NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

TO THE GREATEST LENGTHS: AL QAEDA, PROXIMITY, AND RECRUITMENT RISK

by

Ismael R. Rodriguez

December 2010

Thesis Co-Advisors:	Sean Everton
	Kristen Tsolis
Second Reader:	Heather Gregg

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** December 2010	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE To the Greatest Lengths: Al Qaeda, Proximity and Recruitment Risk	**5. FUNDING NUMBERS**
6. AUTHOR(S) Rodriguez, Ismael R.	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number _____________.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

In October 2007, a raid in the town of Sinjar, Iraq produced a large trove of foreign fighter personnel records. In the years since this discovery, researchers have used this data in an effort to illuminate the places from which recruits joined Al Qaeda and associated movements. While that research is important, it has placed little emphasis on the particular hometowns of these fighters. Thus, building upon social movement theory, environmental criminology, and geospatial analysis techniques, this research will build and test several spatial regression models of the factors potentially contributing to Al Qaeda recruitment patterns in North Africa. Moreover, this study also applies a new spatial crime analysis technique that maps risk terrain in a process using environmental factors to calculate the risk of recruitment. In all, these spatially integrated social science techniques hold great potential for improving intelligence support to ongoing contingency operations.

14. SUBJECT TERMS Sinjar Records, Geospatial Analysis, GIS, Proximity, Risk Terrain, Regression	**15. NUMBER OF PAGES** 145
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

TO THE GREATEST LENGTHS: AL QAEDA, PROXIMITY, AND RECRUITMENT RISK

Ismael R. Rodriguez
Major, United States Army
B.S., United States Military Academy, 1996

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN DEFENSE ANALYSIS

from the

NAVAL POSTGRADUATE SCHOOL
December 2010

Author: Ismael R. Rodriguez

Approved by: Sean Everton
Thesis Co-Advisor

Kristen Tsolis
Thesis Co-Advisor

Heather Gregg
Second Reader

Gordon McCormick
Chairman, Department of Defense Analysis

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

In October 2007, a raid in the town of Sinjar, Iraq produced a large trove of foreign fighter personnel records. In the years since this discovery, researchers have used this data in an effort to illuminate the places from which recruits joined Al Qaeda and associated movements. While that research is important, it has placed little emphasis on the particular hometowns of these fighters. Thus, building upon social movement theory, environmental criminology, and geospatial analysis techniques, this research will build and test several spatial regression models of the factors potentially contributing to Al Qaeda recruitment patterns in North Africa. Moreover, this study also applies a new spatial crime analysis technique that maps risk terrain in a process using environmental factors to calculate the risk of recruitment. In all, these spatially integrated social science techniques hold great potential for improving intelligence support to ongoing contingency operations.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AQAM	Al Qaeda and Associated Movements
AQI	Al Qaeda in Iraq
CIESIN	Center for International Earth Science Information Network
CTC	Combating Terrorism Center
GIS	Geographic Information Systems
GNS	GEONet Names Server
GWR	Geographically Weighted Regression
IAU	International Association of Universities
NGA	National Geospatial-Intelligence Agency
NGC	National Geospatial-Intelligence College
OLS	Ordinary Least Squares
RTM	Risk Terrain Modeling
WHED	World Higher Education Database

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

While there are several people I would like to thank, above all I reserve the highest gratitude for my wife, Rosalynda Uy, and our children, Soledad and Lake. They provided me a refuge from the academic environment without which I could never have completed this course of study. I cherish the time we had to explore Monterey, the San Francisco Bay area, and to the many points beyond. Still, this degree program and the thesis writing process in particular left me away from my family for much more time than I had ever wished. For their shared sacrifice, I dedicate this paper to them.

This paper would not have been possible without the support of the NPS Defense Analysis Department, and the CORE Lab in particular. The Lab is simply superb in its support to students in terms of both its course offerings and its access to leading edge technology. The lab is truly a collaborative environment, and a friendly place to spend long hours. It also hosted excellent ESRI provided training sessions in cartography and spatial statistics. Furthermore, I could not have completed this project without the exposure to new techniques that I received by attending the ESRI User Conference, a trip funded by the lab. I wish to thank professors Sean Everton and Nancy Roberts, GIS lecturer Kristen Tsolis, research associates Dan Cunningham, Joyce Hogan, Robert Schroeder, and Nico Smith, IT manager Malcolm Mejia, and office manager Karen Flaherty. More specifically, I thank Robert for his Arabic translation skill, and for introducing me to excellent spatial analysis literature. Moreover, Joyce provided invaluable cartography advice during the map production phase of this project. Last, but certainly not least, my second reader, professor Heather Gregg provided a refreshing perspective for this oftentimes exceedingly technical study.

Finally, I would be remiss if I failed to mention the time I served in Iraq with the 172nd Stryker Brigade Combat Team. Of the many colleagues within the

Brigade's intelligence and operations staffs, Chief Warrant Officer Four Jason Dyals inspired me to pursue formal training in geospatial analysis. That said, the many Soldiers who daily patrolled the streets of Mosul, Baghdad, Tal Afar, and even on occasion Sinjar, deserve credit for contending so well with such a complex, dynamic situation.

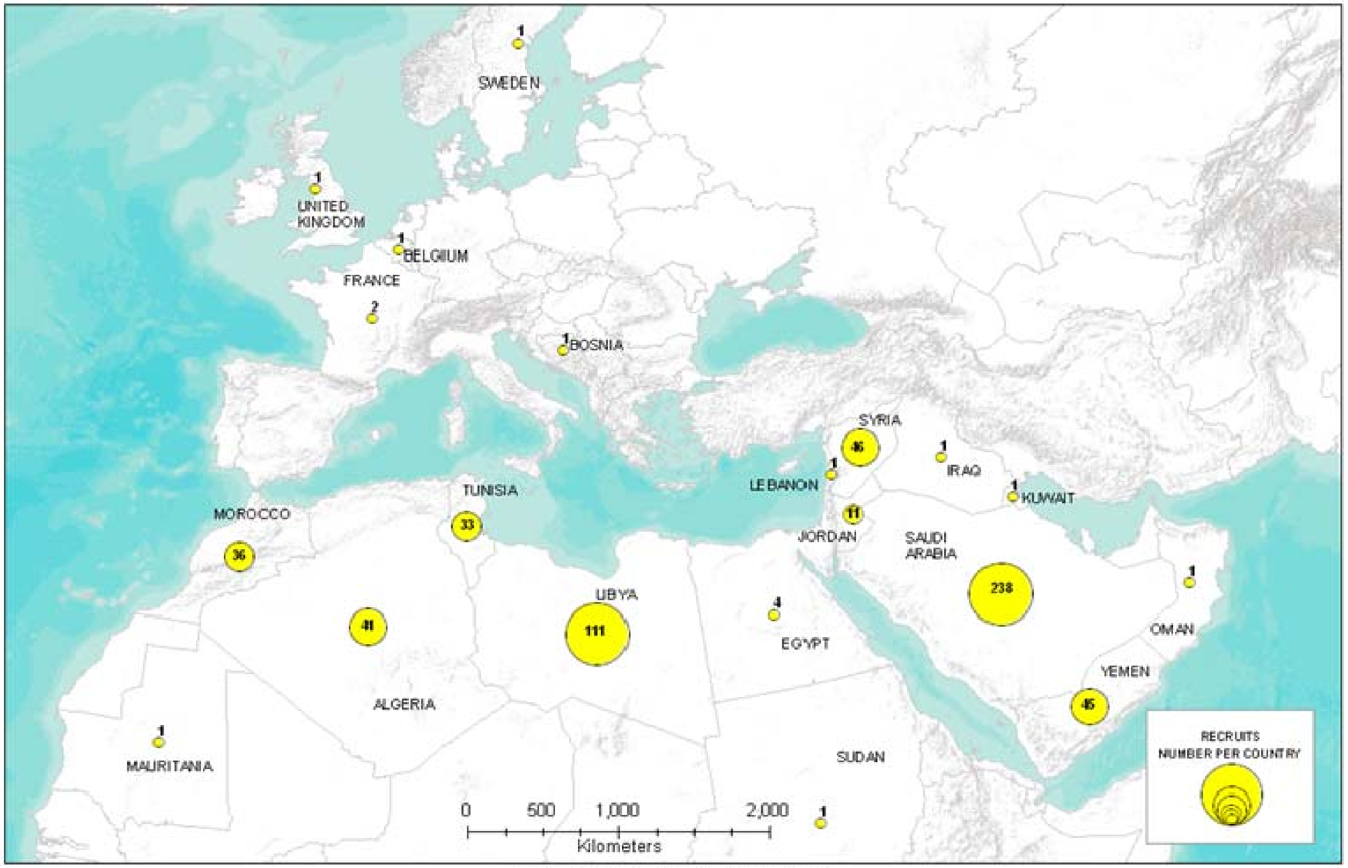
ArcGIS and ESRI are registered trademarks of Environmental Systems Research Institute Incorporated. GeoDa is a registered trademark of Luc Anselin. Google Maps and Google Earth are registered trademarks of Google Incorporated.

I. INTRODUCTION

The world is complex. Fortunately, maps help make sense of this complexity, which explains why to this day, there is a strong emphasis in the U.S. military to map the physical world. The military, and increasingly the commercial sector, has a wealth of tools and techniques to gain a superb understanding of the physical environment. Overhead imagery, precision measurement tools, and rapid developments in computer technology have advanced the realm of mapmaking. Yet, for all of these advances, uncertainty about the human element remains. While it is fairly easy to create a rich map of human terrain in the developed world, this is not the case in the cities, slums, and villages of the developing world. Indeed, in the United States, a researcher can identify down to the city block all variety of useful information pertaining to economics, demographics, politics, or sociology. Yet in Africa or parts of Asia, it can be difficult to identify little more than population density. Still, military leaders insist that human terrain is essential to the contemporary battlefield. For instance, Michael Flynn, Matthew Pottinger, and Paul Batchelor's (2010) critical assessment of intelligence activities in Afghanistan documents both the need for pertinent information about the human environment, as well as the difficulties that the intelligence community has had in compiling that information (pp. 7–10). In any case, social scientists have taken an increasingly important role in explaining the human dynamic. Still, these explanations are not very useful if swamped in the complexity of charts, graphs, tables, and volumes of text. Moreover, this situation reveals a puzzling question. Why has there been such strong emphasis on understanding human terrain, but such weak emphasis on accurately mapping that same human terrain? While maps cannot solve all the problems of fighting irregular wars, they are certainly appropriate tools for providing valuable context and insight. More importantly, maps can form a foundation for high-quality, in-depth explorations of how humans and their environment interact.

One country with many complexities is Iraq. On 11 September 2007, there was a raid in Sinjar, Iraq, a small city in the desert between Mosul and the Syrian border. The target was an alleged al Qaeda in Iraq (AQI) safe house. (Felter & Fishman, 2008b, p. 13). Something incredible emerged from that mission. While a pile of administrative papers might not seem that important, these notes offered a window into the lives of foreign jihadist fighters from across the Middle East and North Africa. There were names, phone numbers, hometowns, and occupations. Some records were thorough, some were rudimentary, but overall, they presented a unique gauge for the underground flow of young jihadists into Iraq. More importantly, the data pointed to the distant sources for the stream of fighters into this war (Felter & Fishman, 2008a).

SINJAR RECRUIT HOME COUNTRIES



COMPILED BY: ISMAEL RODRIGUEZ
DATE: 16 NOVEMBER 2010
DATA SOURCES: ESRI WORLD TERRAIN BASE, ESRI WORLD UN MEMBERSHIP, FISHMAN SINJAR DATA MASTER
COORDINATE SYSTEM: WGS-84, CUSTOM AFRICA EQUIDISTANT CONIC

Figure 1. Home Countries of Sinjar Records Recruits

Soon after the discovery, many researchers rushed to find explanations for the peculiarities of this new data. While quite valuable, the results have largely aimed at answering why certain countries seemed to have generated more jihadists than others. Fascination turned to North Africa. Indeed, at least one intrepid journalist, Kevin Peraino (2008) traveled to Libya looking for answers. His adventure led to a provocative story in *Newsweek* entitled “Destination Martydom: What drove so many Libyans to volunteer as suicide bombers for the war in Iraq? A visit to their hometown the dead-end city of Darnah.” Perhaps most revealing, he documents the towns unique history, and its long tradition of militancy both against the Italian occupation of the early 20th century, and against the Libyan regime of Muammar Kaddafi. Yet, despite this interest in a specific place and the potential role it played in the lives of recruits, there has been little formal research to consider what the impact of cities has had on the recruitment of new jihadists. Indeed, as Joseph Felter and Brian Fishman (2008b), insist there is a need for “[r]esearch that combines qualitative and quantitative methods to predict the local conditions responsible for terrorist 'hot spots'" (p. 62). Nevertheless, the areas from which these recruits came pose many challenges. For one, the recruits emerge from a huge region, as far away as Morocco on the Atlantic Ocean, to Yemen on the Indian Ocean, and to Sweden well to the north. Yet, there were several places apparently central to recruitment activity. Outside of the seemingly obvious locations in and around the holy cities of Saudi Arabia, several regions along the Mediterranean coast of North Africa are of particular interest. What makes these places important? Is it simply a well-placed recruiter feeding off a susceptible local population? Is it a hub of radical thinking? Could there be environmental factors that explain the decision to leave? Questions such of these are not easily answered by studying state-level variables. A new approach is necessary.

At the root of this new approach is a simple research question. What explains the variation in Al Qaeda recruitment patterns in North Africa? In short, the search for an answer to this question underscores the need for an inter-disciplinary approach that rests on three essential premises, which, in turn, forms the primary structure of this thesis.

First, an in-depth review of literature establishes both the theoretical underpinnings of this research while identifying appropriate techniques to analyze relevant information. Social movement theory, place-based policing theory, and spatial statistics methodology form the foundations of this research. While there has been little related academic research specifically using a geospatial perspective to understand the flow of Jihadist recruits, there is a wide array of other research applicable to this problem. Such diverse fields as epidemiology and criminology can offer a useful perspective for framing the problem. With that perspective in mind, the literature review must also consider previous research on the Sinjar records database. While this previous research has never explicitly addressed this study's particular research question, it is nonetheless essential to establishing a baseline of knowledge, and will be very informative in developing appropriate models. Once there is a sufficient theoretical and methodological understanding of the problem, the next stage can begin.

In essence, the second stage is the preparatory effort. It uses a series of proximity based distance calculations, as well as a data extraction technique, to build a matrix of attributes for use in the central portion of the study, setting the foundation for the central focus of the thesis. The result of this third stage is the creation of four types of spatial models. The first set uses ordinary least squared regression analysis techniques to identify potential factors behind recruitment patterns. The second set of models attempts to refine this analysis by adjusting for spatial conditions. The third series conducts a specialized form of regression analysis to identify localized trends within the explanatory variables. In any case, upon completion of these regression techniques, the study turns to a new form of mapping that calculates levels of recruitment risk for North Africa.[1] Essentially, these maps incorporate appropriate variables while also accounting for past recruitment activity in an attempt to explain from where recruits might likely emerge. As a final step, the study uses a small subset of temporal data to compare three different risk maps in order to identify which best explains recruitment patterns. In all, by using the results of these tests, in part, as a proof of concept, the study suggests areas for future research and highlights several implications.

While identifying areas at higher risk of nurturing future foreign fighters is arguably important in itself, this study has a broader set of policy implications. In particular, the results refine the way the intelligence community uses maps to understand complex problems. More specifically, it highlights the inherent difficulty in identifying who within the Army should take responsibility for this type of analysis, considers a possible avenue to teach these techniques within the Army, and suggests the use of risk terrain modeling to improve the Army's ability to assess future activity in a dynamic environment.

[1] These techniques, described in more detail in chapter 2 and 4, derive from the work of the Rutgers University Center for Public Security's Joel Caplan and Leslie Kennedy (2010).

II. LITERATURE REVIEW

The theories and research that apply to this thesis are many and varied. The academic realms of sociology, criminology, geography, and statistical analysis form the foundations of this review. As a start, the tenets of social movement theory are a good lens from which to observe terror recruitment efforts. As such, it is essential to understand the basic principles of this framework. While there has been an overlap between terrorism studies and social movement theory in recent years, the roots of the theory itself are more benign. Nevertheless, social movement theory informs this study's decision to incorporate proximity to national capitals and airports, while other aspects of terrorism studies inform the decision to include proximity to universities, as well as population density.

A. SOCIAL MOVEMENT THEORY

Social movement theory is a robust and evolving area of study. The traditional approach considers four essential elements. Put simply researchers began to emphasize "resource mobilization, political process, repertoires of contention, and framing" (McAdam, Tarrow, & Tilley, 2001, p. 16). Nevertheless, the basic model created by the interaction of these elements has limitations. Indeed, the model tends to work best in a liberal democratic society, while doing little to explain the complexity presented in undemocratic society (pp. 18–19). Use en dash for range of numbers

Social movements have a close association with political activity. Sidney Tarrow (1998) explains this relationship in his work *Power in Movement: Social Movements and Contentious Politics.* Of particular note, he examines the role of authoritarian states on the growth of social movements. He acknowledges that it is easier to operate within a democracy, both for the obvious reason that such movements are permissible, as well as the less noticed ability for a movement to operate at both the national and the grassroots levels as the situation permits. Moreover, within democracies, social movements can lead to a variety of different outcomes. In this sense, the organizational structure of the government enters the equation. A centralized government reacts quite differently than a more localized, multi-faceted democracy such as that seen within the United States. On the contrary, authoritarian states can present different opportunities for a social movement. The centrality of many authoritarian regimes presents a very visible focal point for social movements to aim their attacks (pp. 80–82). Nevertheless, authoritarian regimes have repression as a key tool at their disposal. For a social movement, repression changes the competition significantly. In essence, a government can either suppress a movement's growth or increase the movements organizational and mobilization costs. Over time, this increase in cost can have the greatest effect. As a case in point, cities that suppressed desegregation events fared worse than those that used the court system to delay desegregation efforts. Moreover, suppression has often backfired, with protesters gaining sympathy at the expense of the authorities (p. 83). Perhaps more importantly, Tarrow identifies a paradoxical relationship between authoritarian regimes, harsh responses, and radicalization. Yet, he is quick to point out that not all authoritarian regimes are the same, and that even repressive states can present opportunities for mobilization (pp. 84–85).

The United States civil rights movement formed one of many contexts around which social movement theory came into prominence. Research by Doug McAdam offers a solid explanation of how a social movement functions. His *Freedom Summer* (McAdam, 1988) provides a stirring account of the recruitment efforts the Student Non-violent Coordination Center's Freedom Summer movement. While it is informative, it does not offer an explicit framework for application to other movements. That said, there is a very useful chapter on the composition of Freedom Summer recruits. Of particular note, McAdam compares those who participated in the program with those who applied but chose not to participate. He contends that participants were more likely to have explicitly stated ideological beliefs, ties to organized political parties, and higher levels of previous participation in political movements (pp. 61-64). In effect, McAdam contends "the volunteers enjoyed much stronger social links to the Summer Project than did the no-shows…The practical effect of the this greater 'proximity' to the movement would have been to place the volunteer at considerable 'risk' of being drawn into the project via the application process" (p. 64).

Donatella Della Porta (2002) suggests that recruitment is an important area of social movement research, and summarizes the recruitment-based research into three broad categories. Put simply, the first category concerns the efforts to influence recruits, the second considers the process of becoming active participants in a movement, while the third considers how participants sustain and eventually end their activities (pp. 324-326).

More recently, the social movement approach has gained prominence in terrorism and Islamic studies. Muhammad Hafez (2003) adopts this point of view in his book *Why Muslims Rebel*. Hafez applies the theory to Islamist activities.

In particular, he focuses much attention on the resources necessary to promote societal change. He further divides this broad category down into three distinct groups. First, he distinguishes internal traditional resources such as people, finances, and weaponry. Next, he separates the more esoteric resources based on ideology such as common historic narratives and established systems of morality. Finally, he recognizes that external resources could be opportunistically used to propel the movement (p. 19). Summarizing, he notes, “[e]ach of these resources is a reservoir of power from which Islamists could draw to exert pressure against opponents, including an incumbent regime” (p. 20). However, for Hafez, resources are only one piece to the puzzle of Islamist social movement puzzle. Hafez contends that the political environment is an essential element of the dynamic process of social movement growth. He supports this view by considering potential variations in government response to a growing movement. While in a democracy there may be legitimize outlets for the activities of a movement, in an authoritarian regime, the state may respond by locking up activists and dissolving agitating groups. Thus, social movements must contemplate strategic choices about the best way to adapt to whatever political climate is present (pp. 20–21). This adaption is part of a broader contest that “[r]ather than being an outcome of fixed circumstances…treats social and political struggles as a dynamic of interaction, adaptation, and intended and unintended consequences that are likely to shape the strategies of movements over time” (p. 21). In fact, as Hafez summarizes, Islamist movements have grown because of the restrictive access to legitimate political outlets, and despite the repressive responses of the state. Such conditions compel Islamic activists to become radicalized, which in turn creates secretive organizations, bent on spreading ideological justifications for their radicalization and violent activities (p. 22).

A number of other authors work along the intersection of social movement theory, terrorism, and Islamic studies. Quintan Wiktorowicz (2003), author of *Islamic Activism: a Social Movement Theory Approach,* is of particular note. He adeptly weaves together social movements, their required resources, and local geographies. For instance, he considers a mosque to be a "religiospatial mobilizing structure" (p. 10). As such, he relates the role of mosques to the similar role that churches played during the civil rights movement, in which participants can organize, indoctrinate, and network with other like-minded institutions. However, for Wiktorowicz, this is only one available option. He also lists the role of charitable organizations, and of both student and professional organizations. Within in Islamic societies, religiously oriented members have taken on prominent roles within such organizations, filling a vacuum left by the diminishing influence of socialism (pp. 10–11).

Wiktorowicz also examines the organizational structures that facilitate the growth of resources. While acknowledging the thorough research addressing the impact of formal institutions of social movement growth, he highlights the importance of informal structures. This is especially the case in difficult political environments where formal structures can draw undue attention to a cause. As he draws from a variety of studies to note:

> [i]n such contexts, formal resources are inviting targets for regime repression and may actually make it easier for security services to undermine the institutional capacity of the movement. As a result, movements may instead use informal institutions and networks for activism, since they are embedded in everyday relationships and thus more impervious to state control. (p. 12)

As such, Wiktorowicz asserts that Islamic activism is a useful subject for the examination of informal structures as they pertain to social movement theory, especially given the repressive environment in which Islamist movements exist (p. 13).

Bruce Hoffman (2006) also provides an insightful understanding of terrorism. However, while there are a few parallels, he conceptualizes terrorism in a way that does not fit neatly into social movement theory. Instead, he addresses the tactical use of political violence, by an organization or group of ideologically motivated individuals in order to reap a specific psychological effect (p. 40). More specifically, he makes an important observation with relevancy to the study of social movements. For him:

> [t]he terrorist is fundamentally an altruist: he believes that he is serving a 'good' cause designed to achieve a greater good for a wider constituency…that the terrorist or his organization purport to represent…The terrorist is fundamentally a *violent intellectual*, prepared to use and, indeed, committed to using force in the attainment of his goals. (p. 37)

Distressingly, Hoffman sees terrorism as entering a new dimension. Instead of a clearly defined organizational dimension characteristic of past terror groups, there is now a situation in which individuals may have ideological connections to a broader movement, but act autonomous of those movements. This concept is one that even Al Qaeda considers a potent weapon in its fight against Israel and the United States (pp. 38–39).

Another clear parallel exists between social movement theory and Hoffman's understanding of terrorist resource and operational requirements. Hoffman also contemplates the transfer of tactical and operational methods from one group to another. He recognizes that the influential role of the Palestinian Liberation Organization as a trainer for some forty different terrorist groups from around the world. More so, he argues that the PLO emphasized the cultivation of political and financial resources (pp. 78–79).

Marc Sageman (2008) lends another prominent voice to the study of terrorism. Indeed, the argument in his recent work, *Leaderless Jihad,* falls within the perspective of a social movement approach. However, Sageman approaches terrorism studies in his own unique way. He offers a clear explanation of the prevalent levels of terrorism analysis. He identifies two prominent trends. Analysts often focus attention on either the micro level analysis of individual terrorists, or the macro level analysis of the causes of terrorism (pp. 16-23). Still, he eschews exclusively approaching terrorism from either the individual or the societal perspective, arguing that the two approaches have significant flaws on their own and cannot be merged together to form a coherent understanding of terrorism (p. 23). Instead of these approaches, Sageman contends that there should be a third approach focusing on the dynamic processes of terrorism as they relate to the larger environment in which they take place (p. 24).

Sageman also takes a nuanced view of Al Qaeda. For him, it is not just a social movement or an organization but instead is a mix of both (p. 29). While the organization known as Al Qaeda has diminished in capability, it has been surpassed by an informal social movement, which has grown well beyond the dimensions of a typical organization. Constructed of a fabric of small networks, Al Qaeda is in a sense of social movement of individual organizations. For Sageman, the social movement dimension of Al Qaeda is more important than the remnants of the original remnants of the Al Qaeda organization (p. 31).

How then do social movement theories relate to other approaches to terrorism research? D.K Gupta (2006) in "Tyranny of Data: Going Beyond Theories" offers a succinct, well-organized review of how social movement approaches fit into the broader research on terrorism. In essence, he divides research into studies that apply theory and studies that exclude theory. From the theoretical approach, he distinguishes primarily between psychological and social theories on one hand, and rational actor approaches on the other hand. However, it is outside the theoretical realm that most terrorism studies reside. This is true for both historical approaches to terrorism, as well as the terrorism studies approaches of Hoffman and Sageman (p. 39). Gupta's framework presents a useful tool for identifying the theoretical roots of previous research as they relate to research applied to the Sinjar database.

B. SPATIAL ANALYSIS THEORY

The theories behind geospatial analysis fall within the broad discipline of geography. That said, geography itself has a distinctively interdisciplinary nature. Applied geography is a case in point. Michael Pacione (1999) explains in his work *Applied Geography: Principles and Practice,* that applied geography is essentially the use of geography for a specific purpose, and generally a purpose that addresses real world concerns, not simply the issues of academia. In other words, "applied geography may be defined as the application of geographic knowledge and skills to the resolution of social, economic and environmental problems" (pp. 3–4). Pacione argues that applied geography gains strength from its ability to pull from both geographic theory, as well as the theories of a diverse range of academic disciplines (p. 4).

However, Waldo Tobler deserves credit for enunciating the concept upon which geospatial analysis and geospatial information systems have grown (Miller, 2004, p. 284). As Tobler (2004) states, the first law of geography is that "everything is related to everything else, but near things are more related than distant things" (p. 304). While even Tobler acknowledges that there is debate as to whether such a statement is truly a law, the statement itself deserves attention. Harvey Miller (2004), writing in the *Annals of the Association of American Geographers*, provides a practical explanation while defending the usage of the law (p. 288). Of particular note, He unpacks the law's concept of relation, noting "there is a positive or negative correlation between [geographic] entities…Although correlation is not causality, it provides evidence of causality that can (and should) be assessed in light of theory and/or other evidence" (p. 284). More importantly, Miller describes how the law plays an essential role in a wide variety of spatial statistics and spatial analysis techniques, while he also suggests that those processes that do not tend to follow the law may simply follow an atypical, non-Euclidean measure of nearness (pp. 284–285). Thus, Miller contends "[n]earness is a central organizing principle of geo-space, but it is not required to be a function of Euclidean, metric, or even an empty space"(p. 286).

Proximity analysis is an essential capability of a GIS. As such, proximity is intrinsically associated with distance and can include analysis of areas, networked routes, or pure numerical distances (Honeycutt, Murray, & Prince, 2010, p. 9). Distance though can be problematic. Depending on the scale used, a maps projection can have dramatic effects. Since the earth is not flat, there will always be some level of distortion in measurement. For instance, a Mercator map creates landmasses at the higher latitudes that are far larger than reality. Thus, distances measures using such maps will also display greater distortion (p. 16). Fortunately, the use of equidistant map projections can mitigate the effects of distance distortion (p. 17).

C. SPATIAL STATISTICS THEORY AND METHODS

Geospatial statistics, dependant as they are on the first law of geography, are powerful. These techniques offer a useful tool to conduct more in-depth analysis. In particular, this type of analysis can not only more rigorously identify clusters, but also consider complex sets of independent variables to explain why those clusters exist (Mitchell, 2005, pp. 2–12). Used in a wide variety of academic and policy disciplines, these inter-related processes of cluster analysis, spatially-based regression analysis techniques, and spatial proximity analysis may offer unique insight into the specific question of where human conditions are conducive for AQAM growth.

Cluster analysis is a technique that identifies groups of features that occur in close proximity to one another. Geospatial information systems allow an analyst to calculate precisely whether a cluster has occurred randomly. With improved confidence that the cluster is not random, the analyst can further investigate other spatial features to identify causal factors (pp. 148–149). Thus, the initial analytical step will be to create a foreign fighter overlay that places a point for every fighter on his hometown. With this layer created, it is then possible to run cluster analysis using GIS software.

With clusters identified, the next analytical step is to conduct an exploratory analysis of those areas near statistically significant clusters. The heart of this analysis is the use of GIS software to conduct multivariate regression analysis of the relationship between the dependent and independent variables (pp. 202–203, 215). By identifying these co-varying relationships, the theoretical relationship can then be refined, and ultimately, the theory's explanatory and predictive power will improve (pp. 192–195). Thus, the exploratory analysis will begin with the compilation of data layers for each indicator under consideration. With these overlays in place, the regression analysis can then begin.

Ordinary Least Squared (OLS) Regression is a powerful process adapted for use in geospatial analysis. Andy Mitchell (2005), in his *work The ESRI Guide to GIS Analysis, Volume 2,* describes how this process works. (See Appendix A for a detailed explanation of OLS regression.) However, he also presents a refined regression technique, known as geographically weighted regression (GWR), as a tool for contending with local level variation. In essence, this technique conducts an OLS regression for each occurrence of the spatially attributed dependent variable. At each location, both coefficients and residuals can then be mapped (p. 219). Specifically, "[t]he coefficient for a location depends on the influence of the surrounding data points. The influence is based on how far the particular data point is from the location you're calculating the coefficient for--the closer the point, the greater the influence" (p. 220). When would it be useful to use geographically weighted regression? Suppose that a hypothetical explanatory variable tends to vary across a study area. While a global solution may do a good job of explaining an outcome overall, by considering local variations, it may be possible to improve the fit of a model. More importantly, the procedure allows the analyst to determine regions where specific explanatory factors carry the most weight (pp. 220–221).

Spatial autocorrelation is also a concern for spatial regression analysis. By definition, spatial autocorrelation occurs when "[g]eographic features that are near each other are likely to be more similar than distant features." (Mitchell, 2005, p. 200). As a value, spatial autocorrelation depends on the scale of analysis. In other words, it may exist in extremely small levels of analysis but may dissipate when considering broader levels of analysis. Moreover, its existence suggests that geography is an important factor to consider. As a result, there are a number of techniques to isolate the phenomenon, or to incorporate the phenomenon into more accurate models (p. 201).

D. CRIMINOLOGY THEORY AND SPATIAL CRIME ANALYSIS

Criminology offers a theoretical basis that can easily incorporate a spatial approach to problem solving. Rachel Boba (2005), in her work *Crime Analysis and Crime Mapping* provides an overview of the theory behind spatial approaches to criminology. Considered environmental criminology, it is distinct from traditional criminology because it does not search for a root cause to crime, and instead attempts "to understand the various aspects of a criminal event in order to identify patterns of behavior and environmental factors that create opportunities for crime" (pp. 59–60). Central to this approach is the concept of the crime triangle that considers the offender, the target or victim, and the place where the crime takes place. Moreover, there is a dynamic relationship between each of these aspects and those who can control events, and the theory rests on the argument that a lack of such controls result in criminal behavior. As such, this theory offers the analyst a framework with which to analyze criminal activities in order to identify patterns of criminal activity and to suggest specific prevention techniques (p. 60–61). Furthermore, environmental criminology has a close association with several other theories. Take, for instance, its relationship with rational choice theory. The environmental approach assumes that the criminal makes decisions based on a calculation of risk and opportunity. Thus, by identifying the factors at play in a crime, it is possible to understand the dynamics involved and incorporate techniques that specifically target known opportunities (p. 62).

At the social level, the theory of crime patterns also has a close association with environmental criminology. This theory suggests that in a given area, the likelihood of crime increases when there is an overlap in the zones of daily activity between victims and criminals. In other words, crimes are most likely where the daily lives of victims and offenders overlap. Finally, the theory of routine activities also influences environmental criminology. This theory suggests that crime patterns are a result of changes a society’s routines. For instance, in the decades after the Second World War, homeowners increasingly began working outside the home, leaving their homes without someone present during the day. The result was an increased opportunity for thieves to steal from unguarded residences. Fortunately, there is also an upside to the theory as habitual changes can also increase the risk to an offender (p. 63–64).

Hot spot mapping techniques have gained considerable prominence in recent years. Put simply, a hot spot map shows where crimes have most regularly occurred over a set period. The rigor involved in this process can vary.

Using an analog map, an analyst could simply eyeball clusters. However, with digital mapping, an analyst could apply increasing levels of complexity to determine clusters of activity (Boba, 2005, pp. 218–219). On the complex end of this spectrum, density mapping uses mathematical formulas to determine degrees of criminal density. Yet, this process is fraught with challenges. Not only are density maps deceivingly simple, but they can misrepresent criminal activities, suggesting that crimes have taken place in areas where they actually have not (pp.222–225).

Risk Terrain Modeling (RTM) is a relatively new application of geospatial analysis that has great potential. Developed by Joel Caplan and Leslie Kennedy (2010) and described in their work *Risk Terrain Modeling Manual*, the technique stems from a theoretical foundation in environmental criminology. Put simply, RTM uses a geospatial information system to layer different aspects of risk in order to calculate an overall level of risk and ultimately to create an overall picture of risk within an area. These calculations "combines actuarial risk prediction with environmental criminology to assign risk values to places according to their particular attributes" (p. 24). From a theoretical perspective, there is an emphasis on the variable role of opportunity as it relates to crime. As such, Caplan and Kennedy argue that risk assessments are well suited to incorporate several different factors while also aiding police strategic and tactical activities. Moreover, they suggest that criminals, victims and police officers understand that there is a spatial component inherent to an individual's calculation of risk (p. 14). The authors also distinguish between current geospatial analysis techniques and the potential offered by mapping risk terrain. Hot spot mapping receives a close examination. While largely complimentary, Kennedy and Caplan nevertheless expose the limitations of the approach. In particular, academic studies have suggested that hot spot mapping is an effective means of predicting criminal activity, while other studies have pointed a variety of ways to improve the technique. More importantly, the limitations of hot spot mapping are very real. The emphasis on hot spots is essentially a reactive process that bases prediction purely on past activity and despite the intervention of law enforcement. Indeed, there is a tendency for criminal activity to evolve as police respond to hot spots. (pp. 27-28). On the contrary, Kennedy and Caplan argue in favor of the approach's ability to forecast criminal activity. As they note:

> Forecasting is more advantageous to practitioners because it does not rely on a crime to actually occur, or for the event to occur at an exact location. Predictions are deterministic in that an event is assumed to happen unless proper actions are taken; any occurrence of the predicted event connotes a failure of the public safety practitioners, while any absence of the predicted event connotes either an adequate practitioner response or a failed predictive event. (p. 29)

Even though the authors are clearly in favor of their approach, they still see utility in hot spots maps. More importantly, they propose incorporating hotspot analysis into the RTM process. This allows police departments to selectively target criminal activities while also grounding analytical activities in solid environmental criminology theory. In simpler terms, law enforcement gains a view of past criminal activity, as well as a sense of the environmental factors that might affect that same activity. For them, the use of both techniques could aid police department strategic management. Thus, police departments can base their resource decisions on the levels of risk across their area of operation instead of simply putting resources on hot spots (p. 36–39).

Caplan and Kennedy lay out a simple step-by-step method for completing a risk terrain map. The initial four steps lay the groundwork. An analyst must decide what specific criminal activity to study, where specifically to study the activity, and over what timeframe to observe the activity (p. 42). With these three tasks accomplished, the analyst can move on to more complicated requirements. Gathering appropriate map data begins the next leg of the process. The analyst then reviews available literature to identify the essential factors that impact risk, focusing on those elements with a spatial character. In other words, the analyst considers where criminals might sleep, eat, or congregate. Upon identifying the factors, the analyst can then decide which factors to include in the map (pp. 43–44). This leads to the very intensive step of turning these factors into usable map layers (pp. 45–56). Yet once these layers exist, it is a somewhat simpler process to create the map of overall risk (pp. 56–57). At last, this map can then form the basis for a visual demonstration of criminal risk in the given area (pp. 58–64).

E. SINJAR DATABASE AND RELATED RESEARCH

The Combating Terrorism Center's first report, *Al Qaida's Foreign Fighters in Iraq,* is a preliminary assessment of the Sinjar Records dataset. The CTC received over 700 records from the United States Special Operations Command. This initial set was then reduced to 606 specific files (Felter and Fishman, 2008a, p. 6). The authors clearly warn of the risks of accepting the results of studies based purely upon the Sinjar records dataset. Nevertheless, the records were placed into the open academic environment in the hope that the database would be used to produce new scholarship to either complement or challenge the conclusions of the West Point Combating Terrorism Center (pp. 3–4). The report itself is essentially review of who these recruits are in terms of age, occupation and social connections, and a snapshot of where they come from in terms of countries and cities. What is noticeably lacking from the initial report are maps. There is not a single descriptive map in the report. Instead, locations are depicted using pie charts, tables, and bar graphs. That said, the report uncovers several previously unknown trends. In particular, it notes that within the sample, there is a much higher than expected level of recruits emerging out of North Africa. Libya is the primary source of this activity, but Tunisia, Algeria, and Morocco also produce significant numbers, while Egypt is barely represented in the sample (pp. 8–9).

The follow up to the first report came with the release of *Bombers, Bank Accounts, and Bleedout: Al Qaida's Road in and out of Iraq.* This report is indeed a more rigorous examination of the phenomena that produced the Sinjar dataset. Of the many findings of this second report, the most portentous is the suggestion that there could be a bleed out effect where foreign fighters return to other conflict areas. In other words, veterans of the Iraq Jihad might fight again in another time and place (Felter & Fishman, 2008b, p. 7). Moreover, while the process is similar to the international Islamic response to the Soviet invasion of Afghanistan, those returning from Iraq appear to have better skill-sets than those who fought in the 1980s. Still, there were profound consequences following the first Afghan conflict that could again reappear following the Iraqi conflict (p. 9). Furthermore, the report also contends that foreign recruits join because of local social relationships and not from the efforts of internet recruiting (p. 8). Thus, of the many recommendations offered in the report, perhaps the most important for the military may be the need to cooperate on counter-terrorism efforts with the countries of the Arab world, and North Africa in particular (pp. 10-11).

In the first chapter Vahid Brown (2008) dives into the history of foreign fighter activity in Afghanistan. More specifically, the nucleus of foreign muhajideen leadership in the Soviet-Afghan conflict came from the Islamist thinkers of AL Azhar University in Cairo, Egypt. It was there that these future jihadists adopted a Qutbist ideology and built ties with the Muslim Brotherhood (pp. 18–19).[2] Moreover, following the Soviet invasion, money and jihadist recruits flowed from the previously built local networks of the Muslim Brotherhood (p. 20). The recruitment process included a variety of formal and informal means. Some countries exported their locally troubling islamists off to fight in Afghanistan, while others such as Syria, Kuwait, and Jordan applied repressive pressure on Islamist groups pushing fighters into the Afghan conflict (pp. 22–23). Nevertheless, Brown argues that the role of foreign fighters in Afghanistan was not decisive in the defeat of the Soviets. However, the event presented Arab

[2] As Marc Sageman (2008) explains, in Egypt a violent philosophy, rooted in Salafi Islam, arose in response to the harsh measures taken against the Muslim Brotherhood. It turned away from peaceful solutions and called for the violent downfall of the government. A leading proponent of this philosophy was Sayyid Qutb (p. 37).

fighters with an opportunity to build strong informal bonds while developing a unique strategic and fundamentalist perspective to further the fight against anti-Islamic forces (pp. 30–31).

In the second chapter Joseph Felter and Brian Fishman (2008b), the authors of the initial report, provide a more careful analysis of the Sinjar dataset. First, this new look further refined the dataset down to 590 entries (p. 32). Of particular note, it also includes a geographic perspective that had been largely inadequate in their first attempt. That said, the mapping effort focuses exclusively on the regional level, providing a snapshot of the Middle East, North Africa, and a small subset of Europe. While one map shows a by country breakdown of foreign fighters, the other normalizes the data for population, depicting the number of fighters per million citizens for each country (pp. 34–35). Beyond these broad depictions, this new examination is more detailed in its city level analysis. Libya, Morocco, Tunisia and Algeria each get a city-by-city breakdown of foreign fighters per million residents. Of these, the bulk of attention goes to Libya, with a small fraction of analysis devoted to the other countries (pp. 38–42). While not considered part of the geographical analysis, the report also considers the routes that recruits take. It identifies distinct regional preferences. For instance, many of the Libyans listed that they traveled through Egypt, while Moroccans often traveled through Turkey on their trips (p. 46).

The remainder of the chapter examines the profile of the Jihadist recruits. Particularly insightful is a review of the different means of in which recruits linked to the travel network that brought them to Iraq. The authors suggest that the links underscore the very local nature of recruitment through close family and friends (p. 45). In considering why the internet might not be as prominent as a recruitment tool, the authors suggest that it may be a result of security measures in place to improve the level of trust between facilitator and recruit (p. 46). Finally, the report suggests the clustering of recruits into groups for the trip into Iraq. While this claim is made with some degree of uncertainty, the timing of reports shows that there were large numbers of entrants in both November 2006 and July 2007, while there was little activity in the spring of 2007. Still, the data specifically shows on a single day, 9 May 2007, there were five recruits who arrived from Darnah, Libya (pp. 51–52).

Felter and Fishman conclude with a number of suggestions. Within these, the advice to focus efforts on terrorist clusters stands out. In particular, their suggestion to conduct "[r]esearch that combines qualitative and quantitative methods to predict the local conditions responsible for terrorist 'hot spots,'"(p. 62), is an acknowledgement that more can and should be done to understand the phenomena driving Jihadist recruitment.

Perhaps the best study to date also has a close association with the Combating Terrorism Center (CTC). Clinton Watts, a former member of the CTC released his examination of the material in "Beyond Iraq and Afghanistan: What Foreign Fighter Data Reveals About the Future of Terrorism." That study looked at both the countries and the cities from which these recruits originated. Indeed, the analysis of state-level factors provides strong evidence of causal relationships (pp. D-1–11). However, the city level analysis is not nearly as comprehensive. In particular, that analysis focuses on population size and the number of recruits from the various cities indentified in the Sinjar Records (pp. C-1–5). In essence, that study provides only a look at potential clusters of recruits, without thoroughly testing what makes those specific locations unique. Above all, Watts recommends to "[f]ocus counterterrorism efforts on cities and nodes, not nations and regions" (p. 1–6).

The challenge with the Sinjar data set is to find a creative approach to the data. Temporal and basic social network analysis has been the hallmark of previous analysis. While there has been a spatial component, it has been limited to a very broad scale. In essence, there has not been an attempt to use a theoretical lens to consider the emergence of recruits. Moreover, there have been no systematic examinations of the spatial recruitment patterns at the city level. This thesis attempts to fill a gap in the previous research.

THIS PAGE INTENTIONALLY LEFT BLANK

III. KEY VARIABLES AND DATA PREPARATION

Spatial data is central to this thesis. However, this data requires extensive preparation to use it for visualization and analysis. Taking basic information from a multitude of sources and transforming it into a useful database and ultimately producing a map is a time-consuming, deliberate process. Before unleashing the power of geospatial analysis, it is essential to have confidence in the data being mapped. Questionable data is certainly easy to come by in the information age. While it can seem that there is too much data available, often times, there is a deep geographic divide in the quality, availability, and detail of pertinent information. Take, for instance, the United States. A spatial analyst has access to a vast catalog of geospatial knowledge. If free sources do not meet requirements, then there is also a wealth of commercial, academic, and other sources geared to understanding political, social, economic, and demographic factors of virtually any city block in the country. As soon as an analyst looks beyond the borders of the developed world, the ability to gain a similar degree of understanding diminishes. While there is a significant body of knowledge that compares the many countries of the developing world, there is no equivalent that compares their associated cities. Thus, to compare the 28 different entities identified in this study requires a fair amount of creativity in order to work with the information that is available.

A. THE SINJAR DATASET

The primary dataset for this study is the Sinjar Dataset. Discovered in the fall of 2007, it is a panoramic snapshot of the flow of al Qaeda recruits into Iraq. The West Point Combating Terrorism Center (CTC) led the effort to make the dataset accessible to the academic community. However, this is only part of the story. The United States Special Operations Command released the material to the center (Felter & Fishman, 2008a, p. 3). Yet even before this, the record set exists because someone in an al Qaeda affiliated facility in Iraq thought it important enough to track how and from where recruits entered the country. In all, the Sinjar Data Master lists over 590 records at the individual level. The data are not perfect. Some recruits were very detailed; some were not. (pp. 6-7). This presents several problems for an analyst. From a spatial perspective, 581 recruits list the country from which they came. A smaller portion, 429, also listed a hometown (Fishman, n.d). These broad patterns are quite easy to map.

From a wide angle, the regions of the Arabian Peninsula, North Africa, the Levant, and Europe all generated recruits. Upon closer review, Saudi Arabia and Libya stand out with the highest number of recruits. Following close behind were the countries of Syria, Jordan, Algeria, and Morocco.

Analyzing country level data is a relatively simple process. Not only is country level spatial data readily available, but there is also an immense number of national level statistics from which to identify correlations. Indeed, there is already extensive analysis of recruiting patterns at the national level. Alan Krueger (2007), in his short work *What Makes a Terrorist,* actually takes into consideration one spatial component in analyzing foreign fighters patterns within Iraq . Of note, he suggests "[d]istance to Baghdad has a significant effect…in that countries closer to Iraq are greatly overrepresented among the captured foreign nationals" (p. 85). Moreover, Clinton Watts (2008), building upon the research of Krueger, identified several significant variables. Of these, three stand out. A nation's human development index score, in addition to its Freedom House Political Rights and Civil liberties scores, do much to explain the variation in recruiting patterns (pp. D-2, D-6). While not devoting much attention to spatial dynamics, these previous efforts also identify a relationship in the distance from the home country to Iraq. In other words, more recruits emerged from countries closer to Iraq.

The process of analyzing cities is much harder. Thus far, analysis has focused solely on population levels (Felter & Fishman, 2008b, pp. 36-42) (Watts, 2008, Appendix C). Watts, in particular, conducted statistical analysis in an attempt to identify places where the ratio of recruits to population levels were significantly higher than expected (p. C-5). Why is it so difficult to proceed beyond this level of analysis? Foremost is the issue of identifying hometown locations. Without a recognizable city, it is impossible to assign a location, let alone assign attributes for that location. Although the vast majority of records are straightforward, there are several places with transliteration issues. Moreover, there are also some places that do not exist in spatial databases. Mitigating this problem requires a deliberate process.

While the CTC studies do not specify the source of population data, past analysis by Watts (2008) depended upon the online citypopulation.de database (A-5). However, from a geospatial perspective, the formats used were not very useful. In particular, preparation involved downloading non-tabular files structured for Google Earth. While these files included population and location information, creating a spatial layer acceptable for analysis would necessitate the use of more comprehensive tables. For this thesis, the initial data preparation relied on a commercially compiled database. The data, purchased from GeoDataSource (2010), offered a massive table of cities with alternative spellings in addition to associated locations and populations. Despite this, there were still many incomprehensible hometown references. To whittle down this subset, it was necessary to cross-reference listed city names with several other data sources, and sometimes with online searches. The best of these was the National Geospatial Intelligence Agency (NGA) GEONet Names Server (GNS) Dataset (2010a-d). While this resource did not include detailed population information, it did offer an exhaustive list of potential spellings, in addition to incredibly precise latitude and longitude coordinates. Ultimately, instead of using the commercial data for analysis, the NGA-based location tables form the default for this study.

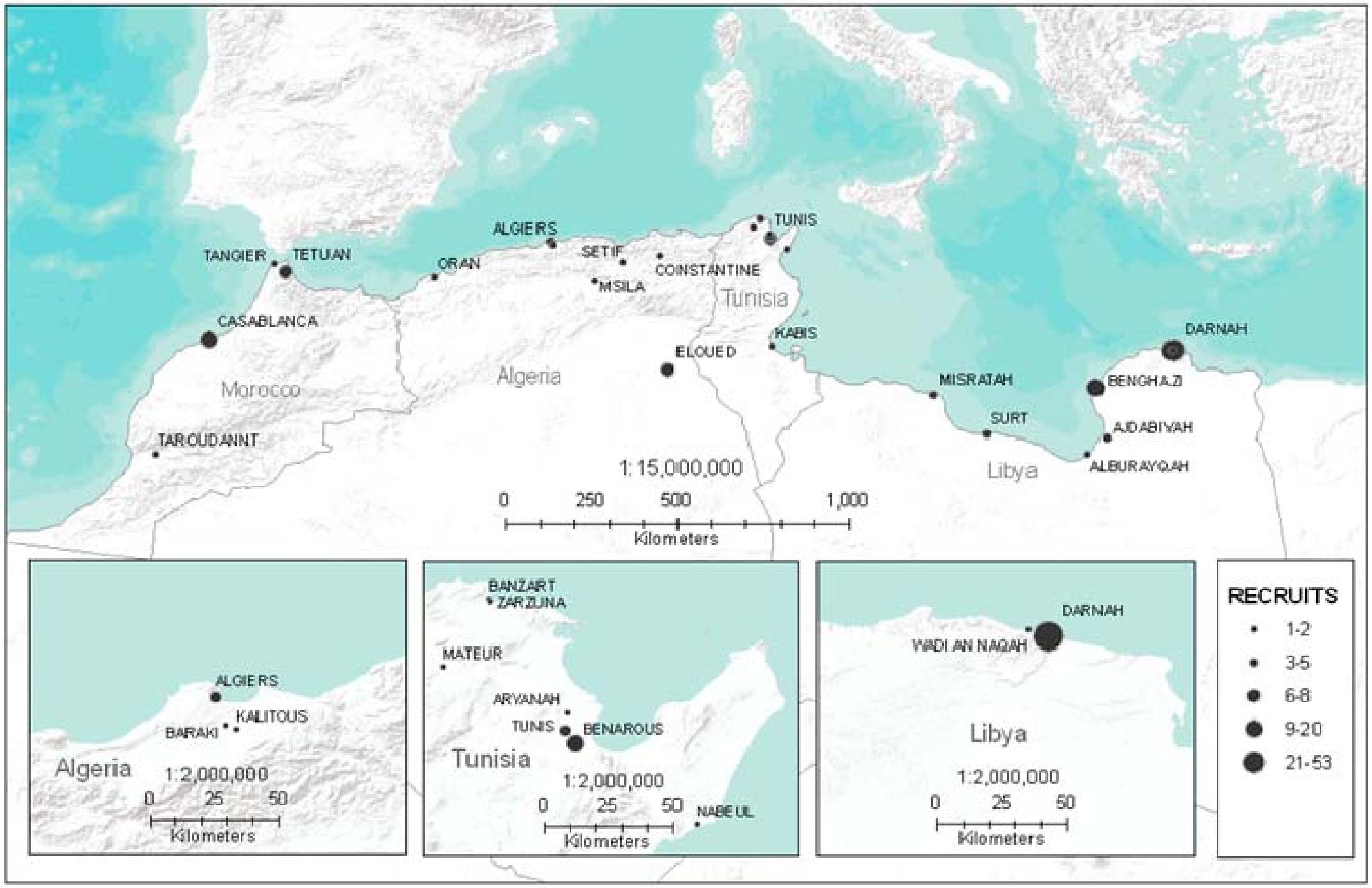
The detailed Study of North African cities resulted in a table of 27 separate entities. Of these, the three locations of Jabal Rarsah, Morocco, Kalitous, Algeria, and Wadi al Naqah, Libya presented the most consternation. Without access to the original Arabic versions, it was finally possible to assign locations by the deliberate process of cross-referencing search engine results. Of the three, Kalitous was the easiest to identify, since there was a French Media reference to the city (Le Point, 2007). On the other hand, the most uncertain location is Jabal Rarsah. A re-examination of the original Arabic version of the Sinjar record, NMEC-2007-658026 (CTC, n.d.a, p. 821) (CTC, n.d.b, p. 598), produces a translation of Jebel Darsa.[3] According to the NGA (2010c) GNS Dataset Jebel Darsa is, when plotted using Google Maps (Google, 2010), a mountain that stands above the city of Tetuan. Thus, the Jabal Rarsah record gains the same spatial coordinates as those for Tetuan. Finally, the name of Wadi al Naqah presents a similar challenge. It is a common feature name within Libya, but NGA (2010b) does not classify any of those as populated places. Therefore, it took a review of online aerial imagery to identify one of those locations that actually had human habitation. Upon review, Wadi al Naqah gains the location assigned to a valley west of Darnah in which there is a small groupings of buildings (Google, 2010).

Once there was a viable table of city spatial coordinates, it was then possible to marry it to a table of individual Sinjar Records for North Africa. The result of this work was an incident map of recruit hometowns.

While geographical space is the primary area of interest for this thesis, it is nonetheless useful to consider the temporal nature of the dataset. Specifically, 204 records included an arrival date. The earliest of these began in September 2006, and ended ten months later in July 2007. Unfortunately, the data were noticeably smaller for specific North African locations. In all, only 58 of these records had country, city, and arrival data. Of these, 38 arrived in the first five months and 30 arrived in the final five months (Fishman, n.d).

[3] CORE Lab research associate Robert Scroeder translated this record.

NORTH AFRICAN RECRUIT HOMETOWNS



COMPILED BY: ISMAEL RODRIGUEZ
DATE: 16 NOVEMBER 2010
DATA SOURCES: ESRI WORLD TERRAIN BASE, ESRI WORLD UN MEMBERSHIP, FISHMAN SINJAR DATA MASTER, NGA GNS COUNTRY FILES.
COORDINATE SYSTEM: WGS-84, CUSTOM AFRICA EQUIDISTANT CONIC

Figure 2. North African Recruit Hometowns

B. PATTERN ANALYSIS OF THE SINJAR DATASET IN NORTH AFRICA

Several notable features emerge from a thematic map of recruit hometowns. By using the ArcGIS *Collect Events* tool, it is possible to summarize the number of recruits for each of the 27 locations in the Four North African countries of Morocco, Algeria, Tunisia, and Libya. Darnah, Libya, stands out as the home of the single largest contingent, with 53 recruits. Also within the Libya, the city of Benghazi has a large share with 20 recruits. Within the other countries, there appear to be groupings near Casablanca, Morocco, Algiers, Algeria, El Oued, Algeria, Tunis, Tunisia, and Banzert, Tunisia.

While any clustering begs further examination, a quick study of the history of Darnah provides a solid context as to why so many people felt moved to join al Qaeda. In particular, the area has long been a hub for fervent jihadi activity, both against Italy in the colonial era, as well as against the Qaddafi regime in the last few decades (Peraino, 2008). Thus, historical context alone may go a long way to explaining the odd results for such a small city. Still, could other structural forces be at play within the broader region? The answer to this question demands additional data.

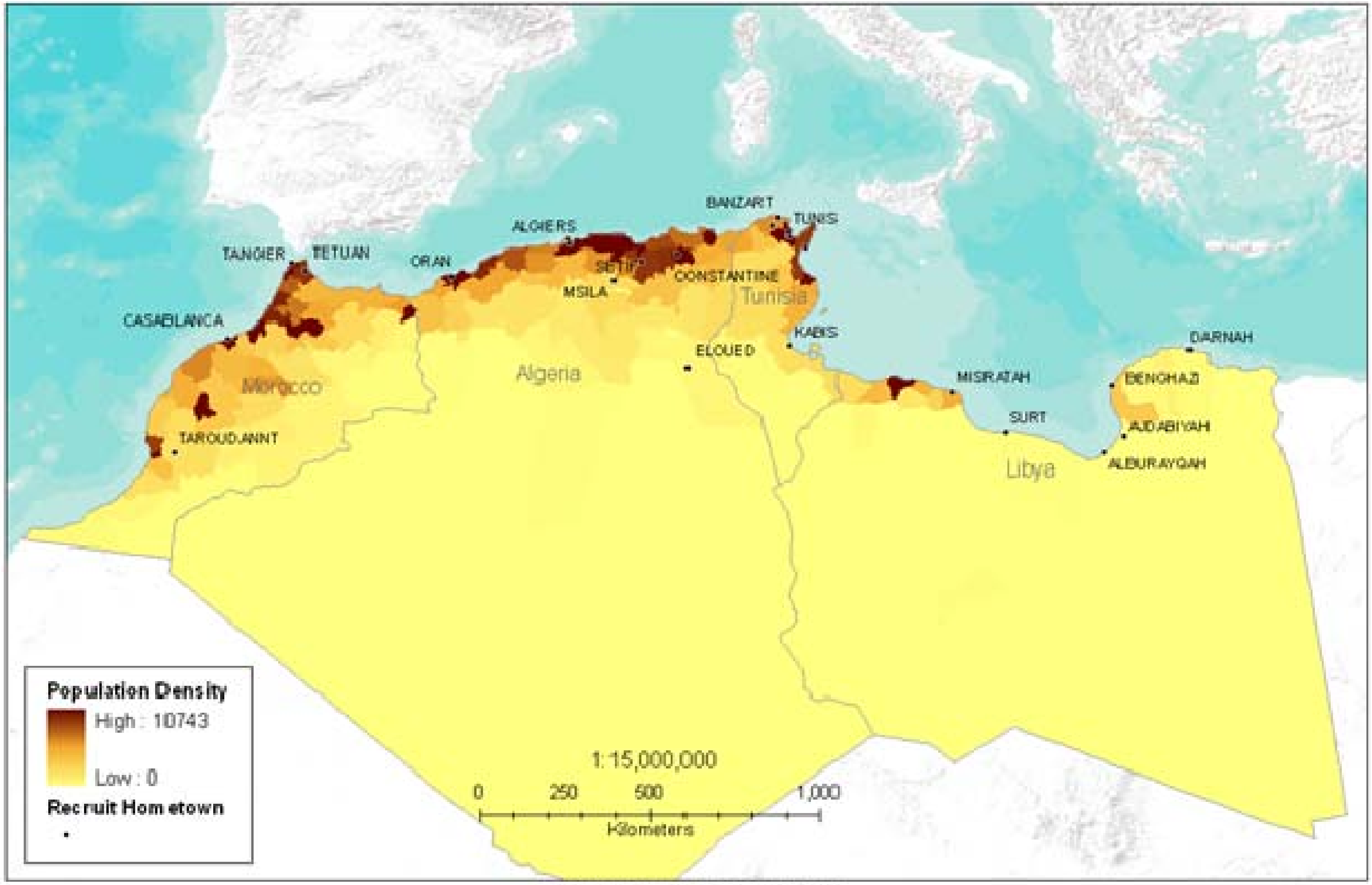


Figure 3. North African Population Density

C. POPULATION DENSITY DATA

The population of hometowns is one variable already examined in previous research on the Sinjar dataset. From a theoretical standpoint, there is not a foundation in social movement theory with which to explain a link between recruitment and population or population density. In terms of previous geospatial research, Angel Rabasa et al. (2007), writing in *Ungoverned Territories*, claim that the complexity of an urban area can provide a terror organization with concealment. Specifically, they note that "[b]eing invisible to the local authorities…and to international counterterrorist forces is therefore a survival requirement for terrorists…invisibility may be a consequence of the anonymity provided by modern, cosmopolitan mass society" (pp. 20–21).

Population levels vary dramatically in North Africa. Indeed, the population tends to stay very close to the coast. The vast Saharan desert is in many ways an ocean devoid of people. While specific population data is non-existent in the NGA dataset (2010a-d), it is possible to turn to other sources to estimate population density. The Columbia University Center for International Earth Science Information Network hosts a particularly useful application known as the Gridded Population of the World (CIESIN, 2005). This data covers the entire world, and estimates population density using a grid of values in the form of a raster map (CIESIN, 2010). With the use of geospatial analysis tools, it is possible to sample the population density at each of the 27 known hometowns. Since these are estimates, there are actually some samples with a value of zero. However, the higher densities do correspond with the national capitals and such large cities as Benghazi and Casablanca. In any case, the results can then become an attribute for later analysis of hometowns.

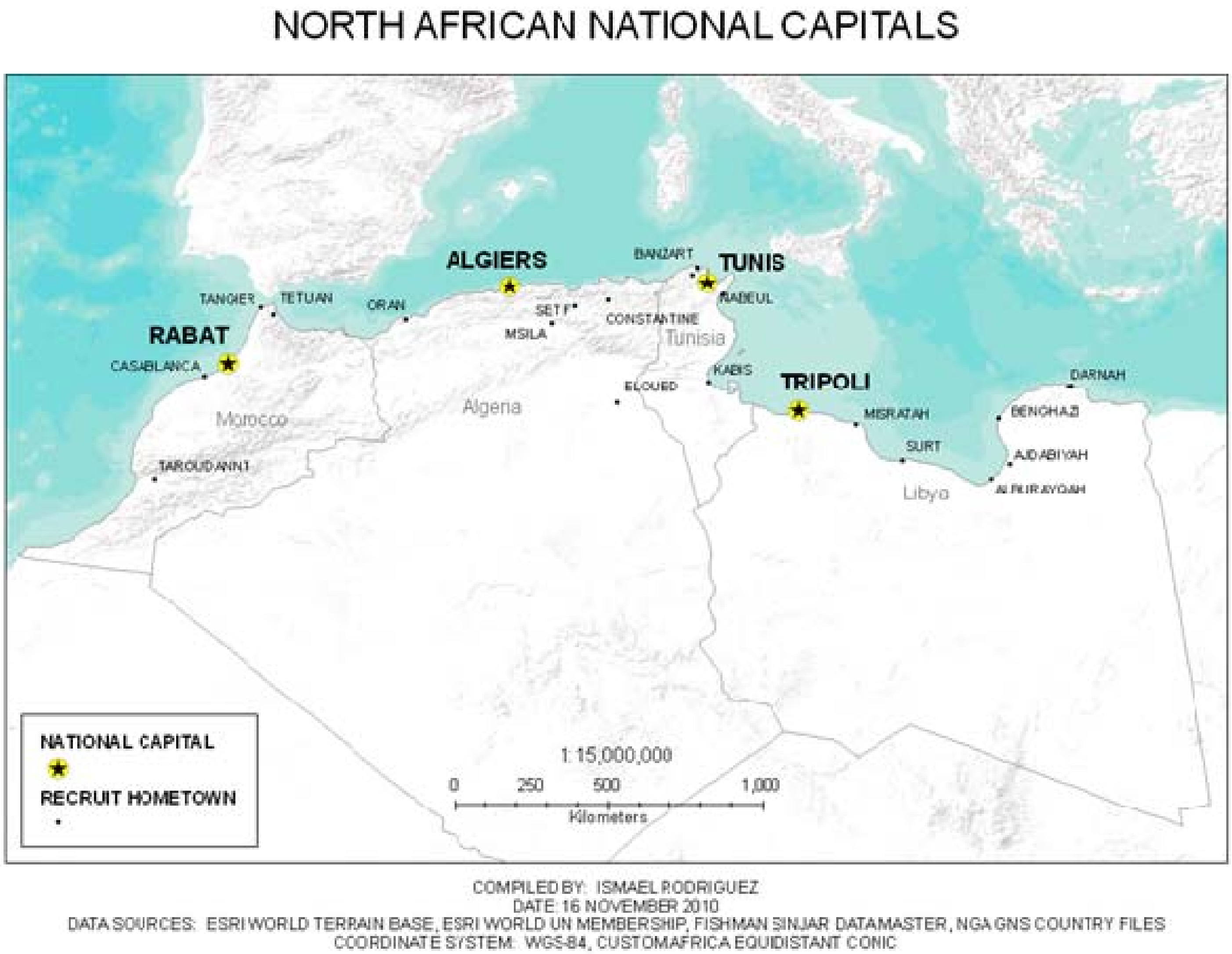


Figure 4. North African National Capitals

D. NATIONAL CAPITALS

The location of national capitals is the easiest data to prepare. The theoretical foundations for this choice of data fall within the realm of social movement theory. In particular, the notion of repression factors into this choice. Mohammad Hafez (2003) in *Why Muslim's Rebel* contends that repression is central to the growth of Islamist movements, despite attempts by the state to check such activities (p. 22). Quintan Wictorowicz (2003) also considers the role in which repression plays in the development of informal organizations meant to counter state applied pressure (p. 12). Each state within North Africa displays varying degrees of authoritarianism. This is quite apparent in the paltry Freedom House (2008) scores for civil liberties and political rights, which taken together depict levels of repression around the world (p. 120). Of the four countries, only Morocco rates as partly free, while the others fall into the category of not free, with Libya receiving a place on the organization's list of poorest performers for 2008.

Table 1. North African Freedom House Scores

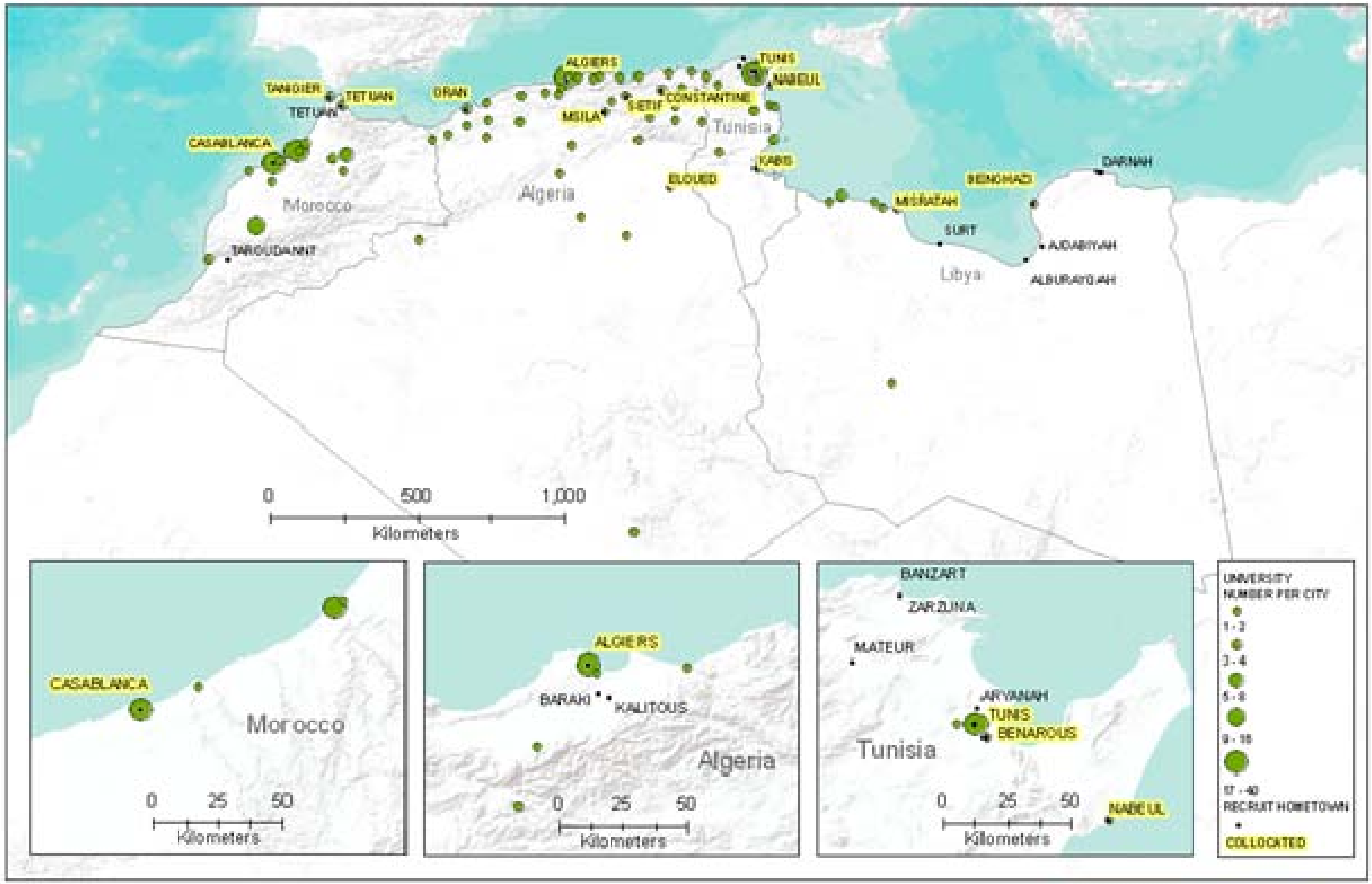
Country	Political Rights	Civil Liberties	Freedom Rating
Algeria	6	5	Not Free
Libya	7	7	Not Free
Morocco	5	4	Partly Free
Tunisia	7	5	Not Free

(Compiled from Freedom House, 2008, pp. 113,115,116, 118)

As previous research suggests, there appears to be a causal link between national level recruitment trends and the freedom house scores (Watts, 2008). Since it is unlikely to find spatial measures of repression internal to these countries, this thesis assumes the national capital as a proxy for the center of repressive power within the state. In other words, this study expects that hometowns further away from national capitals are more likely to produce recruits.

Identifying the national capitals is a simple process of selecting the listed national capitals from the NGA GNS datasets for each country (NGA, 2010a-d)(MIT, n.d). This table of capitals forms the basis for a simple map layer. Once plotted, it is then possible to measure the distance from each hometown to the nearest capital. The results can then form another column of attributes for analysis of those hometowns. Six of the hometowns fall within 15 kilometers of a capital, while the remaining 21 towns are greater than 50 kilometers away. Only ten of the recruits come from capital cities with eleven more coming from nearby suburbs. Darnah was the most distant hometown at 885 kilometers.

NORTH AFRICAN UNIVERSITY CITIES



COMPILED BY: ISMAEL RODRIGUEZ
DATE: 16 NOVEMBER 2010
DATA SOURCES: ESRI WORLD TERRAIN BASE, ESRI WORLD UNI MEMBERSHIP, FISHMAN SINJAR DATA MASTER, NGA GNS COUNTRY FILES,
IAU 2009 WORLD HIGHER EDUCATION DATABASE
COORDINATE SYSTEM: WGS-84, CUSTOM AFRICA EQUIDISTANT CONIC

Figure 5. North African University Cities

E. UNIVERSITIES

Locating North African universities presents several challenges. The theoretical underpinnings of this choice of variable come from both social movement theory, as well as the writings of Marc Sageman. As Sidney Tarrow (1998) explains, "[I]nstitutions are particularly economical 'host' settings in which movements can germinate" (p. 22). Additionally, Sageman (2008) identifies a relationship between membership in al Qaeda and a tendency for those members to have technical training in such fields as engineering or medicine (p. 59). In essence, universities are distinct, identifiable institutions. Thus, while it would be wonderful to have a thorough database of other conducive facilities, this simply is not something readily available in an open academic environment. Nevertheless, the process of putting together a comprehensive list of universities is not an easy endeavor.

There are several online resources that list universities in the developing world. In the case of North Africa, many of these sites seem geared for a general audience. Determining the quality of such sites is difficult. There are, however, more authoritative resources. The World Higher Education Database (WHED) meets such a standard. Authored by the UNESCO affiliated International Association of Universities, this data set includes only institutions that offer four year diplomas or post graduate education (IAU, 2009). In all, there were a total of 230 different institutions listed for the four states of North Africa. Still, this school data set required additional preparation for use in spatial analysis. Specifically, each university location was matched with a corresponding city from the NGA GNS dataset (NGA, 2010a-d). Unlike the Sinjar data set, there were far fewer transliteration issues within the university dataset. With the combined data from NGA and WHED, it was a simple process to plot the locations and measure distances from recruit hometowns.

There are noticeable patterns within the university data layer. Each country tends to group large numbers of universities in a small number of towns.

For instance, the largest cluster occurs in Casablanca, Morocco, with a total of 40 universities. Furthermore, the capitals of Morocco, Tunisia, Libya, and Algeria also host comparatively large numbers of schools, with a total of 75 schools located in these national capitals. On the other end of the spectrum, there are 49 towns that host a single institution and nine towns that host two schools.

On average, the hometowns were 48 kilometers from a college town, with 15 hometowns coinciding with a university town. The most distant hometown was Darnah, Libya, which was 235 kilometers from the nearest institution.

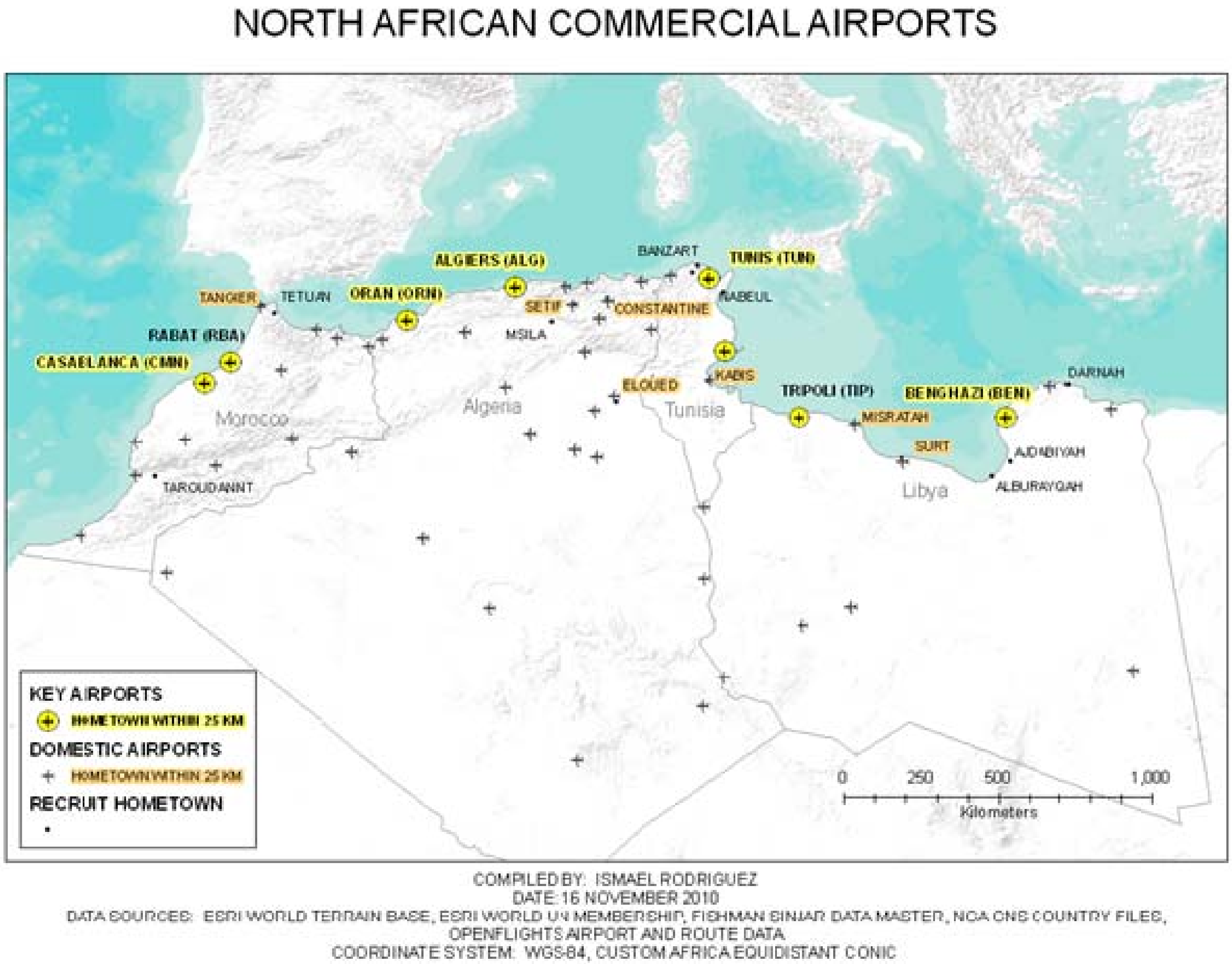


Figure 6. North African Commercial Airports

F. COMMERCIAL AIRPORTS

Commercial airport locations, while very easy to identify, were actually the most difficult to prepare for analytical use. Theoretically, such transportation requirements fall within the realm of necessary resources. As such, these fit most tightly within social movement theory. Hafez (2003) is most explicit about such necessities, differentiating movement resources into the categories that incorporate not only those necessary for group identification, and institutional support, but also include organizational and infrastructure requirements (p. 19). While not explicit about transportation infrastructure, he suggests that "[m]aterial and organizational resources provide Islamists with the capacity to mobilize people." (p. 20). Within insurgency studies, research has also shown a relationship between the density of transportation networks and the occurrence of insurgent violence. Of note, Yuri Zhukov (2010), a graduate researcher at the Harvard Department of Government, has identified a linkage between the spread of violence and the availability of road networks. Moreover, his research suggests that it is possible to predict the diffusion of violence in a manner similar to that used to predict the spread of communicable diseases within a social network (pp. 1–2). Moreover, Zhukov notes that the absence of infrastructure can prohibitively increase the cost of operations for a terrorist or insurgent organization (p. 4).

There are a multitude of resources available to identify air hubs worldwide. While the Federal Aviation Administration provided a worldwide dataset known as the DAFIF database, access to the data ended in 2006 (OpenFlights, 2009). In its place, OpenFlights created a collaboratively compiled dataset. This data builds upon 2006 DAFIF data, adding public domain data from OurAirports. The resulting attributes include the airport name in addition to IATA three letter airport designator codes, ICAO four letter airport codes, and latitude and longitude coordinates (OpenFlights, 2010). That much is easy.

From this point, it is important to identify airports that actually have commercial links to countries surrounding Iraq. Because actual data for activity in 2006 and 2007 are not readily available, this process involved two essential steps. First, a review of the Sinjar Dataset indicates some of the popular air routes used on trips to Iraq. Of the 41 North African recruits who admitted to air travel, 18 flew through Egypt, nine through Turkey, eight through Syria, and the remainder through airports in Saudi Arabia, Spain, and Tunisia.[4] While most of the trips concluded after a single stop, five also made an additional stop in such countries as Libya, Jordan or Turkey (Fishman, n.d)

With this knowledge in hand, there is a wealth of online material to piece together possible flight routes between the airports of North Africa and the airports of Syria. Using the OpenFlights (2010) interactive website, it is possible to explore the network of current routes. This process expanded possible routes to include travel through known hubs in Spain, France, Italy, Germany, Morocco, Algeria, Tunisia, Libya, Turkey, Greece, Jordan, and Saudi Arabia. In terms of connectivity, the North African airports at Casablanca (CMN), Algiers (ALG), Tunis (TUN), Benghazi (BEN), and Tripoli (TIP) have strong links between regional airports and Syria. Additionally, Cairo (CAI), Istanbul (IST), Damascus (DAM), and Amman (AMM) also have many routes into North Africa. Between Europe and North Africa, the airports of Paris (ORY, CDG), Madrid (MAD), Rome (ICO), and Athens (ATH) also have good connections to North Africa (Openflights, 2010). From the initial analysis, it is possible to select two primary airports per country. The major international airport for each state is simple to identify. These have excellent connectivity both regionally and to Europe and the Levant. The secondary airports either had connections to European and domestic flights, or displayed hub like tendencies as in the case of Benghazi.

[4] The number of recruits who listed how they arrived in Syria is quite small. Less than one quarter, or 55 of the 221 North African recruits, described the type of transportation used. Of these, air travel was much more common than ground travel into Syria with only 13 listing some form of ground transportation (Fishman, n.d).

With the location of these eight airports plotted, distance calculations are then possible. On average, the hometowns were 134 kilometers from the nearest major airport. Nine hometowns, with a total of 59 recruits, were within 25 kilometers, while El Oued, Algeria, was the farthest from a major airport at 385 kilometers.

Refining the airport network requires a better understanding of regional flights. To complete this task requires data to model domestic flights into hub airports. In particular, this subset depends upon the domestic routes of the four national carriers, as well as al Buraq Airlines, a private carrier with connections between North Africa and Aleppo, Syria (OpenFlights, 2010)(Kaminski-Morrow, 2005). The result is a list of airports with connections to Casablanca, Algiers, Oran, Tunis, Benghazi, and Tripoli. With this information plotted, a second set of distance calculations are possible. On average, hometowns were 40 kilometers from the nearest domestic airport. 16 hometowns were within 25 kilometers, while Al Bariqah, Libya, was farthest at 199 kilometers.

In summary, the primary result of this extensive data preparation is a table of variables. Pivoting around the number of recruits from each location, it also includes the calculated population density, as well as the distances to the national capital, closest university, closest key airport, and closest domestic airport.

Table 2. Recruit Hometowns and Associated Distances

Hometown	Country	Recruits	Population Density	Capital Distance	University Distance	Key Airport Distance	Domestic Airport Distance
Algiers	Algeria	5	7503	0.00	0.00	16.86	16.86
Baraki	Algeria	2	1533	11.47	7.60	11.09	11.09
Constantine	Algeria	2	411	324.37	0.00	308.72	9.89
El Oued	Algeria	8	13	514.45	0.00	384.92	19.01
Kalitous	Algeria	1	8976	14.96	10.60	7.42	7.42
M'Sila	Algeria	1	57	178.79	0.00	162.37	88.25
Oran	Algeria	1	630	354.58	0.00	7.69	7.69
Setif	Algeria	1	235	222.50	0.00	206.06	8.21
Ajdabiyah	Libya	4	0	706.47	151.25	148.38	148.38
AlBurayqah	Libya	1	0	665.19	195.34	198.76	198.76
Benghazi	Libya	20	82	652.33	0.00	19.25	19.25
Darnah	Libya	53	7	885.12	252.03	234.55	63.02
Misratah	Libya	3	125	188.04	0.00	184.10	6.54
Surt	Libya	5	3	371.78	192.05	361.80	16.14
Wadi Al Naqah	Libya	1	7	878.06	246.41	229.11	56.08
Casablanca	Morocco	17	3816	85.14	0.00	24.75	24.75
Tangier	Morocco	2	1010	217.60	0.00	210.22	11.12
Taroudannt	Morocco	1	49	437.13	68.49	343.52	64.56
Tetuan	Morocco	6	240	219.32	0.00	210.81	52.48
Aryanah	Tunisia	1	2377	6.48	6.48	3.21	3.21
Banzart	Tunisia	2	157	59.10	55.66	56.68	56.68
Benarous	Tunisia	7	1205	6.55	0.00	10.90	10.90
Kabis	Tunisia	1	47	324.20	0.00	107.71	0.69
Mateur	Tunisia	1	157	53.19	46.88	54.65	54.65
Nabeul	Tunisia	1	230	63.41	0.00	63.71	63.71
Tunis	Tunisia	5	2676	0.00	0.00	6.85	6.85
Zarzuna	Tunisia	1	157	57.98	54.51	55.60	55.60

All distances in kilometers.
(Derived from Fishman (n.d); CIESIN (2005); OpenFlights (2010); NGA GNS (2010a-d); and IAU WHED (2009))

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RESULTS

Given the data prepared for this study, the next and most important step is to determine whether these various factors actually impact recruitment patterns in North Africa. Using ArcGIS analytical tools and OpenGeoDa, an open source geospatial analysis package (GeoDa Center, n.d.), it is possible to perform a series of regression tests. Specifically, this section of the study compares the results of simple ordinary least square regression models, spatially lagged ordinary least square regression models, and geographically weighted regression models. The interpretation of these results will then feed into a set of two recruiting risk terrain maps. These examples go head to head with a recruitment density map to see which one best predicts recruitment patterns.

A. ORDINARY LEAST SQUARES REGRESSION

The purpose of ordinary least square regression is to test for correlation between variables (Mitchell, 2005, pp. 212–214). The dependent variable for this study has always been the number of recruits that hail from a given hometown. That said, it is no simple endeavor to develop a set of explanatory variables.

1. Assumptions

This basic model assumes that activity in each location is independent. In other words, there is no influence from one hometown to the next. More importantly, it assumes that the spatial recruitment patterns for the entire region reflect those in the limited sample size. Finally, this model assumes that all data in the original records, the translated records, and the compilation of distance variables are correct.

2. Model

Mathematically, the formula for this test is straightforward.

$$y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + X_4\beta_4 + X_5\beta_5 + \varepsilon$$

y	=	Number of Recruits
β_0	=	Intercept Coefficient
X_1	=	Population Density
β_1	=	Population Density Coefficient
X_2	=	Distance to Capital
β_2	=	Capital Coefficient
X_3	=	Distance to University
β_3	=	University Coefficient
X_4	=	Distance to Domestic Airport
β_4	=	Domestic Airport Coefficient
X_5	=	Distance to Key Airport
β_5	=	Key Airport Coefficient
ϵ	=	Error Term

(Adapted from Scott, Rosenshein & Janikas, 2010, p. 6)

3. Calculations and Results

In essence, OpenGeoDa is a spatial calculator capable of performing a wide variety of spatial statistics processes. (GeoDa Center, n.d.). Moreover, the tool presents a simple user interface to assign dependent and independent variables and provides a thorough set of diagnostic statistics (Anselin, 2005, pp. 172–175).

Table 3. OLS Model Diagnostic Statistics

Criteria	OLS Model 1	OLS Model 2	OLS Model 3	OLS Model 4
Dependent Variable	Recruits	Recruits	Recruits	Recruits
Independent Variables	Population Density	Population Density		
	Capital Distance	Capital Distance	Capital Distance	Capital Distance
	University Distance	University Distance	University Distance	University Distance
	Domestic Airport Distance	Domestic Airport Distance	Domestic Airport Distance	Domestic Airport Distance
		Key Airport Distance		Key Airport Distance
Degrees of Freedom	22	21	23	22
R-Squared	0.289	0.327	0.274	0.324
Adjusted R-Squared	0.159	0.166	0.179	0.201
Akaike Info Criterion (AIC)	203.771	204.288	202.338	202.410
Multicollinearity Condition Number	5.419	6.106	4.694	5.306
Jarque-Bera Test Probability	0.000	0.000	0.000	0.000
Koenker-Basset Test Probability	0.001	0.001	0.000	0.000

(Compiled from OpenGeoDa Regression Results)

4. Interpretation of Results

While these results have some promising elements, the models themselves leave much to be desired. That said, for a more in-depth analysis, it is important to test the model itself. ESRI has developed a six part test of spatial OLS regression results to do just that. Accordingly, ESRI's Lauren Scott, Lauren Rosenstein, and Mark Janikas (2010) list the conditions as:

1 Coefficients have the expected sign.
2 No redundancy among explanatory variables.
3 Coefficients are statistically significant.
4 Residuals are normally distributed.
5 Strong Adjusted R-Square value.
6 Residuals are not spatially autocorrelated. (p. 11)

While quite useful, Scott et al. also offer a pair of more specific suggestions. First, by using the Akaike's Information Criterion (AIC), it is possible to compare different regression models (p. 15). Second, if the Koenker test is statistically significant, then there is room for improvement by implementing a geographically weighted regression (p. 19).

Overall, this framework lays a foundation for reviewing the results produced by OpenGeoDa. As such, it fits closely with the specific procedures described by Luc Anselin (2005) in his workbook *Exploring Spatial Data with GeoDa™*. The software package provides diagnostics that examine the same conditions described by ESRI. In particular, it uses a number of statistics to measure model fit to include R-squared, Adjusted R-squared, and AIC. Anselin also emphasizes that lower AIC values indicate better model performance (p. 175). The regression diagnostics also examine a model for residual related issues, as identified by the Jarque-Bera test, as well as multicollinearity and heteroskedasticity (pp. 193–195).[5] Moreover, in addition to the other residual tests, the package provides a Moran's I statistic to test for spatial autocorrelation (pp. 196–197).

[5] The ESRI (2010a) ArcGIS Desktop 10.0 online help file "Interpreting OLS results" offers a more detailed discussion of the Koenker's studentized Breusch-Pagan statistic used for heteroskedasticity. The GeoDa specific Koenker-Bassett test, as described by Anselin (2005, p. 195) appears to be the same as the ArcGIS Koenker's studentized Breusch-Pagan test.

Table 4. OLS Model 4 Characteristics

	Coefficient	Std. Error	z-value	Probability
Constant (Intercept)	4.39	3.13	1.401	0.175
Capital Distance	2.11E-05	1.07E-05	1.966	0.062
University Distance	3.55E-05	3.58E-05	0.992	0.332
Domestic Airport Distance	-7.87E-05	4.89E-05	-1.612	0.121
Key Airport Distance	-2.32E-05	1.82E-05	-1.276	0.215

(Compiled from OpenGeoDa Regression Results)

Using these guidelines, a comparison indicates that the fourth model is the best, due to its high Adjusted R-squared and low AIC. Moreover, the residuals for this model do not appear to show statistically significant signs of spatial autocorrelation (See Appendix B).[6] Superficially, there appears to be statistically significant relationships between recruitment levels and both the variables for national capital and domestic airport distance. However, there does not appear to be a statistically significant relationship for key airport or university distance. Moreover, population density does not factor into the selected model. As such, these results may suggest a more prominent impact of state repression, and less prominence attribution to the educational, transportation, and high population density associated with many modern urban areas. Nevertheless, Model 4 does still have concerns. Of particular note are the Jarque-Bera test of residuals and the Koenker-Bassett tests for heteroskedasticity. While it may be possible to disregard the Jarque-Bera test (Anselin, 2005, p. 195), the issue of heteroskedasticity warrants contemplating the use of a geographically weighted regression model.

B. SPATIALLY LAGGED ORDINARY LEAST SQUARED REGRESSION

The next iteration of tests actually considers the impact of space on the regression model. In essence, it extracts this value from the error term of a basic

[6] Scott et al. (2010) suggest testing residuals for spatial autocorrelation on models that otherwise meet their listed criteria (p. 34). Borrowing from this notion, this study only tests for spatial autocorrelation on the model chosen.

OLS model. As Michael Ward and Kristian Gleditsch explain, spatially lagged models incorporate the influence of nearby dependent variable values into the overall formula for a dependent variable. However, they also warn that such models are appropriate when the dependent variable is not binary but instead continuous (p. 29). Adjusting for a continuous variable requires additional preparation. This involves setting up a contiguous surface. OpenGeoDa can convert point files into a Theissen polygon file (Anselin, 2005, p. 40). With the polygon file established, one last step is necessary. Known as a spatial weights file, this information takes into consideration a given entities bordering entities (p. 106).

1. Assumptions

While no longer assuming independence between variables, this model still assumes that the sample data reflects actual recruitment patterns. Moreover, the model rests upon the assumption that all data, locations, and data processes are accurate.

2. Model

Mathematically, the new formula appears as:

$$y = \rho W y + \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + X_4\beta_4 + X_5\beta_5 + \varepsilon$$

y = Number of Recruits
ρWy = Spatially Lagged Variable

ρ = Spatial Autoregressive Parameter
W = Spatial Weights Matrix
y = Number of Recruits

β_0 = Intercept Coefficient
X_1 = Population Density

β_1 = Population Density Coefficient
X_2 = Distance to Capital
β_2 = Capital Coefficient
X_3 = Distance to University
β_3 = University Coefficient
X_4 = Distance to Domestic Airport
β_4 = Domestic Airport Coefficient
X_5 = Distance to Key Airport
β_5 = Key Airport Coefficient

ϵ = Error Term

(Adapted from Scott, Rosenshein & Janikas, 2010, p. 6; and Anselin, 2005, p. 201)

3. Calculations and Results

OpenGeoDa again offers an easy interface to calculate results. The only real difference between calculations is the specification of spatial weights. Specifically, these models use a queen contiguity weights matrix. As Anselin (2005) notes, "[t]he *queen* criterion determines neighboring units as those that have *any* point in common, including common boundaries and common corners" (p. 112). Once complete, it is a simple matter of assigning the dependent and independent variables (pp. 204–207).

Table 5. OLS-Lag Model Diagnostic Statistics

Criteria	OLS-Lag Model 1	OLS-Lag Model 2	OLS-Lag Model 3	OLS-Lag Model 4
Dependent Variable	Recruits	Recruits	Recruits	Recruits
Independent Variables	Population Density	Population Density		
	Capital Distance	Capital Distance	Capital Distance	Capital Distance
	University Distance	University Distance	University Distance	University Distance
	Domestic Airport Distance	Domestic Airport Distance	Domestic Airport Distance	Domestic Airport Distance
		Key Airport Distance		Key Airport Distance
	Spatial Lag	Spatial Lag	Spatial Lag	Spatial Lag
Degrees of Freedom	21	20	22	21
R-Squared	0.571	0.590	0.554	0.584
Akaike's Info Criterion (AIC)	196.625	197.237	195.568	195.565

(Compiled from OpenGeoDa Lagged Regression Results)

4. Interpretation of Results

Anselin also emphasizes that the interpretation of spatially lagged results does not use quite the same criteria as those necessary for spatial OLS interpretation. Instead of focusing on r-squared values, he suggests that a model's AIC, as well as its Schwartz criterion and log likelihood, are better indicators of fit (pp. 207–208). For comparison purposes, this study uses AIC to identify the best option among the OLS and OLS-Lag models. Therefore, OLS-Lag Model 4 appears to have the best fit.

Table 6. OLS-Lag Model 4 Characteristics

	Coefficient	Std. Error	z-value	Probability
Constant (Intercept)	5.629	2.285	2.464	0.014
Capital Distance	2.656E-05	7.801E-06	3.405	0.001
University Distance	4.931E-05	2.547E-05	1.936	0.053
Domestic Airport Distance	-6.777E-05	3.481E-05	-1.947	0.052
Key Airport Distance	-1.859E-05	1.295E-05	-1.436	0.151
Spatial Lag	-0.790	0.206	-3.841	0.000

(Compiled from OpenGeoDa Lagged Regression Results)

A closer look at the OLS-Lag Model 4 reveals a much-improved set of statistically significant variables. Still, the very small coefficients call into question the degree of explanatory power for each of the independent variables. In all, the spatially lagged variable, in addition to capital distance, university distance, and domestic airport distance appear to be most statistically significant. Put another way, once the effects of nearby recruitment activity are taken into consideration, proximity to transportation and distance from both capitals and universities come into play. Of particular note is the role of university proximity. Its negative coefficient is not in the direction expected. While it would seem that being close to a university would make a person more likely to become a recruit, the opposite appears to be the case. While speculative, this could be a result of a regime's reaction to the potential threat posed by such locations.

C. GEOGRAPHICALLY WEIGHTED REGRESSION

Geographically weighted regression (GWR) is another technique to account for spatial variation in data. This series of models underscore some interesting trends. The first model considers the four explanatory variables of population density, capital distance, university distance, and domestic airport distance. The second model considers five variables, adding the key airport distance to the original mix. The final model uses four explanatory variables, dropping population density, but keeping all the distance variables. The table below summarizes the results of these iterations.

1. Assumptions

This set of models uses the same assumptions identified for the previous OLS models.

2. Model

A GWR model calculates a regression for the specified locations under examination (ESRI, 2010b). In other words, it determines a specific set of coefficients for each of the 27 locations in the study area. The basic formula for the model is otherwise the same.

$$y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + X_4\beta_4 + X_5\beta_5 + \varepsilon$$

y	=	Number of Recruits
β_0	=	Intercept Coefficient
X_1	=	Population Density
β_1	=	Population Density Coefficient
X_2	=	Distance to Capital
β_2	=	Capital Coefficient
X_3	=	Distance to University
β_3	=	University Coefficient
X_4	=	Distance to Domestic Airport
β_4	=	Domestic Airport Coefficient
X_5	=	Distance to Key Airport
β_5	=	Key Airport Coefficient
ϵ	=	Error Term

(Adapted from Scott, Rosenshein & Janikas, 2010, p. 6)

3. Calculations and Results

ArcGIS provides the platform to estimate GWR models. The process is far more involved than the OLS analysis using OpenGeoDa. For instance, ArcGIS provides a choice between default calculation parameters and a variety of user identified parameters (ESRI, 2010c). For this study, the models each use an adaptive kernel type, cross validation bandwidth methods, distance of six, and number of neighbors of 30.

Table 7. GWR Results

Criteria	GWR Model 1	GWR Model 2	GWR Model 3
Dependent Variable	Recruits	Recruits	Recruits
Independent Variables	Population Density	Population Density	
	Capital Distance	Capital Distance	Capital Distance
	University Distance	University Distance	University Distance
	Domestic Airport Distance	Domestic Airport Distance	Domestic Airport Distance
		Key Airport Distance	Key Airport Distance
R-Squared	0.37897	0.4088	0.40496
Adjusted R-Squared	0.13918	0.1129	0.1572
Corrected Akaike's Information Criterion (AICc)	214.90500	220.0278	215.2262

(Compiled from ArcGIS 9.31 GWR Results)

4. Interpretation of Results

Making sense of GWR results can be difficult. Fortunately, the ArcGIS Resource Center website provides a thorough explanation. In this reference, there is an emphasis to examine the Adjusted R-squared value, since it allows for the comparison of models with differing numbers of explanatory variables. More importantly, the primarily comparison diagnostic is the corrected Akaike's Information Criterion (AICc), which allows for comparison with other regression models (ESRI, 2010c). Therefore, while the best R-squared value occurs in the second model, it is actually quite similar to the third model, which has a significantly improved adjusted R-Squared value, and a smaller AICc. Thus, of these three options, Model 3 seems to provide the best fit.

Still, it is essential to examine the residuals for signs of spatial autocorrelation (ESRI, 2010c). Based on these simple criteria, it is possible to examine the specific results of the third model.

One of the more useful results from the ArcGIS Geographically Weighted Regression Analysis is a series of raster images that depict variation in coefficient values (ESRI, 2010c). These images provide a visualization of where and to what degree an explanatory variable impacts the dependent variable.

The University Distance coefficient indicates that there is a changing relationship largely dependent upon the country in question. In Morocco, there is a small negative relationship while in Libya there is a small positive relationship. Thus, In the case of Morocco, the large pockets of recruits did indeed emerge in or very near the university towns of Casablanca and Tetuan. On the other hand, the positive relationship near Darnah, corresponded with Darnah's great distance from a listed university.

The Capital Distance coefficient also depicts a changing relationship. In the case of Morocco and Algeria, there appears to be a slight negative relationship, while in Libya there is a weak positive relationship. Looking at this from a national perspective, this pattern seems to fall in line with the differences in repression levels, at least between Morocco, which is considered partially free, and Libya which is considered not Free (Freedom House, 2008, pp. 115–116).

Key Airports also show variation across the continent. There is a positive relationship in the west and a negative relationship in the east. While Benghazi in Libya corresponds with a key airport, the other recruitment pockets tend to be a fair distance away from a key airport. At the other end of the continent, the large numerous recruits in Casablanca also have close access to a key airport, while those recruits in Tetuan must travel a great distance to arrive at such a facility.

Domestic Airports show a slight effect and limited variation across the continent. The strongest impact is in the east where the variable has the greatest impact. In the west, the impact not only lessens, but also shifts to a positive relationship.

In all, the GWR results shed light on the regional variation of recruitment patterns. From the standpoint of interpretation, the ArcGIS help file rounds out its discussion by suggesting that there can be a policy role for the coefficient maps. Whereas regional policies can gain insight from statistically significant global variable coefficients that vary little over an area, local policies can gain insight from statistically significant global variable coefficients that vary to a greater degree. Moreover, a changing relationship may cause a variable not to be significant at the global level (ESRI, 2010c). As such, the coefficients in this study are all quite small, and they shift relationships across the region. Of the four variables, the university coefficient shows the largest variation, while the capital coefficient displays the smallest change across the region. However, the university coefficient is also the least statistically significant of the four variables, a trend possibly exacerbated by the balanced shift from positive to negative coefficients. Otherwise, solutions to mitigate the other trends might be feasible at the regional level. In any case, these outcomes seem a bit disappointing. Fortunately, there is another approach to judging the impact of distance on recruitment patterns.

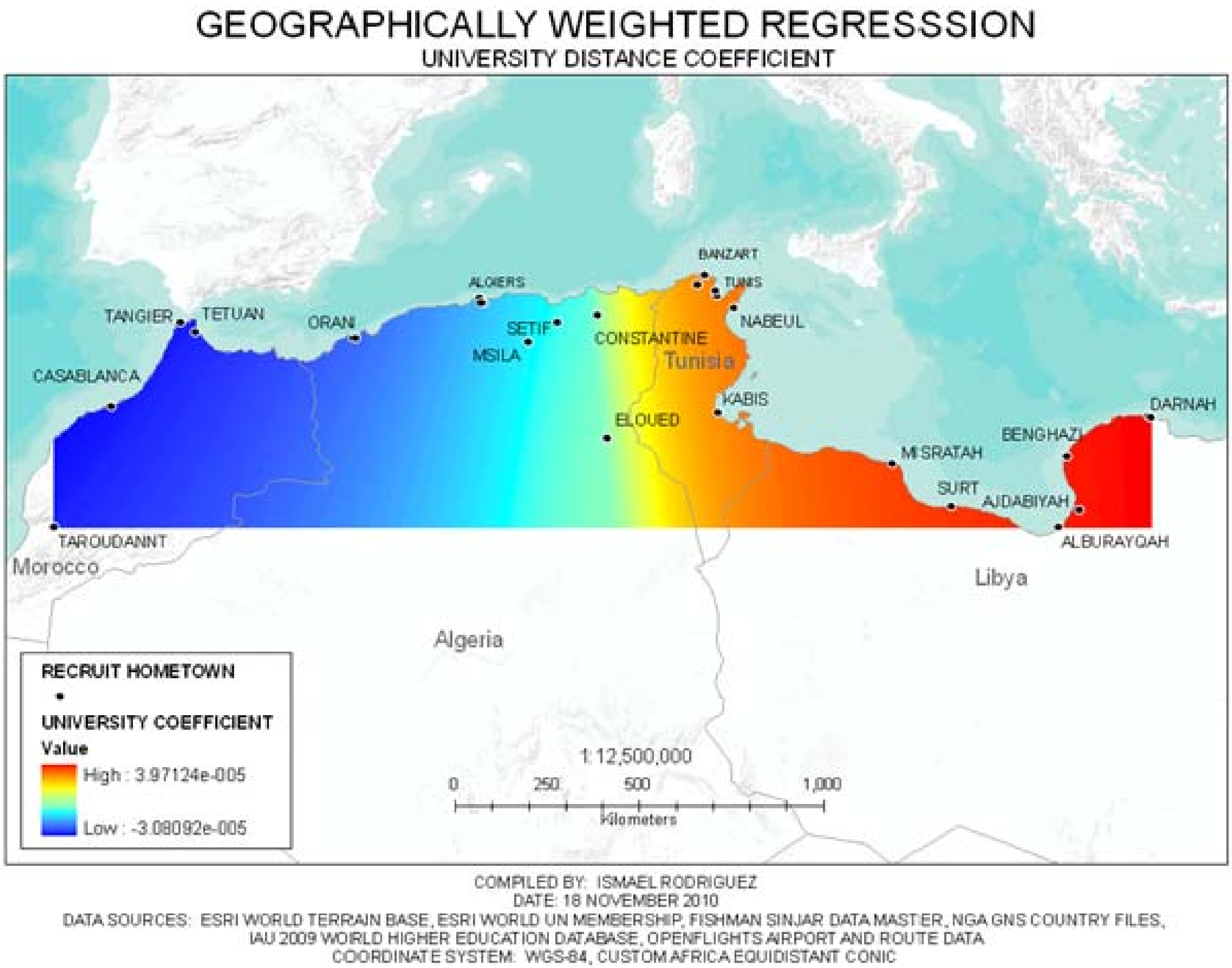


Figure 7. GWR Results University Distance Coefficient

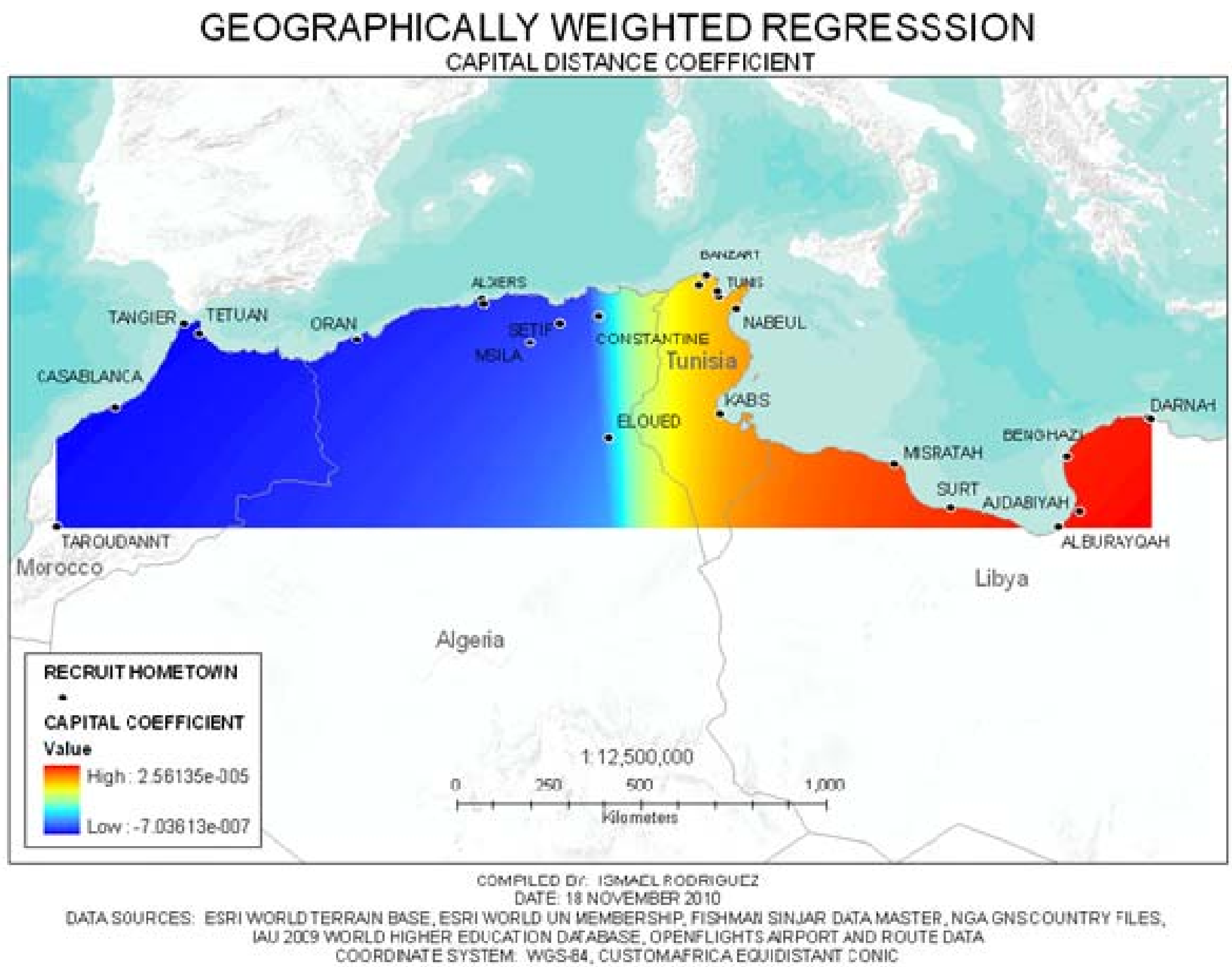


Figure 8. GWR Capital Distance Coefficient

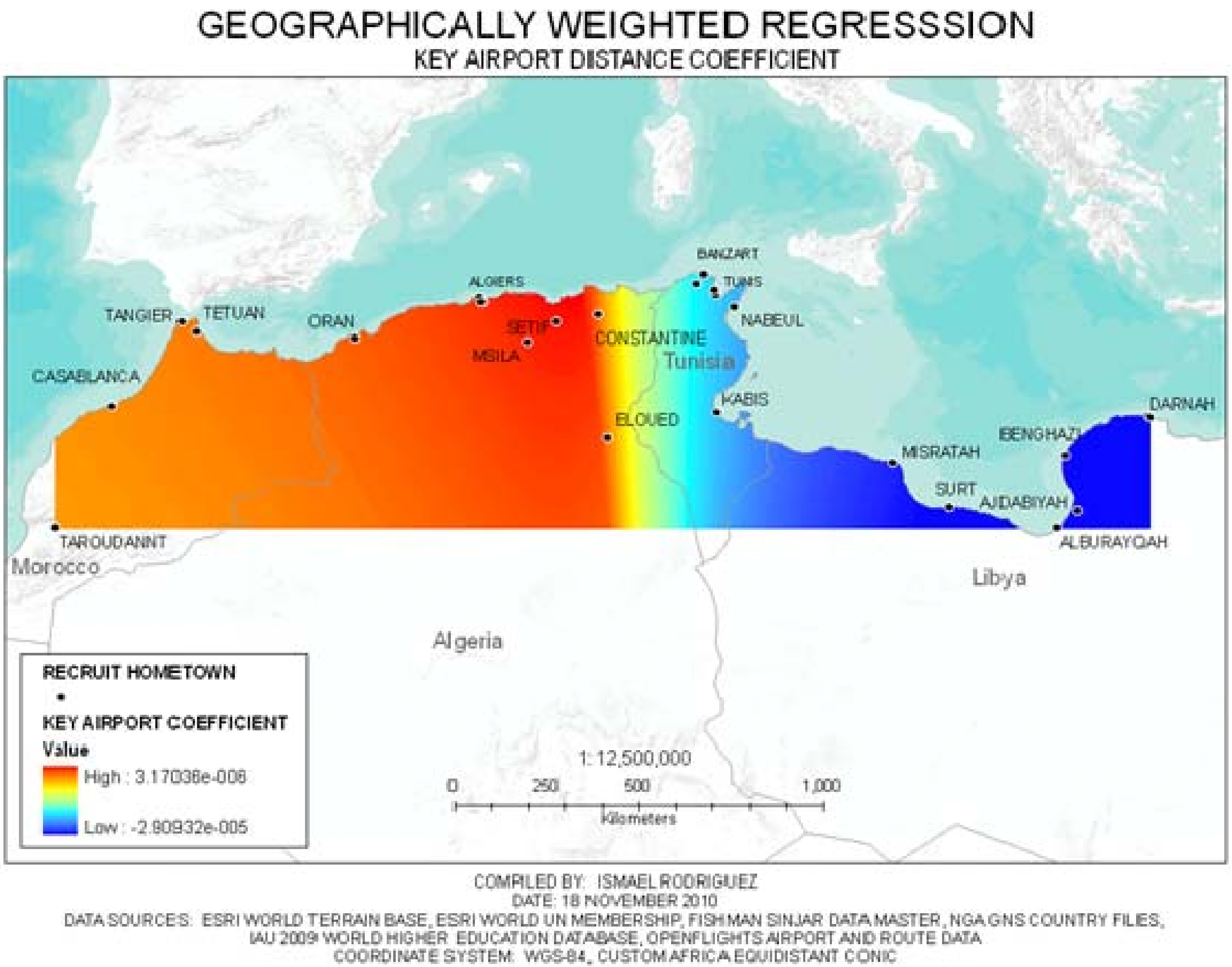


Figure 9. GWR Key Airport Distance Coefficient

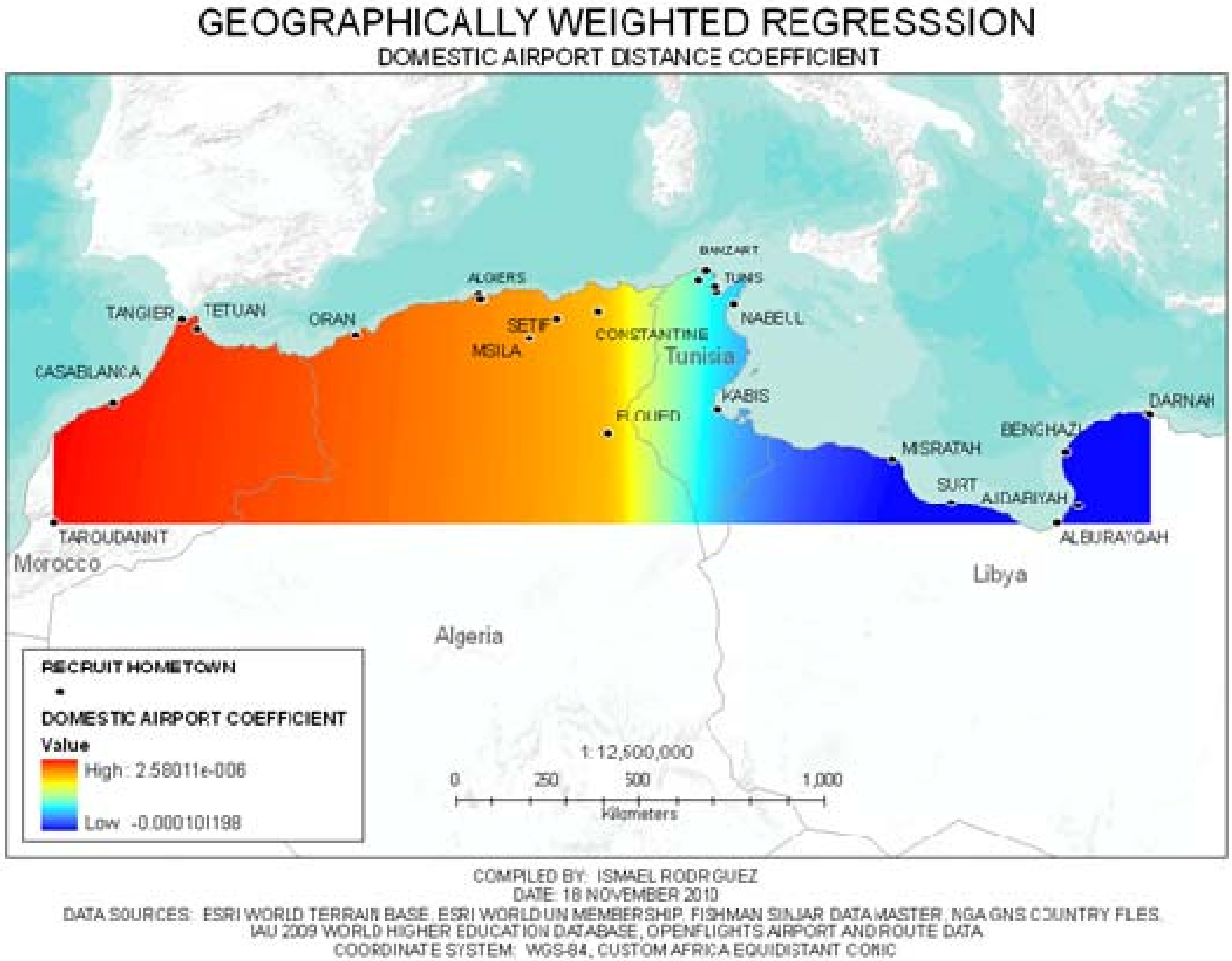


Figure 10. GWR Domestic Airport Coefficient

D. RISK TERRAIN MAP

Joel Caplan and Leslie Kennedy (2010) offer a step-by-step approach to crafting risk terrain maps. In essence, that method standardizes risk factors to common geographic units over a continuous surface. Separate map layers representing the presence, absence, or intensity of each risk factor at every place throughout the terrain is created in a geographic information system (GIS), and then all map layers are combined to produce a composite map with attribute values that account for all risk factors at every place throughout the geography (p. 24).

From a technical standpoint, the choice of variables can come from theory, experience, or study (p. 24). More specifically, the manual suggests that "[a]t the very least, make a reasonable effort to identify as many factors that you believe to be related to the outcome event in your particular study area" (p. 79). Furthermore, it is also possible to incorporate past activity into these maps (pp. 36-39). Finally, risk terrain maps allow for variable weighting. Assigning weights is simply the process of rank ordering variables by degrees of importance. Although Caplan and Kennedy suggest using a logistical regression process to develop weights (pp. 93–94), for purposes of this study, the OLS-lag coefficients identified earlier should form a sufficient weighting scheme.[7]

Thus, it is quite feasible to merge the results of the previous regression analysis into a risk map. In all, this study constructs and examines two distinct composite risk maps. The first risk map considers the same variables as the OLS-Lag model, assigning equal weight for each map. This map uses a simple binary scale to calculate risk for each variable. To account for the spatially lagged dependent variable, it assigns a score to any location from which a recruit

[7] Chapter 8 of the *Risk Terrain Modeling Manual* presents a detailed explanation of the steps necessary to compile a risk terrain map (Caplan and Kennedy, 2010, pp. 72–99).

emerges. The second map builds upon this by creating a weighted map of the same factors. The weights for this map come from the coefficients identified in the OLS-Lag model.[8]

In order to create a risk map it is first essential to create a grid that spans the region under consideration. As the manual suggests, Hawth's Analysis Tools[9] offer an easy means to accomplish this step (p. 83). Diverging from the explicit instructions in the manual, the next step involves assigning attribute values for each grid square that correspond with the attribute values under consideration. In other words, this study uses a grid of 47,069 10 kilometer by 10 kilometer polygons and associated set of centroid locations. From these data points, it is then possible to calculate distances to the airports, universities, and national capitals. This distance data forms the basis for each risk layer.

While there is more than one way to calculate the given risk presented by a distance variable, it is essential to keep the scoring mechanism consistent. In other words, it is feasible to quantify risk either in terms as a simple yes or no for any given location, or as scale based However, for whatever method selected, all the variables should share the same scale (p. 89). Thus, for the purposes of this example, each variable translates into a risk zone and a no risk zone.

Caplan & Kennedy (2010) contend that RTM is a better forecasting tool than a hot spot Map, emphasizing the dynamic perspective that their tool uses. Furthermore, they note that the capability allows for regular revisions to incorporate mitigation efforts (pp. 29–30). They offer a complicated means to validate this claim. Using temporally coded spatial data, they split their sample into two groups they build a modified hot spot map using the same procedures as

---

[8] The use of coefficients for the distance variables posed few problems. However, the scale of the coefficient for spatially lag was much larger than the distance variable. As a result, it was set at 10 times the value of the distance coefficient instead of the actual magnitude of $10.^4$

[9] Hawth's Tools are a set of spatial analysis tools developed for use in ArcGIS (Beyer, n.d.). For a detailed description of the tools and links to follow on capabilities see "Hawth's Analysis Tools for ArcGIS" on the spatialecology.com website <http://www.spatialecology.com/htools/index.php>.

a risk map. In their example, retrospective risk is calculated by using standard deviation of incidents to differentiate levels of risk. They then compare the number of incidents that fall with this modified hot spot map to the number of incidents that fall within the risk terrain map (p. 31). To complete the comparison, Caplan and Kennedy build a comprehensive table that compares the two mapping schemes (pp. 32–33). However, while claiming that the validation step is optional, they also introduce regression as means to test validity. The one necessary ingredient for the procedure is temporal data. Beyond that, this form of regression only requires the risk score for each given location, and the number of events that occur at those same locations (pp. 100-101).

1. Assumptions

These models use a much smaller set of data to develop risk maps. Above all, they assume that the proximity factors identified through regression analysis are valid. Moreover, they also consider the explanatory power of each of these factors to be proportional and related to the OLS coefficients. Finally, the study assumes temporal data to be correct and to correspond closely with the date that each recruit left his hometown.

2. Model

The basic model for this portion of the study is a matter of simple arithmetic (pp. 96–97).

$$R_0=R_1+R_2+R_3+R_4+R_5$$

R_0	=	Composite Risk
R_1	=	Risk from Proximity to Capital
R_2	=	Risk from Proximity to University
R_3	=	Risk from Proximity to Domestic Airport
R_4	=	Risk from Proximity to Key Airport
R_5	=	Risk from Proximity to Past Activity

(Adapted from Caplan & Kennedy, 2010, pp. 96-97)

The second model uses OLS-Lag coefficients as a basis for weighting composite risk. Because the coefficients were very small, each was multiplied by 10^5.

$$R_0 = (10^5)(\beta_1 R_1 + \beta_2 R_2 + \beta_3 R_3 + \beta_4 R_4 + \rho\, W\, R_5)$$

R_0	=	Composite Risk
R_1	=	Risk from Proximity to Capital
β_1	=	Capital Coefficient
R_2	=	Risk from Proximity to University
β_2	=	University Coefficient
R_3	=	Risk from Proximity to Domestic Airport
β_3	=	Domestic Airport Coefficient
R_4	=	Risk from Proximity to Key Airport
β_4	=	Key Airport Coefficient
R_5	=	Risk from Proximity to Past Activity
ρ	=	Spatial Autoregressive Parameter
W	=	Spatial Weights Matrix

(Adapted from Caplan & Kennedy, 2010, p. 94, 96–97; Scott, Rosenshein & Janikas, 2010, p. 6; and Anselin, 2005, p. 201)

3. Calculations and Results

Quite possibly the hardest part of this entire process is the determination of risk zones for each variable. The small sample size of the temporal data set restricts the descriptive statistics for the distances in question. That said, there are 14 different hometowns in the sample. Each of the individual risk models uses standard deviation to set the values for risk. Table 8 shows the mean distances, standard deviations, and calculated risk boundary distance for each variable.

Table 8. October to February Recruit Hometown Distance Statistics

Variable	Mean	Standard Deviation	Risk Boundary
Capital Distance	267.0	258.4	525.4
University Distance	36.5	77.9	114.4
Domestic Airport Distance	26.6	21.4	48.0
Key Airport Distance	144.1	132.9	277.0

All distances in kilometers. (Derived from Fishman (n.d); OpenFlights (2010); NGA GNS (2010a-d); and IAU WHED (2009) data)

For comparison purposes, there are several differences with the descriptive statistics for all 27 cities. Of the four variables, the greatest change is the Domestic airport distance, for which the mean distance increases by over 13 kilometers, and its standard deviation expands by 24 kilometers. Otherwise, the two sample sizes are actually rather similar.

Accounting for past activity forms the final leg of this analysis. Using the default search setting of 20.9 kilometers, the study creates a kernel density estimate map based on the hometown location of each of the 37 recruits known to have arrived in Iraq between October 2006 and February 2007.[10] The resulting map is then symbolized into a risk vs. no-risk map, where risk is set using the standard deviation of values.[11] The density values range from 0.0 to 0.016, and the standard deviation is 0.0003. Thus, the no risk zone is anything less than the standard deviation, while the risk zone is anything higher.

Although there are several products from this analysis, this study focuses on the spatial depictions of composite risk. (See Appendix E for maps of the component risk factors). As Caplan and Kennedy (2010) suggest, the composite risk map is the eventual end product. However, for it to be useful, the map must

[10] See ESRI (2010d) "How Kernel Density Works" for an explanation of kernel density estimates. Once the Kernel Density Estimate raster is set, it is possible to reclassify it to reflect binary scores. This raster can then be converted into a polygon file, spatially joined with the 10 km grid set, and then converted into a binary map for use in the composite risk map.

[11] This choice of boundary emerged as a result of a discussion with Professor Sean Everton.

clearly convey risk. As such, the choice of classification and color schemes can impact its effectiveness. Moreover, while visual inspection of a map may reveal seemingly high-risk areas, statistical hot spot analysis can yield a more rigorous assessment (pp. 97–98). Specifically, for an area "[t]o be statistically significant, a group of cells must have high values and be surrounded by other cells with high values" (p. 98). That said, there appears to be a stark difference between the un-weighted and weighted risk maps. For the first map, the only area with a score of four or five falls in the eastern section of Libya. Otherwise, there are small pockets with a score of three scattered throughout the region. These fall primarily along the coast but also occur in some portions of the interior. Statistically speaking, the only significant areas are in a large swath of eastern Libya, and a small sector of eastern Algeria. As for the second map, there are essentially two risk zones. The first includes scores of 17 and under, while the second includes scores from 81 to 96. This differentiation shows great levels of variation for both zones. Of particular concern are high-risk areas in the east of Libya, with other areas of interest along the Mediterranean coast and on to the Atlantic. The lower risk scores occur in areas where there has been no past activity. Of these, the highest risk areas are again in eastern Libya, but also scattered throughout the Sahara and the southwest corner of Morocco. From a statistical standpoint, small significant clusters near Benghazi, and Darnah, Libya, as well as in Nabeul, Tunis, and Banzart, Tunisia exist.

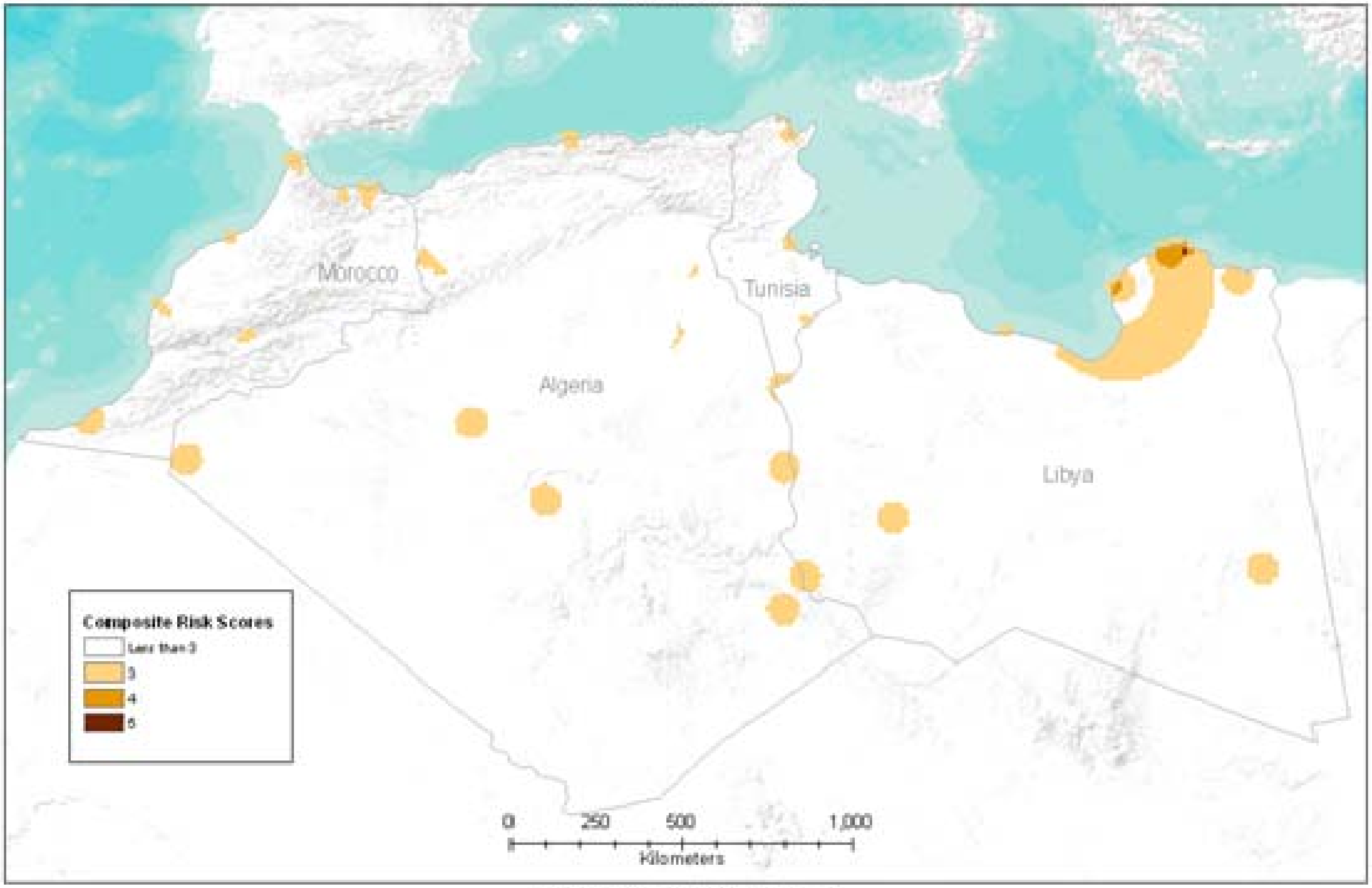
Table 9. March-July Recruit Hometowns and associated risk scores

Hometown	Country	Recruits	Unweighted Risk Score	Weighted Risk Score	KDE Risk Score
Benghazi	Libya	10	4	91	1
Misratah	Libya	1	2	9	0
Aryanah	Tunisia	1	3	88	1
Tetuan	Morocco	1	2	81	1
Darnah	Libya	18	4	89	1

All distances in kilometers. (Derived from Fishman (n.d); OpenFlights (2010); NGA GNS (2010a-d); and IAU WHED (2009))

UNWEIGHTED COMPOSITE RISK

HIGHEST RISK AREAS



COMPILED BY: ISMAEL RODRIGUEZ
DATE: 23 NOVEMBER 2010
DATA SOURCES: ESRI WORLD TERRAIN BASE, ESRI WORLD UN MEMBERSHIP, FISHMAN SINJAR DATA MASTER, NGA GNS COUNTRY FILES, IAU 2009 WORLD HIGHER EDUCATION DATABASE, OPENFLIGHTS AIRPORT AND ROUTE DATA
COORDINATE SYSTEM: WGS-84, CUSTOM AFRICA EQUIDISTANT CONIC

Figure 11. Unweighted Composite Risk

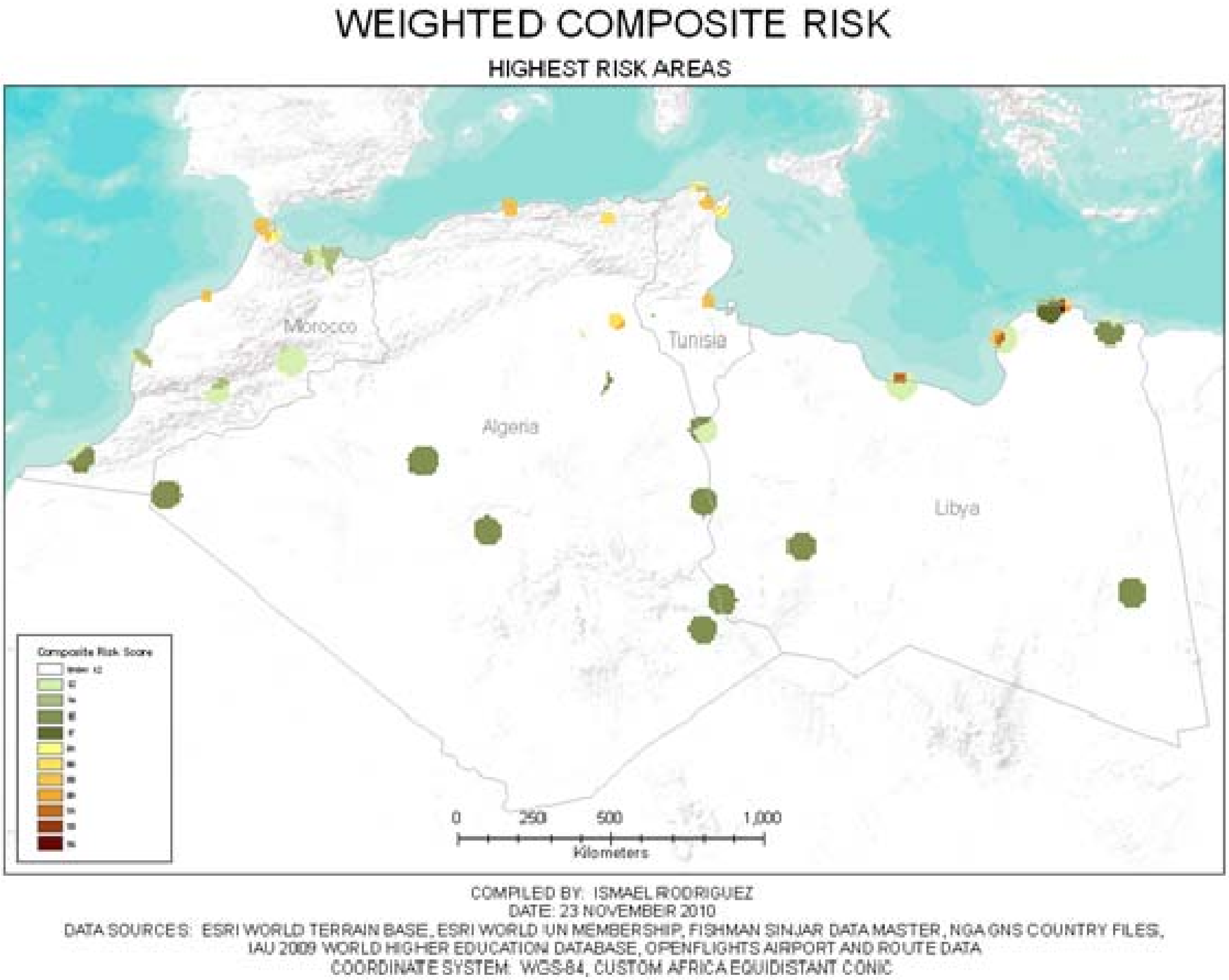


Figure 12. Weighted Composite Risk

4. Interpretation of Results

At first glance, there does appear to be some correlation between high-risk zones and the emergence of the 31 recruits who arrived in Iraq during the second timeframe. While this appears to be a decent sample size, a plot of their hometowns reveals that they came from only five different locales.

Nevertheless, this sets the stage for a comparison between three predictive mapping tools. The availability of temporal data presents an opportunity to test the predictive validity of each map (p. 100). Adapting the process described by Caplan & Kennedy to do just that (p. 101-102), the results of OLS regression analysis, comparing risk to recruitment activity, suggest the unweighted risk map is the best option.

Table 10. Risk Model Comparison

Criteria	Risk Model 1	Risk Model 2	Risk Model 3
Dependent Variable	Recruits	Recruits	Recruits
Independent Variable	Unweighted Score	Weighted Score	KDE Score
IV Probability	0.069	0.463	0.529
R-Squared	0.720	0.190	0.144
Adjusted R-Squared	0.626	-0.0798	-0.141
Corrected Akaike's Information Criterion (AICc)	31.0753	36.381	36.659

(Compiled from OpenGeoDa Regression Results)

The unweighted map performs significantly better than both the weighted map and the basic kernel density map of past activity. In other words, these results indicate that risk mapping may be a better predictive tool than a hot spot map of the same area. That said, while probably quite realistic in terms of data availability, the small sample size does raise the question of the reliability of those results.

The question remains as to whether this technique is valuable for terrorism research or counter-terrorism policy. As this comparison suggests, risk mapping may afford an opportunity for security organizations to depict and track the dynamic interaction between illicit activity and the environment from which it emerges. Thus, in this sense, it could become a worthwhile strategic tool.

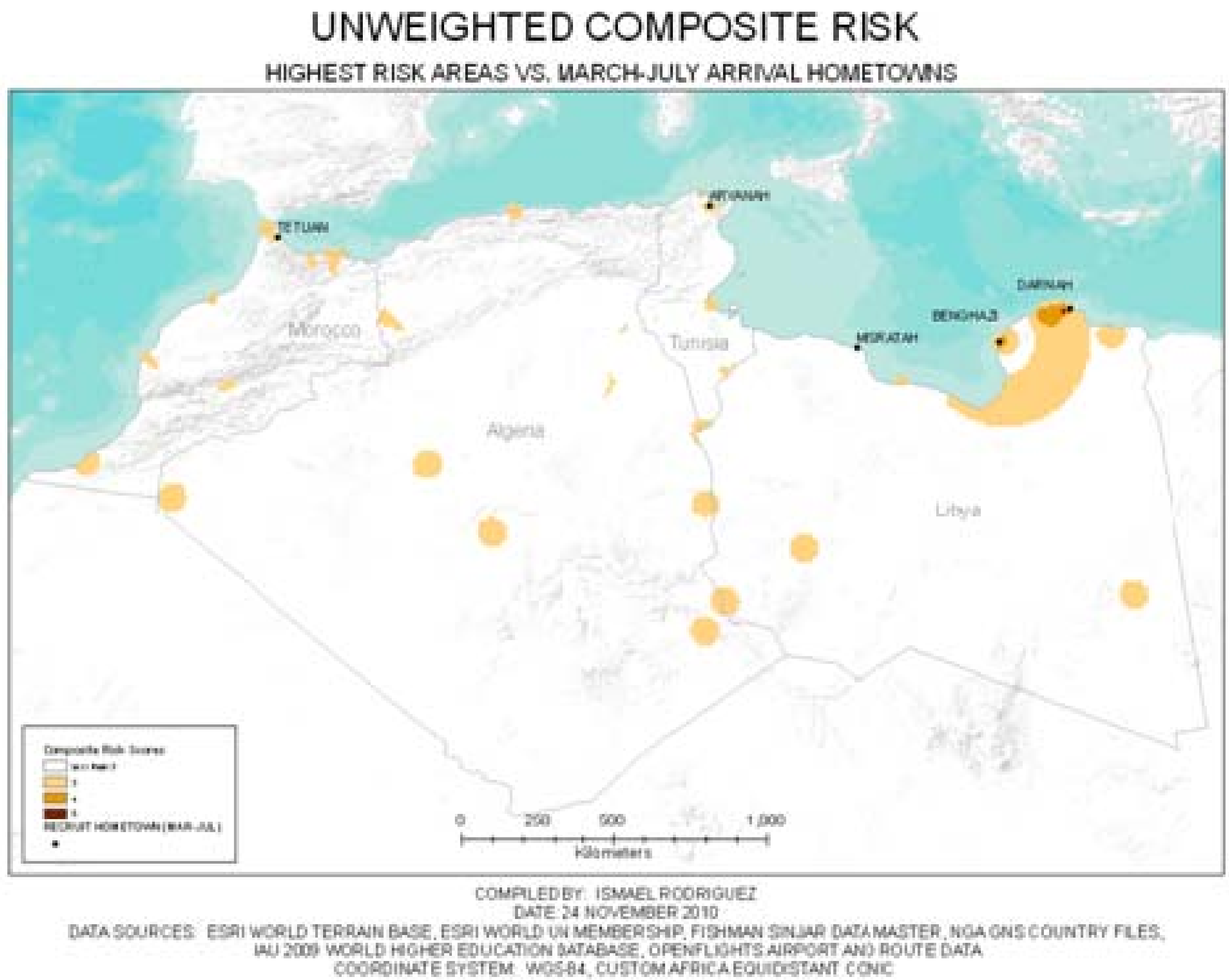


Figure 13. Unweighted Highest Risk Areas vs. March-July Arrival Hometowns

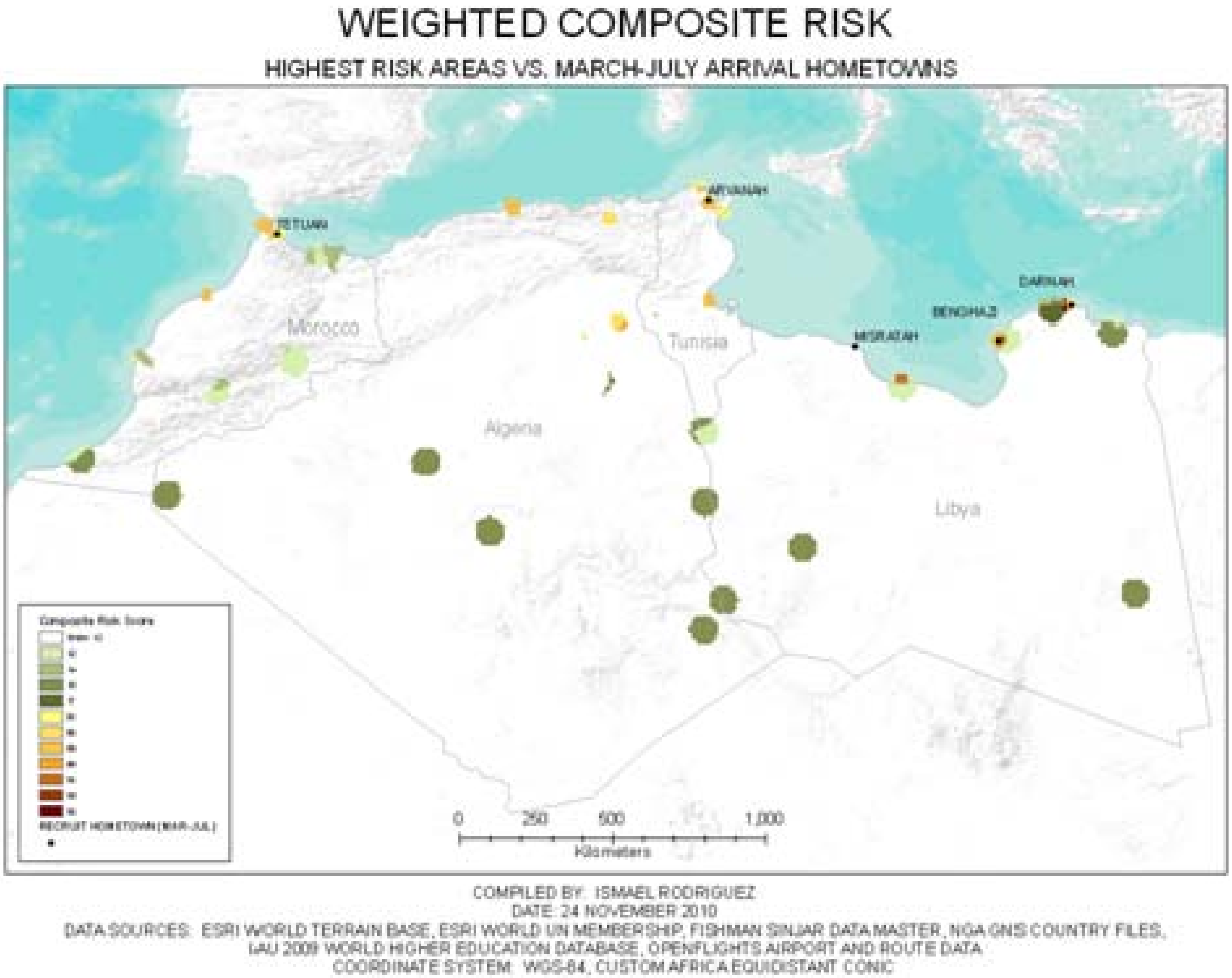


Figure 14. Weighted Highest Risk Areas vs. March-July Arrival Hometowns

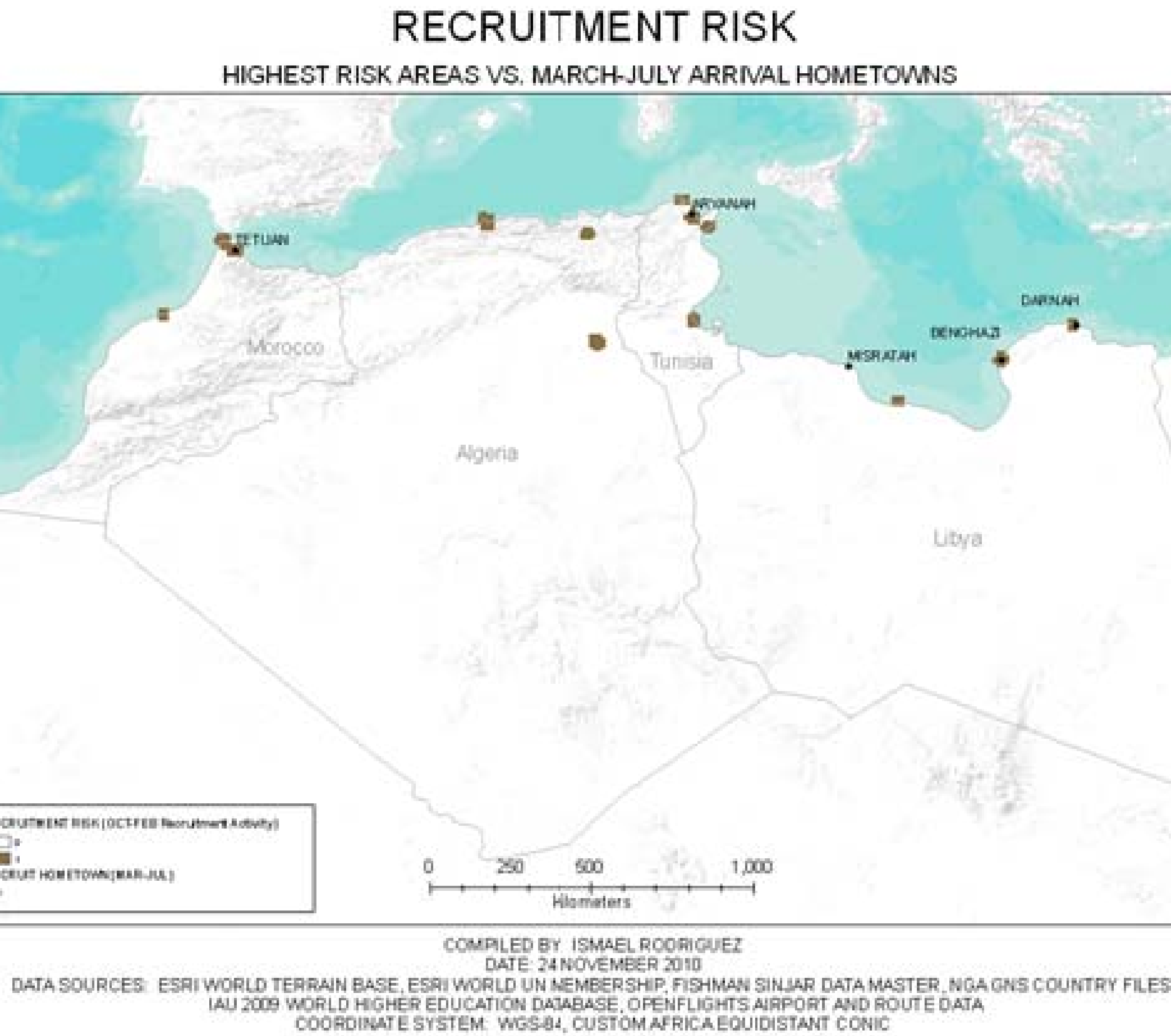


Figure 15. Past Activity Highest Risk Areas vs. March-July Arrival Hometowns

V. CONCLUSION

In the fall of 2006, a man left the world he knew to travel to a distant place. He, and hundreds like him, would eventually pass through Sinjar, a city that he might never have recognized nor might never see again (Fishman, n.d). His final assignment would probably take him hundreds of miles away. Who knows how long that man remained in Iraq, whether he lived or died, whether he failed or succeeded in his mission? Nevertheless, that man went to great lengths to find himself in a distant place on that autumn day.

The records to which this recruit contributed offer only a glimpse into the lives of these recruits. While much has been made of what the records revealed, perhaps more should be made of what the records do not expose. Yes, most were quite detailed in listing hometowns, next of kin, occupational skills and the like. Still, many others listed little more than a name and a country of origin. That said, it is remarkable to see what additional information might be gleaned. The crossroads of social movement theory, criminology, and spatial statistics offer a unique vantage point with which to examine the patterns that did emerge. In other words, these findings correspond relatively well with the theoretical framework of social movement theory. In particular, the study reinforces the importance of repression and resources to the sustainment of a movement interested in terrorism. While the results emerge from a small sample set, they suggest that access to infrastructure in addition to distance from the watchful eye of repressive regimes factored into these observed recruitment patterns.

A. FUTURE RESEARCH

While the theory and processes discussed in this study appear sound, the data preparation still has room for refinement. Surprisingly, the results suggest population density did not factor into the explanation of recruitment patterns. A reliable set of population data remains elusive for this study. Demographic databases are not easy to come by in the countries of North Africa.

Nevertheless, the consolidation of available government population data would make spatial analysis more meaningful, allowing for a more authoritative examination of recruitment rates normalized for population. Beyond this preferred solution, a population model, such as the Oakridge National Laboratory Land Scan population dataset (Oak Ridge National Laboratory, n.d) could provide a valid proxy.

While population and demographic data do not factor into the final results, distances are quite significant. However, these distances are estimates at best. While obviously useful, Euclidean distances are not nearly as realistic as road distances. However, to calculate road distances requires the establishment of a functional road network dataset. Moreover, a cursory glance at recruit hometowns, overlaid on a road map of North Africa (ESRI, 2009e), suggests that proximity to primary road routes might also factor into recruitment patterns. Furthermore, an examination of commercial bus stations throughout the region might also yield useful results. Finally, future study could expand proximity analysis to other regions within the dataset. Of the different possibilities, the Arabian Peninsula would be an obvious choice.

In terms of difficulty, neither transportation infrastructure nor population characteristics should generate many problems for future research. On the contrary, identifying and mapping the spatial dimensions of social networks presents a significant challenge. Such an effort would require a level of detail, experience, and understanding not readily accessible to an outside researcher. However, this type of information could emerge through cooperation with local security organizations. Moreover, such an effort could also aid the Consolidation social information, such as known locations of radical activity, offering yet another angle from which to measure proximity.

In all, the Sinjar records are a fascinating dataset with much room for further study. The real test for this study would be to transfer the theory and techniques to an altogether different dataset. Using activity at a given location as a dependent variable opens an array of proximity, demographic, and economic data to set as independent variables. Although proximity variables may form solid explanations, data from other regions of the world may offer better demographic or economic details at the local level.

Still, the Sinjar dataset does not offer any clear insight into the motivation of the recruits. This study does not attempt to uncover the roots of terrorism in North Africa. Instead, its aim is rather to identify where conditions are most conducive to recruitment. Metaphorically, if terrorist recruitment does have roots, then those roots would require certain conditions to flourish. By identifying what those conditions could be, it is then possible to search the region for other similar places. Just as certain crops thrive in the right mix of soil, nutrients, and climate, terrorist recruitment appears to take hold in certain places. While not entirely conclusive, this study offers an idea of what those conditions might be. In any case, future research and geospatial analysis could do much to refine this understanding.

Figure 16. North African Recruit Hometowns as Compared to Road Network

B. IMPLICATIONS

The Sinjar records are only one part of the story. More important is the impact that new techniques might have on the American military and its allies around the world. Rather than emphasizing how a convergence of theory, data, and techniques could explain past activity, this study should be seen as a viable framework for approaching complex problems of the human environment.

Maps can and should be part of this approach. There has long been a tradition of map making and map interpretation in the American Army. Over the past decade, the military has taken great strides to incorporate cutting edge technology into intelligence, operations, and planning processes. Despite this investment in time, infrastructure, and talent, there are still significant deficiencies. Michael Flynn, Matthew Pottinger, and Paul Batchelor (2010) underscore these issues, noting:

> Having focused the overwhelming majority of its collection efforts and analytical brainpower on insurgent groups, the vast intelligence apparatus is unable to answer fundamental questions about the environment in which U.S. and allied forces operate and the people they seek to persuade…U.S. intelligence officers and analysts can do little but shrug in response to high level decision-makers seeking the knowledge, analysis, and information they need to wage a successful insurgency. (p. 7)

In other words, the human environment remains an elusive, often uncharted, realm. To overcome these obstacles, the military should actively seek innovative ways to use the tools it already has available. GIS may not be a silver bullet, but it is a proven tool, used regularly in the academic, commercial, and government sectors to make sense of all variety of complex issues. Spatially integrated social sciences and the refined spatial analysis techniques of the crime analysis community offer the Army a solid foundation upon which to build. Unfortunately, this is easier said than done. In particular, the Army would need to decide who has responsibility for implementing these processes.

GIS have now had a role in both military planning and military intelligence analysis for many years. Well before this current usage, staffs have relied on paper maps and acetate overlays to analyze terrain, determine possible enemy routes, and decipher complex urban settings. In other words, geospatial analysis has long had a home in the American Army. That said, within the Army there is a somewhat disjointed approach to geospatial intelligence (GEOINT), and little discussion of advanced spatial analysis responsibilities. In effect, Army GEOINT is a collaborative effort between intelligence and engineer functions. Even though intelligence organizations share responsibility for spatial analysis, topographic engineers have responsibility for the provision of spatial data while intelligence organizations have responsibility for providing imagery (U.S. Army, 2008a, p. 1–25). More specifically, Army topographic engineering doctrine explicitly emphasizes the engineering community's responsibility to describe physical terrain (U.S. Army, 2010, p. 1–8). What is largely missing from both sets of doctrine is an explicit delineation of responsibility for human spatial data and analysis. However, based on the topographic doctrine, the engineering community should have some responsibility to assist the intelligence community in compiling and analyzing that information (p. 1-9). Furthermore, despite the introduction of GIS capability to the intelligence community, the engineering community is home to the Army's designated GIS specialists. These technicians have a broad array of responsibilities, primarily geared to physical terrain analysis and map production (pp. 2–28–2–29). Still, geospatial engineers have the best skills to provide the analytical support envisioned by this study. Unfortunately, with their many other responsibilities, it is quite possible that this risk terrain analysis could get lost in the shuffle. Moreover, barring specific doctrinal guidance, there is a distinct chance that spatial analysis of human social, political, or economic patterns could become marginalized within the broader Army GEOINT community.

The introduction of new spatial analysis techniques to the Army poses its own problems. The determination of how best to approach the training, organization, and implementation of advanced geospatial analysis techniques is a legitimate area of study in its own right. That said, there are already organizations within the Army and the broader Department of Defense that could easily adopt these techniques. For instance, with little additional modification, organizations such as the Division-level GEOINT Cell would have the capacity to adopt these methods (Cromer, McDonough, & Conway, 2009, pp. 10–12). Thus, in the short term, these techniques could readily take root. However, over the long term, the Army should consider how best to disseminate these new techniques to its intelligence Soldiers. Fortunately, the Army's Foundry Intelligence Training Program provides a venue with which to offer this type of training (p. 16). Created in 2006, this program gives intelligence organizations the opportunity to train with national level intelligence organizations (U.S. Army, 2008b). The National Geospatial-Intelligence College (NGC) has taken a prominent role in offering GIS instruction to the military. Of the courses offered by the school's mobile training teams, the most popular have been Geospatial Information and Services 101 and Geospatial Information and Services for the Warrior (NGA, 2008, p. 8). Nevertheless, there is room for improvement. The Army should recognize the advancements in geospatial analysis taking placing outside of the realm of military operations. As such, the Army should consider building immersion training programs within the commercial, academic, and law enforcement sectors to improve geospatial analysis capabilities. At the tactical and operational level, there has long been an overarching, often elusive, goal to predict when and where enemy actions might occur. U.S. Army (2008a) Intelligence Capstone Doctrine, Field Manual 2-0,[12] sums up this tendency, noting:

> [o]ne of the most significant contributions that intelligence personnel can accomplish is to accurately predict future enemy events. Although this is an extremely difficult task, predictive

[12] The Army published a new edition of FM 2-0 in 2010. However, unlike the 2008 edition, this manual is not available for public release.

> intelligence enables the command and staff to anticipate key enemy events or reactions and develop corresponding plans or counteractions. (p. 1–2)

However, given the modern operational environment, it is little wonder that this goal has been so hard to achieve. More specifically, as Walter Perry and John Gordon (2008) of the RAND National Defense Research Institute argue, current operations are dynamic actions between enemy and friendly actions which cannot be forecast using the predictive techniques of conventional military operations (p. 31).

From a tactical and operational perspective, there is much to learn from the tenets of environmental crime analysis, and the specific techniques offered in Risk Terrain Modeling (RTM). As Caplan and Kennedy (2010) contend, this technique could support decision-making, and more specifically, resource management, while also providing a mechanism to revise risk assessments over time (pp. 29-30). In terms of difficulty, these techniques would require a fair amount of additional training, but could also yield a refined understanding of any variety of human environments. More importantly, the results would be relatively uncomplicated to decipher and simple to explain.

Comparatively, RTM is a more viable option for a tactical or operational field staff than the more rigorous regression analysis techniques currently available. It is a rare opportunity to establish a new technique for forecasting future activity. Perry and Gordon (2008) suggest:

> Although several predictive methods exist, very few are currently being used in Iraq or Afghanistan…There are several reasons for this: Some of the predictive methods are extremely complex requiring knowledge of sophisticated software packages; some simply do not work in the environment in which they are required to perform some provide information at a level of resolution that is simply too coarse for commanders to take action; and most cannot adapt to rapidly changing enemy tactics. (p. 32)

They go on to list several measures with which to gauge the effectiveness of new prediction tools. Not only should the effort realize that the enemy does not act in a random fashion, but it should also have rigorous means to study clustering within patterns, present a means to adjust for enemy adaptation, adjust for local settings, allow for the inclusion of a unit's local knowledge, be set at an appropriate scale, and be better than the tools already in use (pp. 33-34). While additional proof of concept studies may indeed be in order, the risk modeling approach appears to meet these conditions. Above all, as both this study and the rigorous efforts of Caplan and Kennedy (2010) suggest, it is arguably an improvement upon the techniques currently in use. Still, in the current operation environment, the use of either regression analysis or RTM would require some appreciation for the theoretical roots of insurgency, environmental criminology, and terrorism. Thus, gauging this level of understanding and developing an optimal strategy to improve this familiarity presents another area of potential research.

Overall, the Sinjar Dataset offers far more than a spatial and temporal snapshot of recruitment activity in the Muslim world. It is by no means perfect, but it offers a comprehensive base of information with which to build upon. In the end, this study indicates that when theory is solid, procedures useful and data adequate, it is quite possible to produce relevant analysis.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A

A. ORDINARY LEAST SQUARED REGRESSION

In simple terms, regression is an equation with at least two variables. On one end of the equation is the dependent variable. It is a function of any number of independent variables, which form the other end of the equation. The independent variables, also known as explanatory variables, are quantifiable measurements related to known quantifiable measurements of the dependent variable. The purpose of these measurements is to calculate a formula that explains not only known relationships between variables, but also determines the value of the dependent variable given different values for the independent variable. In other words, the purpose of the calculated formula is prediction. The predictive power depends on the number of measures, in addition to how well the formula fits the given measurements. In the simplest two-variable format, the equation creates a line. The line has two central features, the coefficient of the independent variable, which provides the slope of the line, and the intercept coefficient, which explains where the line would intercept the y axis. However, the line is only as good as its fit. For each given independent variable, the fit is determined by measuring the distance from the line created by the formula and the actual measurement of the dependent variable. The result is the residual. While in a perfect situation, the line would fall exactly along each of the measurements and the residual would be a value of one, in reality the value is a normally a fraction of that amount. The higher the value of that fraction, the better the formula is at modeling the relationship and ultimately predicting additional outcomes (Mitchell, 2005, pp. 212–214). So how does spatial analysis fit into this process? The basic regression process can expand to include more than one independent variable. This expanded process is known as multivariate regression, and is well adapted to spatial analysis. In a spatial process, feature types, whether point, linear, or polygon can have a number of attributes. In addition to the spatial data overlay, these feature types also include a table of associated attributes. These attributes form a readily available pool of variables from which to select a dependent variable, as well as any amount of independent variables (p. 215). In other words, a spatial feature, say a group of cities, may have an associated set of attributes, such as population, number of crimes committed, number of households, or number of businesses. If a hypothesis suggests a relationship between the number of crimes committed as they relate to any or all the other variables, then the table simplifies the process of testing for relationships between the variables. Mitchell warns that regression analysis does not always work within the spatial perspective. For the approach to work, a regression model should accommodate six key assumptions. Not only should the relationship be linear for each of the independent variables, but also the residuals should average zero and vary at a constant rate. Moreover, the residuals should be both randomly spaced and distributed across a normal curve. Finally, the independent variables should not be redundant, displaying a high degree of correlation when compared against one another (p. 217). Fortunately, even if a spatial ordinary least squared regression model does not meet these assumptions, other approaches may still work (p. 218).

APPENDIX B

A. REGRESSION RESULTS CLASSIC OLS MODELS

1. OpenGeoDA OLS Results for Model 1

```
Regression 5 VARIABLE
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : Fishman_Variables
Dependent Variable  :     RECRUITS  Number of Observations:   27
Mean dependent var  :      5.66667  Number of Variables   :    5
S.D. dependent var  :       10.378  Degrees of Freedom    :   22

R-squared           :     0.288635  F-statistic           :      2.23162
Adjusted R-squared  :     0.159296  Prob(F-statistic)     :    0.0985533
Sum squared residual:      2068.65  Log likelihood        :     -96.8853
Sigma-square        :      94.0295  Akaike info criterion :      203.771
S.E. of regression  :      9.69688  Schwarz criterion     :       210.25
Sigma-square ML     :      76.6166
S.E of regression ML:      8.75309
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT       1.098267       3.741516      0.2935352    0.7718636
     POP_DEN   0.0006582144   0.0009631886      0.6833702    0.5015073
    CAP_DIST   1.817281e-005  1.079833e-005      1.682928    0.1065288
   UNIV_DIST   2.994802e-005  3.703041e-005     0.8087413    0.4273238
    DOM_DIST  -6.687027e-005  5.060014e-005     -1.321543    0.1998996
-----------------------------------------------------------------------


DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   5.419296
TEST ON NORMALITY OF ERRORS
TEST                  DF         VALUE             PROB
Jarque-Bera            2           34.4405         0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF         VALUE             PROB
Breusch-Pagan test     4          69.17065         0.0000000
Koenker-Bassett test   4          18.99514         0.0007877
SPECIFICATION ROBUST TEST
TEST                  DF         VALUE             PROB
White                 14          23.12655         0.0582414
```

COEFFICIENTS VARIANCE MATRIX

CONSTANT	POP_DEN	CAP_DIST	UNIV_DIST	DOM_DIST
13.998944	-0.002261	-0.000025	0.000041	-0.000075
-0.002261	0.000001	0.000000	-0.000000	0.000000
-0.000025	0.000000	0.000000	-0.000000	-0.000000
0.000041	-0.000000	-0.000000	0.000000	-0.000000
-0.000075	0.000000	-0.000000	-0.000000	0.000000

OBS	RECRUITS	PREDICTED	RESIDUAL
1	1.00000	-0.08217	1.08217
2	1.00000	7.44236	-6.44236
3	5.00000	2.40162	2.59838
4	5.00000	4.90972	0.09028
5	8.00000	9.18460	-1.18460
6	1.00000	6.97445	-5.97445
7	1.00000	-1.85843	2.85843
8	1.00000	7.09947	-6.09947
9	1.00000	-1.51632	2.51632
10	1.00000	20.68928	-19.68928
11	7.00000	1.28160	5.71840
12	3.00000	4.16079	-1.16079
13	17.00000	3.50215	13.49785
14	20.00000	11.71996	8.28004
15	2.00000	6.60239	-4.60239
16	2.00000	0.15238	1.84762
17	1.00000	6.80848	-5.80848
18	1.00000	2.76035	-1.76035
19	53.00000	20.52202	32.47798
20	2.00000	1.80194	0.19806
21	2.00000	4.97420	-2.97420
22	5.00000	12.52920	-7.52920
23	4.00000	8.54441	-4.54441
24	1.00000	4.74770	-3.74770
25	1.00000	5.74563	-4.74563
26	6.00000	1.73251	4.26749
27	1.00000	0.16972	0.83028

========================= END OF REPORT ==============================

2. OpenGeoDA OLS Results for Model 2

Regression
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION

Data set	: **Fishman_Variables**			
Dependent Variable	: **RECRUITS**	Number of Observations	:	27
Mean dependent var	: 5.66667	Number of Variables	:	6
S.D. dependent var	: 10.378	Degrees of Freedom	:	21
R-squared	: 0.326656	F-statistic	:	2.03753
Adjusted R-squared	: 0.166336	Prob(F-statistic)	:	0.11462
Sum squared residual	: 1958.08	Log likelihood	:	-96.1438
Sigma-square	: 93.2421	Akaike info criterion	:	204.288
S.E. of regression	: 9.65619	Schwarz criterion	:	212.063
Sigma-square ML	: 72.5216			
S.E of regression ML	: 8.51596			

Variable	**Coefficient**	**Std.Error**	**t-Statistic**	**Probability**
CONSTANT	3.491265	4.325616	0.8071139	0.4286481
POP_DEN	0.0003117907	0.00101053	0.3085418	0.7607099
CAP_DIST	2.195023e-005	1.129871e-005	1.942721	0.0655786
AIR_DIST	-2.132572e-005	1.958396e-005	-1.088938	0.2885206
UNIV_DIST	3.364659e-005	3.703113e-005	0.9086028	0.3738697
DOM_DIST	-7.547534e-005	5.100371e-005	-1.479801	0.1537728

DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER 6.105510
TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	33.72598	0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	76.71245	0.0000000
Koenker-Bassett test	5	20.53847	0.0009899

SPECIFICATION ROBUST TEST

TEST	DF	VALUE	PROB
White	20	24.76283	0.2106557

COEFFICIENTS VARIANCE MATRIX

CONSTANT	POP_DEN	CAP_DIST	AIR_DIST	UNIV_DIST	DOM_DIST
18.710953	-0.002942	-0.000017	-0.000043	0.000048	-0.000092
-0.002942	0.000001	0.000000	0.000000	-0.000000	0.000000
-0.000017	0.000000	0.000000	-0.000000	-0.000000	-0.000000
-0.000043	0.000000	-0.000000	0.000000	-0.000000	0.000000
0.000048	-0.000000	-0.000000	-0.000000	0.000000	-0.000000
-0.000092	0.000000	-0.000000	0.000000	-0.000000	0.000000

OBS	RECRUITS	PREDICTED	RESIDUAL
1	1.00000	0.99512	0.00488
2	1.00000	10.72631	-9.72631
3	5.00000	3.66257	1.33743
4	5.00000	4.19900	0.80100
5	8.00000	5.14422	2.85578
6	1.00000	8.27293	-7.27293
7	1.00000	-1.21253	2.21253
8	1.00000	6.25659	-5.25659
9	1.00000	-2.68967	3.68967
10	1.00000	21.93955	-20.93955
11	7.00000	2.95568	4.04432
12	3.00000	3.23854	-0.23854
13	17.00000	4.15397	12.84603
14	20.00000	15.97263	4.02737
15	2.00000	3.40944	-1.40944
16	2.00000	1.22369	0.77631
17	1.00000	3.20775	-2.20775
18	1.00000	4.28241	-3.28241
19	53.00000	21.64407	31.35593
20	2.00000	3.40349	-1.40349
21	2.00000	3.26063	-1.26063
22	5.00000	9.18132	-4.18132
23	4.00000	9.72443	-5.72443
24	1.00000	3.43473	-2.43473
25	1.00000	5.42478	-4.42478
26	6.00000	-0.07642	6.07642
27	1.00000	1.26475	-0.26475

========================= END OF REPORT ==============================

3. OpenGeoDa Results for OLS Model 3

Regression
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION

```
Data set            : Fishman_Variables_17NOV
Dependent Variable  :    RECRUITS  Number of Observations:   27
Mean dependent var  :     5.66667  Number of Variables   :    4
S.D. dependent var  :      10.378  Degrees of Freedom    :   23

R-squared           :    0.273535  F-statistic           :     2.88672
Adjusted R-squared  :    0.178779  Prob(F-statistic)     :   0.0574891
Sum squared residual:     2112.56  Log likelihood        :    -97.1689
Sigma-square        :     91.8504  Akaike info criterion :     202.338
S.E. of regression  :     9.58386  Schwarz criterion     :     207.521
Sigma-square ML     :      78.243
S.E of regression ML:     8.84551
```

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	2.702704	2.87923	0.9386899	0.3576441
CAP_DIST	1.541557e-005	9.899479e-006	1.557211	0.1330749
UNIV_DIST	3.350288e-005	3.62359e-005	0.924577	0.3647852
DOM_DIST	-7.274434e-005	4.928351e-005	-1.476038	0.1534942

DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER 4.693555
TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	34.9472	0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	3	67.75884	0.0000000
Koenker-Bassett test	3	18.66329	0.0003209

SPECIFICATION ROBUST TEST

TEST	DF	VALUE	PROB
White	9	22.43691	0.0075928

COEFFICIENTS VARIANCE MATRIX

CONSTANT	CAP_DIST	UNIV_DIST	DOM_DIST
8.289965	-0.000015	0.000028	-0.000054
-0.000015	0.000000	-0.000000	-0.000000
0.000028	-0.000000	0.000000	-0.000000
-0.000054	-0.000000	-0.000000	0.000000

OBS	RECRUITS	PREDICTED	RESIDUAL
1	1.00000	1.11792	-0.11792
2	1.00000	7.60928	-6.60928
3	5.00000	2.20444	2.79556
4	5.00000	1.47656	3.52344
5	8.00000	9.25035	-1.25035
6	1.00000	7.65000	-6.65000
7	1.00000	-0.95447	1.95447
8	1.00000	2.74862	-1.74862
9	1.00000	-0.96075	1.96075
10	1.00000	20.41465	-19.41465
11	7.00000	2.01081	4.98919
12	3.00000	5.12609	-2.12609
13	17.00000	2.21468	14.78532
14	20.00000	11.35873	8.64127
15	2.00000	6.98387	-4.98387
16	2.00000	1.35545	0.64455
17	1.00000	7.03966	-6.03966
18	1.00000	2.78655	-1.78655
19	53.00000	20.20714	32.79286
20	2.00000	2.32761	-0.32761
21	2.00000	5.24858	-3.24858
22	5.00000	13.69450	-8.69450
23	4.00000	7.86702	-3.86702
24	1.00000	5.53576	-4.53576
25	1.00000	5.04286	-4.04286
26	6.00000	2.26599	3.73401
27	1.00000	1.37812	-0.37812

========================= END OF REPORT ==============================

4. OpenGeoDa Results for Model 4

```
Regression
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : Fishman_Variables_17NOV
Dependent Variable  :    RECRUITS  Number of Observations:   27
Mean dependent var  :     5.66667  Number of Variables   :    5
S.D. dependent var  :      10.378  Degrees of Freedom    :   22

R-squared           :    0.323604  F-statistic           :     2.63133
Adjusted R-squared  :    0.200623  Prob(F-statistic)     :   0.0618221
Sum squared residual:     1966.96  Log likelihood        :    -96.2048
Sigma-square        :     89.4073  Akaike info criterion :      202.41
S.E. of regression  :     9.45554  Schwarz criterion     :     208.889
Sigma-square ML     :     72.8504
S.E of regression ML:     8.53524
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT      4.389406       3.13312       1.40097    0.1751717
    CAP_DIST  2.111055e-005   1.073819e-005     1.965932    0.0620519
    AIR_DIST -2.322798e-005   1.820192e-005    -1.276128    0.2152189
   UNIV_DIST  3.549352e-005   3.578475e-005    0.9918618    0.3320495
    DOM_DIST -7.874965e-005   4.885083e-005    -1.612043    0.1212071
-----------------------------------------------------------------------


DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   5.306456
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2           33.71329        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test     4           76.25531        0.0000000
Koenker-Bassett test   4           20.44218        0.0004084
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                 14           24.47792        0.0400852
```

COEFFICIENTS VARIANCE MATRIX

CONSTANT	CAP_DIST	AIR_DIST	UNIV_DIST	DOM_DIST
9.816440	-0.000008	-0.000024	0.000030	-0.000059
-0.000008	0.000000	-0.000000	-0.000000	-0.000000
-0.000024	-0.000000	0.000000	-0.000000	0.000000
0.000030	-0.000000	-0.000000	0.000000	-0.000000
-0.000059	-0.000000	0.000000	-0.000000	0.000000

OBS	RECRUITS	PREDICTED	RESIDUAL
1	1.00000	1.60335	-0.60335
2	1.00000	11.09047	-10.09047
3	5.00000	3.69090	1.30910
4	5.00000	2.67053	2.32947
5	8.00000	4.81188	3.18812
6	1.00000	8.67704	-7.67704
7	1.00000	-0.76916	1.76916
8	1.00000	4.32470	-3.32470
9	1.00000	-2.55724	3.55724
10	1.00000	21.93388	-20.93388
11	7.00000	3.41620	3.58380
12	3.00000	3.56821	-0.56821
13	17.00000	3.66270	13.33730
14	20.00000	16.19782	3.80218
15	2.00000	3.28742	-1.28742
16	2.00000	1.83266	0.16734
17	1.00000	2.98521	-1.98521
18	1.00000	4.42936	-3.42936
19	53.00000	21.60978	31.39022
20	2.00000	3.77067	-1.77067
21	2.00000	3.22487	-1.22487
22	5.00000	9.37997	-4.37997
23	4.00000	9.54061	-5.54061
24	1.00000	3.65392	-2.65392
25	1.00000	5.09625	-4.09625
26	6.00000	-0.01012	6.01012
27	1.00000	1.87811	-0.87811

========================= END OF REPORT ==============================

5. OpenGeoDa OLS Model 4 Residual Results

```
Regression
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : Fishman_Variables_17NOV
Dependent Variable  :    RECRUITS  Number of Observations:   27
Mean dependent var  :     5.66667  Number of Variables   :    5
S.D. dependent var  :      10.378  Degrees of Freedom    :   22

R-squared           :    0.323604  F-statistic           :     2.63133
Adjusted R-squared  :    0.200623  Prob(F-statistic)     :   0.0618221
Sum squared residual:     1966.96  Log likelihood        :    -96.2048
Sigma-square        :     89.4073  Akaike info criterion :      202.41
S.E. of regression  :     9.45554  Schwarz criterion     :     208.889
Sigma-square ML     :     72.8504
S.E of regression ML:     8.53524
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT      4.389406        3.13312        1.40097    0.1751717
    CAP_DIST  2.111055e-005   1.073819e-005      1.965932    0.0620519
    AIR_DIST -2.322798e-005   1.820192e-005     -1.276128    0.2152189
   UNIV_DIST  3.549352e-005   3.578475e-005     0.9918618    0.3320495
    DOM_DIST -7.874965e-005   4.885083e-005     -1.612043    0.1212071
-----------------------------------------------------------------------

DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   5.306456
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2           33.71329        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test     4           76.25531        0.0000000
Koenker-Bassett test   4           20.44218        0.0004084
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                 14           24.47792        0.0400852
```

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX : Fishman_Variables_17NOV_6V.gwt (row-standardized weights)

TEST	**MI/DF**	**VALUE**	**PROB**
Moran's I (error)	-0.148314	-1.0860402	0.2774613
Lagrange Multiplier (lag)	1	1.4197296	0.2334479
Robust LM (lag)	1	0.0045186	0.9464061
Lagrange Multiplier (error)	1	1.7817559	0.1819339
Robust LM (error)	1	0.3665449	0.5448936
Lagrange Multiplier (SARMA)	2	1.7862745	0.4093694

COEFFICIENTS VARIANCE MATRIX

CONSTANT	CAP_DIST	AIR_DIST	UNIV_DIST	DOM_DIST
9.816440	-0.000008	-0.000024	0.000030	-0.000059
-0.000008	0.000000	-0.000000	-0.000000	-0.000000
-0.000024	-0.000000	0.000000	-0.000000	0.000000
0.000030	-0.000000	-0.000000	0.000000	-0.000000
-0.000059	-0.000000	0.000000	-0.000000	0.000000

OBS	RECRUITS	PREDICTED	RESIDUAL
1	1.00000	1.60335	-0.60335
2	1.00000	11.09047	-10.09047
3	5.00000	3.69090	1.30910
4	5.00000	2.67053	2.32947
5	8.00000	4.81188	3.18812
6	1.00000	8.67704	-7.67704
7	1.00000	-0.76916	1.76916
8	1.00000	4.32470	-3.32470
9	1.00000	-2.55724	3.55724
10	1.00000	21.93388	-20.93388
11	7.00000	3.41620	3.58380
12	3.00000	3.56821	-0.56821
13	17.00000	3.66270	13.33730
14	20.00000	16.19782	3.80218
15	2.00000	3.28742	-1.28742
16	2.00000	1.83266	0.16734
17	1.00000	2.98521	-1.98521
18	1.00000	4.42936	-3.42936
19	53.00000	21.60978	31.39022
20	2.00000	3.77067	-1.77067
21	2.00000	3.22487	-1.22487
22	5.00000	9.37997	-4.37997
23	4.00000	9.54061	-5.54061
24	1.00000	3.65392	-2.65392
25	1.00000	5.09625	-4.09625
26	6.00000	-0.01012	6.01012
27	1.00000	1.87811	-0.87811

========================= END OF REPORT ==============================

APPENDIX C

A. REGRESSION RESULTS SPATIALLY LAGGED OLS MODELS

1. OpenGeoDa OLS Lagged Results Model 1

```
Regression
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set             : Fishman_Variables_17NOV_Theissen
Spatial Weight       : Fishman_Variables_17NOV_Theissen_Queen.gal
Dependent Variable   :     RECRUITS  Number of Observations:   27
Mean dependent var   :      5.66667  Number of Variables   :    6
S.D. dependent var   :       10.378  Degrees of Freedom    :   21
Lag coeff.   (Rho)   :    -0.811581

R-squared            :     0.571068  Log likelihood        :    -92.3125
Sq. Correlation      : -             Akaike info criterion :     196.625
Sigma-square         :      46.1976  Schwarz criterion     :       204.4
S.E of regression    :      6.79688

-----------------------------------------------------------------------
    Variable    Coefficient      Std.Error    z-value      Probability
-----------------------------------------------------------------------
  W_RECRUITS    -0.8115812       0.2057606       -3.944297    0.0000801
    CONSTANT      2.707795        2.636574        1.027013    0.3044145
     POP_DEN  0.0006620405   0.0006774379       0.9772711    0.3284349
    CAP_DIST  2.495764e-005    7.835511e-006       3.185197    0.0014467
   UNIV_DIST  4.452768e-005    2.596984e-005       1.714592    0.0864199
    DOM_DIST -5.678624e-005    3.567087e-005       -1.59195    0.1113960
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE          PROB
Breusch-Pagan test                       4        33.79067      0.0000008

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL      LAG       DEPENDENCE       FOR      WEIGHT      MATRIX      :
Fishman_Variables_17NOV_Theissen_Queen.gal
TEST                                     DF        VALUE        PROB
Likelihood Ratio Test                    1        9.145596      0.0024932
```

COEFFICIENTS VARIANCE MATRIX

CONSTANT	POP_DEN	CAP_DIST	UNIV_DIST	DOM_DIST	W_RECRUITS
6.951520	-0.001096	-0.000012	0.000021	-0.000036	-0.055863
-0.001096	0.000000	0.000000	-0.000000	0.000000	-0.000011
-0.000012	0.000000	0.000000	-0.000000	0.000000	-0.000000
0.000021	-0.000000	-0.000000	0.000000	-0.000000	-0.000000
-0.000036	0.000000	0.000000	-0.000000	0.000000	-0.000001
-0.055863	-0.000011	-0.000000	-0.000000	-0.000001	0.042337

OBS	RECRUITS	PREDICTED	RESIDUAL	PRED ERROR
1	1	0.02150	0.07952	0.97850
2	1	10.35413	-5.26235	-9.35413
3	5	2.75402	3.34429	2.24598
4	5	1.52242	-0.63582	3.47758
5	8	10.37564	-5.50244	-2.37564
6	1	8.61856	-6.81497	-7.61856
7	1	-2.11397	2.61006	3.11397
8	1	8.65408	-6.24813	-7.65408
9	1	-3.17788	0.91373	4.17788
10	1	22.58072	-10.58379	-21.58072
11	7	0.88980	5.41079	6.11020
12	3	0.94641	1.36565	2.05359
13	17	-0.68053	13.07538	17.68053
14	20	9.37499	4.32256	10.62501
15	2	5.93973	-6.28211	-3.93973
16	2	2.86493	-0.73486	-0.86493
17	1	13.58598	1.76321	-12.58598
18	1	4.02883	-1.11500	-3.02883
19	53	20.76338	22.58192	32.23662
20	2	0.19690	-0.09453	1.80310
21	2	8.24403	3.15713	-6.24403
22	5	14.51103	-8.13131	-9.51103
23	4	6.21160	0.56856	-2.21160
24	1	3.52910	-4.51577	-2.52910
25	1	8.57847	-7.87541	-7.57847
26	6	0.51302	6.05031	5.48698
27	1	1.65795	-1.44661	-0.65795

========================= END OF REPORT ==============================

2. OpenGeoDa OLS Lagged Results Model 2

Regression

SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION

```
Data set            : Fishman_Variables_17NOV_Theissen
Spatial Weight      : Fishman_Variables_17NOV_Theissen_Queen.gal
Dependent Variable  :     RECRUITS  Number of Observations:   27
Mean dependent var  :      5.66667  Number of Variables   :    7
S.D. dependent var  :       10.378  Degrees of Freedom    :   20
Lag coeff.   (Rho)  :    -0.798175

R-squared           :     0.590196  Log likelihood        :    -91.6184
Sq. Correlation     : -             Akaike info criterion :     197.237
Sigma-square        :      44.1374  Schwarz criterion     :     206.308
S.E of regression   :       6.6436
```

Variable	Coefficient	Std.Error	z-value	Probability
W_RECRUITS	-0.7981749	0.2061277	-3.872235	0.0001079
CONSTANT	4.488103	3.004035	1.494025	0.1351691
POP_DEN	0.000400401	0.0006965415	0.5748415	0.5653984
CAP_DIST	2.769781e-005	8.006631e-006	3.459359	0.0005416
UNIV_DIST	4.707955e-005	2.557368e-005	1.840938	0.0656306
DOM_DIST	-6.34503e-005	3.534337e-005	-1.795254	0.0726132
AIR_DIST	-1.610254e-005	1.350601e-005	-1.19225	0.2331632

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	5	35.49057	0.0000012

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : **Fishman_Variables_17NOV_Theissen_Queen.gal**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	9.05071	0.0026259

COEFFICIENTS VARIANCE MATRIX

CONSTANT	POP_DEN	CAP_DIST	UNIV_DIST	DOM_DIST	AIR_DIST
9.024225	-0.001375	-0.000007	0.000024	-0.000042	-0.000021
-0.001375	0.000000	0.000000	-0.000000	0.000000	0.000000
-0.000007	0.000000	0.000000	-0.000000	-0.000000	-0.000000
0.000024	-0.000000	-0.000000	0.000000	-0.000000	-0.000000
-0.000042	0.000000	-0.000000	-0.000000	0.000000	0.000000
-0.000021	0.000000	-0.000000	-0.000000	0.000000	0.000000
-0.084272	-0.000009	-0.000000	-0.000000	-0.000001	0.000000

W_RECRUITS
-0.084272
-0.000009
-0.000000
-0.000000
-0.000001
0.000000
0.042489

OBS	RECRUITS	PREDICTED	RESIDUAL	PRED ERROR
1	1	0.60070	-0.71735	0.39930
2	1	13.53857	-7.76148	-12.53857
3	5	3.14750	2.37985	1.85250
4	5	0.00454	-0.08718	4.99546
5	8	6.97853	-2.38033	1.02147
6	1	10.29836	-7.78153	-9.29836
7	1	-2.02128	2.12646	3.02128
8	1	8.16538	-5.60923	-7.16538
9	1	-4.18950	1.82616	5.18950
10	1	22.11099	-11.67825	-21.11099
11	7	1.59554	4.15181	5.40446
12	3	-0.23121	2.02028	3.23121
13	17	0.67924	12.59019	16.32076
14	20	13.46658	1.17684	6.53342
15	2	3.67551	-3.84345	-1.67551
16	2	3.26600	-1.50113	-1.26600
17	1	9.66986	4.35696	-8.66986
18	1	4.82165	-2.27494	-3.82165
19	53	21.74490	21.89816	31.25510
20	2	1.39953	-1.29899	0.60047
21	2	7.26958	4.34972	-5.26958
22	5	11.32483	-5.59346	-6.32483
23	4	6.33486	-0.40691	-2.33486
24	1	3.98691	-3.51169	-2.98691
25	1	8.01573	-7.58144	-7.01573
26	6	-1.78248	7.38674	7.78248
27	1	1.98846	-2.23583	-0.98846

========================= END OF REPORT ==============================

3. OpenGeoDa OLS Lagged Results Model 3

Regression

SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION

```
Data set            : Fishman_Variables_17NOV_Theissen
Spatial Weight      : Fishman_Variables_17NOV_Theissen_Queen.gal
Dependent Variable  :    RECRUITS  Number of Observations:   27
Mean dependent var  :     5.66667  Number of Variables   :    5
S.D. dependent var  :      10.378  Degrees of Freedom    :   22
Lag coeff.   (Rho)  :   -0.800716

R-squared           :    0.553734  Log likelihood        :    -92.7838
Sq. Correlation     : -            Akaike info criterion :     195.568
Sigma-square        :     48.0645  Schwarz criterion     :     202.047
S.E of regression   :     6.93286
```

Variable	Coefficient	Std.Error	z-value	Probability
W_RECRUITS	-0.8007157	0.2044389	-3.91665	0.0000898
CONSTANT	4.299885	2.123604	2.024806	0.0428872
CAP_DIST	2.209376e-005	7.364525e-006	3.000025	0.0026997
UNIV_DIST	4.790774e-005	2.623429e-005	1.826149	0.0678277
DOM_DIST	-6.2829e-005	3.580616e-005	-1.754698	0.0793109

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	3	33.7461	0.0000002

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : **Fishman_Variables_17NOV_Theissen_Queen.gal**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	8.770133	0.0030620

COEFFICIENTS VARIANCE MATRIX

CONSTANT	CAP_DIST	UNIV_DIST	DOM_DIST	W_RECRUITS
4.509694	-0.000007	0.000015	-0.000027	-0.084695
-0.000007	0.000000	-0.000000	-0.000000	-0.000000
0.000015	-0.000000	0.000000	-0.000000	-0.000000
-0.000027	-0.000000	-0.000000	0.000000	-0.000001
-0.084695	-0.000000	-0.000000	-0.000001	0.041795

OBS	RECRUITS	PREDICTED	RESIDUAL	PRED ERROR
1	1	1.10266	-1.11403	-0.10266
2	1	10.71397	-5.44603	-9.71397
3	5	2.22892	3.53261	2.77108
4	5	-1.20663	2.82675	6.20663
5	8	9.93733	-5.51076	-1.93733
6	1	8.95978	-7.48314	-7.95978
7	1	-1.42055	1.70424	2.42055
8	1	3.94152	-1.87034	-2.94152
9	1	-2.39387	0.37642	3.39387
10	1	22.53147	-10.42949	-21.53147
11	7	1.44204	4.68151	5.55796
12	3	1.57445	0.36099	1.42555
13	17	-2.62997	14.37589	19.62997
14	20	8.93424	4.73884	11.06576
15	2	5.79246	-6.64329	-3.79246
16	2	3.23682	-1.91026	-1.23682
17	1	15.28869	1.42933	-14.28869
18	1	3.60209	-1.15000	-2.60209
19	53	20.59952	23.03110	32.40048
20	2	2.00776	-0.61930	-0.00776
21	2	8.93982	2.79909	-6.93982
22	5	15.84461	-9.29523	-10.84461
23	4	5.87103	1.18139	-1.87103
24	1	5.24164	-5.29807	-4.24164
25	1	7.68634	-7.12670	-6.68634
26	6	1.30445	5.48991	4.69555
27	1	2.57943	-2.63145	-1.57943

4. OpenGeoDa OLS Lagged Results Model 4

Regression_Queen_Theissen
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION

```
Data set             : Fishman_Variables_17NOV_Theissen
Spatial Weight       : Fishman_Variables_17NOV_Theissen_Queen.gal
Dependent Variable   :    RECRUITS  Number of Observations:   27
Mean dependent var   :     5.66667  Number of Variables   :    6
S.D. dependent var   :      10.378  Degrees of Freedom    :   21
Lag coeff.   (Rho)   :   -0.790067

R-squared            :    0.583752  Log likelihood        :    -91.7826
Sq. Correlation      : -            Akaike info criterion :     195.565
Sigma-square         :     44.8315  Schwarz criterion     :      203.34
S.E of regression    :     6.69563
```

Variable	Coefficient	Std.Error	z-value	Probability
W_RECRUITS	-0.7900671	0.2057052	-3.840774	0.0001227
CONSTANT	5.628776	2.284747	2.463633	0.0137537
CAP_DIST	2.656352e-005	7.800635e-006	3.405303	0.0006610
AIR_DIST	-1.859299e-005	1.295044e-005	-1.435704	0.1510868
UNIV_DIST	4.930959e-005	2.547362e-005	1.935712	0.0529029
DOM_DIST	-6.776785e-005	3.480861e-005	-1.94687	0.0515502

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	4	35.46441	0.0000004

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : **Fishman_Variables_17NOV_Theissen_Queen.gal**

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	8.84449	0.0029398

COEFFICIENTS VARIANCE MATRIX

CONSTANT	CAP_DIST	AIR_DIST	UNIV_DIST	DOM_DIST	W_RECRUITS
5.220067	-0.000003	-0.000013	0.000016	-0.000027	-0.112258
-0.000003	0.000000	-0.000000	-0.000000	-0.000000	-0.000000
-0.000013	-0.000000	0.000000	-0.000000	0.000000	0.000000
0.000016	-0.000000	-0.000000	0.000000	-0.000000	-0.000001
-0.000027	-0.000000	0.000000	-0.000000	0.000000	-0.000001
-0.112258	-0.000000	0.000000	-0.000001	-0.000001	0.042315

OBS	RECRUITS	PREDICTED	RESIDUAL	PRED ERROR
1	1	1.27905	-1.48934	-0.27905
2	1	14.21030	-8.24804	-13.21030
3	5	2.92843	2.33296	2.07157
4	5	-1.70019	1.88030	6.70019
5	8	6.22902	-1.90130	1.77098
6	1	10.73081	-8.29416	-9.73081
7	1	-1.62601	1.55923	2.62601
8	1	5.53148	-3.13030	-4.53148
9	1	-3.91313	1.67541	4.91313
10	1	22.03153	-11.76506	-21.03153
11	7	2.01282	3.56065	4.98718
12	3	-0.05828	1.57493	3.05828
13	17	-0.17156	13.22226	17.17156
14	20	13.84276	0.91726	6.15724
15	2	3.25304	-3.66238	-1.25304
16	2	3.53783	-2.25827	-1.53783
17	1	9.99128	4.57541	-8.99128
18	1	4.71561	-2.47346	-3.71561
19	53	21.79343	22.03816	31.20657
20	2	2.56006	-1.77053	-0.56006
21	2	7.48222	4.33855	-5.48222
22	5	11.56243	-5.83365	-6.56243
23	4	6.17766	-0.22539	-2.17766
24	1	4.95928	-3.78160	-3.95928
25	1	7.44822	-7.12843	-6.44822
26	6	-1.68904	7.28848	7.68904
27	1	2.54470	-3.00171	-1.54470

========================= END OF REPORT ==============================

5. OpenGeoDa OLS Lagged Model 4 Residual Results

```
Regression
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : Fishman_Variables_17NOV_Theissen
Spatial Weight      : Fishman_Variables_17NOV_Theissen_Queen.gal
Dependent Variable  :     RECRUITS  Number of Observations:   27
Mean dependent var  :      5.66667  Number of Variables   :    6
S.D. dependent var  :       10.378  Degrees of Freedom    :   21
Lag coeff.   (Rho)  :    -0.790067

R-squared           :     0.583752  Log likelihood        :    -91.7826
Sq. Correlation     : -             Akaike info criterion :     195.565
Sigma-square        :      44.8315  Schwarz criterion     :      203.34
S.E of regression   :      6.69563

-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    z-value      Probability
-----------------------------------------------------------------------
  W_RECRUITS    -0.7900671      0.2057052      -3.840774    0.0001227
    CONSTANT      5.628776       2.284747       2.463633    0.0137537
    CAP_DIST  2.656352e-005   7.800635e-006       3.405303    0.0006610
    AIR_DIST -1.859299e-005   1.295044e-005      -1.435704    0.1510868
   UNIV_DIST  4.930959e-005   2.547362e-005       1.935712    0.0529029
    DOM_DIST -6.776785e-005   3.480861e-005       -1.94687    0.0515502
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE         PROB
Breusch-Pagan test                       4        35.46441     0.0000004

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL      LAG      DEPENDENCE      FOR      WEIGHT      MATRIX      :
Fishman_Variables_17NOV_Theissen_Queen.gal
TEST                                     DF       VALUE        PROB
Likelihood Ratio Test                    1         8.84449     0.0029398
```

COEFFICIENTS VARIANCE MATRIX

CONSTANT	CAP_DIST	AIR_DIST	UNIV_DIST	DOM_DIST	W_RECRUITS
5.220067	-0.000003	-0.000013	0.000016	-0.000027	-0.112258
-0.000003	0.000000	-0.000000	-0.000000	-0.000000	-0.000000
-0.000013	-0.000000	0.000000	-0.000000	0.000000	0.000000
0.000016	-0.000000	-0.000000	0.000000	-0.000000	-0.000001
-0.000027	-0.000000	0.000000	-0.000000	0.000000	-0.000001
-0.112258	-0.000000	0.000000	-0.000001	-0.000001	0.042315

OBS	RECRUITS	PREDICTED	RESIDUAL	PRED ERROR
1	1	1.27905	-1.48934	-0.27905
2	1	14.21030	-8.24804	-13.21030
3	5	2.92843	2.33296	2.07157
4	5	-1.70019	1.88030	6.70019
5	8	6.22902	-1.90130	1.77098
6	1	10.73081	-8.29416	-9.73081
7	1	-1.62601	1.55923	2.62601
8	1	5.53148	-3.13030	-4.53148
9	1	-3.91313	1.67541	4.91313
10	1	22.03153	-11.76506	-21.03153
11	7	2.01282	3.56065	4.98718
12	3	-0.05828	1.57493	3.05828
13	17	-0.17156	13.22226	17.17156
14	20	13.84276	0.91726	6.15724
15	2	3.25304	-3.66238	-1.25304
16	2	3.53783	-2.25827	-1.53783
17	1	9.99128	4.57541	-8.99128
18	1	4.71561	-2.47346	-3.71561
19	53	21.79343	22.03816	31.20657
20	2	2.56006	-1.77053	-0.56006
21	2	7.48222	4.33855	-5.48222
22	5	11.56243	-5.83365	-6.56243
23	4	6.17766	-0.22539	-2.17766
24	1	4.95928	-3.78160	-3.95928
25	1	7.44822	-7.12843	-6.44822
26	6	-1.68904	7.28848	7.68904
27	1	2.54470	-3.00171	-1.54470

========================= END OF REPORT ==============================

APPENDIX D

A. REGRESSION RESULTS FOR RISK TERRAIN COMPARISON

1. OpenGeoDa OLS Results for Unweighted Risk Model

```
Regression
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set              : RTM_Regression_Late_Final
Dependent Variable    :      ICOUNT  Number of Observations:    5
Mean dependent var    :         6.2  Number of Variables   :    2
S.D. dependent var    :     6.85274  Degrees of Freedom    :    3

R-squared             :    0.719761  F-statistic           :     7.70517
Adjusted R-squared    :    0.626349  Prob(F-statistic)     :   0.0692316
Sum squared residual:          65.8  Log likelihood        :    -13.5376
Sigma-square          :     21.9333  Akaike info criterion :     31.0753
S.E. of regression    :      4.6833  Schwarz criterion     :     30.2942
Sigma-square ML       :       13.16
S.E of regression ML:       3.62767
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT          -13.3       7.33053       -1.81433    0.1672651
      UWRISK            6.5      2.341652       2.775818    0.0692316
-----------------------------------------------------------------------


DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   6.854102
                                      (Extreme Multicollinearity)
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE             PROB
Jarque-Bera            2          0.7002892         0.7045862

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE             PROB
Breusch-Pagan test     1          0.6325794         0.4264108
Koenker-Bassett test   1           1.536219         0.2151814
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE             PROB
White                  2           3.483323         0.1752290
```

COEFFICIENTS VARIANCE MATRIX

CONSTANT	UWRISK
53.736667	-16.450000
-16.450000	5.483333

OBS	ICOUNT	PREDICTED	RESIDUAL
1	10.00000	12.70000	-2.70000
2	1.00000	-0.30000	1.30000
3	1.00000	6.20000	-5.20000
4	1.00000	-0.30000	1.30000
5	18.00000	12.70000	5.30000

========================= END OF REPORT ==============================

2. OpenGeoDa OLS Results for Weighted Risk Model

```
Regression
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : RTM_Regression_Late_Final
Dependent Variable  :      ICOUNT  Number of Observations:    5
Mean dependent var  :         6.2  Number of Variables   :    2
S.D. dependent var  :     6.85274  Degrees of Freedom    :    3

R-squared           :    0.190185  F-statistic           :    0.704548
Adjusted R-squared  :   -0.079754  Prob(F-statistic)     :    0.462878
Sum squared residual:     190.145  Log likelihood        :    -16.1906
Sigma-square        :     63.3816  Akaike info criterion :     36.3811
S.E. of regression  :     7.96125  Schwarz criterion     :        35.6
Sigma-square ML     :     38.0289
S.E of regression ML:     6.16676
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT    -0.5970294       8.845887    -0.06749232    0.9504362
       WRISK    0.09493058      0.1130969      0.8393736    0.4628783
-----------------------------------------------------------------------


DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   4.758937
                                      (Extreme Multicollinearity)
TEST ON NORMALITY OF ERRORS
TEST                     DF          VALUE            PROB
Jarque-Bera               2          0.7239992        0.6962826

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF          VALUE            PROB
Breusch-Pagan test        1          0.5798596        0.4463673
Koenker-Bassett test      1            1.22383        0.2686103
SPECIFICATION ROBUST TEST
TEST                     DF          VALUE            PROB
White                     2           1.293823        0.5236607

COEFFICIENTS VARIANCE MATRIX
   CONSTANT       WRISK
  78.249714   -0.915830
  -0.915830    0.012791


  OBS          ICOUNT        PREDICTED        RESIDUAL
    1        10.00000          8.04165         1.95835
    2         1.00000          0.25735         0.74265
    3         1.00000          7.75686        -6.75686
    4         1.00000          7.09235        -6.09235
    5        18.00000          7.85179        10.14821
========================= END OF REPORT ==============================
```

3. OpenGeoDa OLS Results for KDE Risk Model

```
Regression
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set             : RTM_Regression_Late_Final
Dependent Variable   :      ICOUNT  Number of Observations:    5
Mean dependent var   :         6.2  Number of Variables   :    2
S.D. dependent var   :     6.85274  Degrees of Freedom    :    3

R-squared            :    0.143952  F-statistic           :    0.504478
Adjusted R-squared   :   -0.141397  Prob(F-statistic)     :    0.528774
Sum squared residual:          201  Log likelihood        :    -16.3294
Sigma-square         :          67  Akaike info criterion :     36.6587
S.E. of regression   :     8.18535  Schwarz criterion     :     35.8776
Sigma-square ML      :        40.2
S.E of regression ML:      6.34035
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT              1      8.185353      0.1221694    0.9104889
       KRISK            6.5      9.151503      0.7102659    0.5287738
-----------------------------------------------------------------------


DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   4.236068
                                      (Extreme Multicollinearity)
TEST ON NORMALITY OF ERRORS
TEST                     DF           VALUE            PROB
Jarque-Bera               2           0.7413172        0.6902796

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                     DF           VALUE            PROB
Breusch-Pagan test        1               0.625        0.4291953
Koenker-Bassett test      1            1.314444        0.2515917
SPECIFICATION ROBUST TEST
TEST                     DF           VALUE            PROB
White                     2                   5        0.0820850

COEFFICIENTS VARIANCE MATRIX
   CONSTANT       KRISK
  67.000000  -67.000000
 -67.000000   83.750000


  OBS         ICOUNT        PREDICTED        RESIDUAL
    1        10.00000         7.50000         2.50000
    2         1.00000         1.00000         0.00000
    3         1.00000         7.50000        -6.50000
    4         1.00000         7.50000        -6.50000
    5        18.00000         7.50000        10.50000
========================= END OF REPORT ==============================
```

4. OpenGeoDa OLS Unweighted Risk Model Residual Results

```
Regression
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : RTM_Regress_Results_Arc
Dependent Variable  :      ICOUNT  Number of Observations:    5
Mean dependent var  :         6.2  Number of Variables   :    2
S.D. dependent var  :     6.85274  Degrees of Freedom    :    3

R-squared           :    0.719761  F-statistic           :     7.70517
Adjusted R-squared  :    0.626349  Prob(F-statistic)     :   0.0692316
Sum squared residual:        65.8  Log likelihood        :    -13.5376
Sigma-square        :     21.9333  Akaike info criterion :     31.0753
S.E. of regression  :      4.6833  Schwarz criterion     :     30.2942
Sigma-square ML     :       13.16
S.E of regression ML:     3.62767
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT         -13.3        7.33053       -1.81433    0.1672651
      UWRISK           6.5       2.341652       2.775818    0.0692316
-----------------------------------------------------------------------


DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   6.854102
                                      (Extreme Multicollinearity)
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2          0.7002892        0.7045862

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test     1          0.6325794        0.4264108
Koenker-Bassett test   1           1.536219        0.2151814
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                  2           3.483323        0.1752290

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR  WEIGHT  MATRIX  :  RTM_Regress_Results_Arc.gwt     (row-standardized
weights)
TEST                           MI/DF       VALUE          PROB
Moran's I (error)            -0.250000    -0.0000000      1.0000000
Lagrange Multiplier (lag)        1         0.6250000     0.4291953
Robust LM (lag)                  1         0.0000000     0.9999999
Lagrange Multiplier (error)      1         0.6250000     0.4291953
Robust LM (error)                1         0.0000000     0.9999998
Lagrange Multiplier (SARMA)      2         0.6250000     0.7316156
========================= END OF REPORT ==============================
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX E

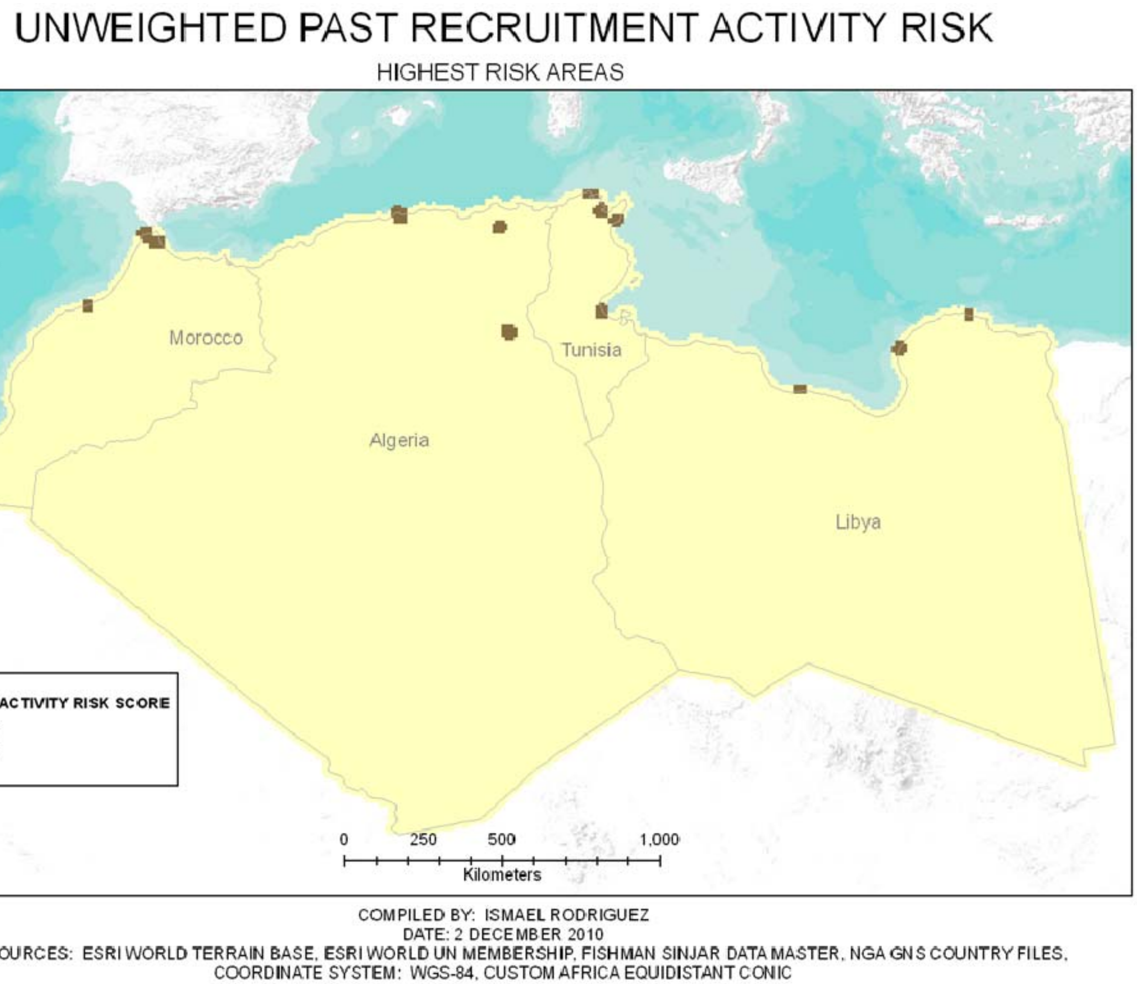


Figure 17. Unweighted Recruitment Activity Risk

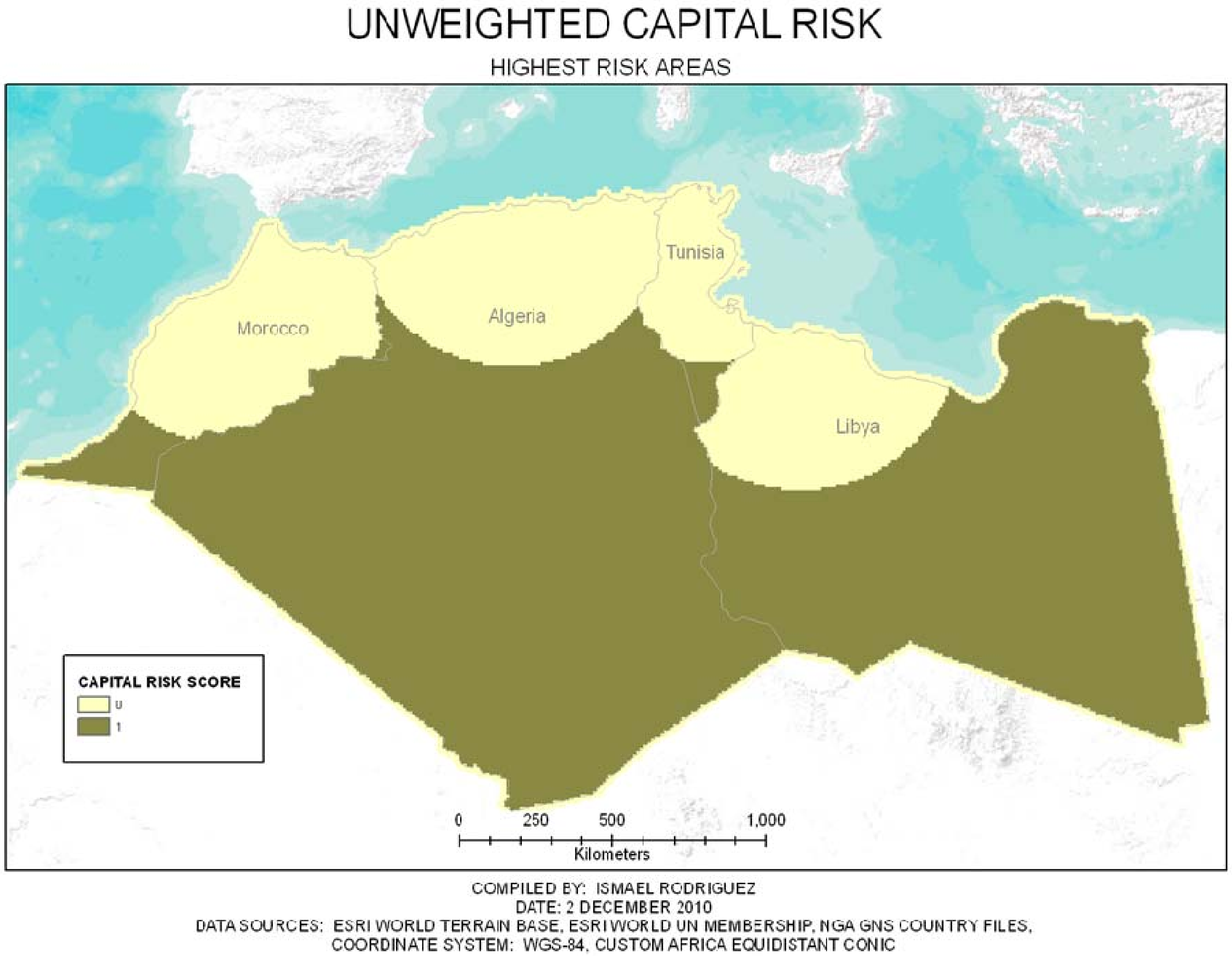


Figure 18. Unweighted Capital Risk

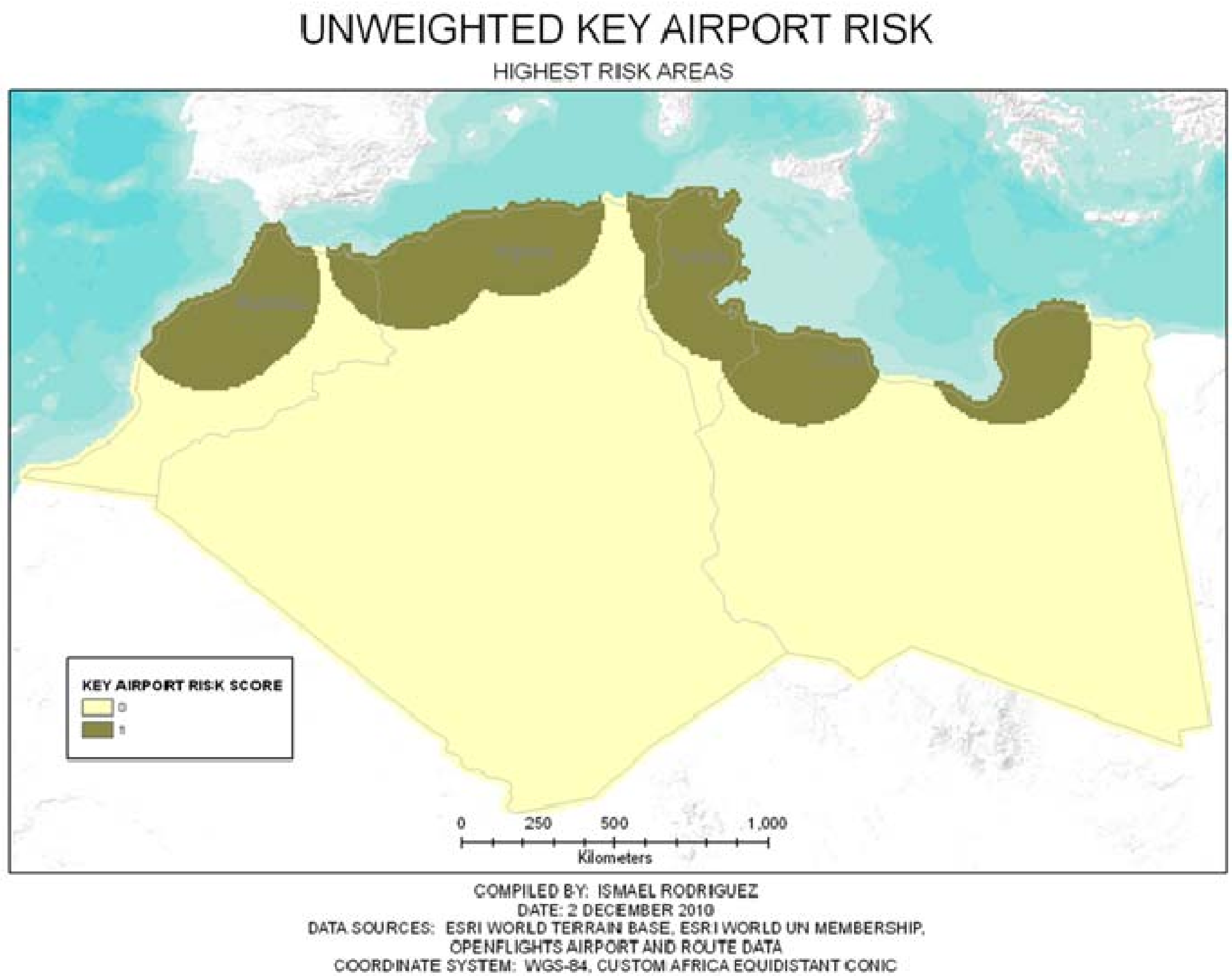


Figure 19. Unweighted Key Airport Risk

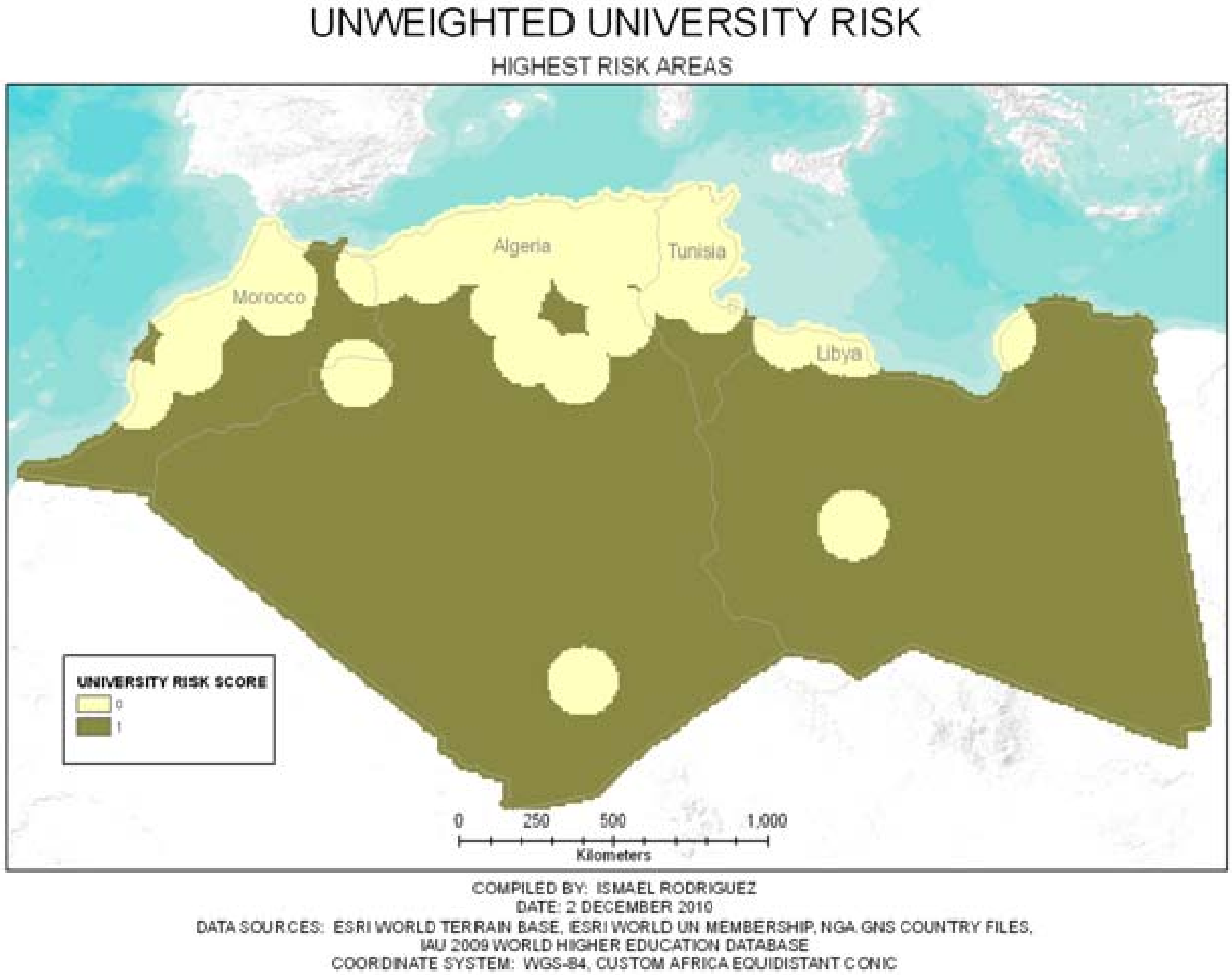


Figure 20. Unweighted University Risk

LIST OF REFERENCES

Anselin, L (2005). Exploring spatial data with GeoDa: A workbook. Retrieved November 2010 from ASU GeoDa Center at <http://geodacenter.asu.edu/learning/tutorials>

Beyer, H. (n.d.). Hawth's analysis tools for ArcGIS. Accessed 24 November 2010 from <http://www.spatialecology.com/htools/>

Boba, R (2005). *Crime Analysis and Crime Mapping,* Thousand Oaks, CA: Sage.

Brown, V (2008). Foreign Fighters in Historical Perspective: The Case of Afghanistan, in J. Felter (Ed.), *Bombers, bank accounts, and bleedout: Al Qaida's road in and out of Iraq* (pp. 16-31). WestPoint: Combating Terrorism Center. Retrieved March 2010 from West Point CTC at <http://ctc.usma.edu/harmony/harmony_menu.asp>

Caplan, J. M., and Kennedy, L. W. (2010). *Risk Terrain Modeling Manual: Theoretical Framework and Technical Steps of Spatial Risk Assessment*. Newark, NJ: Rutgers Center on Public Security. Retrieved July 2010 from < http://rutgerscps.org>

Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). (2005). Gridded population of the world version 3 (GPWv3): Population density grids. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Retrieved on February 2010 from Columbia University website at <http://sedac.ciesin.columbia.edu/gpw.>

Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). (2010). Meta data: Gridded population of the world version 3 (GPWv3). Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Accessed 10 November 2010 from Columbia University website at <http://sedac.ciesin.columbia.edu/gpw>.

Combating Terrorism Center (CTC) at West Point (n.d.a). Foreign fighter bios orig. Retrieved from Harmony Database at <http://www.ctc.usma.edu/harmony/Foreign_Fighter_Bios-Orig.pdf>

Combating Terrorism Center (CTC) at West Point (n.d.b). FF bios trans. Retrieved from Brian Fishman at brianfishman@gmail.com. Also available at Harmony Database at <http://www.ctc.usma.edu/harmony/FF-Bios-Trans.pdf>

Cromer, M., McDonough, W., & Conway, J. (2009, July-September). Leading the way in Geospatial Intelligence, *Military Intelligence Professional Bulletin*. Retrieved December 2010 from the Intelligence Knowledge Network website at <https://ikn.army.mil/apps/dms/search/getDocument.cfm?id=30530>

Della Porta, D. (2002). Comparative Politics and Social Movements in B. Klandermanns and S. Staggenborg (Eds.), *Methods of Social Movement Research*, Minneapolis, MN: University of Minnesota Press.

Della Porta, D. (2006). *Social Movements: An Introduction.* Malden, MA: Blackwell. Retrieved on 16 September 2010 from books.google.com.

ESRI (2010a). Interpreting OLS results. Accessed 22 November 2010 from <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Interpreting_OLS_results/005p00000030000000/>

ESRI (2010b) How GWR works. Accessed 24 November 2010 from ArcGIS Resource Center website at <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_GWR_works/005p00000031000000/>

ESRI (2010c) Interpreting GWR results. Accessed 22 November 2010 from ArcGIS Resource Center website at <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Interpreting_GWR_results/005p00000032000000/>

ESRI (2010d). How kernel density works. Accessed 29 November 2010 from ArcGIS Resource Center at <http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_Kernel_Density_works/009z00000011000000/>

ESRI. (2009e). World Street Map. Accessed 2 December 2010 from ArcGIS Online website at <http://www.arcgis.com/home/item.html?id=3b93337983e9436f8db950e38a8629af>

Felter, J. & Fishman B. (2008a). Al Qaida's foreign fighters in Iraq: A first look at the Sinjar Records. West Point, NY: Combating Terrorism Center. Retrieved March 2010 from CTC at <http://ctc.usma.edu/harmony/harmony_menu.asp>

Felter, J. & Fishman B. (2008b). Becoming a foreign fighter: A second look at the Sinjar records, in J. Felter (Ed.), Bombers, bank accounts, and bleedout: Al Qaida's road in and out of Iraq (pp. 32-65). WestPoint, NY: Combating Terrorism Center. Retrieved March 2010 from CTC at <http://ctc.usma.edu/harmony/harmony_menu.asp>

Flynn, M., Pottinger, M., and Batchelor, P. (2010). *Fixing intelligence: A blueprint for making intelligence relevant in Afghanistan.* Washington: Center for a New American Security. Retrieved January 2010 from CNAS at <http://www.cnas.org>

Fishman, B. (n.d). Sinjar data master [Electronic Spreadsheet]. Retrieved April 2010 from <brianfishman@gmail.com>

Freedom House (2008). The worst of the worst: The world's most repressive societies 2008. Washington, DC: Freedom House. Retrieved August 2010 from Homeland Security Digital Library Database <https://www.hsdl.org/?abstract&doc=98780&coll=documents>

GeoDa Center (n.d.). OpenGeoDa project page, Accessed 24 November 2010 from < http://geodacenter.asu.edu/projects/opengeoda>

GeoDataSource (2010). World cities database platinum edition (Africa Region). Retrieved September 2010 from <www.Geodatasource.com>

Google (2010). Google Maps. Accessed 29 November 2010 at <http://maps.google.com/>\

Gupta, D.K. (2006). Tyranny of Data: Going Beyond Theories, in J. Victoroff (Ed), *Tangled Roots: Social and Psychological Factors in the Genesis of Terrorism.* Amsterdam: IOS Press, 37–50. Retrieved September 2010 from ebrary.com.

Hafez, M. (2003). *Why Muslims Rebel: Repression and resistance in the Islamic world,* Boulder, CO: Lynne Rienner Publishers.

Hoffman, B. (2006). *Inside Terrorism*. New York: Columbia University Press.

Honeycutt, D., Murray, S., and Prince, G. (2010). Fundamentals of GIS Analysis--Proximity [Technical Workshop Presentation]. ESRI. Retrieved September 2010 from <http://resources.arcgis.com/gallery/file/geoprocessing/details?entryID=1546D757-1422-2418-7F77-34F13DBFD075>

International Association of Universities (IAU) (2009). World higher education database [CD-ROM]. Houndmills, UK: MacMillan.

Kaminski-Morrow, D. (2005). Buraq unveils international plan. Flightglobal.com. Accessed November 2010 at <http://www.flightglobal.com/articles/2005/02/15/193886/buraq-unveils-international-plan.html>

Krueger, A (2007). *What makes a terrorist*? *Economics and the roots of terrorism*, Princeton, NJ: Princeton University Press. Accessed 12 November 2010 from < http://books.google.com>

Le Point (2007, January 17). Daily terror area, Eucalyptus, a suburb of Algiers [Translation]. Accessed 12 November 2010 from , <http://translate.google.com/translate?hl=en&sl=fr&u=http://www.lepoint.fr/archives/article.php/9486&ei=6-abTPrEFpKosQPb4JSbCQ&sa=X&oi=translate&ct=result&resnum=4&ved=0CCQQ7gEwAw&prev=/search%3Fq%3Dles%2Beucalyptus%2Bkalitous%26hl%3Den%26prmd%3Div>

McAdam, D. (1988). *Freedom Summer*, New York: Oxford University Press.

McAdam, D., Tarrow, S., and Tilley, C. (2001). *Dynamics of Contention*. New York: Cambridge University Press. Retrieved from ebrary.com.

Massachusetts Institute of Technology (n.d). NGA GeoNet designations: Their codes and definitions. Accessed 29 November 2010 from <http://libraries.mit.edu/gis/data/findingaids/geonet-names/designations.html>

Miller, H. (2004, June) Tobler's First Law and Spatial Analysis in Annals of the Association of American Geographers, *94:2*, 284–289. Retrieved September 2010 from EBSCOHOST database.

Mitchell, A. (2005). *The ESRI Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics*. Redlands,CA: ESRI Press.

National Geospatial Intelligence Agency (2008, July-August). Mobile training teams take GEOINT on the road. *Pathfinder.* Retrieved December 2010 from NGA website at <https://www1.nga.mil/Newsroom/Pathfinder/0604/Documents/0604.pdf>

National Geospatial Intelligence Agency (2010a). Country files (GNS): Algeria [Data updated through 30 September 2010]. Retrieved October 2010 from <http://earth-info.nga.mil/gns/html/namefiles.htm>

National Geospatial Intelligence Agency (2010b). Country files (GNS): Libya [Data Updated through 14 October 2010]. Retrieved October 2010 from <http://earth-info.nga.mil/gns/html/namefiles.htm>

National Geospatial Intelligence Agency (2010c). Country files (GNS): Morocco [Data updated through 15 October 2010]. Retrieved October 2010 from <http://earth-info.nga.mil/gns/html/namefiles.htm>

National Geospatial Intelligence Agency (2010d). Country files (GNS): Tunisia [Data Updated through 4 October 2010]. Retrieved October 2010 from <http://earth-info.nga.mil/gns/html/namefiles.htm>

Oak Ridge National Laboratory (n.d). LandScan. Accessed 2 December 2010 from Oak Ridge National Laboratory website at <http://www.ornl.gov/sci/landscan/>

OpenFlights (2009). Airport and airline databases released [blog]. Accessed 29 November 2010 from <http://openflights.org/blog/2009/02/03/airport-and-airline-databases-released/>

OpenFlights (2010). Airport, airline and route data. Retrieved October and November 2010 from <http://openflights.org/data.html>. Airport Data, <http://openflights.svn.sourceforge.net/viewvc/openflights/openflights/data/airports.dat>; Route Data <http://openflights.org>; Data Description <http://openflights.org/data.html>

Pacione, M. (Ed.). (1999). *Applied Geography: Principles and Practice*, London: Routledge. Retrieved September 2010 from ebrary.com.

Peraino, K. (2008, April). Destination Martyrdom :What drove so many Libyans to volunteer as suicide bombers for the war in Iraq? A visit to their hometown the dead-end city of Darnah.. *Newsweek,* 151(17), Retrieved October 2010, from ABI/INFORM Global. (Document ID: 1466166181).

Perry, W. L. & Gordon, J. (2008). *Analytic support to intelligence in counterinsurgencies* Santa Monica, CA: RAND Corp. Retrieved December 2010 from RAND website at < http://www.rand.org/pubs/monographs/2008/RAND_MG682.pdf>

Rabasa, A. et al. (2007). *Ungoverned territories : Understanding and reducing terrorism risks*. Santa Monica, CA: RAND. Retrieved on February 2010 from <http://www.rand.org/pubs/monographs/MG561/>

Sageman, M. (2008). *Leaderless Jihad,* Philadelphia, PA: University of Pennsylvania Press.

Scott, L., Rosenshein, L. & Janikas, M. (2010) Regression Analysis for Spatial Data [Presentation Slides]. Retrieved July 2010 from ESRI ArcGIS Resource Center website <http://resources.arcgis.com/gallery/file/geoprocessing/details?entryID=F15A382D-1422-2418-348C-6D2A80463896>

Tarrow, S. (1998). *Power in Movement: Social Movement and Contentious Politics*. New York: Cambridge University Press.

Tobler W. (2004, June). On the First Law of Geography: A Reply, *Annals of the Association of American Geographers*, *94:2*, 304–310. Retrieved September 2010 from EBSCOHOST database.

United States Army (2008a) *FM 2-0: Intelligence*. Washington, DC: Headquarters, Department of the Army. Retrieved February 2010 from USAPA database. This version no longer available at that site.

United States Army (2008b). Revitalizing intelligence training [Information Paper] in 2008 Army posture statement. Accessed 6 December 2010 from Army website at <http://www.army.mil/aps/08/information_papers/transform/Intelligence_Training.html>

United States Army (2010). *ATTP 3-34.80: Geospatial engineering*. Washington, DC: Headquarters, Department of the Army. Retrieved December 2010 from Army Publications website at <https://armypubs.us.army.mil/doctrine/DR_pubs/dr_aa/pdf/attp3_34x80.pdf>

Watts, C. (2008). Beyond Iraq and Afghanistan: What foreign fighter data reveals about the future of terrorism: Part I & II. Small Wars Journal. Retrieved March 2010 from <http://smallwarsjournal.com/blog/journal/docs-temp/49-watts.pdf> and <http://smallwarsjournal.com/blog/journal/docs-temp/69-watts.pdf>.

Watts, C. (2008). Beyond Iraq and Afghanistan: What foreign fighter data reveals about the future of terrorism: Appendices. Retrieved from PJ Sage Inc. website March 2010 from <http://www.pjsage.com/products.htm>

Wiktorowicz, Q. (Ed.). (2003). *Islamic activism: A social movement theory approach,* Bloomington, IN: Indiana University Press. Retrieved September 2010 from ebrary.com

Zhukov, Y. (2010, March 12). *Roads and the diffusion of insurgent violence.* [Unpublished Manuscript]. Retrieved November 2010 from zhukov.yuri@gmail.com. Previous version also available at <http://www.fas.harvard.edu/~postcomm/papers/2009-10/Zhukov_Roads.pdf>

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF DATA REFERENCES

Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). (2005). Gridded population of the world version 3 (GPWv3): Population density grids. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Retrieved on 25 February 2010 from Columbia University website at <http://sedac.ciesin.columbia.edu/gpw.>

ESRI. (2009a). World terrain base. Accessed initially March 2010 from ArcGIS Online website at <http://goto.arcgisonline.com/maps/World_Terrain_Base>

ESRI. (2009b). World UN membership. Accessed initially March 2010 from ArcGIS Online website at <http://www.arcgisonline.com/home/item.html?id=c3ee6cbf54aa46e3979ee234140a955e >

ESRI. (2009c). World Street Map. Accessed 2 December 2010 from ArcGIS Online website at <http://www.arcgis.com/home/item.html?id=3b93337983e9436f8db950e38a8629af>

Fishman, B. (n.d.). Sinjar data master [Electronic Spreadsheet]. Retrieved April 2010 from brianfishman@gmail.com

International Association of Universities (IAU) (2009). World higher education database [CD-ROM]. Houndmills, UK: MacMillan.

National Geospatial Intelligence Agency (2010a). Country files (GNS): Algeria [Data updated through 30 September 2010]. Retrieved October 2010 from <http://earth-info.nga.mil/gns/html/namefiles.htm>

National Geospatial Intelligence Agency (2010b). Country files (GNS): Libya [Data Updated through 14 October 2010]. Retrieved October 2010 from <http://earth-info.nga.mil/gns/html/namefiles.htm>National Geospatial Intelligence Agency (2010c). Country files (GNS): Morocco [Data updated through 15 October 2010]. Retrieved October 2010 from <http://earth-info.nga.mil/gns/html/namefiles.htm>

National Geospatial Intelligence Agency (2010d). Country files (GNS): Tunisia [Data Updated through 4 October 2010]. Retrieved October 2010 from <http://earth-info.nga.mil/gns/html/namefiles.htm>

OpenFlights (2010). Airport, airline and route data. Retrieved October and November 2010 from <http://openflights.org/data.html >. Airport Data, <http://openflights.svn.sourceforge.net/viewvc/openflights/openflights/data/airports.dat>; Route Data <http://openflights.org>; Data Description <http://openflights.org/data.html>

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, VA

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, CA

3. JSOU
   Attn: James Anderson
   Tampa, FL

4. ASD/SOLIC
   Attn: Timothy Strabbing
   Washington, DC

5. US SOCOM J-7
   Attn: Thomas Walton
   Tampa, FL

6. HQ US SOCOM Library
   Tampa, FL