Cao, L., Stow, D., Kaiser, J., and Coulter, L. "Monitoring Cross-border Trails Using Airborne Digital Multispectral Imagery and Interactive Image Analysis Techniques." Geocarto International (2007): 107-125.

Coalición de Derechos Humanos ("The Human Rights Coalition"). (n.d.). List of Recovered Bodies for Arizona-Mexico Border, 2004-2005. Retrieved February 1, 2011, from Coalición de Derechos Humanos: http://derechoshumanosaz.net/images/pdfs/list%20of%20dead%202004-2005.pdf

—. (n.d.). List of Recovered Bodies for Arizona-Mexico Border, 2006-2007. Retrieved February 1, 2011, from Coalición de Derechos Humanos: http://derechoshumanosaz.net/images/pdfs/list%20of%20dead%202006-2007%20%28web%20page%29.pdf

—. (2011, March - May). Staff Interviews. (W. D. Grannis, Interviewer)

Cornelius, W. A. (2001, Dec). Death at the Border: Efficacy and Unintended Consequences of US Immigration Control Policy. Population and Development Review, 27(4), 661-685.

Coulter, Lloyd L., Stow, Douglas A. "Monitoring Habitat Preserves in Southern California Using High Spatial Resolution Multispectral Imagery." Environmental Monitoring Assessments (2009): 343-356.

Desert Research Institute. (n.d.). Climate of Arizona. Retrieved September 24, 2011, from Desert Research Institute Web site: http://www.wrcc.dri.edu/narratives/ARIZONA.htm

Greenwald, Michael J., Boarnet, Marion G. “Built environment as determinant of walking behavior: analyzing nonwork pedestrian travel in Portland, Oregon.” Journal of the Transportation Research Board (2001): 33-41

Handy, Susan L., Cao, Xinyu, and Mokhtarian, Patricia L. “Self-Selection in the Relationship between the Built Environment and Walking: Evidence from Northern California.” Journal of the American Planning Association (2006): 55-74

Harner, John P. “Continuity Amidst Change: Undocumented Mexican Migration to Arizona.” The Professional Geographer (1995): 399-411.

i Bort, H. M., Ausserer, K., Barker, M., Basbas, S., Cabello, A. B., Cabello, E., et al. (2010). Cost 358 - PQN Final Report. Paris: COST.

Kaiser, John V. Jr., Stow, Douglas A., and Cao, Lina. "Evaluation of Remote Sensing Techniques for Mapping Transborder Trails." Photogrammetric Engineering and Remote Sensing (2004): 1441-1447.

McIntyre, David L., Weeks, John R. "Environmental Impacts of Illegal Immigration on the Cleveland National Forest in California." The Professional Geographer (2002): 392-405.

Rossmo, D. Kim, Thurman, Quint C., Jamieson, J.D., and Egan, Kristine. "Geographic Patterns and Profiling of Illegal Crossings of the Southern U.S. Border." Security Journal (2008): 29-57.

Smith, William, Bond, John W., and Townsley, Michael. "Determining How Journeys-to-Crime Vary: Measuring Inter- and Intra-Offender Crime Trip Distributions." Weisburd, D., et al. (eds). Putting Crime in its Place. Springer Science+Business Media, LLC, 2009. 217-236.

U.S. Census Bureau. (2010). American FactFinder. Retrieved May 22, 2011, from U.S. Census Bureau FactFinder: http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=DEC_10_PL_GCTPL2.ST13&prodType=table

United States Geological Society. (n.d.). The National Map Seamless Server Viewer. Retrieved January 28, 2011, from United States Geological Society: http://seamless.usgs.gov/website/seamless/viewer.htm

Van Daele, Stijn, Vander Beken, Tom. "Out of Step? Mobility of 'Itinerant Crime Groups'." (eds), P. Van Duyne et al. Crime, Money, and Criminal Mobility in Europe. Nijmegen: Wolf Legal Publishers, 2009. 43-70.

Witztum, E.R., Stow, D.A. "Analysing Direct Impacts of Recreation Activity on Coastal Sage Scrub Habitat with Very High Resolution Multi-spectral Imagery." International Journal of Remote Sensing (2004): 3477-3496.

VITA

Warren Douglas Grannis was born in Redwood City, California, on May 10, 1979. After completing his work at Calaveras High School, San Andreas, California, in 1997, he married Amber Lauren Hammel of Austin, Texas in 2000 and began his college coursework in 2001. After transferring to Texas State University-San Marcos in 2006, he received the degree of Bachelor of Arts in August 2007. In August 2008, he entered the Graduate College of Texas State University-San Marcos to pursue a Master of Science degree in Geography, simultaneously working as a historian and geographer with the Texas Department of Transportation in Austin, Texas, in addition to pursuing freelance geospatial analysis projects.

He and his wife are the parents of three children.

Permanent email: wgrannis@gmail.com

This thesis was typed by Warren D. Grannis