



Do All Roads Lead to Violence? A Geospatial Analysis of Insurgent Diffusion in Iraq

Citation

Griffin, Patrick. 2025. Do All Roads Lead to Violence? A Geospatial Analysis of Insurgent Diffusion in Iraq. Master's thesis, Harvard University Division of Continuing Education.

Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37380080>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

Do All Roads Lead to Violence?
A Geospatial Analysis of Insurgent Diffusion in Iraq

Patrick Griffin

A Thesis in the Field of International Relations
for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

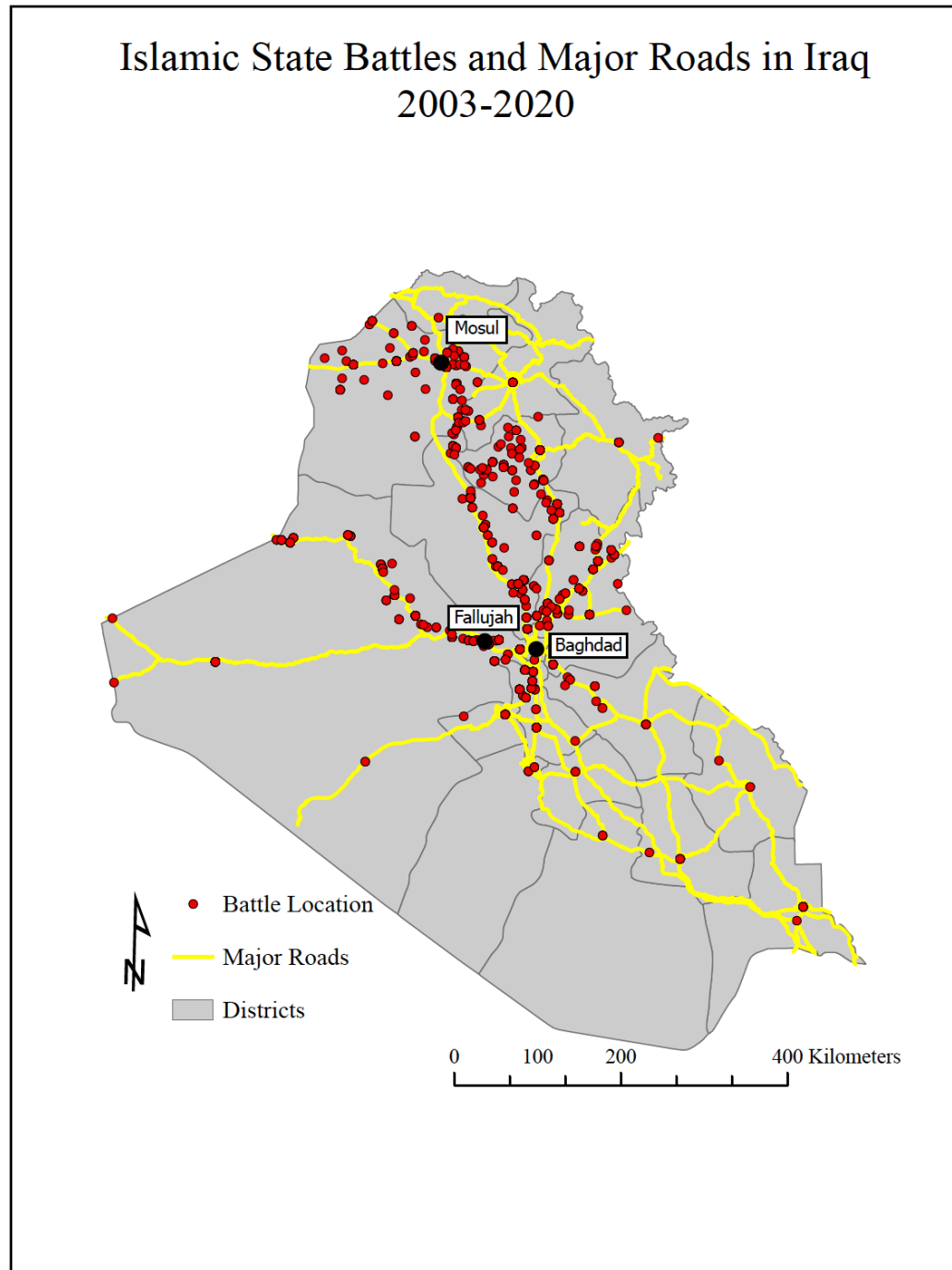
February 2025

Abstract

“The art of war is the art of the logistically feasible” - Admiral Rickover. The practical constraints of geography are often thoroughly examined before the undertaking any conventional military campaign. These factors set boundaries. They are sometimes obscured by the blind ambition of military commanders yet are attributable to victory or defeat after the guns fall silent. Despite considerable research on geography for traditional military purposes, the application of similar analysis on the spread of insurgent violence remains sparse. In 2012, Dr. Yuri Zhukov categorized insurgent logistics into two theories; cellular and self-sufficient, or extremely mobile and flexible.¹ In response to these two theories, he proposed that insurgency was instead most significantly reliant on modern infrastructure and networks of travel (primarily roads). To test his hypothesis, Dr. Zhukov used data from insurgent violence in the Russian Caucasus from 2000 to 2008. Dr. Zhukov found significant quantitative results that confirmed his theory that a network-based model is a more consistent predictor of the spread of insurgent violence.² This thesis looks at these same theories in the case of the Islamic State in Iraq from 2003 to 2020. In aggregate, the results concur with Dr. Zhukov’s findings that roads are the best predictor for the spread of insurgent violence. However, by temporally disaggregating the data, this thesis finds that the strength of these models change over time as an insurgency learns, fights, and evolves.

¹ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

² Zhukov, 144-156.



Frontispiece map created by the author using ArcGIS Pro, September 2024

Author's Biographical Sketch

Patrick Griffin earned a Bachelor of Arts in International Affairs at Florida State University in 2015. Afterwards, Patrick commissioned as a Marine Corps Officer and served five years active duty as a Ground Supply Officer. In 2020, Patrick left active duty to pursue a graduate degree in International Relations at the Harvard Extension School. In 2022, Patrick earned a graduate certificate from the Harvard Extension School in International Security. Patrick enjoys professionally writing on national security and geopolitical topics. His published work can be found in the Marine Corps *Gazette*, U.S. Naval Institute *Proceedings*, the Modern War Institute at West Point, and *Task & Purpose*.

Dedication

To my children.

Acknowledgments

First, I owe a great deal of gratitude to Jeff Blossom, Geospatial Information Systems Service Manager at Harvard, for his constant consultations and steadfast support of my thesis. I also want to thank Dr. Kanglin Chen, Preceptor in Geospatial Analysis at Harvard University's Government Department, and Dr. Scott Bell, lead instructor at the Harvard GIS institute and professor of Geography & Planning at the University of Saskatchewan. Dr. Chen and Dr Bell both helped to teach me the way of geospatial analysis. Their consultation and help in mastering geospatial tools were critical for the completion of this thesis. Thank you to Dr. Doug Bond, Professor in the International Studies Department at Harvard Extension School. My journey at the extension school began and ends with your mentorship. It was a pleasure having you as my professor during my first semester at the Extension School, and a gift to have you again as my research advisor for my thesis.

I also want to thank my many family members and friends that listened patiently to me as I struggled to articulate the thesis in my head into a comprehensible and complete thought (a vain attempt more often than naught). Special thanks to my mother and my friend Kirk for reading my unpolished material whenever I felt it needed a second and third pair of eyes. Finally, and most importantly thank you to my wife. Without your inexhaustible support for my education and my dreams none of this would have ever been possible.

Table of Contents

Frontispiece.....	iv
Author's Biographical Sketch.....	v
Dedication.....	vi
Acknowledgments.....	vii
List of Tables	xii
List of Figures	xiii
Chapter I Introduction.....	1
Definition of Terms.....	1
Background.....	4
Research Question	7
Objectives of the Study	7
Significance of the Study	8
Overview of the Thesis Structure	8
Chapter II Literature Review	10
Military Logistics as the Pacing Function for Operations	10
Classics of Guerrilla Warfare and Arming the Insurgency.....	13
Geography, Supply, and Insurgent Violence	18
Current Gaps	25
Justification for Research.....	26
Chapter III Data and Methodology	28

Research Design and Methods.....	28
Three Theories of Insurgent Logistics	29
Linking Violence to Populations	30
Accounting for Evolution	31
Measuring Violence Diffusion.....	33
Model Breakdown.....	34
Model 1: Geodesic	35
Model 2(a): Network.....	35
Cellular Model	36
Additional Variables	36
Justification for Chosen Methodology.....	37
Data Collection and Analysis.....	38
ArcGIS Pro Software	38
Uppsala Data.....	38
Geofabrik and Open Street Map Data.....	40
WorldPop Data.....	41
DIVA-GIS Data	42
Food and Agricultural Organization of the United Nations.....	43
Description of the Data	44
Iraq Country Data	44
Islamic State Conflict Data	45
Chapter IV Results.....	48
Temporally Aggregated Correlation Results	48

Model 1: Geodesic	49
Model 2: Network	49
Model 3: Cellular	49
Additional Variables	50
Logistics Regression Results Temporally Aggregated	50
Model 1: Geodesic	51
Model 2: Network	51
Model 2a: Network (Major Roads).....	53
Model 3: Cellular	54
Additional Variables	55
Logistics Regression Temporally Disaggregated Annual Results.....	58
Model 1: Geodesic	59
Model 2: Network	59
Model 3: Cellular	60
Additional Variables	60
Logistics Regression Temporally Disaggregated Phased Results	63
Model 1: Geodesic	64
Model 2: Network	64
Model 3: Cellular	64
Additional Variables	65
Chapter V Discussion	67
The Evolution of the Islamic State.....	67
Early Beginnings.....	68

The Rise	71
The Fall	74
Interpreting the Results	78
Correlation Analysis	78
Temporally Aggregated Model Analysis.....	81
Annual Temporally Disaggregated Model Analysis.....	83
Phased Temporally Disaggregated Model Analysis	87
Comparing to Previous Research.....	89
Implications of Findings	90
Limitations of the Study.....	91
Chapter VI Conclusion	92
Summary of Main Findings	92
Answers to Research Questions.....	96
Recommendations for Future Research	96
Appendix 1 Codebook and Technical Data	98
Appendix 2 List of Acronyms.....	102
Appendix 3 Additional Tables and Figures	104
Bibliography	114

List of Tables

Table 1.	Point Biserial Correlation Results.....	48
Table 2.	Bivariate Correlation Results.....	48
Table 3.	Temporally Aggregated Logistics Regression Results.	104
Table 4.	Temporally Disaggregated Area Under the Curve (AUC) Results.	109

List of Figures

Figure 1.	Violent Events in Iraq by Year.	46
Figure 2.	Violent Events by Populated Area.	47
Figure 3.	Annual Model AUC Competition.	58
Figure 4.	Annual AUC Competition All Variables.	61
Figure 5.	Phase Model AUC Competition.	63
Figure 6.	Phase AUC Competition All Variables.	65
Figure 7.	Islamic State Battles, Population Centers, and Elevation in Iraq.	80
Figure 8.	Model 1 (Geodesic) Temporally Aggregated ROC Curve	110
Figure 9.	Model 2 (Network) Temporally Aggregated ROC Curve.	111
Figure 10.	Model 2a (Major Roads Only) Temporally Aggregated ROC Curve.....	112
Figure 11.	Model 3 (Cellular) Temporally Aggregated ROC Curve.	113

Chapter I

Introduction

This chapter will focus on laying the groundwork for this thesis by providing a clear definition of important terms, outlining the research problem, the objectives of the study, and its potential significance to peace and security. The background section of this chapter will provide the reader with a foundational understanding of the context for this research.

Definition of Terms

Guerrilla Warfare/ Irregular Warfare / Unconventional Warfare: “Guerrilla(s)” is a term used to describe an individual or group participating in a style of warfare (guerrilla warfare). Guerrilla warfare defined by the *Guide to the Analysis of Insurgency* is “a form of warfare in which small, lightly armed groups use mobile tactics against a stronger opponent. Guerrillas employ small-scale attacks, such as ambushes and raids, to harass their enemy rather than to win a decisive victory in battle.”³ Terms like “guerrilla warfare” and “terrorism” apply to insurgents when they engage in these tactics. For the purposes of this thesis, the terms “irregular warfare” and “unconventional warfare” are interchangeable with the term “guerrilla warfare.”

³ US Government, *Guide to the Analysis of Insurgency*, 1.

Insurgency: The U.S. Department of Defense Joint Publication 3-24

Counterinsurgency, published in 2018, defines insurgency as “the organized use of subversion and violence to seize, nullify, or challenge political control of a region.

Insurgency can also refer to the group itself.”⁴

The U.S. government 2012 Guide to the Analysis of Insurgency, breaks insurgencies into five primary types: revolutionary, reformist, separatist, resistance, and commercialist. The report also lists four primary structures of an insurgency; politically organized (focused on independent governance), militarily organized (focused on military victory against the government), traditionally organized (centered around tribal/ethnic roots), and urban-cellular (decentralized).⁵

The Islamic State (IS) and the Islamic State in Iraq (ISI): The case subject of this thesis is the insurgent group known primarily as the Islamic State (IS). However, through the evolution of its lifecycle this organization changed official titles as it matured, expanded, broke, and combined with other groups. Chapter V covers a brief history of the Islamic State. Ultimately this group was born from a sub-sect of Al-Qaeda in Iraq (AQI) in 2003 and formed into the Islamic State in Iraq (ISI) in 2006. Then in 2013 ISI shifted to the Islamic State in Iraq and Syria/the Levant (ISIS/L). In 2014 ISIS/L changed its name to the Islamic State (IS). In the scope of this thesis, we will refer to this organization throughout its lifecycle simply as the Islamic State in Iraq (ISI) when denoting operations specific to the internal geography of Iraq, and as the Islamic State (IS) when speaking to the larger international activities of the same group. Although the

⁴ Department of Defense, *Counterinsurgency*, ix.

⁵ US Government, *Guide to the Analysis of Insurgency*, 3-4.

group's official name changes throughout its lifecycle, referring to it in this geo-specific manner will simplify this aspect for the reader.

Logistics: This thesis will use the definition of logistics published in the Marine Corps Doctrinal Publication 4 *Logistics* (MCDP 4): “Logistics is the aspect of military operations that deals with the procurement, transportation, and maintenance of military materiel, facilities, and personnel. It provides for the physical needs of a force.”⁶

According to marine corps doctrine, there are six functions of logistics: “supply, maintenance, transportation, general engineering, health service support, and services (e.g., exchange, food, billeting, mortuary affairs).”⁷

Supply: Supply is a critical function of logistics. Defined by MCDP 4, “supply is the procurement, distribution, maintenance, while in storage, and salvage of supplies.”⁸

Terrorists: Terrorists are agents who partake in the tactics of terrorism. The U.S. legal definition of terrorism under U.S. Code Title 22 is a “premeditated, politically motivated violence perpetrated against noncombatant targets by subnational groups or clandestine agents.”⁹ Insurgents often use terrorism tactics, and thus can be dual labeled as terrorists (or terrorist groups). However, organized groups actively participating in an insurgency are not labeled as terrorists by default, unless they engage in terrorist tactics.

⁶ United States Marine Corps, *Logistics*, 1-2.

⁷ United States Marine Corps, 2-12.

⁸ United States Marine Corps, 2-13.

⁹ U.S. Department of State, *Glossary*.

Violent Event: This term is used to determine whether an event is coded within the dataset. This thesis uses the Uppsala University UCDP Georeferenced Event Dataset Codebook Version 5.0 definition: “An incident where armed force was used by an organized actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a specific date.”¹⁰

Background

“Political power comes out of the barrel of a gun.”¹¹ These words coined by insurgent leader and the future ruler of communist China, Mao Tse-Tung, captures the axiom that political power must carry with it the capacity for violence. When organized groups rise to challenge the authority of dominant political regimes, violence is often the natural accomplice where options for political contestation are absent. While the necessary relationship between political power and violence is well understood, getting the barrel of the gun into the hands of the insurgent is both critically important, and critically understudied.

Cultivating and maintaining the capacity for violence is largely a matter of logistics. Simply put, “[l]ogistics make organized violence possible. One cannot kill without the means to reach a target.”¹² Che Guevara, Cuban revolutionary and insurgent leader, understood the profound importance of logistics, saying: “Always in guerrilla tactics it is necessary to keep in mind the grave problem of procuring the war materiel necessary for continuing the fight.”¹³ The matter of logistics is a practical one. It deals

¹⁰ Croicu, *UCDP GED Codebook version 5.0*, 9.

¹¹ Tse-Tung, *On Guerilla Warfare*, 12.

¹² Zhukov, Y. M. *On the Logistics of Violence*, 399.

¹³ Guevara, *Guerrilla Warfare*, 25.

with constraints from the enemy, but also constraints of the physical environment. On this, researchers agree that structural geographic conditions are instrumental for insurgencies.¹⁴ Civil war researchers James Fearon and David Laitin empirically conclude that geographic conditions like mountainous, rural, and dispersed terrain favors insurgency.¹⁵ In *The Logic of Violence*, Stathis Kalyvas finds that Greek insurgency is strongest in rural, mountainous environments.¹⁶ The importance of the “countryside” is further echoed in the classic works of insurgent warfare from Mao, Guevara, and others.¹⁷ Scholars and practitioners alike find the refuge of rough terrain key for insurgent groups to anchor themselves outside of incumbent controlled ground. Meanwhile, empirical analysis from Dr. Yuri Zhukov provides a robust case for the importance of roads for the diffusion of insurgent violence.¹⁸

Roads, Dr. Zhukov argues, are a strong predictor of the breadth and intensity of insurgent violence - particularly in places that feature harsh terrain that is otherwise impassable.¹⁹ Dr. Zhukov’s empirical work attempts to improve the prediction of violence in civil war based on practical logistical constraints. This in turn, drives the capacity for insurgent action. However, these findings run contrary to T. E. Lawrence’s theoretical vision of guerrilla warfare that focuses on extreme mobility, and the “independence of arteries of supply.”²⁰

¹⁴ Owens, *Sanctuary: The Geopolitics of Terrorism and Insurgency* & Arnold, *Jungle of Snakes*, 29, 79.

¹⁵ Fearon, *Ethnicity, Insurgency, and Civil War*, 75-90.

¹⁶ Kalyvas, *The Logic of Violence*, 249-272.

¹⁷ Guevara, *Guerrilla Warfare*, 19-20 & Tse-Tung, *On Guerilla Warfare*, 108-109, & Aboul-Enein, *A Theoretical Exploration of Lawrence*, 6-9.

¹⁸ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

¹⁹ Zhukov, 144-156.

²⁰ Aboul-Enein, *A Theoretical Exploration of Lawrence*, 8.

Colloquially known as “Lawrence of Arabia,” some credit T. E. Lawrence as the first modern western scholar to work seriously on the theory of guerrilla warfare.²¹ The crux of this divergence between Dr. Zhukov and Lawrence rests on their works' separation across both time and space. Where Dr. Zhukov uses data from political unrest in Russia's North Caucasus region in the 1990s, T. E. Lawrence's theoretical approach to guerrilla warfare stems from his combat experience in early 20th century Ottoman badlands. Here, the vastness of the desert worked to Lawrence's advantage. This terrain is undisputedly different from the Caucasus regional geography featured in Dr. Zhukov's empirical work. Furthermore, the limits of early 20th century technology supported Lawrence's reliance on camels and other pack animals naturally attuned to traversing undeveloped terrain. Today, even in similar desert geography, the demand for motor vehicles as transportation creates practical constraints that Lawrence's insurgents were not subjected to. In Dr. Zhukov's own words, “newer literature suggests that organized violence generally – whether government or insurgent, in flat terrain or mountain valleys – is constrained by road networks.”²²

In Lawrence's view, insurgency must “[drift] about like a gas” while their traditional military adversaries are “like plants, immobile as a whole, firm-rooted, and nourished through long stems to the head.”²³ While this traditional view still embodies the spirit of a modern insurgency, Dr. Zhukov proposes that this perspective of maximum flexibility does not fully account for the practical realities of logistical limitations.²⁴ As a result, Dr. Zhukov argues that this often leads to an overestimation of insurgent violence

²¹ Abdol-Encin, 9.

²² Zhukov, *Roads and the diffusion of insurgent violence*, 4.

²³ Britannica, *T.E. Lawrence on guerrilla warfare*.

²⁴ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

diffusion.²⁵ While Dr. Zhukov comprehensively tests his hypothesis on empirical data, it remains tied to a single conflict. His theory relies heavily on the premise that all modern insurgencies, by the nature of modern logistics, are operationally dependent upon roads.

This thesis seeks to test the limits of Dr. Zhukov's theory on the relationship between roads and insurgent violence by expanding independent empirical testing to the Islamic State Insurgency in Iraq from 2003 to 2020. This thesis will also test these theories in a temporally disaggregated way not attempted by any other researcher known to the author at this point. To better understand the impact of the statistical results this thesis will interpret these results in the context of contemporary research and historical reporting on the Islamic State as a whole.

Research Question

This thesis seeks to understand the significance of established road networks on the diffusion of insurgent violence by analyzing the Islamic State insurgency in Iraq from 2003 to 2020.

Objectives of the Study

Today's modern infrastructure is built and maintained by local, regional, and national government bodies. This creates an assumption that because this infrastructure is a product of the government, it is inherently a power-projection enabler for government control.²⁶ The objective of this study is two-fold. It intends to challenge this assumption and argue that road network infrastructure is figuratively and literally a "two-way street"

²⁵ Zhukov, 25.

²⁶ Kilcullen, *Political Maneuver in Counterinsurgency*.

that enables both insurgents and incumbent forces. It also seeks to explore in level of prominence road networks play for an insurgency over its lifecycle.

Significance of the Study

The implications of these findings are intended to help shape future policy for all levels of counterinsurgency operations. By providing a clearer understanding of how insurgent violence spreads, planners can devise more targeted approaches to prevent its spread, and practitioners can reap more effective results. These results ultimately intend to better counterterrorism strategies for the reduction of conflict, human suffering, and save to lives.

Overview of the Thesis Structure

This thesis is divided into six parts. Chapter I, Introduction, provides definitions, context, and the purpose for this thesis. Chapter II, Literature Review, provides an overview of the existing relevant scholarly work surrounding this topic. This chapter also identifies where gaps in this literature exist and how this thesis seeks to remedy this. Chapter III, Methodology, discusses the designs of the research, the data sources and collection process, and answers why these data and methods were selected. Chapter IV, Results, presents the findings of the statistical portion of this thesis. Chapter V, Discussion, puts these results into context through qualitative analysis on the case selected for this thesis. In this chapter the results are also compared to previous insurgency research. Chapter V will conclude with some of the limitations of this thesis. Chapter VI, Conclusion, will summarize the main findings of this thesis, answer the research question posed in Chapter I, and provide some recommendations for future

research. Appendix 1 provides the reader with codebook and technical data from the quantitative portion of this thesis. Appendix 2 lists the relevant acronyms used in this thesis. Appendix 3 provides additional tables and graphs for the data used in the thesis. For a complete list of all works cited, see the Bibliography.

Chapter II

Literature Review

This chapter provides a comprehensive literature review and is broken into three parts. Part I begins by setting a foundational understanding of the role and importance of logistics systems in conflict using classical military literature. Part II discusses the classics of insurgent revolution literature and the relevance of these irregular warfare theorists on insurgent logistics. Part III brings in contemporary research on the relationship between roads, geography, and insurgent violence. After the literature review, I will discuss current gaps in the contemporary literature and justification for my research.

Military Logistics as the Pacing Function for Operations²⁷

For as long as armies have formed on the field of battle, logistics dictated their capabilities and set their limitations. Roman roads were built and maintained to facilitate the rapid deployment of legions across the empire.²⁸ The timing of medieval European campaigns relied on seasonal grain harvest projections.²⁹ Campaign routes were predetermined by a daisy chain of countryside magazines, or the flow of local rivers that could ferry food and materials on barges downstream.³⁰ The latter of these drew the proximity and direction of military operations to key geographic features.³¹ Eventually, Napoleon Bonaparte developed a method of moving smaller, lighter armies, only to

²⁷ United States Marine Corps, *Logistics*, 1-1.

²⁸ Keegan, *Intelligence in War*, 10.

²⁹ Crevel, *Supplying War*, 11.

³⁰ Crevel, 11-17.

³¹ Crevel, 11-17.

converge in a single time and place where they could overwhelm the enemy.³² This higher level of maneuverability came about through the diffusion of forces which lessened the logistical burden on any single unit on the march. While this innovation dramatically increased mobility, Napoleon still suffered from logistical constraints of poor road conditions and fuel shortages (in the form of animal fodder).³³ These structural parameters ultimately set limits to the success and reach of Napoleon's campaigns.³⁴

Advancements in technology over time increased speed and ease of travel. However, alongside these advancements roads and rail still tie modern logistics to geography. Roads and railways work in the same way that rivers controlled the diffusion, direction, and intensity of medieval military campaigns. Mackinder's Heartland Theory proposed that the development in trans-continental rail networks would facilitate enough logistical power for a central European nation to consolidate authority over the "World Island."³⁵ In his own words, "railways work the greater wonders in the steppe because they directly replace horse and camel mobility, the road stage here having been omitted."³⁶ However, Mackinder's theory that a heartland power would come to dominate the world ultimately failed to materialize. Despite their advantages, railways restricted ground-based logistics to narrow corridors over land. The further forces ventured from these supply corridors, the weaker their combat power became.³⁷ In essence, land logistics became a prisoner to its own efficiency.

³² Jelineo, *Napoleon's Logistics*.

³³ Jelineo.

³⁴ Jelineo.

³⁵ Mackinder, *The geographical pivot of history*, 434.

³⁶ Mackinder, 434.

³⁷ Fettweis, *Sir Halford Mackinder, Geopolitics, and Policymaking*, 58-71.

Meanwhile, Alfred Thayer Mahan's *Influence of Sea Power Upon History* emphasized the protection of sea-based logistics through sea power dominance.³⁸ The control of the seas freed coastal militaries from the geographic limitations of Mackinder's World Island. While moving equipment over land faces the obstacles of mountains, jungles, deserts, swamps, rivers, and the like, the ocean is a vast and open space where logistics moves in any direction to support a fight. The freedom of the sea gave Mackinder's Rimland Nations the capability to project power unabated by land geography so long as they could maintain operational security. Despite the sea being a larger space to maneuver, people still live on the land. Therefore, just as relative power for Heartland Nations diminished the farther they strayed from their narrow rail and road corridors - the farther the Rimland Nations ventured inland from the coast, the weaker military operations became due to a diminishing supply chain. This is evident in Great Britain's colonization efforts of Southeast Asia where British control waned the further one traveled inland from the ports of India, until it finally dissipated along the Hindu-Kush mountains of the Afghan border.³⁹ This paradoxical balance trapped heartland powers like Germany and Russia within their own geographical fortresses. Here the relationship between geography and logistics acts as both a shield from sea-power incursion and a slough for heartland expansion.

Even with the advent of airpower, supply hubs remain tethered to where aircraft can land and takeoff. Move away in any direction from that point, and military operations become strained from diminishing rates of supply-support capacity. This diminishing rate affects both commercial and military logistics operations and dictate methods and

³⁸ Mahan, *The Influence of Sea Power Upon History*, 25-29.

³⁹ Goraya, *The Great Game*, 81-94.

locations of hub-and-spoke distribution systems.⁴⁰ These practical restrictions set hard limitations on military adventurism and create boundaries for operations based on diminishing capability of sustainment from a point. Even where belligerents were assumed to rely on unconventional logistics or supply themselves from local sources or captured enemy material like the communist forces during the Korean War: “Contrary to the common opinion, the NKPA [North Korean People’s Army] and CCF [Chinese Communist Forces] in Korea were dependent on formal, albeit relatively lean, logistical systems.”⁴¹ While most modern and classical scholarship that exists on logistics and supply remains focused on conventional military operations, it is equally important to consider these effects on irregular warfare and insurgency. Even when faced with one another, an insurgency and a military conduct different operations based on their overall strategy. These differences both create, but more importantly are based upon differences in supply and logistical conditions.

Classics of Guerrilla Warfare and Arming the Insurgency

Guerrilla warfare is a practice older than the term itself. In the Old Testament, Israelites under the leadership of Joshua conducted hit-and-run attacks against their stronger enemies in and around Canaan.⁴² Since then, many recorded events describe local peasants partaking in guerrilla tactics to resist occupying traditional armies. The term guerrilla warfare comes from the Spanish word for war; “guerra.”⁴³ In the early 1800’s a series of small but intense insurrections swept across the Spanish peninsula.

⁴⁰ Chopra, *Designing the Distribution Network in a Supply Chain*, 123-140.

⁴¹ Shrader, *Communist logistics in the Korean War*, 224.

⁴² Central Intelligence Agency, *Guerrilla Warfare*.

⁴³ Oxford English Dictionary, *Definition of Guerrilla*.

These little wars taxed Napoleon's troops, eventually forcing the capitulation of Napoleon's army and the expulsion of his eldest brother, Joseph Bonaparte, from Spain.⁴⁴ To classify the collection of these “small wars” that ultimately defeated the more conventional occupying force, French stenographers coined the term “petite guerra” from the Spanish term “guerra.”⁴⁵ Spanish resistance fighters blurred the front lines for French troops, while British envoy Arthur Wellesley, the First Duke of Wellington, coordinated with “guerrilla” efforts to assist British forces in eventually defeating Napoleon's army.⁴⁶ Wellesley’s experience coined the term “guerrilla’s” in the English as it is used today.⁴⁷

Outside of Spain, Clausewitz observed firsthand guerrilla tactics used against the French in Russia. Based on his own observations Clausewitz determined that insurgency could strain a conventional military’s supply lines and overcommit an army’s forces to operational security.⁴⁸ Yet for Clausewitz, these tactics only served to weaken a conventional military so that it could be defeated by another traditional army on the field of battle.⁴⁹ Clausewitz categorized this type of warfare as “popular uprisings” from local residents upset by invading French troops moving across the Russian Caucasus.⁵⁰ In German, Clausewitz defined these wars as “kleinkriegs” or “small wars.”⁵¹ Clausewitz points out the importance of favorable geography in guerrilla warfare - characterized by rough terrain, a large theater of operations, and a deep interior to which insurgent forces can take refuge.⁵²

⁴⁴ Britannica, *Peninsular War*.

⁴⁵ Walker, *Guerrilla warfare's epic history*.

⁴⁶ Wilkins, *Guerrilla Warfare*, 613.

⁴⁷ Oxford English Dictionary, *Definition of Guerrilla*.

⁴⁸ Hahlweg, *Clausewitz and Guerrilla Warfare* 127-133.

⁴⁹ Hahlweg, 127-133.

⁵⁰ Hahlweg, 127-133.

⁵¹ Hahlweg, 127-133.

⁵² Young, *The Guerrilla and the Peninsular War*, 67-80. f

Antoine-Henri Jomini, a Prussian staff officer, also observed the same type of hybrid warfare during this time. Instead of suggesting the adoption of guerrilla forces to aid in the destruction of your enemy as Clausewitz does, Jomini views irregular warfare through a counterinsurgency lens.⁵³ Jomini recommends the use of overwhelming force to occupy and pacify a region experiencing insecurity from guerrilla action.⁵⁴ Based on Jomini's writing, it is reasonable to believe he would view government-built road networks as an asset for government forces to occupy and sustain themselves against local insurgencies. For Jomini, this government infrastructure would likely only present as net-positive asset, and not be considered something to facilitate a dual-purpose for insurgents.

In fact, early theorists do not place much emphasis on the supply and logistics that allow these "popular uprisings" and "small wars" to coalesce. Despite being an essential factor for creating combat capability, early theorists only allude to the arming of insurgents via a persistent attack on enemy supplies lines, and favorable support from the local population.⁵⁵ The only significant external military support provided to the Spanish Guerrillas came from a small British contingent led by Arthur Wellington, who in Clausewitz style, leveraged local resistance fighters to exhaust French troops before deploying his own.⁵⁶

It was not until T. E. Lawrence published a series of articles that would culminate in his books *Revolt in the Desert*, and *The Seven Pillars of Wisdom*, that guerrilla warfare began to stand apart from conventional war in literature. In 1916, T. E. Lawrence

⁵³ Foerster, *Reassessing Jomini on Hybrid Warfare and Counterinsurgency*.

⁵⁴ Foerster.

⁵⁵ Young, *The Guerrilla and the Peninsular War*, 67-80.

⁵⁶ National Army Museum, *Peninsular War*.

departed Britain for the Arabian Peninsula where he would advise and assist local Hashemite forces in waging a guerrilla war against the Ottoman empire. From this experience, Lawrence developed core concepts for the successful conduct of irregular warfare that was independent from any conventional military campaign. Lawrence emphasized speed, maneuver, and using a small force to hit a weak point along an enemy's supply line.⁵⁷ Lawrence viewed the extreme ability to maneuver in the deserts as akin to open-ocean naval warfare.⁵⁸ In this way, Lawrence saw the enemy as constrained by Mackinder's rail, while he and his camel-bound insurgents operated with the complete logistical freedom of Mahan's open-sea warfare. While the Ottomans operated with superior equipment, it was the asymmetry in logistical conditions that gave insurgents the advantage for Lawrence. Using traditional forms of desert travel guerrilla forces could strike and retreat into the unforgiving heartland of the Arabian desert. The harsh geography limited the ability for Ottoman troops to venture beyond their strongholds or supply lines in pursuit. This gave Lawrence total initiative during his campaigns. While Lawrence does not provide detailed insight into the provisioning of his guerrilla bands, he does emphasize the importance of using "the smallest force in the quickest time at the farthest place."⁵⁹

During the Second World War, communist guerrilla leader Mao Tse-Tung borrowed both from Clausewitz and Lawrence when waging insurgent warfare against Imperial Japan. In Mao's self-published instruction manual, *On Guerrilla Warfare*, he echoes Clausewitz's belief that this style of warfare must be paired with conventional

⁵⁷ Aboul-Enein, *A Theoretical Exploration of Lawrence*, 8.

⁵⁸ Aboul-Enein, 3, 7.

⁵⁹ Aboul-Enein, 6.

military battles to bring about a decisive victory.⁶⁰ Mao also emphasizes the need for insurgents to strike and then withdraw into rough geographic terrain beyond the reach of conventional forces.⁶¹ This concept echoes the writing of both Clausewitz and Lawrence. However, Mao's manual is more pragmatic than the theoretical writings of earlier authors. Unlike Clausewitz, Jomini, or Lawrence, Mao's manual is designed for the guerrillas in the field - not the staff-officer at the war college. Therefore, practical concerns of supply and logistics are addressed directly. Mao states that "the enemy is the principle source of supply," and beyond this the guerrilla is responsible for maintaining and sustaining themselves at the individual level.⁶² Mao's emphasis on the individual-initiative of the Guerrilla acting in the enemy's rear areas does not lend to the need for strong logistic lines feeding an insurgent force. However, Mao still maintains the need for an insurgent stronghold where leadership directs the flow of resistance efforts from the safety of rough terrain.⁶³ The lack of importance that Mao places on strong logistical networks to sustain insurgent effort stemming from a stronghold allude to Clausewitz's influence on Mao's thinking. Mao does not see insurgency as a decisive strategy for victory, rather like Clausewitz, "[g]uerrilla hostilities are but one phase of the war of resistance against Japan."⁶⁴

Ernesto "Che" Guevara, Cuban revolutionary and author of *Guerrilla Warfare*, saw the role of the insurgent as more central to revolution than Mao. In his writing, Guevara saw the evolution of an insurgency beginning from self-sustaining cells that

⁶⁰ Tse-Tung, *On Guerilla Warfare*, 50, 66-70.

⁶¹ Tse-Tung, 46, 68

⁶² Tse-Tung, 83.

⁶³ Tse-Tung, 108-109.

⁶⁴ Tse-Tung, 66.

lacked centrality, and eventually growing into a concentrated political and military force that overtakes an incumbent.⁶⁵ Guevara's book, *Guerrilla Warfare*, speaks little on supply and logistics beyond the high praise of it always being a "grave problem" for the insurgent to procure proper ammunition and weapons.⁶⁶ In his book, Guevara clearly warned that the establishment of strongholds and other physical centers of gravity for insurgencies must come under conditions that ensure their survival from enemy attack.⁶⁷ This evolution of insurgency from decentralized cells into a formidable force alludes to an increasing need for more sustained and controlled supply networks to support such developments.

Geography, Supply, and Insurgent Violence

Classic Guerrilla warfare theorists like T. E. Lawrence exhort the extreme mobility of the insurgent. Meanwhile Mao Tse-Tung draws from Clausewitz theory and demands that insurgency must be cellular and self-sufficient in terms of supply and logistics. Che Guevara argues that once these cellular structures gain footholds through leveraging harsh terrain, they can build the power to sufficiently challenge incumbent forces without the need of outside conventional military assistance. Together all these authors acknowledge the importance of mastery over the local geography as a precondition for victory. Yet none of them mention the utilization of conventional supply chains and logistic infrastructure. Classic insurgency writing characterizes operations as free as open-ocean warfare - while their opponents remain helplessly tethered to their

⁶⁵ Guevara, *Guerrilla Warfare*, 11-19, 71-73.

⁶⁶ Guevara, 25.

⁶⁷ Guevara, 114-116.

supplies and bases. However, these guides and theoretical writings often lack detail on the practical constraints of logistics with its relationship to violence. On this, Dr. Zhukov writes that “[m]ass violence can only occur where it is logistically feasible...logistics can help us predict how much violence a group can conceivably generate, and – by extension – whether and how that group’s capacity for violence may be curtailed.”⁶⁸ Dr. Zhukov’s research applies equally to conventional military and insurgent groups, narrowing the gap between an insurgents “extreme mobility” and the mobility of their conventional counterpart. Quantitative research using Global Terrorism Database and the Uppsala Conflict Data Program (UCDP) Georeferenced Event Dataset conclude that the further a conflict region is from its capital city, the lower the probability of violence.⁶⁹ Separate research using geo-located data from the Chechen conflict from 1999 to 2007 also found that for all eight years of the conflict, “clustering of events along the Caucasian Federal Highway in central Chechnya...showed the importance of the main transportation artery and the targets in Chechnya in explaining the distribution of violence.”⁷⁰ This analysis challenges the notion that insurgent supply-chains are only relevant in later stages of an insurgency’s lifecycle.⁷¹

Modern and classic writers on insurgency agree that rough terrain where guerillas can take refuge is positively associated with the potential for insurgent violence.⁷² Geospatial analysis on the Afghanistan conflict from 2008-2009 places the highest

⁶⁸ Zhukov, *On the Logistics of Violence*, 399-424.

⁶⁹ Carter, *Geography of Separatist Violence*, 1-14.

⁷⁰ O’Loughlin, *The Localized Geographies of Violence*, 178-201.

⁷¹ McColl, *The Insurgent State*, 613-631. & Guevara, *Guerrilla Warfare*, 74.

⁷² Fearon, *Ethnicity, Insurgency, and Civil War*, 75-90. & Kalyvas, *The Logic of Violence*, 38-39. & Guevara, *Guerrilla Warfare*, 19-20. & Tse-Tung, *On Guerilla Warfare*, 108-109. & Aboul-Enein, *A Theoretical Exploration of Lawrence*, 6. & Gordillo, *Terrain as insurgent weapon*, 55. & Hendrix, *Head for the Hills?*, 345-370. & Tollefsen, *Insurgency and Inaccessibility*, 6-25.

concentration of insurgent violence in the Hindu-Kush mountains between Afghanistan and Pakistan where the terrain is historically rural and austere.⁷³ However, broader geospatial research on the scope of civil conflict finds that forests and mountains do not affect the relative size of the conflict area after violence has broken out.⁷⁴ Instead, other aspects like access to resources, and proximity to both the state's capital or the international border are strongly correlated to the relative scope of conflict during civil war.⁷⁵ One particularly well researched aspect is the “periphery-factor” where insurgencies tend to trend around a state’s periphery regions.⁷⁶

According to an in-depth study done on the Tamil Tiger insurgency in Sri Lanka, poor control of border areas allowed Tamil insurgents to receive a significant number of weapons and amount training from India’s military.⁷⁷ This phenomenon of external state support is not unique to the Tamil Tigers. According to Daniel Byman and the other authors *Trends in Outside Support for Insurgent Movements*, “Of the 74 insurgencies active since 1991 that were surveyed... 44 received state support that, in our judgment, was significant or critical to the survival and success of the movement.”⁷⁸ In Weidmann’s 2010 research on *Predicting Conflict in Space and Time*, he finds a statistical significance for a direct relationship between border proximity and probability to violence.⁷⁹ However, in Tollefsen’s 2015 research on *Insurgency and Inaccessibility*, he finds that not all international borders are beneficial to insurgency grounds.⁸⁰ Meanwhile, Buhaug

⁷³ O’Loughlin, *The Afghanistan-Pakistan Wars*, 437-471.

⁷⁴ Buhaug, *The Geography of Civil War*, 417-433.

⁷⁵ Buhaug, *The Geography of Civil War*, 417-433. & Raleigh, *Population size, concentration, and civil war*, 224-238. & Buhaug, *Insurgency and Inaccessibility*, 6-25.

⁷⁶ Buhaug, *Dude, Where’s My Conflict?*, 107-128.

⁷⁷ Beehener, *The Taming of the Tigers*, 10.

⁷⁸ Byman, *Trends in Outside Support for Insurgent Movements*, 2.

⁷⁹ Weidmann, *Predicting Conflict in Space and Time*, 883-901.

⁸⁰ Tollefsen, *Insurgency and Inaccessibility*, 6-25.

finds that these border regions can be infectious to peaceful nations, often influencing instability and conflict upon neighboring states.⁸¹

Where external support to insurgents from a conventional military power is available, it not only has the potential to affect the practical performance of an insurgency, but also an insurgency's processes. According to Lionel, during the final phases of the Tamil Tiger insurgency the group abandoned Guerrilla tactics and functioned as a conventional military so that it would appear as a more legitimate administrator of captured territory.⁸² It is possible that this shift from irregular warfare to conventional tactics was not only done to appear legitimate, but also influenced by prior support from outside conventional actors. Insurgencies grow over time. According to Janet Lewis, author of *How Insurgency Begins*, several insurgent groups are born and die before they can even arise to a level of committing violence.⁸³ However, as they grow, insurgencies take on lifecycle phases that can sometimes influence their level of violence, as well as their ways and means of spreading violence to new territory.⁸⁴

After visiting Kunar Province, Afghanistan, the highest concentrated area of insurgent violence in the country, political and military advisor David Kilcullen published his field notes on the importance of road development in the region for counterinsurgency efforts. Kilcullen offered several points that lend road development to improving incumbent security in insurgent territory.⁸⁵ This included improving economic development, cutting off insurgents from the local population, and better facilitation of

⁸¹ Buhaug, *Contagion or Confusion? Why Conflicts Cluster in Space*, 215-233.

⁸² Beehener, *The Taming of the Tigers*, 5.

⁸³ Lewis, *How Insurgency Begins*, 69-70.

⁸⁴ US Government, *Guide to the Analysis of Insurgency*, 3.

⁸⁵ Kilcullen, *Political Maneuver in Counterinsurgency*.

troop movement in the area.⁸⁶ Kilcullen argues that there is inherent benefit that road networks bring to logistical sustainment of government troops from areas beyond their natural stronghold. However, where Kilcullen promotes roadbuilding as a strong counterinsurgency tool, the geospatial analysis of the role of roads in multiple African civil conflicts confirm that they act like a double-edged sword for states.⁸⁷ While road networks ease the travel-times for government forces, Boulding's loss-of-strength gradient (LSG) still argues that state control diminishes as power travels across distance.⁸⁸ Even in instances of intra-state conflict "[t]he projection of power across distance comes at a cost."⁸⁹

How then does the variable of government control relate to insurgent violence when not being measured by government infrastructure? By looking at tax collection data as another form of government control, Hendrix was able to establish that "both state capacity and geography exert a significant effect on civil war onset."⁹⁰ Therefore, while state control has an inverse relationship to insurgent violence, road networks (even when they are the product of state development) are at best a neutral factor.

Geography is equally nuanced in this case. In Afghanistan, researcher Gaston Gordillo finds that while insurgent activity is more prevalent in rough terrain, it rises and falls alongside seasonal changes.⁹¹ Gordillo finds that this is because seasonal weather patterns make the conduct of combat operations difficult and thus dampen the intensity of violence.⁹² Others have found that "the conclusion of poppy cultivation, improved

⁸⁶ Kilcullen.

⁸⁷ Mu'ller-Crepon, *Roads to Rule, Roads to Rebel*, 563-590.

⁸⁸ Boulding, *Conflict and defense: a general theory*, 230-233.

⁸⁹ Buhaug, Dude, *Where's My Conflict?*, 113.

⁹⁰ Hendrix, *Head for the Hills?* 345-370.

⁹¹ Gordillo, Terrain as insurgent weapon, 53-62.

⁹² Gordillo, 53-62.

weather conditions and recesses in madrassas in neighboring Pakistan – have made spring Afghanistan’s ‘fighting season.’”⁹³ One study done on the intensity of violence and the harvest of wheat cycles in Iraq, Afghanistan, and Pakistan found that “the onset of harvest reduces the average number of monthly attacks by around 6% to 22% in Iraq, 20% in Pakistan, and 8% to 18% in Afghanistan.”⁹⁴ This research draws a line between insurgent violence intensity and seasonal geography, yet stops short of exploring how these seasonal changes affect the spread of insurgent violence.

Based on road and conflict data gathered from the Democratic Republic of Congo, the Central African Republic, and Senegal, researchers conclude: “For states, building roads is literally a two-way street. It extends their ability to govern, but also provides their challengers with an infrastructure for mobilization.”⁹⁵ In a comprehensive study done on FARC insurgents in Columbia, researchers found that in the absence of roads, insurgents built their own network of “mobility corridors” to outmaneuver Colombian forces.⁹⁶ Here, researchers conclude that the construction of road networks do not inherently benefit the incumbent force.⁹⁷ Instead, the context and conditions of the conflict and the environment can cause roads to benefit insurgent forces more than incumbents.⁹⁸ In the case of FARC, these insurgents constructed roads to leverage the natural terrain including jungle cover and local rivers - giving them stronger logistical capability than traditional incumbent road networks that only act to connect urban areas to one another.⁹⁹ This compliments geospatial research in Afghanistan that list Kunar

⁹³ Koven, *The End of Afghanistan’s Spring Fighting Seasons*.

⁹⁴ Guardado, *The Seasonality of Conflict*, 4.

⁹⁵ Mu’ller-Crepon, *Roads to Rule, Roads to Rebel*, 565.

⁹⁶ Otero-Bahamon, *Seeing like a Guerrilla*, 199.

⁹⁷ Otero-Bahamon, 205-206.

⁹⁸ Otero-Bahamon. 203-205.

⁹⁹ Otero-Bahamon, 201-202.

Province as the epicenter of insurgent violence. Here the Khyber Pass remains the main artery of supply through the region and is subject to consistent insurgent harassment.¹⁰⁰ Meanwhile, “it is estimated that there are 111 illegal and unregulated crossings along the northern portion of the border and at least 229 in the south.”¹⁰¹

In a study done on conflict density inside Iraq in 2004, Townsley finds that “like property offenders who do not travel farther than necessary to steal property (the least effect principle), insurgents are unlikely to stray too far from a weapon haul due to the increased risk of interdiction prior to attack.”¹⁰² When violence does spread from one village to another, research is consistent that insurgents often choose existing road networks. Using geospatial information systems and social network analysis, Hammond finds that in “46 civil conflicts across 32 states between 1989 and 2015... cities, towns, and villages that are strategically located on the road network are significantly more likely to experience deadly violence during civil conflict.”¹⁰³ When studying the population size and concentration during civil war across 14 countries in Central Africa, Raleigh finds that “conflict events are 47% less likely to happen in squares with secondary roads...and 75% less likely to happen in squares with no roads or only informal roads than in squares with primary roads.”¹⁰⁴

In his thesis, Dr. Zhukov uses geo-referenced data from the Russian Caucasus conflict in the 1990s to assess the relationship between roads and the diffusion of insurgent violence. Here, Dr. Zhukov finds strong evidence that roads facilitate the

¹⁰⁰ O'Loughlin, *The Afghanistan-Pakistan Wars*, 437-471.

¹⁰¹ O'Loughlin, 443.

¹⁰² Townsley, *Space Time Dynamics of Insurgent Activity in Iraq*, 139-146.

¹⁰³ Hammond, *Maps of mayhem*, 32-46.

¹⁰⁴ Raleigh, *Population size, concentration, and civil war*, 224-238.

supply-support needed for insurgency to spread.¹⁰⁵ In his own words, “this paper demonstrates that a failure to account for logistical constraints in empirical models of civil war can produce misleading predictions about the scope of conflict diffusion.”¹⁰⁶ Dr. Zhukov’s thesis ties the conclusion that violence in civil war follows a “pattern of diffusion” to practical realities of geography and logistics.¹⁰⁷

Current Gaps

The study of conventional military logistics is a well-covered topic, ranging over several centuries.¹⁰⁸ The literature on insurgency and intra-state conflict has gained the attention of more researchers since the terror attacks of September 11th, 2001. Even the attention of on the relationships between roads, geography, and insurgency have already received some attention.¹⁰⁹ Then what remains to be understood about this area of study?

“War is both timeless and everchanging.”¹¹⁰ So far, the handful of studies that have produced results with findings of significance relating to road networks and insurgent violence have been tied to the Caucasus region of southern Russia,¹¹¹ various countries within Central Africa,¹¹² South America,¹¹³ or viewed at the macro level across

¹⁰⁵ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

¹⁰⁶ Zhukov, 144-156.

¹⁰⁷ Schutte, *Diffusion patterns of violence in civil wars*, 143-152.

¹⁰⁸ Creveld, *Supplying War*.

¹⁰⁹ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156. & Hammond, *Maps of mayhem*, 32-46. & Raleigh, *Population size, concentration, and civil war*, 224-238. & Otero-Bahamon, *Seeing like a Guerrilla*, 198-207. & Müller-Crepon, *Roads to Rule, Roads to Rebel*, 563-590. & O'Loughlin, *The Localized Geographies of Violence*, 178-201.

¹¹⁰ United States Marine Corps, *Warfighting*, 17.

¹¹¹ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156. & O'Loughlin, *The Localized Geographies of Violence*, 178-201.

¹¹² Müller-Crepon, *Roads to Rule, Roads to Rebel*, 563-590. & Raleigh, *Population size, concentration, and civil war*, 224-238.

¹¹³ Otero-Bahamon, *Seeing like a Guerrilla*, 198-207.

several nations.¹¹⁴ Similarly, these studies used cases of insurgencies from the 1990s and earlier. There is no study I could find that attempted to explore this relationship in the Middle East using a post-9/11 insurgency.

There is also a temporal and spatial gap in the current literature. Temporally, research is sparse on how various structural conditions and variables affect the total lifecycle of an insurgency. Current research has focused on factors that initiate insurgent violence, but do not stick around to see how these factors affect a single insurgent group over the decades of its existence.¹¹⁵ Finally, there is a gap in exploring the spatial spread of violence. Current literature primarily seeks to answer the question of what makes insurgence violence fertile but falls short of considering how it spreads once it takes hold.¹¹⁶

Justification for Research

Looking at the factors that encourage the diffusion of insurgent violence has immense potential value for the peace and security of fragile states. The Islamic State (IS) is one of the most infamously prolific insurgencies in the 21st century and is still active in several countries today. At its peak, the speed at which the Islamic State took Iraq surprised the world. At the height of the insurgency, the BBC reported that the Islamic State insurgency controlled roughly 88,000 kilometers, a territory home to almost 8 million people.¹¹⁷ The book *The Jihadist Threat: ISIS, al-Qaeda, and Beyond* called the

¹¹⁴ Hammond, *Maps of mayhem*, 32-46.

¹¹⁵ Lewis, *How Insurgency Begins*. & Weidmann, *Predicting Conflict in Space and Time*, 883-901. & Tollefsen, *Insurgency and Inaccessibility*, 6-25. & Gleditsch, *Contagion or Confusion?*, 215-233.

¹¹⁶ Guardado, *The Seasonality of Conflict*. & Hendrix, *Head for the Hills?*, 345-370.

¹¹⁷ BBC, *The rise and fall of the Islamic State*.

Islamic State the “world’s richest terrorist group.”¹¹⁸ Spreading far and fast, the Islamic State transitioned through multiple lifecycle stages over nearly two decades. Selecting this insurgent group as a case study allows me to examine key theories of insurgent violence diffusion using a modern insurgency, in unique geography, and draw from a rich source of contextual background of help bring clarity of the results of my data.

Exploring how this insurgency infected the nation-state of Iraq, and the level to which existing infrastructure was used to ferry insurgent violence throughout the country, is of grave importance for tailoring future counterinsurgency strategy. Previous work provides foundational understanding of the importance of road networks to insurgent groups. However, this thesis humbly seeks to expand the growing literature on insurgent logistics to create even better predictive models for diffusion, and more measured responses in future conflicts.

¹¹⁸ Levallios, *The Financing of the 'Islamic State' in Iraq and Syria (ISIS)*, 4.

Chapter III

Data and Methodology

Chapter III discusses the designs and methods used to explore the central question of this thesis. It begins by explaining the three major theories for insurgent logistics. Then it lays out the methods used to explore the relationship between these theories and insurgent violence diffusion. The chapter closes with a description of the sources used for data collection, along with a rundown of the data itself. For additional details on creation of the variables for this data set see the Codebook in Appendix 1.

Research Design and Methods

Approaching a topic as ambiguous and everchanging as insurgency, is as author John A. Nagl describes it, like “*Learning to Eat Soup with a Knife*. ”¹¹⁹ Any attempt to draw order from a naturally disordered phenomenon will inevitably be a slow and messy process. To organize and isolate the subject of insurgent violence diffusion this thesis will rely on three theories of insurgent logistics presented by Dr. Zhukov in his paper *Roads and the Diffusion of Insurgent Violence*.¹²⁰ These three theories are competing supply and logistics strategies. When looking at a single insurgency, one can sometimes see a blend of these three theories. However, scholars and practitioners have posited at one point or another that each of these three strategies is the dominant one that best supports insurgency as a whole.

¹¹⁹ Nagl, *Learning to Eat Soup with a Knife*.

¹²⁰ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

Three Theories of Insurgent Logistics

The first theory of insurgent logistics is that it is cellular and self-sustaining. This theory is supported from guerrilla scholars like Mao, who emphasize the isolation and self-reliance of insurgent cells.¹²¹ The presumption that insurgencies are home-grown and rely on locally sourced supplies is echoed by scholars and practitioners alike.¹²² However, this limited resource-pool makes the projection of insurgent violence difficult. In a purely cellular environment, insurgent violence can only spread ideologically, with each new cell forced to forage locally for their own means to produce violence.

The second theory of insurgent logistics is that it is highly flexible and mobile. T. E. Lawrence champions this theory in his descriptions that insurgency must “be an influence, an idea, a thing intangible, invulnerable, without front or back, drifting about like a gas.”¹²³ The goal of this maximum mobility was to limit the Ottoman presence “to an atoll of armed islands amid an Arab-liberated sea.”¹²⁴ Using this strategy, insurgents must be prepared to sustain a highly flexible supply-chain that gives them total freedom of movement. Dr. Zhukov classifies this second model as one with geodesic logistics characteristics that do not consider the human or physical terrain.

To challenge these traditional views, Dr. Zhukov presents a third theory of insurgent logistics. One that is bound to traditional networks. In this case insurgents utilize the same logistical networks that conventional militaries use to project power: roads. While rough terrain offers insurgents shelter from government interference, road networks offer the most efficient form of transportation over difficult ground. Dr. Zhukov

¹²¹ Tse-Tung, *On Guerilla Warfare*, 44.

¹²² Tse-Tung, *On Guerilla Warfare*, 44. & Arnold, *Jungle of Snakes*. & Guevara, *Guerrilla Warfare*, 88.

¹²³ Anderson, *Lawrence in Arabia*, 284.

¹²⁴ Anderson, 284.

argues that the exploitation of traditional road networks is the dominant form of insurgent logistics between all three theories.¹²⁵ For the purposes of this thesis, these three theories can be summarized as (1) Cellular, (2) Geodesic, and (3) Network.

Linking Violence to Populations

In his thesis, Dr. Zhukov emphasizes that “the diffusion of insurgency requires: (1) a location currently experiencing insurgent violence, (2) a location susceptible to insurgent violence, and (3) a channel of communication between the two.”¹²⁶ To measure the spread of insurgent violence, violent events are geographically linked to populated areas to measure the relationship between them. To associate violent events to populated areas, this thesis set parameters of a 15-kilometer radius for cities (populations >100,000), a 10-kilometers for towns (populations 10,000 - 100,000), and a 5-kilometers for villages (populations <10,000). In total there are 11,292 populated areas coded in Geofabrik’s Iraq database. From the Uppsala Georeferenced Event Dataset, only seven of the 4,552 violent events (VEs) fell outside of these parameters; ID 267814, 138820, 120503, 232884, 227325, 238464, and 207493. This leaves 4,545 violent events that are linked to populated areas.

Of the 4,545 violent events in the data set, violence occurred in only 247 populated areas. This is a relatively small 2.19% of the 11,292 villages, towns, and cities categorized by Geofabrik. Within these 247 locations, violence occurs a total of 1,823 times monthly, and there are 703 incidents where violence repeated in a location during a single month. To measure the spread of insurgent violence these 247 populated areas are

¹²⁵ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

¹²⁶ Zhukov, 144-156.

binary categorized as either in a state of peace (P) or violence (V) for each month from August 2003 to December 2020 (209 total months). This binary categorization is coded in the dataset under “outcome” with a “0” representing peace and a “1” representing violence. Temporally disaggregating the 247 locations across 209 months, and binary coding them for peace ($P = 0$) or violence ($V = 1$), gives us a total sample size of $n = 51,623$.

Accounting for Evolution

Some researchers like O’Loughlin and Witmer conclude that the dependence on road networks for insurgent logistics is persistent throughout their lifecycle.¹²⁷ This presumption is also implicitly made in Dr. Zhukov’s analysis through the omission of any significant analysis on the dependence of road networks changing throughout the course of the eight-year scope of his study.¹²⁸ However, when looking at long-term cases of insurgency that span decades, the evolution of means and methods of their supply and logistics is clear.

In El Salvador, Jose Bracamonte and David Spencer study the Farabundo Marti National Liberation Front (FMLN) insurgency from 1970 to 1992. According to Bracamonte and Spencer, “[t]he guerrilla effort in El Salvador was without comparison the best-funded, best-organized, and best-supported guerrilla war ever fought in Latin America.”¹²⁹ Initial military support from Cuba and Nicaragua influenced the insurgents to fight conventionally against the El Salvador military. However, after facing battlefield

¹²⁷ O’Loughlin, *The Localized Geographies of Violence*, 178-201.

¹²⁸ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

¹²⁹ Moroni Bracamonte, *Strategy and Tactics of the Salvadoran FMLN Guerrillas*, 7.

losses the group transitioned to irregular warfare in 1985, then eventually into a legal political opposition party in 1992. While the FMLN failed to overthrow the established government, their flexibility allowed for their survival to the point where “around twenty of the former guerrilla commanders are now members of the Salvadorian National Assembly.”¹³⁰ After suffering conventional military defeat, the FMLN divided their functioning military divisions down into small guerrilla companies and platoons. These dispersed groups were used both as cellular domestic units, and highly mobile attack units. By their nature these units required different logistics and supply sources, sustained by a fusion of methods.

The Liberation Tigers of Tamil Eelam (LTTE) insurgency in Sri Lanka from 1976 to 2009 had a near polar-opposite experience. Beginning as political movement that escalated into an insurgency, the LTTE eventually gained enough momentum to win material support from India.¹³¹ However, once the LTTE achieved the strength to rival the Sri Lankan military conventionally, they abandoned guerrilla tactics and refused to return to them even after external military support was withdrawn.¹³² This led to their eventual erosion of their movement and total annihilation by Sri Lankan forces.¹³³

When looking at the entire lifespan of an insurgent group, strategy and tactics are both shaped by, and shape logistics and supply methods. Unlike other studies on insurgent logistics whose scope only spans a few months,¹³⁴ this thesis seeks to uncover the nuanced balance between the three major theories of insurgent logistics over 18 years.

¹³⁰ Moroni Bracamonte, 37.

¹³¹ Weiss, *The Cage*, 49-50.

¹³² Weiss, 226-227.

¹³³ Weiss, 219.

¹³⁴ Townsley, *Space Time Dynamics of Insurgent Activity in Iraq*, 139-146.

This thesis also aims to determine the significance that each of the three theories of insurgent logistics have on the Islamic State as it evolves over this time.

Measuring Violence Diffusion

This thesis will use two statistical tests to explore the significance of geodesic, cellular, and network models on ISI diffusion: correlation and logistics regression.

Point Biserial Correlation will measure the strength, direction, and significance of correlation between the spread of insurgent violence (*Outcome*) and the independent variables gathered in the data set. Where binary independent variables are used, correlation will be analyzed using the Phi-Coefficient to measure strength and direction of the correlation.

Logistics Regression is a statistical method that can predict the probability of a binary outcome. In this case it is used to predict the probability that a populated area will experience peace (P) or violence (V) depending on one or multiple predictor variables. The variables in this dataset were selected to the model geodesic, cellular, and network theories of insurgent logistics. Logistics regression will tell us how accurately these variables predict the outcome of peace or violence.

The standard temporal measurement of this thesis is months. Running these tests over all 209 months will give us a general idea as to the relationship between ISI violence diffusion and logistics. However, by temporally disaggregating the dataset, the strength of the three models can give better insight into if/how these three models change over time. This thesis breaks up the data set into annual sets of $n=2,964$ to run the models of logistic regression for each year. In these annual tests the first year, 2003, is omitted because it only has partial data for this year. As a form of validity between the temporally

aggregated and annually disaggregated models, the data set was also split into four phases that roughly correspond to major phases in the lifecycle of the ISI.¹³⁵ Phase 1 (2003 – 2006, n=10,127) is considered the “early formation” phase. Phase II (2007 – 2011, n=14,820) is considered the “decline and resurgence” phase. Phase III (2012 – 2015, n=11,856) is considered the “rise and peak” phase. Phase IV (2016 – 2020, n=14,820) is considered the “final decline” phase.

On all models with more than one independent variable a multicollinearity test was run to ensure that no two variables are too-highly correlated which could result in unstable estimates, inflated standard errors, or general interpretation issues. To do this, all variables used in a single model were tested against one another using linear regression. Their resulting R-squared value was used to calculate a variance inflation factor (VIF). A threshold of a VIF of 5 was used to determine whether there was any multicollinearity between variables.

A baseline model that tested the dependent variable *Outcome* against no independent variables was used to establish a baseline -2 Log Likelihood (15771.2) so that other models could be run against the baseline for comparison. A fourth model containing the aggregate of all variables used in Models 1, 2, and 3 was attempted but not considered valid due to multicollinearity between the variables in Model 1 and 2.

Model Breakdown

The following section will describe the independent variables used to indicate the three models of insurgent logistics and describe justification for selection. All models

¹³⁵ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

will be tested against the dependent variable “outcome.” The outcome is a binary variable that represents whether a location experienced a violent event for a given month with “0” representing peace and “1” representing violence. The model for the prediction will be “1” so that the logistics regression test will examine the influence of the independent variables on predicting whether violence will occur.

To measure these models, this thesis recreated several variables used by Dr. Zhukov in his study on the insurgency in the Russian Caucasus from 2000-2008. While recreating these variables will allow me to test Dr. Zhukov’s theories under new conditions, these variables were sourced from different data using different methods and carry some additional features for analysis. For more information on the comparison between Dr. Zhukov’s independent variables and the independent variables used in this thesis, see Appendix 1.

Model 1: Geodesic

To measure the theory of a highly flexible/mobile insurgent logistics model in Iraq this thesis uses the following variable:

- *KM Geo – Nearest Recent VE* – This independent variable measures the distance a populated area is (in kilometers) from the nearest populated area that experienced a violent event in the previous month using geodesic distance.

Model 2(a): Network

To measure the theory of an insurgent logistics model that utilizes traditional road networks I created the following variables:

- *KM Net – Nearest Recent VE* – This independent variable measures the distance a populated area is (in kilometers) from the nearest populated area that experienced

a violent event in the previous month using road networks from the ArcGIS Data and Maps package by Esri.

- *Prox to Major Roads* – This is the proximity (in kilometers) each populated area is to a major road network.
- *Prox to All Roads* – This is the proximity (in kilometers) each populated area is from major, national, regional, and local road networks.

For Model 2 the variables *Prox to Major Roads* and *Prox to All Roads* expressed multicollinearity. As a result, the temporally aggregated Model 2 was split into “2” and “2a.” Model 2 uses *Prox to All Roads*, and Model 2a uses *Prox to Maj Roads*. For simplicity, all temporally disaggregated models only use the predictors *KM – Net Nearest Recent VE*, and *Prox to All Roads*.

Cellular Model

To measure the theory that insurgents source their supplies locally and rely on cellular, independent units, this thesis uses the following variable:

- *Repeat VE* – This is a binary coded variable where “0” indicates that the previous month the populated area did not experience a violent event, and “1” indicates that the populated area experienced a violent event in the previous month and is now experiencing a violent event again.

Additional Variables

Several additional variables were collected as part of the data set to better understand the conditions surrounding this case study. For a complete review of the variables of this dataset including their sources and geoprocessing steps, refer to Appendix 1. Not all variables collected are used in the data set are used in this thesis,

leaving opportunity for further research in future studies. Specifically, data collected on the province-level is omitted because the city/town/village-level focus of the dependent variable (Outcome). The following variables are tested individually using logistics regression and correlation analysis against the dependent variable Outcome.

- *Population Density* – The average population density within 10-kilometers of a populated area (measured annually).
- *Avg Location Elevation m ASL* – The average elevation within a 10-kilometer radius of a populated area in meters above sea level.
- *Harvest Season* – A binary coded variable where “0” indicates a non-harvest month and “1” indicates a harvest month.
- *Election Month* – A binary coded variable where “0” indicates a non-election month and “1” indicates an election month. The variable *Election Month* was not included in the annual temporally disaggregated models because it was not consistently present during each year.

Justification for Chosen Methodology

The purpose of this thesis is not to be a replication, but rather an extension of Dr. Zhukov’s and other’s research on road networks and the diffusion of insurgent violence. Where Dr. Zhukov uses a series of similar variables to test the same three theories of insurgent logistics, this thesis draws from new sources and injects additional nuance into the dataset. Where Dr. Zhukov uses Markov Chain models to simulate the spread of insurgency, this thesis will use correlation and logistics regression in both a temporally aggregated and disaggregated manner. Viewing these results in an aggregated form gives

us an overview of Islamic State violence diffusion and logistics model relationship in Iraq. Breaking the results into annual sections allows us to see if these models compete for dominance in predictive ability, or if the collective predictive strength of the models rise and fall over time. To the extent of my knowledge, no existing research has attempted to test similar variables or models in both a temporally aggregated and disaggregated way to examine not only the strength of the models against one another, but also if and how they change over time.

Data Collection and Analysis

The following sections list the sources and data collected for use in this thesis.

ArcGIS Pro Software

For all geospatial processing I used the desktop version of ArcGIS Pro version 3.2.2. ArcGIS Pro was run using a Windows 10 operating system. This Geospatial Information System is responsible for processing all the Uppsala conflict data, Geofabrik road and population center data, WorldPop population density data, and DIVA-GIS administrative, boundary, and elevation data. Detailed steps on the geoprocessing of this data for the creation of this data set is available in Appendix 1.

Uppsala Data

For conflict data, this thesis used Uppsala georeferenced conflict data. The Uppsala Conflict Data Program (UCDP) is “the world’s main providers of data on organized violence and the oldest ongoing data collection project for civil war, with a

history of almost 40 years.”.¹³⁶ The program is a part of Uppsala University in Sweden. The data is individually indexed and organized geographically and temporally. The UCDP Georeferenced Event Dataset Version 23.1 used in this thesis was downloaded from the Department of Peace and Conflict Research at Uppsala University on 11 December 2023, (<https://ucdp.uu.se/downloads/>).¹³⁷

After downloading the data, I filtered it for the only relevant insurgent group, the Islamic State in Iraq (ISI). I filtered the data further to only display between the years 2003 to 2020. I then filtered the data only to display data gathered for the highest level of confidence for the event precision: “1: exact location of the event known and coded.”¹³⁸ I further filtered the data for the highest value of event clarity: “1 (high) for events where the reporting allows the coder to integer identify the event in full. That is, events where the individual happening is described by the original source in a sufficiently detailed way as to identify individual incidents, i.e. separate activities of fighting in a single location.”¹³⁹ After these filtered were applied to the data set the following sample size remained: n =4,552. Linking this conflict data to populated areas further reduced the sample size to n=4,545.

It is important to remember that these data do not encapsulate the entirety of the conflict during this period. Instead, they are a “snapshot” of the conflict based on what information can be determined as accurate following Uppsala data requirements. It is also important to note that this case is conditioned on a specific insurgency. Between 2003 and 2020 the government of Iraq was engaged in fighting multiple insurgent groups.

¹³⁶ Uppsala University, *About The Uppsala Conflict Data Program*.

¹³⁷ Sundberg, *Introducing the UCDP Georeferenced Event Dataset*, 523-532.

¹³⁸ Stina, *UCDP GED Codebook version 23.1*, 9.

¹³⁹ Stina, 10.

However, this thesis is only following incidents involving the Islamic State in Iraq. Furthermore, as in all conflicts civilian casualties are amounted on both sides of the conflict. However, this thesis is focused on insurgent operations so violence against civilians is only captured when the data pertains to ISI actions.

Geofabrik and Open Street Map Data

Geofabrik (<https://download.geofabrik.de>) is a free to use, community maintained geo-database. The download server extracts geospatial data from Open Street Map “normally updated every day.”¹⁴⁰ Open Street Map, the source data for Geofabrik is an open-source, free to use, community driven global mapping database.¹⁴¹

Geofabrik organizes shapefile packages for download by their respective country and include a wide variety of publicly available information. For my thesis I downloaded the shape file package for the country of Iraq on 20 December 2023 at 11:00pm PST.

For my road data I used Geofabrik’s “gis_osm_free_1.shp” file. Inside this package I filtered the data using the “Select Attributes” feature on ArcGIS Pro to only display the following roads: major roads, freeways, national, regional, and local roads (along with their associated connecting entrance and exit ramps). From this data I divided the road networks into two data sets. The first is “major roads” which consists only of the major roads and freeways (and their associated ramps). The second is “total/all roads” which consists of major roads, freeways, national, regional, and local roads (and associated entrance and exit ramps).

¹⁴⁰ Geofabrik, *About Page*.

¹⁴¹ Open Street Map, *About Open Street Map*.

To associate insurgent violence with populated areas, I used the Geofabrik data file “gis_osm_places_free_1.shp” file. I used the “Select Attributes” features on ArcGIS Pro to filter this data to only display cities, towns, and villages. The file displays this data as geo-referenced point data. According to the Geofabrik Codebook, cities consist of places that contain populations of over 100,000. Towns consist of places that contain populations between 10,000 to 100,000, and villages consist of places that contain populations of less than 10,000.¹⁴² I grouped all these data into one set which I refer to as “populated areas.” The following is the sample size of populated areas for Iraq: n=11,292. This data was then used to link violent events to populated areas. See more about this in the “Linking Violence to Populations” section of this thesis.

WorldPop Data

To better understand population centers this thesis used raster data from www.worldpop.org. WorldPop is a collaborative project for mapping annual population information for countries. Contributors include the School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Department de Geographie, University de Namur; and Center for International Earth Science Information Network (CIESIN), Columbia University (2018).¹⁴³ The Global High Resolution Population Denominators Project is funded by the Bill and Melinda Gates Foundation.¹⁴⁴

¹⁴² Ramm, *OpenStreetMap Data in Layered GIS-Format Version 0.7.12*.

¹⁴³ WorldPop, *About*.

¹⁴⁴ WorldPop.

WorldPop displays the estimated total number of people per grid cell (pixel). The data set is available to download in the Geotiff format at a resolution of 3 arc seconds (approximately 100m at the equator). The projection is Geographic Coordinate System, WGS84. The units are the number of people per pixel. These population estimates are produced using a Random Forest-based dasymetric redistribution.¹⁴⁵

The data set downloaded from this database and used in this thesis is called “Population counts / Unconstrained Individual Countries 2000-2020 (100, resolution)” (<https://hub.worldpop.org/geodata/listing?id=29>). This database includes annual population density information in raster form for the country of Iraq. A total of 18 raster data sets were downloaded to provide accurate population density information for each year of the conflict.

This data was uploaded into ArcGIS Pro then used to calculate the population density per square 100 meters within a 10-kilometer radius around each population center, or (listed on the dataset as) “location.”

DIVA-GIS Data

DIVA-GIS (<https://www.diva-gis.org/gdata>) is a database for country related data collected from several sources.¹⁴⁶ Data taken from this source include administrative boundaries (provinces and districts) in polygon format that are collected by DIVA-GIS from the Database of Global Administrative Areas (GDAM) at <https://gdam.org/>. This thesis also used DIVA-GIS elevation data, which is derived from the Consultative Group

¹⁴⁵ WorldPop, *Population Counts*.

¹⁴⁶ During the writing portion of this thesis the DIVA-GIS underwent maintenance and is currently unavailable.

for International Agricultural Research (CGIAR). CGIAR manages a Shuttle Radar Topography Mission (SRTM) dataset. The SRTM30 dataset is aggregated to a grid cell size of 30 arc(1 km) seconds (<https://srtm.csi.cgiar.org/>).¹⁴⁷

The polygon shape file labeled “adm1” displays the physical boundaries of the 19 provinces in Iraq. These administrative areas were used to calculate several different data points for the final data set. The details of this calculation along with the steps used can be found in the codebook (Appendix 1).

Food and Agricultural Organization of the United Nations

The Food and Agricultural Organization (FAO) is a specialized agency of the United Nations that is responsible for the worldwide collection, validation, and dissemination of data and information related to food and agriculture (<https://www.fao.org/home/en/>).

The FAO data was used to create the binary variable “harvest season” for the data set. If the month was a harvest season the column displayed a “1,” otherwise it displayed a “0.” Data for Iraq was collected from the following webpage: <https://www.fao.org/gIEWS/countrybrief/country.jsp?code=IRQ>. FAO Global Information and Early Warning System country page shows a crop calendar that indicates which months Iraq farmers generally harvest their crops.

¹⁴⁷ United Nations, *FAO GIEWS Country Brief on Iraq*.

Description of the Data

This section provides a description of the data collected for Iraq and used in the statistical analysis portion of this thesis. The following data was sourced from the above-mentioned databases and was processed using ArcGIS Pro version 3.2 software on a Windows 10 desktop.

Iraq Country Data

Iraq is a country with a total size of 455,364.92 square kilometers. There are a total of 66,794.95 kilometers of road in Iraq. Of that, approximately 14,530 kilometers are major roads, and 52,265 kilometers are national, regional, and local roads. In Iraq there are approximately 0.13 kilometers of road per square kilometer of land.

There are a total of 19 provinces in Iraq. The largest is Al Anbar province with 139,864 square kilometers (sq. km). The smallest is Baghdad province at 821 sq. km. The average province area in Iraq is 23,967 sq. km. Al Anbar province contains the most roads at 7,551 kilometers of road networks. Karbala province has the least with only 1,433 kilometers of road. The average kilometers of roads per province is 3,515 kilometers.

Population density is measured at the city, village, and town level using raster data set to a 10-kilometer radius. Population data is updated every year for populated areas. The average population density for populated areas in Iraq is 2.09 per 100 square meters. On average the most densely populated area is Baghdad city with an average population density of 42.18. Some populated areas have an average population density of zero. Population density is rounded to the nearest fourth decimal point (e.g. 0.0001). Collecting temporally disaggregated population density over an 18-year period allows for

population shift during the conflict. This means that some low populated places could become completely abandoned during the conflict, or less populated areas can become denser as internally displaced people migrate away from violence.

Elevation is measured at the city, village, and town level using the average within a 10-kilometer radius via raster data. Elevation is also captured at the province-level to determine not only the pure elevation, but how dramatically elevation changes within a given area. The average elevation for populated areas in Iraq is 177.32 meters above sea level (ASL). The highest populated area is Paynjuwayn town at 1,390 meters ASL. The lowest populated area is al Basrah town at 2.32 meters ASL.

Elections and Harvest data is collected at a national level. There was a total of four elections in Iraq from 2003 to 2020. These elections occurred in December 2005, March 2010, April 2013, and May 2018. According to United Nations data, harvest season in Iraq occurs from May until September with the primary crops being barley, millet, potatoes, sorghum, and wheat.¹⁴⁸

Islamic State Conflict Data

The Iraq data set spans from August 2003 to December 2020 (209 months). It lists violent events for the group the Islamic State (IS). After linking violent events to populated areas, there are a total of 4,545 violent events (n=4,545). See the distribution of IS violent events over time in Figure 1 below.

¹⁴⁸ United Nations, *FAO GIEWS Country Brief on Iraq*.

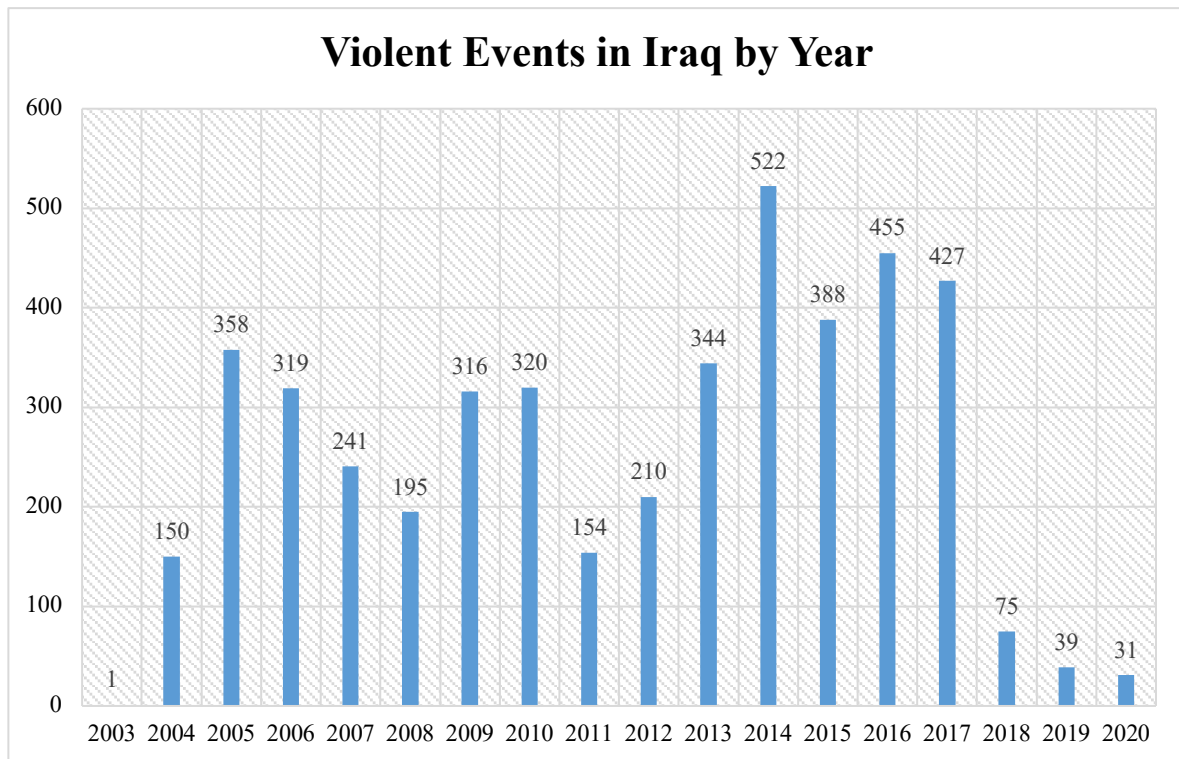


Figure 1. Violent Events in Iraq by Year.

Total number of violent events attributed to the Islamic State in Iraq annually. Source: Uppsala Conflict Data Program GED Dataset.

In Iraq there were an average of 253 violent events each year. Over the years the violent events appear to have three waves that peak in 2005, 2010, and 2014. The year 2014 is the most intense, with 522 battles (almost twice the national average). The final years (2018-2020) show a drastic decline in violence when compared to the rest of the timeframe.

There was an average of 22 attacks every month. March and April of 2017 were the most active months with 78 and 72 violent events respectively. No violent events

were recorded from September 2003 to February 2004, December 2017, April and December of 2019, and September and November of 2020.

Of the 11,292 populated areas, violent events only occurred in 247 (2.19%). Every month an average of 9 locations experienced a violent event. When violence did occur in a location there was a 53% chance it would reoccur, and a 39% chance of reoccurring more than twice. In other words, 118 out of 247 populated areas only experienced one violent event. Thirty-five populated areas experienced two. Of the 4,545 violent events listed, Mosul is the victim of over a third (1,539). Only seven villages listed (Mosul, Baghdad, Fallujah, Ramadi, Baqubah, Telefar, and Kirkuk) are victims to 68% (3,091) of the 4,545 violent events in the dataset. See the distribution of IS violent events by populated area in Figure 2 below.

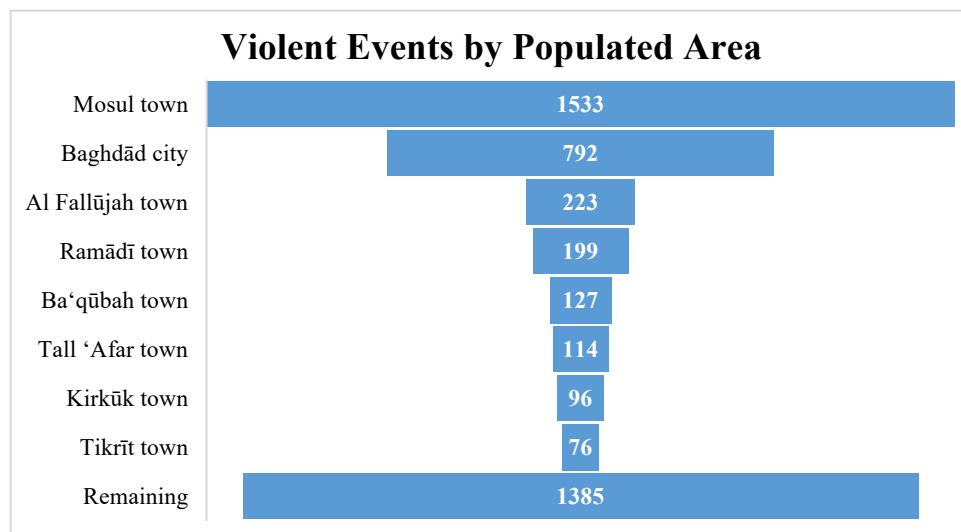


Figure 2. Violent Events by Populated Area.

Total number of violent events conducted by the Islamic State in Iraq. Organized by populated area in descending order. Source: Uppsala Conflict Data Center GED Dataset.

Chapter IV

Results

The following results are broken into four sections: (1) temporally aggregated correlation results, (2) temporally aggregated logistics regression results, (3) temporally disaggregated annual results, and (4) temporally disaggregated phased results.

Temporally Aggregated Correlation Results

The following two tables list the point biserial and bivariate correlation results.

Table 1. Point Biserial Correlation Results.

Point Biserial Correlation	KM - Geo	KM - Net	Prox to Major Roads
Rpb	-0.09	-0.09	-0.07
degrees of freedom	51621	51621	51621
t-statistic	-20.61	-21.57	-15.67
p-value	<0.001	<0.001	<0.001
	Prox to All Roads	Pop Dens	Avg Location Elevation
Rpb	-0.07	0.22	-0.04
degrees of freedom	51621	51621	51621
t-statistic	-15.73	50.5	-8.49
p-value	<0.001	<0.001	<0.001

Point Biserial Correlation Results.

Table 2. Bivariate Correlation Results

Bivariate Correlation	Repeat VE	Harvest Season	Election Month
Phi Coefficient	0.61412222	0	0
Chi-square statistic	19469.412	13.63063	0.004111
p-value	<0.001	0	0
n	51623	51623	51623

Correlation analysis for binary variables.

Model 1: Geodesic

A point-biserial Pearson correlation was run to determine the relationship between the binary variable *Outcome* and the numerical variable *KM – Geo Nearest Recent VE* to determine the strength and direction of the relationship. There was a weak negative correlation, which was significant ($R_{pb} = -0.09$, $n = 51623$, $p < 0.001$, $t = -20.61$).

Model 2: Network

A point-biserial Pearson correlation was run to determine the relationship between the binary variable *Outcome* and the numerical variable *KM – Net Nearest Recent VE* to determine the strength and direction of the relationship. There was a weak negative correlation, which was significant ($R_{pb} = -0.09$, $n = 51623$, $p < 0.001$, $t = -21.57$).

A point-biserial Pearson correlation was run to determine the relationship between the binary variable *Outcome* and the numerical variable *Prox to Major Roads* to determine the strength and direction of the relationship. There was a weak negative correlation, which was significant ($R_{pb} = -0.07$, $n = 51623$, $p < 0.001$, $t = -15.67$).

A point-biserial Pearson correlation was run to determine the relationship between the binary variable *Outcome* and the numerical variable *Prox to All Roads* to determine the strength and direction of the relationship. There was a weak negative correlation, which was significant ($R_{pb} = -0.07$, $n = 51623$, $p < 0.001$, $t = -15.73$).

Model 3: Cellular

To determine the correlation between binary variable *Outcome* and the binary variable *Repeat VE* a Phi Coefficient was calculated along with its associated Chi-square

statistic and p-value. There was a moderately strong positive correlation, which was significant ($\phi = 0.614$, $n = 51623$, $p < 0.001$, $\chi^2 = 19469.4124$).

Additional Variables

Point-biserial Pearson correlation tests were run against the binary variable Outcome and the variables Population Density and Average Location Elevation. The correlation tests resulted in significant findings ($p < 0.001$). Both variables also have the same number of degrees of freedom ($df = 51621$). The variable with the strongest correlation and highest absolute value t-statistic was Population Density ($R_{pb} = 0.22$, $p < 0.001$, $t = 50.5$). Average Location Elevation had the lowest t-statistic absolute value ($R_{pb} = -0.04$, $p < 0.001$, $t = -8.49$).

To determine the correlation between binary variable Outcome and the binary variables Harvest Season, and Election Month, a Phi Coefficient was calculated along with its associated Chi-square statistic and p-value. Harvest Month had no correlation but was significant ($\phi = 0$, $n = 51623$, $p < 0.001$, $\chi^2 = 13.63063$). Election Month had a very weak negative correlation, but was not significant ($\phi = -0.00145$, $n = 51623$, $p = 0.948874$, $\chi^2 = 0.004111$).

Logistics Regression Results Temporally Aggregated

This section displays the logistic regression results temporally aggregated for all three models.

Model 1: Geodesic

Null Hypothesis (H_0): There is no relationship between the predictor variable and the binary outcome variable. $H_0: \beta = 0$.

Alternative Hypothesis (H_A): There is a relationship between the predictor variable and the binary outcome variable. $H_A: \beta \neq 0$.

A binary logistics regression was conducted to examine the effects of the variable *KM – Geo Nearest Recent VE* on the binary dependent variable *Outcome* to predict the value of “1.” The logistic regression model was statistically significant ($\text{Chi}^2(1) = 627.06$, $p < .001$, $n = 51623$). The model correctly classified 96.47% of cases and fits the data reasonably well ($-2 \text{ Log-likelihood} = 15099.14$). However, the model explains a very small amount (1-5%) of the variance in the outcome variable (Cox & Snell $R^2 = 0.01$, Nagelkerke $R^2 = 0.05$, McFadden’s $R^2 = 0.04$). The area under the curve (AUC) value is 0.689. A graph of this AUC equation is included in Appendix 3, Figure 8.

The coefficient of the variable *KM – Geo Nearest Recent VE* is negative ($b = -0.01$). This means an increase in the variable *KM – Geo Nearest Recent VE* is associated with a decrease in the probability that the dependent variable *Outcome* is “1.” The p-value of $< .001$ indicates that this influence is statistically significant. The Odds Ratio of 0.99 indicates that a one unit increase of the predictor variable will increase the odds that the dependent variable is “1” by 0.99 times, with a 95% confidence interval of 0.99 – 0.99. These results reject the null hypothesis and accept the alternative hypothesis.

Model 2: Network

Null Hypothesis (H_0): There is no relationship between the predictor variables and the binary outcome variable. $H_0: \beta = 0$.

Alternative Hypothesis (H_A): There is a relationship between the predictor variables and the binary outcome variable. $H_A: \beta \neq 0$.

A binary logistics regression was conducted to examine the effects of the variables *KM – Net Nearest Recent VE*, and *Prox to All Roads KM* on the binary dependent variable *Outcome* to predict the value of “1.” The logistic regression model was statistically significant ($\text{Chi}^2(2) = 1006.46, p < .001, n = 51623$). The model correctly classified 96.47% of cases and fits the data reasonably well (-2 Log-likelihood = 14764.74). However, the model explains a very small amount (2-7%) of the variance in the outcome variable (Cox & Snell $R^2 = 0.02$, Nagelkerke $R^2 = 0.07$, McFadden’s $R^2 = 0.06$). The area under the curve (AUC) value is 0.736. A graph of this AUC equation is included in Appendix 3, Figure 9.

The coefficient of the variable *KM – Net Nearest Recent VE* is negative ($b = -0.01$). This means an increase in the variable *KM – Net Nearest Recent VE* is associated with a decrease in the probability that the dependent variable *Outcome* is “1.” The p-value of $<.001$ indicates that this influence is statistically significant. The Odds Ratio of 0.99 indicates that a one unit increase of the predictor variable will increase the odds that the dependent variable is “1” by 0.99 times, with a 95% confidence interval of 0.99 – 0.99.

The coefficient of the variable *Prox to All Roads KM* is negative ($b = -0.06$). This means an increase in the variable *Prox to All Roads KM* is associated with a decrease in the probability that the dependent variable *Outcome* is “1.” The p-value of <0.001 indicates that this influence is statistically significant. The Odds Ratio of 0.94 indicates that a one unit increase of the predictor variable will increase the odds that the dependent

variable is “1” by 0.94 times, with a 95% confidence interval of 0.93 – 0.95. These results reject the null hypothesis and accept the alternative hypothesis.

Model 2a: Network (Major Roads)

Null Hypothesis (H_0): There is no relationship between the predictor variables and the binary outcome variable. $H_0: \beta = 0$.

Alternative Hypothesis (H_A): There is a relationship between the predictor variables and the binary outcome variable. $H_A: \beta \neq 0$.

A binary logistics regression was conducted to examine the effects of the variables *KM – Net Nearest Recent VE*, and *Prox to Major Roads KM* on the binary dependent variable *Outcome* to predict the value of “1.” The logistic regression model was statistically significant ($\text{Chi}^2(2) = 1003.68, p < .001, n = 51623$). The model correctly classified 96.47% of cases and fits the data reasonably well ($-2 \text{ Log-likelihood} = 14767.52$). However, the model explains a very small amount (2-7%) of the variance in the outcome variable (Cox & Snell $R^2 = 0.02$, Nagelkerke $R^2 = 0.07$, McFadden’s $R^2 = 0.06$). The area under the curve (AUC) value is 0.727. A graph of this AUC equation is included in Appendix 3, Figure 10.

The coefficient of the variable *KM – Net Nearest Recent VE* is negative ($b = -0.01$). This means an increase in the variable *KM – Net Nearest Recent VE* is associated with a decrease in the probability that the dependent variable *Outcome* is “1.” The p-value of $<.001$ indicates that this influence is statistically significant. The Odds Ratio of 0.99 indicates that a one unit increase of the predictor variable will increase the odds that the dependent variable is “1” by 0.99 times, with a 95% confidence interval of 0.99 – 0.99.

The coefficient of the variable *Prox to Major Roads KM* is negative ($b = -0.06$). This means an increase in the variable *Prox to All Roads KM* is associated with a decrease in the probability that the dependent variable *Outcome* is “1.” The p-value of <0.001 indicates that this influence is statistically significant. The Odds Ratio of 0.94 indicates that a one unit increase of the predictor variable will increase the odds that the dependent variable is “1” by 0.94 times, with a 95% confidence interval of 0.94 – 0.95. These results reject the null hypothesis and accept the alternative hypothesis.

Model 3: Cellular

Null Hypothesis (H_0): There is no relationship between the predictor variable and the binary outcome variable. $H_0: \beta = 0$.

Alternative Hypothesis (H_A): There is a relationship between the predictor variable and the binary outcome variable. $H_A: \beta \neq 0$.

A binary logistics regression was conducted to examine the effects of the variable *Repeat VE* on the binary dependent variable *Outcome* to predict the value of “1.” The logistic regression model was statistically significant ($\text{Chi}^2(1) = 5006.09, p < .001, n = 51623$). The model correctly classified 97.83% of cases and fits the data reasonably well ($-2 \text{ Log-likelihood} = 10765.1$). The model’s pseudo-R-squared tests had a wide level of variance (9-35%) on the outcome variable (Cox & Snell $R^2 = 0.09$, Nagelkerke $R^2 = 0.35$, McFadden’s $R^2 = 0.32$). The area under the curve (AUC) value is 0.693. A graph of this AUC equation is included in Appendix 3, Figure 11.

The coefficient of the variable *Repeat VE* is negative ($b = -3.79$). This means an increase in the variable *Repeat VE* is associated with a decrease in the probability that the dependent variable *Outcome* is “1.” The p-value of 0.996 indicates that this influence is

not statistically significant. The Odds Ratio of 3.48563E+12 with a 95% confidence interval of 0 – infinity indicates that this model's predictive ability is very weak. While the p-value of the indicator is not significant, the overall model is, therefore these results reject the null hypothesis and accept the alternative hypothesis.

Additional Variables

Null Hypothesis (H_0): There is no relationship between the predictor variables and the binary outcome variable. $H_0: \beta = 0$.

Alternative Hypothesis (H_A): There is a relationship between the predictor variables and the binary outcome variable. $H_A: \beta \neq 0$.

A binary logistics regression was conducted to examine the effects of the variables *Population Density*, *Average Location Elevation*, *Harvest Season*, and *Election Month* on the binary dependent variable *Outcome* to predict the value of “1.” Each predictor variable was run using its own model. The isolation of these predictors will show how each interacts with the *Outcome* variable without the potential for interference from other variables.

The logistic regression model for *Population Density* was statistically significant ($\text{Chi}^2(1) = 1059.74, p < .001, n = 51623$). The model correctly classified 96.66% of cases and fits the data reasonably well ($-2 \text{ Log-likelihood} = 14711.46$). However, the model explains a very small amount (2-8%) of the variance in the outcome variable (Cox & Snell $R^2 = 0.02$, Nagelkerke $R^2 = 0.08$, McFadden's $R^2 = 0.07$). The area under the curve (AUC) value is 0.688.

The coefficient of the variable *Population Density* is positive ($b = 0.1$). This means an increase in the variable *Population Density* is associated with an increase in the

probability that the dependent variable *Outcome* is “1.” The p-value of <0.001 indicates that this influence is statistically significant. The Odds Ratio is 1.11, with a 95% confidence interval of 1.1 – 1.2. These results reject the null hypothesis and accept the alternative hypothesis.

The logistic regression model for *Average Location Elevation* was statistically significant ($\text{Chi}^2(1) = 86.3, p < .001, n = 51623$). The model correctly classified 96.47% of cases and fits the data reasonably well ($-2 \text{ Log-likelihood} = 15684.89$). However, the model explains a very small amount (1%) of the variance in the outcome variable (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0.01$, McFadden’s $R^2 = 0.01$). The area under the curve (AUC) value is 0.564.

The coefficient of the variable *Average Location Elevation* is zero ($b = 0$). This means that the variable *Average Location Elevation* has no effect on the dependent variable *Outcome* is “1.” The p-value of <0.001 indicates that this influence is statistically significant. The Odds Ratio is 1, with a 95% confidence interval of 1 – 1. Because this variable’s $b = 0$ these results accept the null hypothesis and reject the alternative hypothesis.

The logistic regression model for *Harvest Season* was statistically significant ($\text{Chi}^2(1) = 14.5, p < .001, n = 51623$). The model correctly classified 96.47% of cases and fits the data reasonably well ($-2 \text{ Log-likelihood} = 15756.66$). However, the model explains virtually none (0%) of the variance in the outcome variable (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0$, McFadden’s $R^2 = 0$). The area under the curve (AUC) value is 0.364.

The coefficient of the variable *Harvest Season* is positive ($b = 0.18$). This means an increase in the variable *Harvest Season* is associated with an increase in the

probability that the dependent variable *Outcome* is “1.” The p-value of <0.001 indicates that this influence is statistically significant. The Odds Ratio is 1.2, with a 95% confidence interval of 1.09 – 1.32. These results reject the null hypothesis and accept the alternative hypothesis.

The logistic regression model for *Election Month* was statistically significant ($\text{Chi}^2(1) = 0.11, p < .001, n = 51623$). The model correctly classified 96.47% of cases and fits the data reasonably well ($-2 \text{ Log-likelihood} = 15771.08$). However, the model explains virtually none (0%) of the variance in the outcome variable (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0$, McFadden’s $R^2 = 0$). The area under the curve (AUC) value is 0.019.

The coefficient of the variable *Election Month* is negative ($b = -0.06$). This means an increase in the variable *Election Month* is associated with a decrease in the probability that the dependent variable *Outcome* is “1.” The p-value of 0.742 indicates that this influence is not statistically significant. The Odds Ratio is 0.94, with a 95% confidence interval of 0.66 – 1.34. These results reject the null hypothesis and accept the alternative hypothesis.

Logistics Regression Temporally Disaggregated Annual Results

The AUC from models 1, 2, and 3 are displayed annually from 2004 to 2020 in

Figure 3 below.

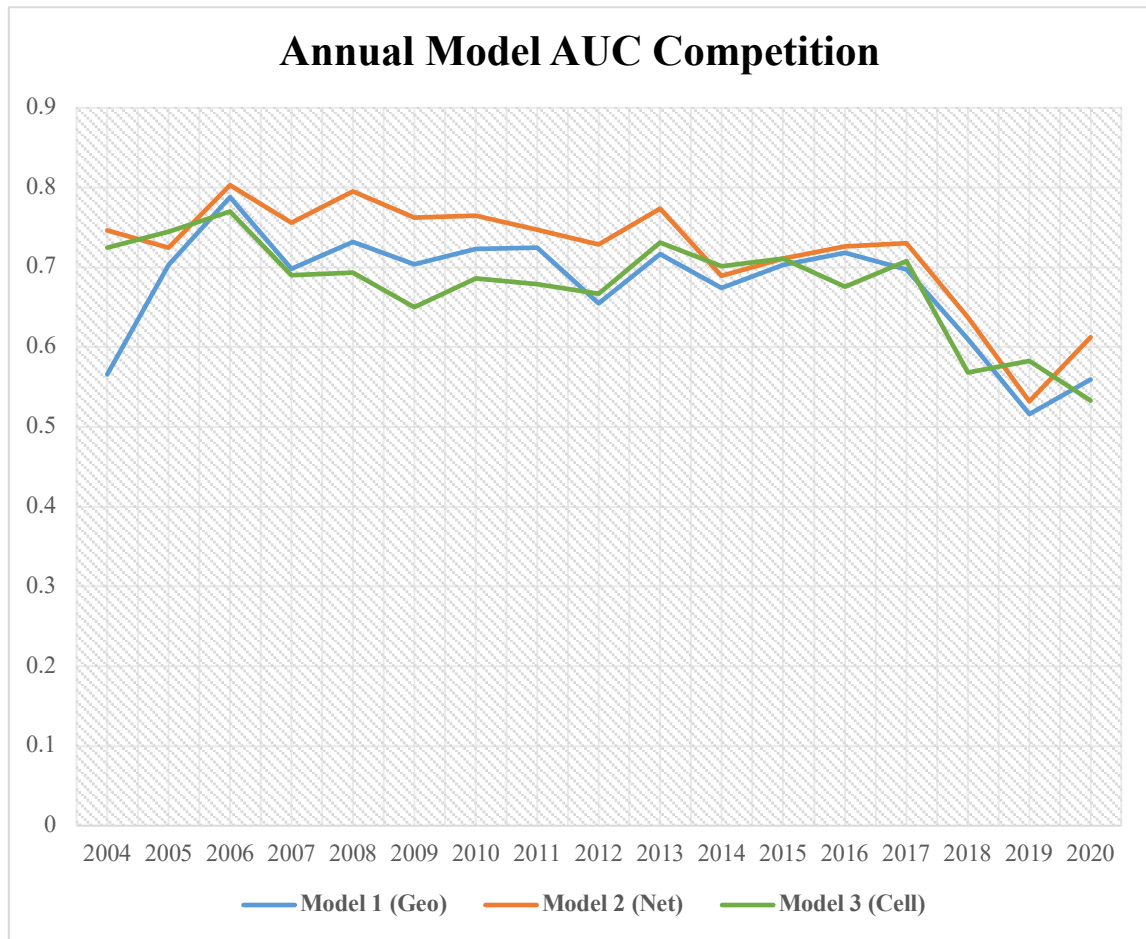


Figure 3. Annual Model AUC Competition.

Temporally disaggregated annual results displaying Area Under the Curve (AUC) for each model.

In these models the number of cases remains constant ($n = 2964$). The primary measurement used for comparing the strength of the models against one another as well as across time is the Area Under the Curve (AUC). The AUC is a metric taken from the Receiver Operating Characteristic (ROC) curve; a graph that depicts how close the model is to predicting the True Positive Rate (TPR) over various threshold settings. The AUC is a metric between 0 and 1 that is representative of the space between the TPR and the False Positive Rate (FPR). The closer the AUC is to 1, the better the model is at distinguishing between binary events (e.g. the likelihood of peace or violence).

Model 1: Geodesic

From 2005 to 2020 the Geodesic model remained relatively consistent. The model was unreliable in 2004, 2019, and 2020 where the model's dropped in significance ($p = 0.04, 0.245, 0.05$) as well as reported a odds ratio of 1, and a coefficient-b of 0. From 2005 to 2019 the model remained significant ($p < .001$), and the coefficient-b remained between -0.01 and -0.02. The AUC peaked in the middle (2012, $AUC = 0.725$) and was lowest on either end (2004, 2019, 2020 $AUC = 0.566, 0.516, 0.559$) Overall the model remained steady throughout the years with no major deviances in variation.

Model 2: Network

Model 2 displayed similar characteristics to Model 1, with a 2004, 2005, 2019, and 2020 seeing a drop in significance ($p = 0.034, 0.005, 0.304, 0.038$) as well as Odds ratios of 1 and coefficient-b's of 0 for the variable *KM – Net Nearest Recent VE*. The variable *Prox All Roads KM* remained significant until 2018, 2019, and 2020 ($p = 0.961, 0.445, 0.411$). Like Model 1, variable *KM – Net Nearest Recent VE* had a coefficient-b

between -0.01 and -0.02. However, the variable *Prox All Roads KM* had larger variance with a coefficient-b between -0.03 and -0.2 in significant model variations (2016 and 2006). The AUC appears higher than Model 1 across all models with the highest reaching 0.803 in 2006, and the lowest at 0.532 in 2019. While there was slightly more variance among results in Model 2 compared to Model 1, the Model 2 also remains steady across the years.

Model 3: Cellular

Like in the Temporally Aggregated Model, Model 3's predictor value *Repeat VE*, shows no significance ($p = 0.996 - 0.998$) with a high Odds ratio, and a 95% Confidence Interval of 0 – infinity that makes the model generally unreliable. These results remained consistent across all 17 years in its performance with only small variations in other metrics. The Coefficient-b was strongly positive and fell between 24.36 and 27.03. The AUC appears roughly bimodal in its distribution, with a peak in 2006 (AUC = 0.803) and 2013 (AUC = 0.774). The lowest AUC is recorded in 2019 (AUC = 0.532).

Additional Variables

The annual AUC results for additional variables are added to the line graph of the annual AUC model results to make Figure 4 below.

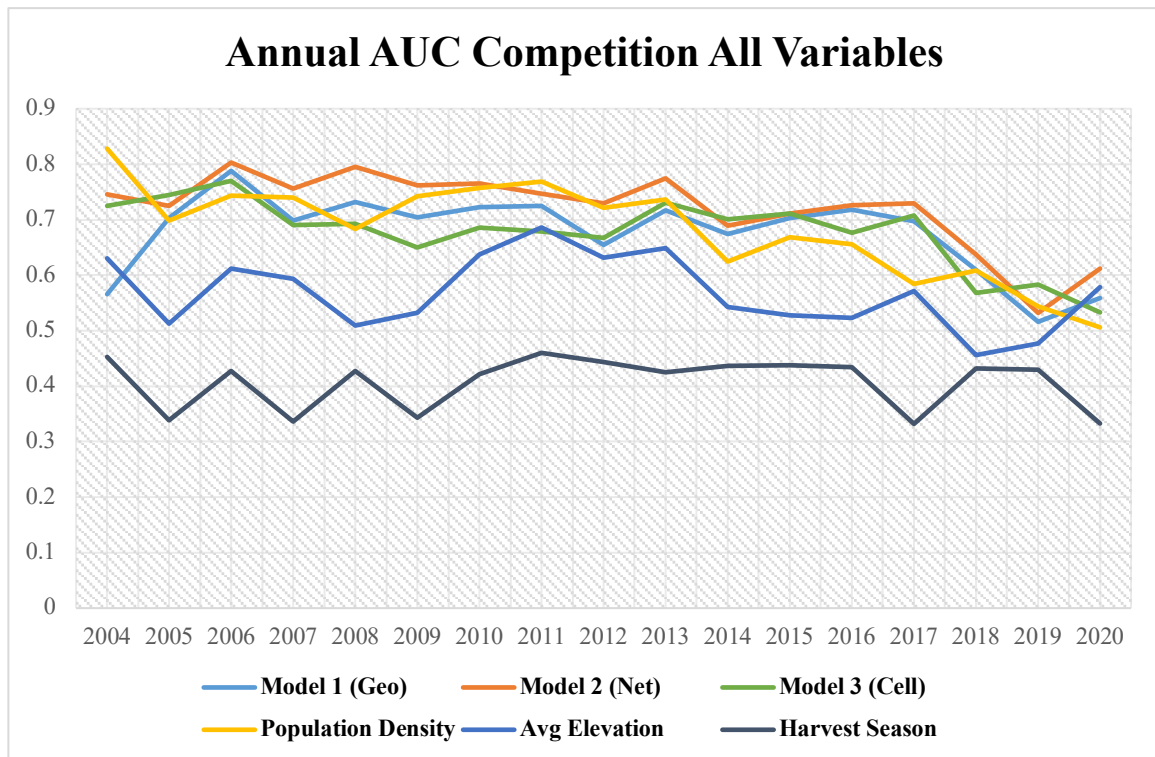


Figure 4. Annual AUC Competition All Variables.

Temporally disaggregated annual results including additional variables.

Three additional variables were placed in individual models and tested annually from 2004 to 2020. These variables included *Population Density*, *Average Location Elevation*, and *Harvest Season*. The variable *Election Month* from the temporally aggregated model was removed due to its irregular appearance on an annual basis.

Of the 3 variables, only one, *Population Density*, was found significant ($P < 0.001$) throughout all 17 years except 2020 ($p = 0.141$). The coefficient-b is positive and fell between 0.03 (2020) and 0.16 (2004). For this variable the Odds ratio fell between 0.99 and 1.19 with 95% Confidence Intervals with variance of less than 1. The pseudo-r-squared test results indicated that the annual models explained between 1-16% of the

variance in the outcome, apart from the year 2020 where these tests indicated that the model explained 0% of the variance in the outcome (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0$, McFadden's $R^2 = 0$).

The variable *Average Location Elevation* was significant in 2004, 2006, 2007, 2010, 2011, 2012, 2013 and 2014 ($p = 0.012, <0.001, 0.025, <0.001, <0.001, <0.001, <0.001, 0.031$). However, only the year 2011 had a coefficient-b that was not zero ($b = -0.01$). For this year the Odds ratio was 0.99 with 95% Confidence Intervals between 0.99 and 1. The pseudo-r-squared test results indicated that the annual models explained between 1-4% of the variance in the outcome, apart from the year 2014 where these tests indicated that the model explained 0% of the variance in the outcome (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0$, McFadden's $R^2 = 0$).

The variable *Harvest Season* was significant in 2010 and 2014 ($p = 0.004, 0.023$). Both times the coefficient-b was positive (0.61 and 0.31). The Odds ratio was 1.85 for 2010, and 1.36 for 2014 with 95% Confidence Intervals of 1.22 – 2.79 for 2010 and 1.04 – 1.78 for 2014. In 2010 the pseudo-r-squared tests indicated that the model accounted for only 1% (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0.01$, McFadden's $R^2 = 0.01$) of the variance in the *Outcome* variable. In 2014 the pseudo-r-squared tests indicated that the model accounted for none (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0$, McFadden's $R^2 = 0$) of the variance in the *Outcome* variable.

Logistics Regression Temporally Disaggregated Phased Results

Figure 5 below compares the AUC for all three models during a phased basis.

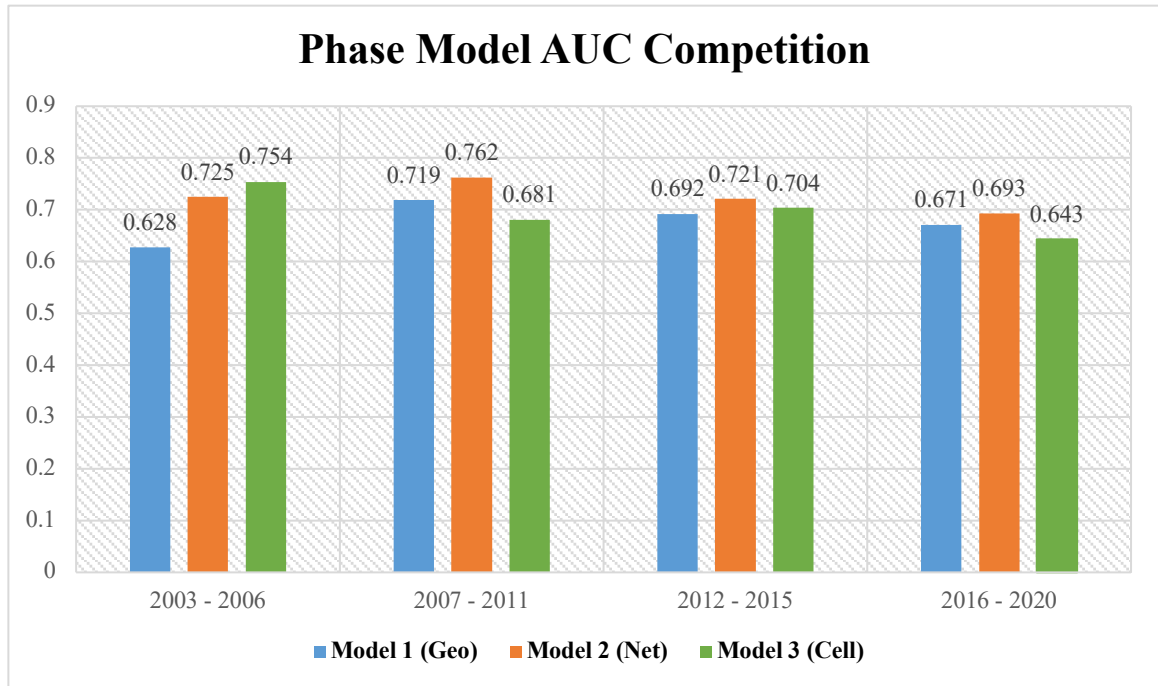


Figure 5. Phase Model AUC Competition.

Temporally disaggregated phased results displaying the Area Under the Curve (AUC) for each model.

Similarly to the annually disaggregated logistics regression models, the phased models are compared using the AUC. However, in the temporally phased section the Additional Variables model is also compared to Model 1, 2, and 3 using its AUC. While the phased models have a larger number of cases to draw from, they are unequal; 2003 – 2006 ($n = 10127$), 2007 – 2011 ($n = 14820$), 2012 – 2015 ($n = 11856$), and 2019 – 2020 ($n = 14820$).

Model 1: Geodesic

Model 1 is statistically significant in all four phases ($p < 0.001$) with Odds ratios between 0.98 and 0.99 and a 95% Confidence Interval between 0.99 and 1. However, the Chi-Squared Test Statistic and z-factor are much higher in phases 2 and 3, than 1 and 4. This means that Model 1 is more reliable between 2007 and 2015 than it is in the front and tail end of the insurgency. This aligns to the annual disaggregation's results.

Model 2: Network

Model 2 is statistically significant in all four phases ($p < 0.001$). However, in phase 1 (2003 – 2006) the Odds ratio is 1 and the Coefficient b is 0, making the model unreliable for this phase. The model performs well in the other three phases with little variance. The coefficient b for phases 2 – 4 is -0.01, with an Odds ratio of 0.99 and 95% Confidence Interval of 0.99 – 0.99. However, phase 4 has a slightly more variant 95%CI at 0.97 - 1. Model 2's performance is reflective of the annual disaggregated models and remains generally consistent to the temporally aggregated model as well.

Model 3: Cellular

Like in all other versions, Model 3 shows no significance ($p = 0.996 – 0.998$) with a high Odds ratio, and a 95% Confidence Interval of 0 – infinity that makes the model generally unreliable. However, the model remained consistent across all four phases in performance, with only small variations in other metrics. Once again, the Coefficient-b was strongly positive and fell between 27.3 and 28.22. The AUC appears bimodal again in its distribution, with a peak in phase 1 (AUC = 0.754) and phase 3 (AUC = 0.704). The trough AUCs in phase 2 and 4 (AUC = 0.681, 0.643).

Additional Variables

In Figure 6 below, the AUC of additional variables are included beside the AUC for the three models during all phases.

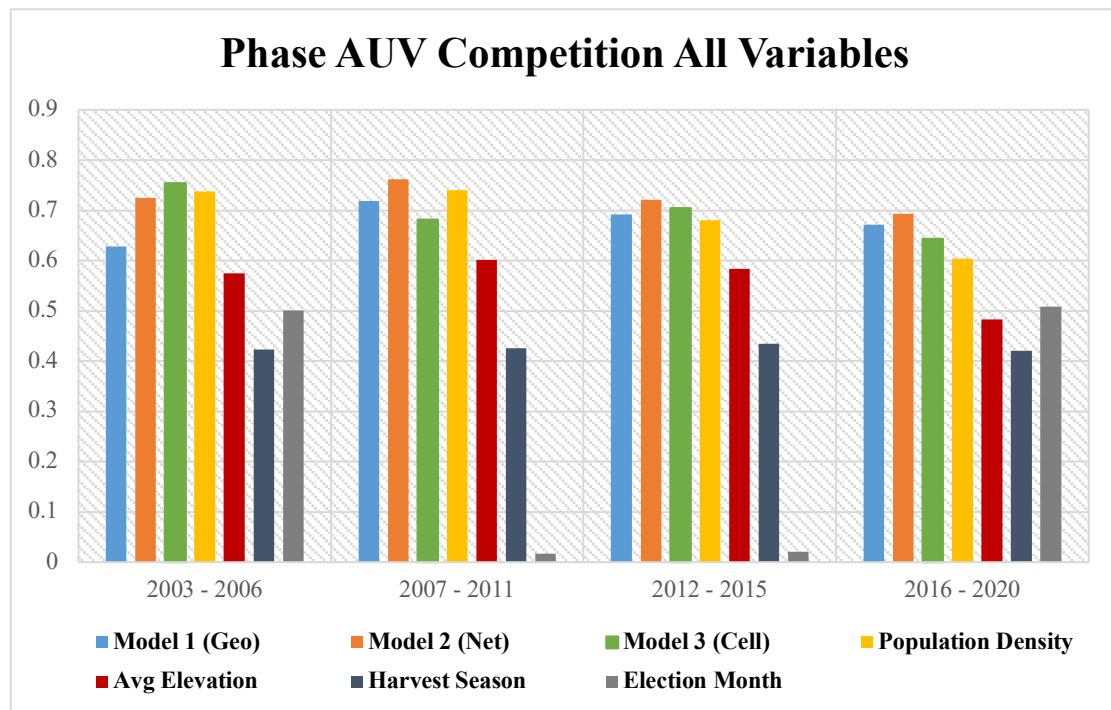


Figure 6. Phase AUC Competition All Variables.

Temporally disaggregated phased variables displaying their associated Area Under the Curve (AUC).

Three additional variables were placed in individual models and tested across all four phases. These variables included *Population Density*, *Average Location Elevation*, *Harvest Season*, and *Election Month*.

Of the four variables, only one, *Population Density*, was found significant throughout all four phases. The coefficient-b is positive and fell between 0.07 (Phase 4)

and 0.14 (Phase 1). For this variable the Odds ratio fell between 1.08 and 1.15 with 95% Confidence Intervals with variance less than 1. The pseudo-r-squared test results indicated that the phased models explained between 1-10% of the variance in the outcome via Cox & Snell R^2 , Nagelkerke R^2 , and McFadden's R^2 test results.

The variable *Average Location Elevation* was significant in Phases 1, 2, and 3 ($p < 0.001$). However, the coefficient-b was zero, meaning the variable has no effect on the outcome. The Odds ratio was 1 with 95% Confidence Intervals of 1 - 1. The pseudo-r-squared test results indicated that the phased models explained 0% of the variance in the outcome (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0$, McFadden's $R^2 = 0$).

The variable *Harvest Season* was significant only in Phase 3 (2012 – 2015) ($p = 0.001$). The coefficient-b was positive (0.26) and the Odds ratio was 1.3 with 95% Confidence Intervals of 1.11 – 1.52.

The variable *Harvest Season* was significant only in Phase 3 (2012 – 2015) ($p = 0.001$). The coefficient-b was positive (0.26) and the Odds ratio was 1.3 with a 95% Confidence Intervals of 1.12 – 1.51. However, The pseudo-r-squared test results indicated that the phased models explained 0% of the variance in the outcome (Cox & Snell $R^2 = 0$, Nagelkerke $R^2 = 0$, McFadden's $R^2 = 0$).

The variable *Election Month* was statistically significant in Phase 3 and 4 ($p = 0.033, 0.009$). In Phase 3 the coefficient-b was negative ($b = -2.71$) with an Odds ratio of 0.44 and 95% Confidence Interval of 0.21 – 0.94. In Phase 4 the coefficient-b was negative ($b = -3.64$) with an Odds ratio of 2.12 and a 95% Confidence Interval of 1.2 – 3.74.

Chapter V

Discussion

The discussion portion of this thesis will begin with an overview of the Islamic State insurgency with a geographic focus on Iraq. Part 2 will interpret the results in Chapter IV against the Islamic State insurgency narrative discussed in Part 1. Part 3 will compare the results in Chapter IV to other academic work on insurgent logistics and the spread of insurgent violence. Part 4 discusses the implication of the findings in this thesis, and Part 5 discusses thesis limitations.

The Evolution of the Islamic State

Insurgent groups are comparatively far less structured than their conventional military counterparts. A military is often a composite of branches that take on defined and organized roles. For example, the U.S. Department of Defense includes the Army, Navy, Air Force, Marines, Space Force, and Coast Guard. Like a conventional military, an insurgency is usually formed from several sub-groups. However, these groups are often loosely defined and come together with competing strategies, tactics, agendas, and leadership. For example, the El Salvadorian FMLN insurgent group was comprised of the Popular Revolutionary Army (ERP), National Resistance (RN), and the Popular Liberation Front (FPL).¹⁴⁹ Where there is an established institutional hierarchy inside conventional militaries, underground and revolutionary insurgencies do not have these same barriers that prevent fluid internal changes in leadership, structure, and allegiance.

¹⁴⁹ Moroni Bracamonte, *Strategy and Tactics of the Salvadoran FMLN Guerrillas*, 14-15.

This makes insurgencies resilient to frequent leadership change but also inherently unstable. The Islamic State in Iraq is no exception.

Early Beginnings

The Islamic State was born as a sub-sect of Al-Qaeda in Iraq (AQI) between 2003 and 2004.¹⁵⁰ AQI formed when Abu Musab Al-Zarqawi merged his insurgent group, Jama 'at al-Tawhid wal-Jihad, with Osama bin Laden's Al-Qaeda.¹⁵¹ AQI quickly grew with a new wave of disillusioned, disenfranchised, and unemployed Sunni Muslims. The ill-fated U.S. policy of de-Baathification gave AQI an initial boost of newly unemployed recruits.¹⁵² Many of these new recruits came from Saddam's former military intelligence officers, who would later lead the Islamic State through its most active years.¹⁵³ The Shia-led government that replaced Saddam drove increasingly neglectful policy against their former ruling-minority Sunni counterparts. This animosity continued to sustain insurgent recruitment for decades.¹⁵⁴

The Mujahideen Shura Council (MSC), a sub-group inside AQI, is where the Islamic State got its start. The MSC itself was an amalgamation of several smaller extremist groups including Jama'at Jaysh Ahl al-Sunna wal-Jama'ah founded by Abu Bakr al-Baghdadi (future first Caliph of the Islamic State).¹⁵⁵ Abu Musab Al-Zarqawi, leader of AQI, held the dual-title of Emir for the MSC and AQI from 2003 until his death in 2006. As the Emir, Zarqawi led the insurgency using the strategic vision of "Bin

¹⁵⁰ Hassan, *How Did ISIS Really Emerge?* & Williams, *ISIS vs. Al Qaeda: Jihadism's global civil war*.

¹⁵¹ U.S. Office of the Director of National Intelligence, *National Counterterrorism Center*.

¹⁵² Zinn, *Consequences of Iraqi De-Baathification*, 80.

¹⁵³ Barfi, *The Military Doctrine of the Islamic State and the Limits of Ba'athist Influence*, 18-20.

¹⁵⁴ Watson, *The Conflict with ISIS*, 9-15.

¹⁵⁵ Orton, *Governing the Caliphate*, 20.

Ladenism.”¹⁵⁶ This strategy focused on the “far enemy” of the United States, while appealing to mass-Muslims during the U.S. occupation of Iraq.¹⁵⁷ “Bin Ladenism” resulted in an operational environment that favored guerrilla warfare with targeted strikes against U.S. Forces, the new Iraqi Government, and Shia Muslims who held onto power as the popular-majority in Iraq. AQI tactics included the frequent use of improvised explosive devices (IEDs), suicide bombings, and assassinations with the end-goal of exhausting the resolve of U.S. Forces and the American public.¹⁵⁸ Bin Laden and al-Zarqawi drew their strategy from their experience fighting the Soviets in Afghanistan. Both leaders saw parallels in the U.S. Invasion of Iraq and believed that a prolonged popular-war would slowly bleed the U.S. into complete economic and political collapse.¹⁵⁹

The death of al-Zarqawi in June of 2006 led to a major strategic pivot within the insurgency. Abu Ayyub al-Masri succeeded al-Zarqawi as Emir for the MSC.¹⁶⁰ By October 2006, the MSC disbanded to become the Islamic State in Iraq (ISI). Shortly thereafter, Abu Ayyub al-Masri stepped aside for Abu Omar al-Baghdadi to be appointed as the ISI Emir.¹⁶¹ Both of these new leaders were influenced by religious predictions of an immediate apocalyptic reckoning. This belief created a new urgency to reestablish an Islamic Caliphate, prepare for the return of the Prophet Muhammad, and fight a final battle between good and evil.¹⁶² To realize these objectives both Abu Ayyub al-Masri and

¹⁵⁶ Kamolnick, *The Al-Qaeda Organization and the Islamic State Organization*, 26.

¹⁵⁷ Kamolnick, 27.

¹⁵⁸ Kamolnick, 27.

¹⁵⁹ Kamolnick, xviii.

¹⁶⁰ Abu Ayyub al-Masri is also known as Abu Hamza al-Muhajir

¹⁶¹ Kamolnick, 96-97.

¹⁶² Kamolnick, 96-97.

Abu Omar al-Baghdadi were pivotal in turning ISI strategy away from AQI “Bin Ladenism.”

As the name suggests, ISI became more focused on establishing a territory for general autonomy inside Iraq. To do this ISI leadership began pivoting away from focusing on the “far enemy” of the West. While doing so they dropped their attempts to appeal to a wider base of Muslims needed to foment a popular war against the United States. Instead, “apostates,” or those Sunni Muslims who did not adhere to the ultra-conservative Salafi-Wahhabi Jihadi beliefs carried by ISI, were denounced as “takfeer” (excommunicated) and targeted with violence.¹⁶³ Violence against other ethnic and religious minorities also deemed “apostates” proliferated, as ISI fighters pivoted to attack the “near enemy.”

Despite this strategic pivot, U.S. counterterrorism efforts in Iraq were in full swing by 2007. Aggressive counterterrorism operations kept ISI from posing a grave threat to regional stability. While these efforts constrained ISI to a continuation of guerrilla warfare, their stature as a sub-group within the larger AQI network meant that they were not the primary focus for counterterrorism operations. During this time, U.S. Forces ultimately cured the intra-group competition inside Iraq as it fought to quell ISI’s larger parent structure AQI.¹⁶⁴

¹⁶³ Kamolnick, 50.

¹⁶⁴ Watson, *The Conflict with ISIS*, 5.

The Rise

On April 18th, 2010, both Abu Omar al-Baghdadi and Abu Ayyub al-Masri were killed by U.S. Forces.¹⁶⁵ In their wake ISI leadership appointed Abu Bakr al-Baghdadi as the new Emir on 16 May 2010.¹⁶⁶ The appointment of Abu Bakr al-Baghdadi could not come at a more pivotal time for ISI. The confounding of several external events would clear the insurgent group's path for growth.

On 17 December 2010, a street vendor in Tunisia set himself on fire outside of a government building in protest over systemic state corruption.¹⁶⁷ This act of rebellion sparked a wave of protests and general instability across the Middle East known as the Arab Spring.¹⁶⁸ Three months later in March 2011, ten school-aged boys no older than fifteen were arrested by local police in Syria for spray painting pro-democracy graffiti.¹⁶⁹ They were tortured in prison. In response the local community came out to protest their arrest.¹⁷⁰ The violent government response to these protests snowballed Syria into a full-blown civil war by mid-2012.¹⁷¹ Meanwhile in neighboring Iraq, the last 500 U.S. combat troops withdrew on 18 December 2011, turning over all internal security to Iraqi Security Forces.¹⁷²

This perfect storm of security degradation in both Syria and Iraq meant a new opportunity for ISI expansion. While the governments of Iraq and Syria recognized the arbitrary line that divides their two nations, ISI saw the Sykes-Picot line drawn by a Brit

¹⁶⁵ Hanna, *Timeline: The Life and Death of Abu Bakr al Baghdadi*.

¹⁶⁶ Hanna.

¹⁶⁷ Gelvin, *The Arab Uprisings*, 27.

¹⁶⁸ Gelvin, 126-127.

¹⁶⁹ Gelvin, 126-127.

¹⁷⁰ Gelvin, 126-127.

¹⁷¹ Simon-Skjodt Center for the Prevention of Genocide, *From a Democratic Uprising to Full-Scale War*.

¹⁷² Watson, *The Conflict with ISIS*, 7.

and a Frenchman after World War I as little more than an extension of colonialism meant to divide and weaken Sunni Muslims.¹⁷³ In July of 2011 Baghdadi sent Abu Muhammad al Julani into Syria to form a new AQI sub-group called al-Nusra Front.¹⁷⁴ By April 2013, Abu Bakr al-Baghdadi announced that ISI will merge with the Syrian AQI sub-group al-Nusra Front to form the Islamic State in Iraq and Syria/the Levant (ISIS/L).¹⁷⁵ This announcement came as a surprise to AQI and al-Nusra Front leadership who both rejected this new merger. Rejection came largely because of ideological differences, but partially out of fear that ISIS/L was starting to eclipse AQI. On December 30th of 2013 ISIS/L took over Fallujah – a major city in Iraq.¹⁷⁶ By January 2014 ISIS/L and al-Nusra Front were in open-conflict with one another over control of parts of Syria, with many al-Nusra Front fighters defecting to join ISIS/L.¹⁷⁷ During this time ISIS/L captures the Syrian city of Raqqa and declared it as their capital.¹⁷⁸ In February of 2014, AQI publicly disowns ISIS/L.¹⁷⁹ The two groups formally split with the AQI still focused on the “far enemy” and ISIS/L focused on the capture of newly destabilized territory from a weakened security state in both Iraq and Syria.¹⁸⁰

During the summer of 2014 ISIS/L captured large swaths of Iraq in a blitzkrieg. The capture of ground occurred linearly. Early access to oil and gas refineries meant that the insurgency had a large stockpile of fuel to support troop movement and logistical

¹⁷³ Al Jazeera, *A century on: Why Arabs resent Sykes-Picot*.

¹⁷⁴ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*. & Blake, *Al Nusra Front*.

¹⁷⁵ ISIS and ISIL are interchangeable acronyms where the last two letters denote the geographic regions the group proclaims to control.

¹⁷⁶ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

¹⁷⁷ Glenn.

¹⁷⁸ Glenn.

¹⁷⁹ Glenn.

¹⁸⁰ Kamolnick, *The Al-Qaeda Organization and the Islamic State Organization*, 95.

sustainment.¹⁸¹ Mosul, the second largest city in Iraq, fell to ISIS/L on June 10th, 2014.¹⁸² Tikrit, the childhood home of Saddam Hussein, fell to ISIS/L on June 11th.¹⁸³ On July 18th the government of Iraq formally began the request for U.S. airstrikes on ISIS/L positions, and before the end of the month the Iraqi cities of Tal Afar, Hawija, Rawa, Qaim, and Baiji all fell to ISIS/L.¹⁸⁴ On the 29th of June 2014, Abu Bakr al-Baghdadi declares ISIS/L a Caliphate under new branding of the all-inclusive “Islamic State” (IS).¹⁸⁵ For this new Caliphate, Abu Bark al-Baghdadi is appointed as the group’s first Caliph.¹⁸⁶

In both Iraq and Syria, IS employed a hybrid strategy of conventional tactics drawn from former Ba’ath Party Intelligence Officers mixed with radical suicidal fanaticism of IS fighters.¹⁸⁷ In Iraq insurgents would launch a coordinated attack with a suicide vehicle-borne improvised explosive device (SVBIED) to create chaos, then follow-through with a combined armed assault usually using small-arms, rockets, and indirect fire systems.¹⁸⁸ While specific tactics varied, insurgents typically stormed positions from multiple directions in small groups, occasionally substituting suicide bombers for artillery.¹⁸⁹ The rapid on-set of ISI’s offensive challenged Iraqi Security Forces on a conventional level for which they were unprepared. Therefore, despite being outnumbered in many of their engagements, ISI attacks caused a near total collapse of Iraqi Security Forces.

¹⁸¹ Levallios, *The Financing of the 'Islamic State' in Iraq and Syria (ISIS)*, 9.

¹⁸² Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

¹⁸³ Glenn.

¹⁸⁴ Glenn.

¹⁸⁵ Glenn.

¹⁸⁶ Byman, *ISIS vs. Al Qaeda: Jihadism's global civil*.

¹⁸⁷ Barfi, *The Military Doctrine of the Islamic State and the Limits of Ba’athist Influence*, 18-20.

¹⁸⁸ Barfi, 18-20

¹⁸⁹ Barfi 19.

After the establishment of the Caliphate IS continued expansion in both Syria and Iraq. In Iraq IS captured the city of Sinjar between August 2nd and 3rd of 2014.¹⁹⁰ Sinjar is home to the ethnic minority Yazidi population, who fled to the highlands with IS fighters in pursuit. On the August 8th, 2014, U.S. military aircraft conducted kinetic strikes on IS artillery positions bombarding Yazidi civilians sheltering in the mountains. This event marked the re-entry of U.S. and Joint-Coalition Forces into Iraq, Operation Inherent Resolve begins.¹⁹¹

The Fall

In 2014 the United States re-enters Iraq with a new operational approach of indirect assistance and air support for local forces on the ground. While the U.S. provided strike and re-supply capability from the air, it also worked to re-train/re-build Iraqi Security Forces.¹⁹² U.S. airstrikes on IS oil refineries as well as IS financial and logistical networks led to an increased stress on the IS supply chain.¹⁹³

With battlefield conditions shifting, IS leadership makes a public call on September 22nd, 2014, for IS supporters abroad to attack U.S. Citizens, France, and other European countries abroad. This marked the beginning of a two-front war where IS fighters would continue to capture and defend territory in Iraq and Syria, but also plan, coordinate, and assist in “far enemy” attacks abroad. After this decision IS activity took place in at least 36 countries around the world including; Bosnia, Honduras, Sweden, Australia, Iran, Argentina, Spain, Lebanon, Brazil, Malaysia, the United Kingdom,

¹⁹⁰ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

¹⁹¹ Watson, *The Conflict with ISIS*, 25.

¹⁹² Watson, 5-6.

¹⁹³ Maizland, *The Islamic State*.

China, Algeria, Libya, Tunisia, Egypt, Belgium, Israel, Germany, Philippines, France, Bangladesh, Russia, Somalia, Yemen, the United States, Kenya, Jordan, Italy, Saudi Arabia, Turkey, Mali, Austria, Indonesia, Kosovo, and Afghanistan.¹⁹⁴

In April of 2015 the town of Tikrit is liberated from IS control. This is the first major re-taking of an Iraqi town.¹⁹⁵ Despite the liberation of Tikrit, the Islamic State would prove resilient and determined to defend their hard-won territory.¹⁹⁶ On May 19th, 2016, Iraq Security Forces take back the city of Rutbah with the help of U.S. air support.¹⁹⁷ By October 10th, 2016, a report from West Point claims that the Islamic State is losing influence in Iraq.¹⁹⁸ On November 3rd, 2016, Baghdadi appears in a public message, calling specifically on foreign fighters in Algeria, Libya, and Tunisia, and addressing them as “pillars of Islam on earth and poles of the Caliphate.”¹⁹⁹ This public address is a change of tone that indicates IS leadership is beginning explore a regression from conventional operations back into a traditional guerrilla insurgency. This announcement also indicates an attempt to open new fronts against Western forces to relieve pressure off IS controlled territory in Iraq. On December 7th, 2016, a message to IS fighters demands they stop using encrypted messaging software like WhatsApp and Telegram. The compartmentalization of communications is another indication of the transition to a more cellular state of insurgency.

After the return of U.S. Forces and the marshalling of the Popular Mobilization Forces (PMF) made up of various armed militias, the Islamic State is forced to take a

¹⁹⁴ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

¹⁹⁵ Parker, *After Iraqi forces take Tikrit*.

¹⁹⁶ Kamolnick, *The Al-Qaeda Organization and the Islamic State Organization*, 195.

¹⁹⁷ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

¹⁹⁸ Milton, *Communication Breakdown: Unraveling the Islamic State's Media Efforts*, IV-V.

¹⁹⁹ Wright, *The Jihadi Threat: ISIS, al-Qaeda, and Beyond*, 16.

defensive stance. While IS continued sporadic counterattacks to take new territory, most of their fighting during this time became defensive. According to the book *The Conflict With ISIS*, “while on the defensive, the Islamic State fought primarily in cities, where dense urban terrain and the presence of civilians afforded some protection from aerial attack.”²⁰⁰ Defensive tactics usually consisted of employing snipers, IEDs, and digging elaborate tunnel systems between houses for troops movement, resupply, and counterattacks from unexpected directions.²⁰¹

On March 8th of 2017 U.S. Forces announce that Baghdadi has fled the IS stronghold in Mosul across the Syrian border against mounting coalition pressure.²⁰² By April, most of the territory IS controlled in Iraq is now lost to coalition forces.²⁰³ In June as Syrian Defense Forces take the IS capital Raqqa, IS expands its cellular networks in the Philippines and Afghanistan.²⁰⁴ By November 21st, 2017, Iraqi Prime Minister Haider al Abadi announced that the Islamic State is “militarily defeated” in Iraq.²⁰⁵

The final phase of U.S. and Joint-Coalition fighting in Iraq began in September of 2018 named *Operation Roundup*.²⁰⁶ On August 23rd, 2018, Baghdadi releases an audio recording that appears to act as a strategic pivot from a conventional insurgency towards cellular guerrilla insurgency.²⁰⁷ The choice to release an audio recording rather than address his organization in a more public setting alludes to a stronger compartmentalization of the insurgency. In his address Baghdadi remarked that “for the

²⁰⁰ Watson, *The Conflict with ISIS*, 81.

²⁰¹ Watson, 52.

²⁰² Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

²⁰³ Glenn.

²⁰⁴ Glenn.

²⁰⁵ Glenn.

²⁰⁶ Watson, *The Conflict with ISIS*, 76.

²⁰⁷ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

Mujahadeen [holy warriors] the scale of victory or defeat is not dependent on a city, or a town being stolen.”²⁰⁸ This messaging contrasts starkly from messaging in their initial rise by downplaying the original emphasis on capturing territory. One month later in September of 2018, the West Point Counterterrorism Center released a report assessing that IS was indeed reversing strategic course from a conventional force to guerrilla tactics in Northern Iraq.²⁰⁹

On October 31st, 2019, Baghdadi is killed by a U.S. Special Forces raid in Syria. He is succeeded by Ibrahim al Hashemi al Qurayshi. In 2019 the BBC estimated IS total strength at “14,000 to 18,000 militants in Iraq and Syria.”²¹⁰ Qurayshi inherited an organization that was broken and splintered. Key territory along the border between Iraq and Syria is lost to U.S. and coalition forces, making travel and coordination difficult. Before taking the position of Caliph, Qurayshi held positions of IS leadership over both funding and ideology.²¹¹ His lack of tactical experience and distance from the conventional frontlines during the height of IS makes al Qurayshi more prone to operating the insurgent group in a cellular manner.²¹² During this time an investigative report noted that IS focused on attacks against relatively weaker Arab tribes that were working alongside Syrian Defense Forces.²¹³ While the Islamic State failed to resurge in Iraq or Syria under Qurayshi’s command, the organization did expand its operations in Africa.²¹⁴ On February 3rd, 2022, the Department of Defense released a statement announcing that al Qurayshi was killed during a U.S. military operation to kill or capture

²⁰⁸ Glenn.

²⁰⁹ Knights, *The Islamic State Inside Iraq*.

²¹⁰ BBC, *The rise and fall of the Islamic State group*.

²¹¹ Kilani, *A Caliph Without a Caliphate*.

²¹² Kilani.

²¹³ Kilani.

²¹⁴ Roggio, *Islamic State's emir dead after U.S. military raid in Syria*.

him in Syria.²¹⁵ By September of 2022 the U.S. National Counterterrorism Center estimated between 8,000 – 16,000 IS fighters still remain active in both Syria and Iraq.²¹⁶

Interpreting the Results

This section interprets the statistical results from each model based on the context of the case study.

Correlation Analysis

Point Biserial Correlation shows that there is a significant negative relationship with the increase of both geodesic and network distances to insurgent violence. These correlations indicate that the further a populated area is from a current violence, the less likely it will experience future violence. This connection to physical geography indicates that cellular insurgent violence diffusion is a less compelling contagion than geodesic or network models.

There is also a significant ($p < 0.001$) negative relationship between the proximity of a populated area to road networks, and the likelihood of insurgent violence. The results indicate that the closer a populated area is to a road network, the more likely it is to experience violence. These results suggest that Model 2 (Network) is ultimately the strongest predictor for how violence will spread.

The strongest performing variable for Point Biserial Correlation was *Population Density* ($R_{pb} = 0.22$, $df = 51621$, $t = 50.5$, $p < 0.001$). This result implies that the larger the populated area is, the more likely it will experience violence. This finding supports the

²¹⁵ Vergun, *Leader of ISIS Dead Following U.S. Raid in Syria*.

²¹⁶ National Counterterrorism Center, *Islamic State of Iraq and Ash-Sham (ISIS)*.

narrative that most fighting took place in and over urban environments.²¹⁷ The data also supports this finding with the three most populous cities in Iraq (Baghdad, Mosul, and Kirkuk) accounting for 53% of the violent events recorded.

The additional variable *Avg Location Elevation* was significant ($p < 0.001$) but had a very weak negative correlation (-0.04), as well as a comparatively low t-statistic ($t = -8.49$). This weak correlation indicates that the higher the elevation of a populated area, the *less* likely it will experience violence. This result is surprising because of the conventional understanding that insurgency favors mountainous terrain.²¹⁸ Therefore, it is expected that a populated area with a higher average elevation would indicate more mountainous terrain, and thus more insurgent activity. However, the most mountainous northern region of Iraq is already home to a strong separatist Kurdish population. The Kurds were some of the most effective domestic forces to stem the expansion of the Islamic State.²¹⁹ Additionally, most of the populated areas in Iraq were distributed in these Northeastern mountains, and thus out of reach from IS violence. Figure 7 below overlays IS violent events in red, and the location of populated areas in dark gray on top of an elevation map of Iraq. Note the lack of violent events along the mountain range in Iraq's northeast – an area traditionally controlled by Iraq's Kurdish population.

²¹⁷ Watson, *The Conflict with ISIS*, 81.

²¹⁸ Fearon, *Ethnicity, Insurgency, and Civil War*, 75-90.

²¹⁹ Payne, *Sinjar: Kurds try to retake key Iraqi town from ISIS*.

Islamic State Battles, Populated Areas, and Elevation Iraq 2003-2020

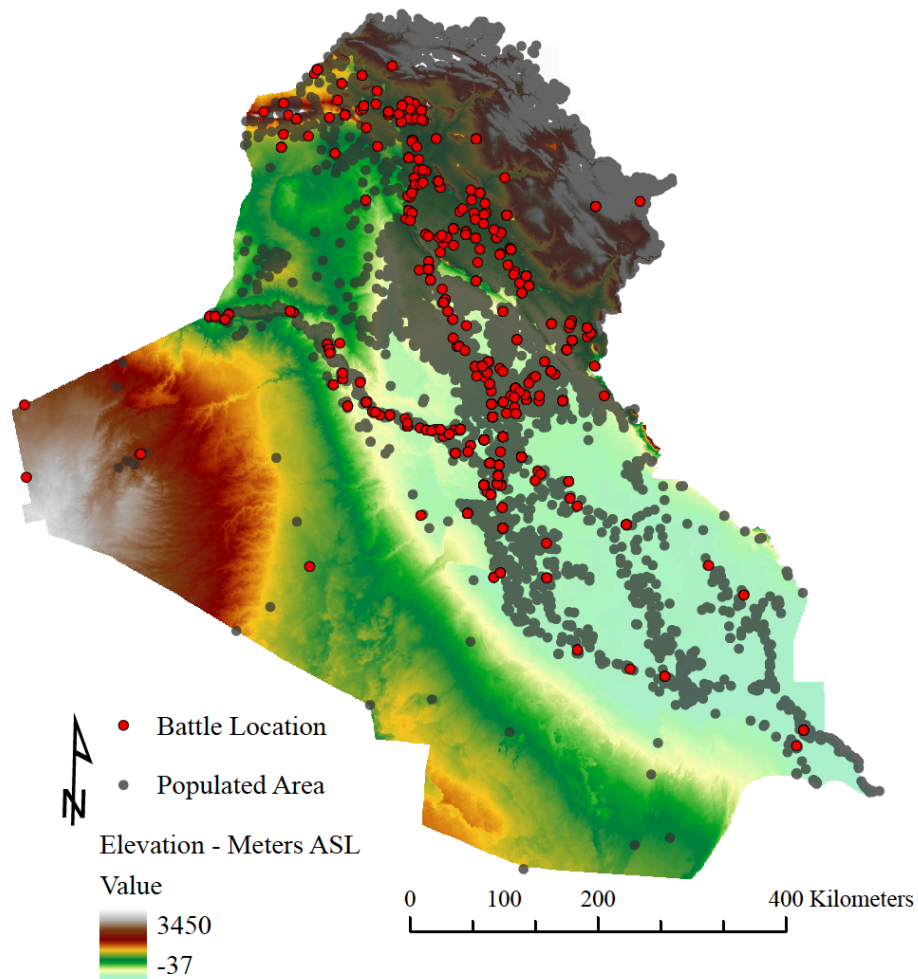


Figure 7. Islamic State Battles, Population Centers, and Elevation in Iraq.

A map of elevation in Iraq with plotted populated areas and violent events. Sources: Uppsala Conflict GED Dataset and DIVA-GIS Iraq Country Data.

The Phi Coefficient, which was used to indicate correlation between two binomial variables, is null for both Election Month and Harvest Season. However, when assessing the correlation between whether a violent event would repeat itself once violence already occurred was a moderate to strong positive association (Phi Coefficient = 0.614, Chi-square = 19469, $p < 0.001$). This suggests that once violence arrives, it becomes “sticky.” This “stickiness” aligns with the self-perpetuating nature of violence explored by other writers.²²⁰ In the case of Iraq, this stickiness did not always correlate to captured Islamic State territory. For example, the Islamic State was never able to capture territory inside the Capital city of Baghdad. However, this populated area was the one of the most targeted locations, suffering 792 violent events between 2003 and 2020. Mostly the victim of bombings, on 3 July 2016 Baghdad also suffered the deadliest attack in IS history – a suicide bombing that killed over 200 civilians.²²¹

Temporally Aggregated Model Analysis

Between 2014 and 2020 the UK investigative organization Conflict Armament Research published a series of reports on the sources and networks that facilitated the armament of the Islamic State.²²² These reports in conjunction with the IS narrative in Part 1 bring important context to the statistical results of this thesis.

To compare models across a common standard, this thesis uses their AUC (Area Under the Curve). This metric measures the model’s discriminatory ability. The closer the value is to “1,” the better the model is at distinguishing between the two outcome

²²⁰ Blattman, *Why We Fight*, 212-214.

²²¹ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

²²² See a full list of publications here: <https://www.conflictarm.com/publications/>

variables (how well it can predict violence versus peace). The highest performing temporally aggregated model in this metric is Model 2 (Network) with an AUC of 0.736. Model 2a, which only measures the effects of major road networks (trunks and freeways), performed slightly weaker with an AUC of 0.727. These results indicate that when predicting the for diffusion of insurgency, it is better to incorporate national, regional, and local road networks alongside major roadways.

Model 1 (Geodesic) had an AUC of 0.689. This was the lowest AUC of the three temporally aggregate models, but still higher than the additional variables tested (Population Density AUC = 0.688, Avg Location Elevation AUC = 0.564, Harvest Season AUC = 0.364, Election Month AUC = 0.019).

Model 3 (Cellular) performed slightly worse than Models 2/a (Network) with an AUC of 0.693. It is important to note that these findings are not significant ($p = 0.996$). However, even when disaggregated Model 3 was consistent across all years with slight variations in AUC. Despite this model's lack of statistical significance, its consistency relative AUC changes can still provide some insight into the comparative performance between all three models.

While there is a lack of significance in Model 3 (Cellular), there is strong anecdotal evidence that supports cellular-aligned activity in the case of ISI. According to a 2017 report that traced the origin of small-arms and ammunition captured from the Islamic State, "IS forces captured significant quantities of NATO weaponry during initial assaults on Iraqi forces in 2014."²²³ The acquisition of these weapons from retreating Iraqi Security Forces likely helped to create the momentum necessary for IS fighter to

²²³ Inowlocki, *Weapons of the Islamic State*, 5.

continue capturing territory. An October 2014 report titled *Islamic State Ammunition in Iraq and Syria* also found that “the sample of ammunition recovered in Iraq is mainly US-manufactured and comprises 5.56 x 45 mm cartridges, which are used in US-supplied M16 and M4 assault rifles of the Iraqi defence and security forces.”²²⁴ Therefore, when looking at the tactical context, it appears that Model 3 (Cellular) also played an important role for IS throughout the conflict.

Annual Temporally Disaggregated Model Analysis

In 2004 while the Islamic State was still forming as a group underneath AQI, Model 2 (Network AUC = 0.746) holds as the dominant model. Model 1 (Geodesic AUC = 0.566) initially performs much worse than the other models in 2004. Comparatively there is only a slight variation between Models 2 (Network AUC = 0.746) and 3 (Cellular AUC = 0.725). In 2004 the additional variable *Population Density* (AUC = 0.828) performs better than all three models. Contextually ISI violence in 2004 was comparatively light. AQI strategic guidance limited ISI to guerrilla attacks on U.S. Forces with no priority of taking and holding territory. To target U.S. Forces, ISI was forced to focus on attacks in populated areas with limited strategic or operational goals beyond simply creating casualties.

In 2005 Model 3 (Cellular) outperformed all other models. The emergence of Model 3 (Cellular AUC = 0.745) over Model 2 (Network AUC = 0.725) and Model 1 (Geodesic AUC = 0.703) could be the result of “function following form,” with the

²²⁴ Conflict Armament Research, *Analysis of small-calibre ammunition recovered from Islamic State forces*, 6.

insurgency adapting over their first years. The low volume of attacks in the early years may also play into how these models performed.

In 2006 all Models peak together in their overall performance with Model 2 (Network) on top with an AUC of 0.803. Model 1 (Geodesic) is next at 0.788, and Model 3 (Cellular) is last with an AUC of 0.77. This switch comes at a major transitional period for ISI. 2006 is the year that al-Zarqawi is killed, and together Abu Ayyub al-Masri and Abu Omar al-Baghdadi worked to transform the MSC into the ISI. This transition is the beginning of a strategic pivot that turns away from focus on the “far enemy” and Guerrilla tactics, to a more aggressive stance on Iraqi civilians with the intent of establishing sovereign territory under ISI control.

After this transition for the next few years (2009-2012) we see a stark separation in performance between all three models with Model 2 (Network) performing better than the rest. During this time all models are generally trending downward in their ability to predict the spread of violence until 2013. Despite ISI establishing itself as an organization with different strategic goals than its parent group AQI, ISI was still embracing guerrilla tactics during these years. An aggressive U.S. counterterrorism campaign was still underway until the end of 2011, which prevented ISI from making a definitive shift to conventional tactics at this time.

In 2013 all Models peak again, this time Model 2 (Network AUC = 0.774) is comparatively far stronger than Models 1 (Geodesic AUC = 0.717) and 3 (Cellular AUC = 0.731). This peak coincides with the Islamic State’s expansion into Syria as well as the capture of the city of Fallujah. In 2013 the departure of U.S. Forces in Iraq and the civil unrest caused by the Arab Spring have created ideal conditions for the ISI insurgency to

operate more conventionally and aggressively. The expansion into Syria and forced merger with al-Nusra Front in 2013 establishes a core supply-line from Turkey, through Syria, and into Iraq.²²⁵ This international overland supply route exported oil and gas, and imported materials to manufacture weapons and ordnance in Iraq.²²⁶ According to a 2016 report titled *Tracing the Supply of Components Used in Islamic State IEDs* “Turkey is the most important choke point for components used in the manufacture of IEDs by IS forces.”²²⁷

In 2014 Model 3 (Cellular) overtook Models 1 (Geodesic) and 2 (Network) for the second time. This is interesting timing for this model to suddenly outperform all others. In 2014 al-Baghdadi announces the establishment of a Caliphate and leads a blitzkrieg into Iraq takes most of the territory away from Iraqi Security Forces. In 2014 the Islamic State is operating at arguably one of its most conventional states. Therefore, one would expect Model 2 (Network) to dominate as the most conventional of the three. However, the performance of all three models slowly begin to converge after 2013, with the next year (2015) as the closest all three perform throughout all 17 years. This convergence may be the result of the watershed of locally sourced weapons and materials that became available during the 2014 campaign that were captured from Iraqi Security Forces.²²⁸ During this push, captured materials would have supported the expansion of Islamic State violence. Meanwhile roads acted as the gateway to further plunder as the Islamic State pushed further and further into Iraq.

²²⁵ Inowlocki, *Procurement Networks Behind Islamic State Improvised Weapon Programmes*, 6-11.

²²⁶ Levallos, *The Financing of the 'Islamic State' in Iraq and Syria (ISIS)*, 16.

²²⁷ Inowlocki, *Tracing the Supply of Components Used in Islamic State IEDs*, 11.

²²⁸ Inowlocki, *Weapons of the Islamic State*, 5-8.

In 2016 Model 3 (Cellular AUC = 0.676) begins to drop in performance compared to the others. Contextually, in 2016 the U.S. decisively re-entered the war in *Operation Inherent Resolve* and in 2016 ISI is transitioning from an offensive campaign to consolidate and defend its gains.²²⁹ Contextually during this time local procurement would have slowed, but external supply-chains would initially remain unharmed.

The year 2017 is the last year before a rapid decline in performance for all three models. It is also the last year the Islamic State held significant territory in Iraq. By March 2016, Baghdadi fled from Mosul into Syria, and by November 2016 the Iraqi Government announced that it had militarily defeated the Islamic State.²³⁰ This drop in performance parallels the decline in the Islamic State. This decline is a sign that the models are generally performing in tandem with contextual realities.

From 2018 to 2019 there is a definite transition from conventional to guerrilla tactics inside the Islamic State.²³¹ This is coupled with a transition in performance from Model 2 (Network) to Model 3 (Cellular) being the dominant performer. In 2019 al-Baghdadi is killed and the Islamic State is highly degraded. However, in 2020 there is another shift with the order of performance turning to Model 2 (Network AUC = 0.612), Model 1 (Geodesic AUC = 0.559), then Model 3 (Cellular AUC = 0.533). These shifts are reminiscent of the performance from the early stage of the insurgency, when attack-volume was low, and the organization was in a highly transitional state. This contextual parallel could suggest that logistics networks during dramatic lifecycle changes can be volatile. Meanwhile, when an insurgency is performing consistently (like in 2007 – 2012

²²⁹ Watson, *The Conflict with ISIS*, 45-47.

²³⁰ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

²³¹ Watson, *The Conflict with ISIS*, 74-76.

or 2014 – 2017 for the Islamic State) their logistics models for violence diffusion are more regular.

Phased Temporally Disaggregated Model Analysis

When looking at the models in temporal phases the model performance is slightly different than annually. Phase 1 (2003-2006), the early formation of the insurgency, shows Model 3 (Cellular AUC = 0.754) as the top performing model. This is followed by Model 2 (Network AUC = 0.725) then Model 1 (Geodesic AUC = 0.628). These results align with the guerrilla tactics imposed by ISI's parent organization AQI, the relatively low-level insurgent action, and high-level of U.S. counterterrorism efforts.

Phase 2 (2007-2011) which saw the establishment of the ISI, and its leadership under Ayyub al-Masri and Abu Omar al-Baghdadi saw a transition both strategically and in model performance.²³² Strategically, this new leadership pivoted from the “far enemy” to the “near enemy” as their primary objective.²³³ In model performance, Model 2 (Network AUC = 0.762) becomes dominant. Model 1 (Geodesic AUC = 0.719) is second, and Model 3 (Cellular AUC = 0.681) performs the worst. While still under the parent organization of AQI and the pressure of U.S. counterterrorism operations, ISI begins to shift its strategic priorities during this phase.²³⁴

In Phase 3 (2012 – 2015) all three models are closely aligned with Model 2 (Network AUC = 0.721) first, followed by Model 3 (Cellular AUC = 0.704), then Model 1 (Geodesic AUC = 0.692). In this phase the power and activity of the Islamic State

²³² Kamolnick, *The Al-Qaeda Organization and the Islamic State Organization*, 87.

²³³ Kamolnick, 87-90.

²³⁴ Kamolnick, 87-90.

peaks. Contextual reports reveal that the Islamic State drew from both a robust over-land logistics network, as well as receive a windfall of locally sourced material from collapsing Iraqi Security Forces.²³⁵ While during this stage all models are closest in performance, Model 2 (Network) remains the best predictor of insurgent diffusion. This is the expected outcome. Even though ISI logistics and supply were varied during this time, the most efficient method for spreading conventionally from one populated area to another during the blitzkrieg was to use road networks. Wide access to fuel and oil refineries also supplies IS vehicles with an abundance of fuel, further encouraging the utilization of roads to move men and materials.²³⁶

Phase 4 (2016 – 2020) is the collapse of the Islamic State after U.S. and other joint-coalition intervention. During this phase the Islamic State attempted to defend its territory, and slowly transitions back into a guerrilla network. In this phase Model 2 (Network AUC = 0.693) remains the most dominant but is also performing at its lowest point. Model 1 (Geodesic AUC = 0.671) is the second-best performer with Model 3 (Cellular AUC = 0.643) as the lowest of the three. While Model 3 is performing the worst inside Iraq, internationally the Islamic State is operating with a new strategic priority of cellular dispersion. Emphasis on foreign expansion in Afghanistan, Indonesia, Africa, and elsewhere are voiced by IS leadership as the Caliphate loses ground in Iraq and Syria.²³⁷ While this thesis remains focused on the parameters of Iraq, the insurgency clearly transitions strategically into a dispersed and cellular phase during its decline.

²³⁵ Inowlocki, *Weapons of the Islamic State*, 5.

²³⁶ Levallios, *The Financing of the 'Islamic State' in Iraq and Syria (ISIS)*, 9.

²³⁷ Glenn, *Timeline: the Rise, Spread, and Fall of the Islamic State*.

Comparing to Previous Research

The results of this thesis reinforce Dr. Zhukov's findings that road networks in general are a better predictor of the spread of insurgent violence than geodesic or cellular models. Not only was Model 2 (Network) the most robust in predicting insurgent violence at the temporally aggregate level, but it was also consistently the best performer in the annual and phased series of models. When compared to Dr. Zhukov's research, this thesis chose a new time-period, new insurgent group, new physical geography, and new statistical processes to test Dr. Zhukov's three theories of insurgent diffusion. With all these differences the results remain consistent – roads tend to deliver violence.

To explore the subject of insurgent diffusion deeper, this thesis temporally disaggregated the data in a way that has never been done before. By looking at the entire lifecycle of an insurgency in annual and phased sections, this thesis was able to uncover shifts in the violence diffusion that suggest that violence diffusion methods can change in over time. Simply put, insurgencies are learning organizations that grow, change, and react to their environment. When looking at model competition along the temporally disaggregated models, the rise and fall of the Islamic State in Iraq is reflected in the results. These results indicate that there is some form of symbiotic relationship between the violence diffusion models and the evolution of an insurgency. While all three models are present, their utility and predictive power shift over time. This suggests that while road networks continue to appear as the best predictors of insurgent violence, their predictive power is not stagnant throughout the entire lifecycle of an insurgency.

Implications of Findings

How can this understanding of insurgent diffusion benefit modern society and the real world? Insurgencies change over time - they are not stagnant. Therefore, a robust counterterrorism policy must be flexible and open to change as well. This is not a novel idea. John Nagl, author of *Learning to Eat Soup With a Knife*, compares the U.S. counterterrorism response in Vietnam to that of the U.K. counterterrorism actions in Malaya.²³⁸ Nagl discovers that the United Kingdom was successful where the United States failed largely because U.K. institutions were capable of learning and adapting to new conditions.²³⁹ According to Nagl, one of the largest hurdles preventing the U.S. military to change was the failure of their institutions to grasp the nature of the war they were fighting. Nagl highlights this in his book with a quote “in the process, we lost sight of one of the cardinal maxims of guerrilla war: the guerrilla wins if he does not lose; the conventional army loses if it does not win.”²⁴⁰ Unlike conventional wars that embody the maxim “survival of the fittest,” irregular warfare appears to be “survival of the most adaptable.” The results of this thesis further support this concept. While the data shows that road networks act as a key facilitator, it also suggests that insurgent groups are capable of rapid reprioritization and change based on external and internal conditions. For these results to be useful, counterinsurgency strategy must be adaptable and prepared for insurgent organizations to shift when conditions change.

²³⁸ Nagl, *Learning to Eat Soup with a Knife*, xii.

²³⁹ Nagl, xxii.

²⁴⁰ Nagl, 171.

Limitations of the Study

Due to the selection of the variables used in the models, not all the models were statistically significant, specifically Model 3 (Cellular). Although this model remained consistent in its temporally disaggregated state, this lack of significance limits the result reliability. Another limitation was the lack of control variables and additional variables. Models 1 and 3 only consisted of one independent variable related to the general theory that the model was meant to represent. Model 2 used two predictor variables. Using limited predictor variables in the models help to isolate and clearly understand the effects these variables have on the outcome variable. However, this method creates relatively low pseudo-R-squared results for each of the models. This limits the model's overall explanatory power as to the results of the outcome variable. Finally, the creation of the data set for this thesis involved collecting a wide range of information. Most of this data was not selected for analysis as part of this thesis. The decision to limit the variables used from the data set was primarily to remain focused on the original research questions. However, this was also done because several of these variables showed no statistical significance when tested, or other signs of unreliability in their results.

Chapter VI

Conclusion

This chapter will begin with a summary of the main findings, followed by a short discussion over the answer to the research questions. This chapter ends with recommendations for future research.

Summary of Main Findings

This thesis intended to explore the relationship between roads and the spread of insurgent violence. To accomplish this, correlation and logistics regression were used to test three models of insurgent diffusion originally conceived by Dr. Zhukov in this research titled *Roads and the Diffusion of Insurgent Violence*.²⁴¹ Model 1 (Geodesic) represents violence diffusion that is not constrained by physical geography. Model 2 (Network) represents violence diffusion that is associated with road networks and existing infrastructure. Model 3 (Cellular) represents violence diffusion that is self-sufficient, locally sourced, and self-contained. The Islamic State insurgency in Iraq from 2003 to 2020 was selected as the case study. A thorough review of the insurgency through government, news, and research institution reports and books was conducted to provide context to the statistical results of the study. Where Dr. Zhukov's research remained temporally aggregated, this thesis disaggregated the data both annually and in phases to explore the possibility of change in model predictive strength over the lifecycle of the insurgent group.

²⁴¹ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

In correlation analysis, this thesis found significant ($p < 0.001$) relationships between all three models. The strongest correlation Model 3 (Cellular) with a moderately strong positive correlation, which was significant ($\phi = 0.614$, $n = 51623$, $p < 0.001$, $\chi^2 = 19469.4124$). Correlation analysis also tested the relationship between insurgent violence and the proximity of a populated area to a road network and found the results to be significant ($R_{pb} = -0.07$, $n = 51623$, $p < 0.001$, $t = -15.73$). Additional variables *Population Density* and *Average Location Elevation* had the strongest and weakest correlation results, respectively. *Population Density* ($R_{pb} = 0.22$, $p < 0.001$, $t = 50.5$).²⁴² *Average Location Elevation* had the lowest t-statistic absolute value ($R_{pb} = -0.04$, $p < 0.001$, $t = -8.49$). Additional variables *Harvest Season* and *Election Month* had no correlation or were not significant. *Harvest Month* ($\phi = 0$, $n = 51623$, $p < 0.001$, $\chi^2 = 13.63063$). *Election Month* ($\phi = -0.00145$, $n = 51623$, $p = 0.948874$, $\chi^2 = 0.004111$).

The results of the temporally aggregated logistics regression found that both Models 1 (Geodesic) and 2 (Network) were significant ($p < 0.001$), while Model 3 (Cellular) predictor was insignificant ($p = 0.996$). The metric AUC (Area Under the Curve) was used to compare the strength of the predictive models to one another. Model 1 (Geodesic) had an AUC of 0.689. Model 2 (Network) was tested twice, with the original model using data from major, local, national, and regional road networks. Model 2a used data only from major road networks, to further explore this relationship. Both models were significant ($p < 0.001$) with Model 2 performing slightly better (AUC = 0.736) than Model 2a (AUC = 0.727). The predictor variable, *Repeat VE*, for Model 3 (Cellular) was not significant ($p = 0.996$) and its AUC was the lowest of the models

²⁴² Population Density correlation was stronger than all variables including those used in Models 1, 2, and 3.

(AUC = 0.693). Additional variables *Population Density*, *Average Location Elevation*, *Harvest Season*, and *Election Month* were also tested individually. All predictors were statistically significant ($p < 0.001$) except for *Election Month* ($p = 0.742$). The predictor *Average Location Elevation* has a coefficient-b of 0, meaning it has no effect on the outcome variable. The predictor *Population Density* has the strongest AUC of 0.688. The Other variables AUC are as follows; *Average Location Elevation* AUC = 0.564, *Harvest Season* AUC = 0.364, *Election Month* AUC = 0.019.

The temporally disaggregated models were largely consistent from year to year and phase to phase in their statistical significance. Over time Model 2 (Network) consistently outperformed the other two models in its predictive performance using the AUC metric. However, variation in performance strength between the models over time suggest that Model 2 is not always better at predicting insurgent violence than other models. Occasionally Model 3 (Cellular) outperformed other models during both the annual and phased tests.²⁴³ Additional variables generally performed worse as predictors for insurgent violence except for *Population Density*. Apart from *Population Density*, the additional variables had no statistical significance ($p > 0.05$) for most of the years/phases. When annually disaggregated variables *Population Density* and *Average Location Elevation* follow a near-parallel temporal path in predictive ability until the year 2014. After 2014 both variables appear to become inverse of one another. This is an interesting result, as 2014 is a pivotal year for the Islamic State where the Caliphate was at its peak. Following 2014 a U.S. led intervention forced the Islamic State into recession and decline.

²⁴³ While this thesis uses Model 3 (Cellular) AUC as a comparative metric, it is important to note that Model 3 was consistently not significant ($p > 0.05$) in its results as a predictor.

Overall, road networks appear to be the strongest predictor for the spread of insurgent violence. This finding concurs with Dr. Zhukov's original research.²⁴⁴ Importantly however, the results of this thesis suggest that the power that road networks have on the diffusion of insurgent violence change over time as an insurgency goes through various phases. The Islamic State insurgency in Iraq saw definitive rise and fall in which it transitioned from guerrilla tactics to take on more conventional military operations as well as attempts of hold and govern autonomous territory. Then, when confronted by overwhelming opposition, it was forced to recede and transform again into a guerrilla insurgency. The variance in the temporally disaggregated performance of each of the models indicate that as an insurgency evolves, so does its ways and means of violence diffusion.

The implication of these findings is of great importance to counterterrorism strategy and to better understanding the growth of insurgent movements. They reveal the importance of flexibility, adaptation, and organizational learning when countering an insurgent force. Insurgencies will adapt when met with resistance or opportunity.

This thesis has both several limitations and opportunities for future study. The models used in this thesis used only one to two predictor variables each. Deciding to use minimal variables to predict the outcome meant that there was a clear understanding of how these variables impacted the model as a whole. However, it also limited the model's predictors to explaining a relatively low proportion of the variance in the outcome variable (as evidenced by low pseudo-r-squared results). The lack of statistical significance for the predictor in Model 3 (Cellular) also limited the usefulness of this

²⁴⁴ Zhukov, *Roads and the diffusion of insurgent violence*, 144-156.

model's results. These limitations open opportunities for further study. Future research should include more comprehensive models along with comparative studies across multiple insurgencies that have different historical outcomes. Additionally, while this thesis indicates that insurgent diffusion models change over time, it does not fully explain what factors drive this change and to what degree. Further research on this matter could lead to new and interesting discoveries on the nature of the spread of insurgency.

Answers to Research Questions

This thesis intended to explore the relationship between road networks and the diffusion of insurgent violence by studying the Islamic State in Iraq from 2003 to 2020. The models tested show a clear association between road networks and the spread of insurgent violence. Interestingly, when looking at this relationship over time it appears that the dependence of road networks and insurgent violence is not consistent. Rather, the performance of the models show that it changes in importance over the lifecycle of an insurgency.

Recommendations for Future Research

There is still a tremendous amount of work left to understanding the spread of insurgent violence.

First, there needs to be more case studies to better understand organizational, historical, and geographical variation. The Islamic State was a preferable insurgent group to isolate and research because of its parabolic trajectory of a clear rise and fall. However, there is much to be gained from comparing this insurgency to others that do not share this same course, and seeing how well each model performs comparatively. The

LTTE in Sri Lanka are known for their guerrilla beginnings, their conventional transition, and their decisive defeat after failing to return to guerrilla tactics once their conventional forces failed.²⁴⁵ The Taliban is an example of a successful insurgency that, through decades of persistent fighting, was able to successfully drive the United States out of Afghanistan.²⁴⁶ Observing the performance of these models across these insurgencies also allow for opportunity to explore how the vast differences in physical and cultural geography affect the outcomes. Each of these insurgencies would make a strong addition to testing the ways in which insurgent violence spreads. This thesis originally intended to explore these cases, but due to limitations on time and scope chose to remain focused on the Islamic State in Iraq.

Second, while this thesis demonstrates that insurgent diffusion models can change over the lifecycle of insurgency, exactly what factors drive this change remains to be answered. This thesis draws parallels between the performance of the models and the narrative of events using reports and assessments on the Islamic State to provide context to the results. This, in addition to measuring the impact of internal and external factors on the change in insurgent diffusion could provide counterterrorism planners with the ability to better plan, forecast, and predict changes. Creating stronger counterterrorism strategy through a deeper understanding of systems that support the violence ultimately saves lives. If political power truly “comes out of the barrel of a gun,”²⁴⁷ then knowing how to effectively take guns out of the hands of insurgents is a critical step towards peace.

²⁴⁵ Weiss, *The Cage*, 227.

²⁴⁶ Maizland, *The Taliban in Afghanistan*.

²⁴⁷ Tse-Tung, *On Guerilla Warfare*, 12.

Appendix 1

Codebook and Technical Data

Outcome – Whether or not a populated area has experienced a violent event

- 0 = No
- 1 = Yes
- Source: Uppsala GED Data Set

Date – The year and month (YYYYMM)

- Source: Uppsala GED Data Set

Location – The name of the populated area

- Source: Uppsala GED Data Set

Province – The name of the province associated with the location

- Source: Uppsala GED Data Set

Latitude – Latitude associated with the populated area

- WGS 1984 Coordinate System
- Source: Uppsala GED Data Set

Longitude - Longitude associated with the populated area

- WGS 1984 Coordinate System
- Source: Uppsala GED Data Set

VE Count – Total number of times a violent event occurred at a location for that month

- Source: Uppsala GED Data Set

Repeat VE – Whether or not a violent event occurred the month prior at a location

- 0 = No

- 1 = Yes
- Source: Uppsala GED Data Set
- This variable is a replication of a variable Dr. Zhukov used in *Roads and the Diffusion of Insurgent Violence*.

KM Geo – Nearest Recent VE – The number of kilometers away each location is from the closest violent event from the previous month using geodesic measurement.

- Source: Uppsala GED Data Set
- To calculate this variable, I first used the “Select Key Attributes” function in ArcGIS Pro to select the locations where a violent event occurred in the previous month. Then I used the “Near” geoprocessing function, calculating the geodesic distance in meters for all the populated areas to those areas selected. This created a new column that I then exported using the “table to excel” geoprocessing function. In excel I then converted the distance from meters to kilometers using the function “=SUM(X/1000).”
- This variable is a replication of the geodesic variable Dr. Zhukov used in *Roads and the Diffusion of Insurgent Violence*.

KM Net – Nearest Recent VE – The number of kilometers away each location is from the closest violent event from the previous month using ArcGIS Pro built-in road network database.

- Source: Violent events taken from Uppsala GED Data Set. Road network taken from the ArcGIS Pro built-in road network database.
- To calculate this field, I first used the “Select Key Attributes” function in ArcGIS Pro to select the locations where a violent event occurred in the previous month.

Then I used the internal road networks from the ArcGIS Data and Maps package by Esri to calculate the distance in kilometers from the selected locations from the previous month to the nearest violent event occurrence of the current month. This process was repeated for each month of the scope of study.

- This variable is a replication of the network variable Dr. Zhukov used in *Roads and the Diffusion of Insurgent Violence*.

Prox to Major Roads KM – The number of geodesic kilometers a violent event is from a major road network.

- Source: Uppsala GED Data Set and GeoFabrik
- To calculate this variable, I used GeoFabrik road network polyline data for Iraq (gis_osm_roads_free_1.shp file). I first used “Select Key Attributes” function in ArcGIS Pro to select on the Trunks (major roads) and Freeways. Then I used the “Create New Layer” function to create a new layer only displaying major road networks and freeways of Iraq. After this I used the “Near” geoprocessing function to give me the geodesic distance in meters for the nearest major roads to a violent event. Then I used the “table to excel” geoprocessing function and converted the meters into kilometers using the excel formula “=SUM(X/1000).”

Prox to All Roads KM – The number of geodesic kilometers a violent event is from major roads, freeways, national, regional, and local roads

- Source: Uppsala GED Data Set and GeoFabrik
- To calculate this variable, I used the same steps used in *Prox to Major Roads KM* (above), only this time the calculation was conducted from a layer that included all roads (Trunks, Freeways, National, Regional, and Local).

- This variable is similar to Dr. Zhukov’s “Accessibility” variable. However, his was binary coded to whether the populated area was within 5 kilometers of a road.

This thesis uses geodesic measured kilometers.

Population Density by Sq100m – The average population density within a 10-kilometer radius of a populated area

- Source: WorldPop
- This variable is a replication of a variable Dr. Zhukov used in *Roads and the Diffusion of Insurgent Violence*. However, it is collected from a different source.

Average Location Elevation m ASL – The average elevation in meters above sea level within a 10-kilometer radius of a populated area.

- Source: DIVA-GIS Country Information

Harvest Season – Whether or not the month is associated with a harvest season

- 0 = No
- 1 = Yes
- Source: Food and Agriculture Organization of the United Nations

Election Month – Whether or not the data correlates with an election month for the state of Iraq

- 0 = No
- 1 = Yes
- Source: Historical data
- This variable is a replication of a variable Dr. Zhukov used in *Roads and the Diffusion of Insurgent Violence*.

Appendix 2

List of Acronyms

ASL – Above Sea Level

AUC – Area Under the Curve

CCF – Chinese Communist Forces

CGIAR - Consultative Group for International Agricultural Research

CI – Confidence Interval

FAO – Food and Agricultural Organization

FMLN – Farabundo Marti National Liberation Front

FPR – False Positive Rate

GDAM - Database of Global Administrative Areas

GDL – Global Data Lab

HDI – Human Development Index

IED – Improvised Explosive Device

IS – Islamic State

ISI – Islamic State in Iraq

LMICs – Low- and Middle-Income Countries

LTTE – Liberation Tigers of Tamil Eelam

NKPA – North Korean People’s Army

PPP – Purchase Power Parity

ROC – Receiver Operating Characteristic

SRTM - Shuttle Radar Topography Mission

sq. km – Square Kilometers

TPR – True Positive Rate

UCDP – Uppsala Conflict Data Program

VIF – Variance Inflation Factor

Appendix 3

Additional Tables and Figures

Table 3. Temporally Aggregated Logistics Regression Results.

Model 1: Geodesic	AUC 0.689	Model 2: Network	AUC 0.736
LR Results		LR Results	
Number of Cases	51623	Number of Cases	51623
Correct Assignments	49800	Correct Assignments	49800
Percentage	96.47%	Percentage	96.47%
Chi-Squared Test		Chi-Squared Test	
Chi2	627.06	Chi2	1006.46
Degrees of Freedom	1	Degrees of Freedom	2
p-value	<0.001	p-value	<0.001
Model Summary		Model Summary	
-2 Log-likelihood	15099.14	-2 Log-likelihood	14764.74
Cox & Snell R2	0.01	Cox & Snell R2	0.02
Nagelkerke R2	0.05	Nagelkerke R2	0.07
McFadden's R2	0.04	McFadden's R2	0.06
Model		Model	
<i>Constant</i>		<i>Constant</i>	
Coefficient B	-2.64	Coefficient B	-2.27
Standard Error	0.03	Standard Error	0.04
z-factor	78.4	z-factor	60.35
p-factor	<0.001	p-factor	<0.001
Odds Ratio	0.07	Odds Ratio	0.1
95% conf. interval	0.07 - 0.08	95% conf. interval	0.1 - 0.11
<i>KM - Geo</i>		<i>KM - Net</i>	
Coefficient B	-0.01	Coefficient B	-0.01
Standard Error	0	Standard Error	0
z-factor	20.81	z-factor	20.85
p-factor	<0.001	p-factor	<0.001
Odds Ratio	0.99	Odds Ratio	0.99
95% conf. interval	0.99 - 0.99	95% conf. interval	0.99 - 0.99
		<i>Prox All Roads KM</i>	
		Coefficient B	-0.06
		Standard Error	0
		z-factor	14.75
		p-factor	<0.001
		Odds Ratio	0.94
		95% conf. interval	0.93 - 0.95

Model 2a: Network	AUC	Model 3: Cellular	AUC
LR Results	0.727	LR Results	0.693
Number of Cases	51623	Number of Cases	51623
Correct Assignments	49800	Correct Assignments	50503
Percentage	96.47%	Percentage	97.83%
		Chi-Squared Test	
Chi2	1003.68	Chi2	5006.09
Degrees of Freedom	2	Degrees of Freedom	1
p-value	<0.001	p-value	<0.001
Model Summary		Model Summary	
-2 Log-likelihood	14767.52	-2 Log-likelihood	10765.1
Cox & Snell R2	0.02	Cox & Snell R2	0.09
Nagelkerke R2	0.07	Nagelkerke R2	0.35
McFadden's R2	0.06	McFadden's R2	0.32
Model		Model	
<i>Constant</i>		<i>Constant</i>	
Coefficient B	-2.27	Coefficient B	-3.79
Standard Error	0.04	Standard Error	0.03
z-factor	60.35	z-factor	125.59
p-factor	<0.001	p-factor	<0.001
Odds Ratio	0.1	Odds Ratio	0.02
95% conf. interval	0.1 - 0.11	95% conf. interval	0.02 - 0.02
<i>KM - Net</i>		<i>Repeat VE</i>	
Coefficient B	-0.01	Coefficient B	28.88
Standard Error	0	Standard Error	6405.4
z-factor	20.86	z-factor	0
p-factor	<0.001	p-factor	0.996
Odds Ratio	0.99	Odds Ratio	3.48563E+12
95% conf. interval	0.99 - 0.99	95% conf. interval	0 - infinity
<i>Prox Major Roads KM</i>			
Coefficient B	-0.06		
Standard Error	0		
z-factor	14.69		
p-factor	<0.001		
Odds Ratio	0.94		
95% conf. interval	0.94 - 0.95		

Baseline	
LR Results	
Number of Cases	51623
Correct Assignments	49800
Percentage	96.47
Chi-Squared Test	
Chi2	0
Degrees of Freedom	0
p-value	<0.001
Model Summary	
-2 Log-likelihood	15771.19
Cox & Snell R2	0
Nagelkerke R2	0
McFadden's R2	0
Model	
<i>Constant</i>	
Coefficient B	-3.31
Standard Error	0.02
z-factor	138.7
p-factor	<0.001
Odds Ratio	0.04
95% conf. interval	0.03 - 0.04

Population Density	AUC 0.688	Avg Location Elevation	AUC 0.564
LR Results		LR Results	
Number of Cases	51623	Number of Cases	51623
Correct Assignments	49900	Correct Assignments	49800
Percentage	96.66%	Percentage	96.47%
Chi-Squared Test		Chi-Squared Test	
Chi2	1059.74	Chi2	86.3
Degrees of Freedom	1	Degrees of Freedom	1
p-value	<0.001	p-value	<0.001
Model Summary		Model Summary	
-2 Log-likelihood	14711.46	-2 Log-likelihood	15684.89
Cox & Snell R2	0.02	Cox & Snell R2	0
Nagelkerke R2	0.08	Nagelkerke R2	0.01
McFadden's R2	0.07	McFadden's R2	0.01
Model		Model	
<i>Constant</i>		<i>Constant</i>	
Coefficient B	-3.66	Coefficient B	-3.06
Standard Error	0.03	Standard Error	0.04
z-factor	129.62	z-factor	85.39
p-factor	<0.001	p-factor	<0.001
Odds Ratio	0.03	Odds Ratio	0.05
95% conf. interval	0.02 - 0.03	95% conf. interval	0.04 - 0.05
Pop Density		Avg Location Elevation	
Coefficient B	0.1	Coefficient B	0
Standard Error	0	Standard Error	0
z-factor	33.78	z-factor	8.56
p-factor	<0.001	p-factor	<0.001
Odds Ratio	1.11	Odds Ratio	1
95% conf. interval	1.1 - 1.2	95% conf. interval	1 - 1

Harvest Season	AUC 0.364	Election Month	AUC 0.019
LR Results		LR Results	
Number of Cases	51623	Number of Cases	51623
Correct Assignments	48900	Correct Assignments	49800
Percentage	96.47%	Percentage	96.47%
Chi-Squared Test		Chi-Squared Test	
Chi2	14.54	Chi2	0.11
Degrees of Freedom	1	Degrees of Freedom	1
p-value	<0.001	p-value	0.74
Model Summary		Model Summary	
-2 Log-likelihood	15756.66	-2 Log-likelihood	15771.08
Cox & Snell R2	0	Cox & Snell R2	0
Nagelkerke R2	0	Nagelkerke R2	0
McFadden's R2	0	McFadden's R2	0
Model		Model	
<i>Constant</i>		<i>Constant</i>	
Coefficient B	3.2	Coefficient B	-3.31
Standard Error	0.04	Standard Error	0.02
z-factor	90.94	z-factor	137.4
p-factor	<0.001	p-factor	<0.001
Odds Ratio	24.64	Odds Ratio	0.04
95% conf. interval	23 - 26.41	95% conf. interval	0.03 - 0.04
Harvest Season		Election Month	
Coefficient B	0.18	Coefficient B	-0.06
Standard Error	0.05	Standard Error	0.18
z-factor	3.82	z-factor	0.33
p-factor	<0.001	p-factor	0.742
Odds Ratio	1.2	Odds Ratio	0.94
95% conf. interval	1.09 - 1.32	95% conf. interval	0.66 - 1.34

Table 4. Temporally Disaggregated Area Under the Curve (AUC) Results.

Year	Model 1 (Geo)	Model 2	Model 3 (Cell)
		(Net)	
2004	0.566	0.746	0.725
2005	0.703	0.725	0.745
2006	0.788	0.803	0.77
2007	0.698	0.756	0.69
2008	0.732	0.795	0.693
2009	0.704	0.762	0.65
2010	0.723	0.765	0.686
2011	0.725	0.747	0.679
2012	0.655	0.729	0.667
2013	0.717	0.774	0.731
2014	0.674	0.689	0.701
2015	0.703	0.711	0.711
2016	0.718	0.726	0.676
2017	0.697	0.73	0.708
2018	0.61	0.637	0.568
2019	0.516	0.532	0.583
2020	0.559	0.612	0.533
Year	Population Density	Avg	Harvest Season
		Elevation	
2004	0.828	0.63	0.453
2005	0.698	0.513	0.339
2006	0.743	0.612	0.427
2007	0.74	0.593	0.336
2008	0.683	0.509	0.427
2009	0.742	0.532	0.343
2010	0.757	0.637	0.421
2011	0.769	0.686	0.46
2012	0.721	0.632	0.443
2013	0.736	0.649	0.425
2014	0.625	0.543	0.436
2015	0.668	0.528	0.438
2016	0.656	0.523	0.434
2017	0.584	0.572	0.332
2018	0.608	0.456	0.432
2019	0.544	0.477	0.43
2020	0.506	0.578	0.333

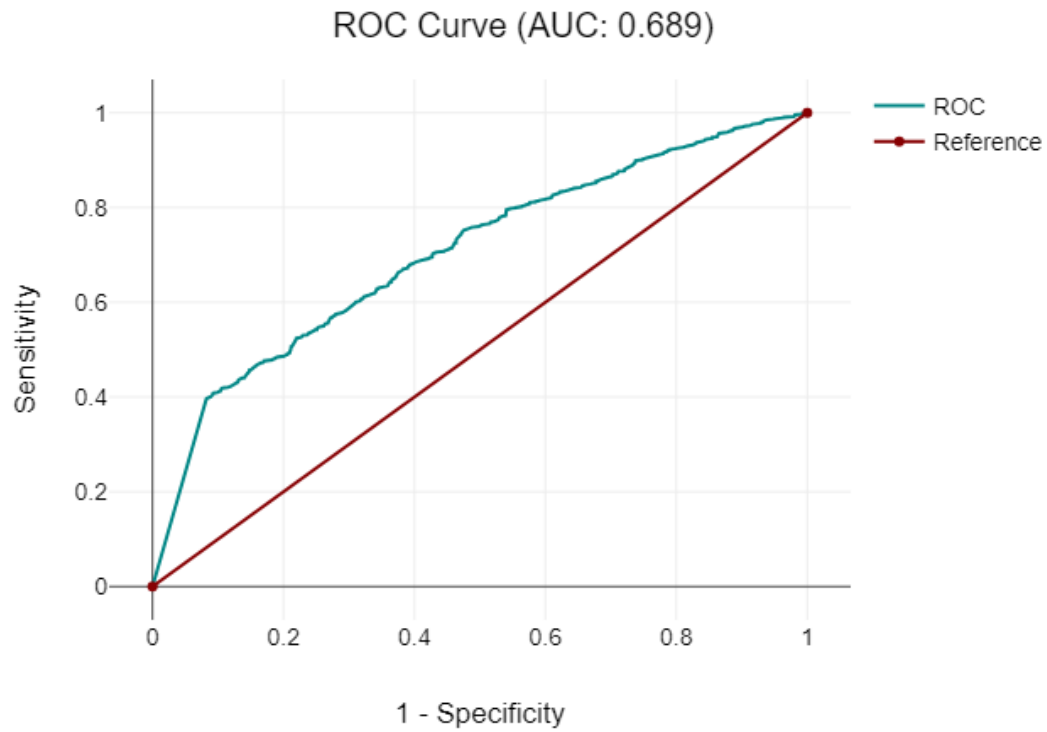


Figure 8. Model 1 (Geodesic) Temporally Aggregated ROC Curve

The ROC (Receiver Operating Characteristic) Curve is the graphical representation of the model's discriminatory ability to predict the true outcome. The space between the ROC Curve and the Reference is used to calculate the AUC (Area Under The Curve).

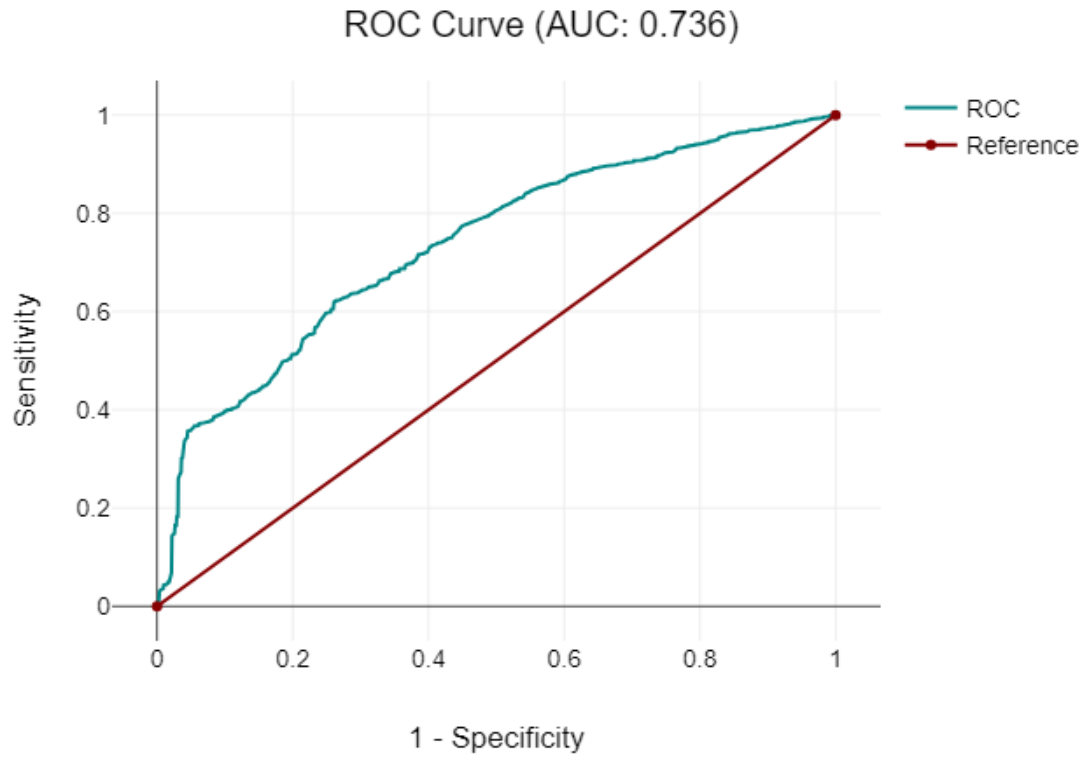


Figure 9. Model 2 (Network) Temporally Aggregated ROC Curve.

The ROC (Receiver Operating Characteristic) Curve is the graphical representation of the model's discriminatory ability to predict the true outcome. The space between the ROC Curve and the Reference is used to calculate the AUC (Area Under The Curve).

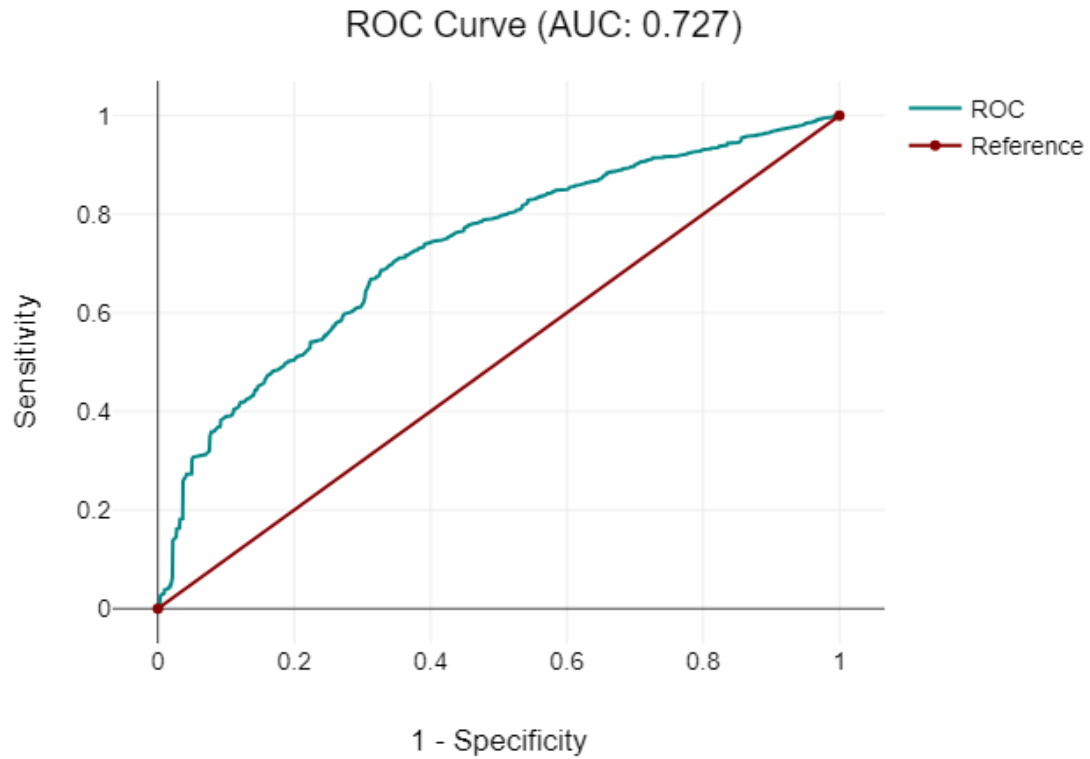


Figure 10. Model 2a (Major Roads Only) Temporally Aggregated ROC Curve.

The ROC (Receiver Operating Characteristic) Curve is the graphical representation of the model's discriminatory ability to predict the true outcome. The space between the ROC Curve and the Reference is used to calculate the AUC (Area Under The Curve).

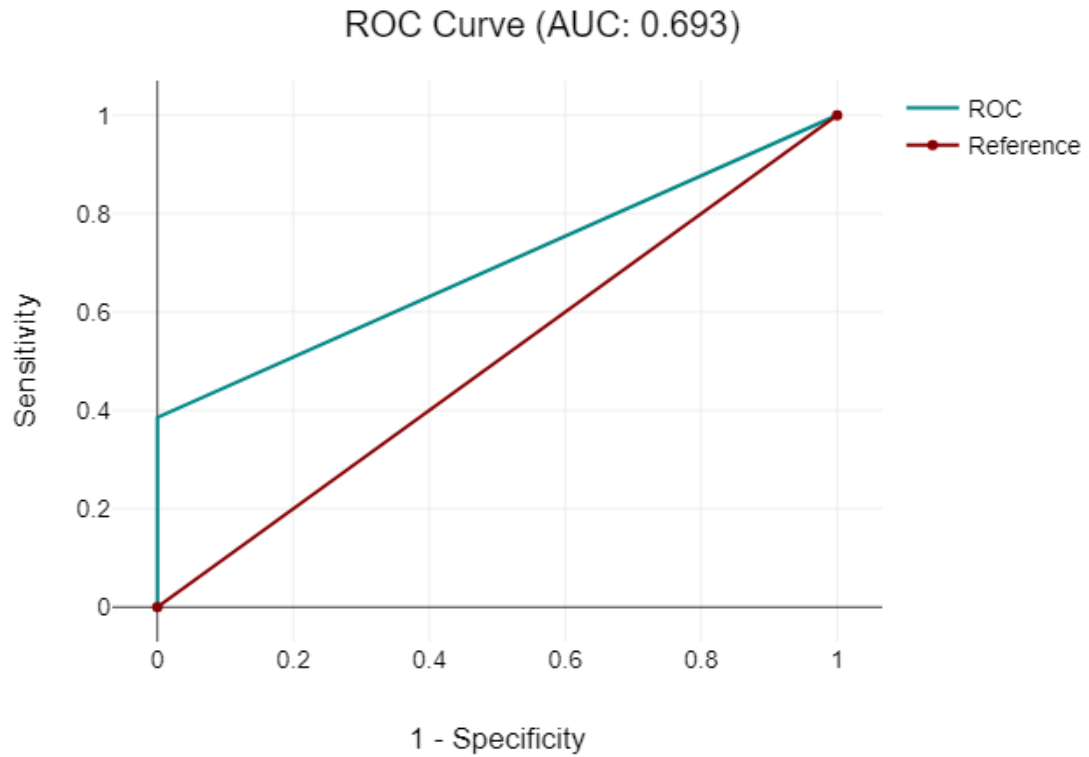


Figure 11. Model 3 (Cellular) Temporally Aggregated ROC Curve.

The ROC (Receiver Operating Characteristic) Curve is the graphical representation of the model's discriminatory ability to predict the true outcome. The space between the ROC Curve and the Reference is used to calculate the AUC (Area Under The Curve).

Bibliography

- Aboul-Enein, Basil, and Youssef Aboul-Enein. 2011. "A Theoretical Exploration of Lawrence of Arabia's Inner Meanings on Guerrilla Warfare." *Small Wars Journal*. <https://apps.dtic.mil/sti/pdfs/ADA546376.pdf>.
- Anderson, Scott. 2013. *Lawrence in Arabia: War, Deceit, Imperial Folly and the Making of the Modern Middle East*. New York: Knopf Doubleday Publishing Group.
- Arnold, James R. 2009. *Jungle of Snakes: A Century of Counterinsurgency Warfare from the Philippines to Iraq*. New York: Bloomsbury USA.
- Barfi, Barak. 2016. "The Military Doctrine of the Islamic State and the Limits of Ba'athist Influence." *CTC Sentinel* 9, no. 2 (February): 18-23. <https://ctc.westpoint.edu/the-military-doctrine-of-the-islamic-state-and-the-limits-of-baathist-influence/>.
- Beehener, Lionel, Liam Collins, Steven Ferenzi, and Mike Jackson. 2020 (revised). "The Taming of the Tigers," An MWI Contemporary Battlefield Assessment of the Counterinsurgency in Sri Lanka. 2. *Modern War Institute*, West Point. https://mwi.westpoint.edu/wp-content/uploads/2020/04/The-Taming-of-the-Tigers_April-2020.pdf.
- Benotman, Noman, and Roisin Blake. 2013. "Al Nusra Front: Syria's Top Islamist Militia." Wilson Center. <https://www.wilsoncenter.org/article/al-nusra-front-syrias-top-islamist-militia>.
- Blattman, Christopher. 2022. *Why We Fight: The Roots of War and the Paths to Peace*. Dublin: Penguin Random House LLC.
- Boulding, Kenneth E. 1962. *Conflict and defense: a general theory*. New York: Harper.
- Buhaug, Halvard. 2010. "Dude, Where's My Conflict? LSG, Relative Strength, and the Location of Civil War." *Conflict Management and Peace Science* 27, no. 2 (January): 107-128. <https://doi.org/10.1177/0738894209343974>.
- Buhaug, Halvard, and Scott Gates. 2002. "The Geography of Civil War." *Journal of Peace Research* 39, no. 4 (July): 417-433. <https://doi.org/10.1177/0022343302039004003>.
- Buhaug, Halvard, and Kristian Gleditsch. 2008. "Contagion or Confusion? Why Conflicts Cluster in Space." *International Studies Quarterly* 52, no. 2 (May): 215-233. <https://doi.org/10.1111/j.1468-2478.2008.00499.x>.

- Byman, Daniel. 2013. "Outside Support for Insurgent Movements." *Studies in Conflict & Terrorism* 36, no. 12 (July): 981-1004.
<https://doi.org/10.1080/1057610X.2013.842132>.
- Byman, Daniel, Peter Chalk, and Bruce Hoffman. 2001. *Trends in Outside Support for Insurgent Movements*. Santa Monica: RAND Corporation.
https://www.rand.org/pubs/monograph_reports/MR1405.html.
- Byman, Daniel L., and Jennifer R. Williams. 2015. "ISIS vs. Al Qaeda: Jihadism's global civil war." Brookings Institution. <https://www.brookings.edu/articles/isis-vs-al-qaeda-jihadisms-global-civil-war/>.
- Carter, David B., Morgan L. Kaplan, and Kenneth A. Schultz. 2022. "The Geography of Separatist Violence." *International Studies Quarterly* 66, no. 3 (September): 1-14.
<https://doi.org/10.1093/isq/sqac030>.
- Chopra, Sunil. 2003. Designing the Distribution Network in a Supply Chain. Vol. 39, no. 2. N.p.: Transportation Research Part E: *Logistics and Transportation Review*.
10.1016/S1366-5545(02)00044-3.
- CIA. n.d. "Guerrilla Warfare." CIA. Accessed October 17, 2023.
https://www.cia.gov/library/abbottabad-compound/4B/4BB2A65F1FF9040B4BDBFBAB15CC01C1_%D8%AD%D8%B1%D8%A8_%D8%A7%D9%84%D8%B9%D8%B5%D8%A7%D8%A8%D8%A7%D8%AA.pdf.
- Conflict Armament Research. 2014. Analysis of small-calibre ammunition recovered from Islamic State forces in Iraq and Syria.
<https://www.conflictarm.com/dispatches/islamic-state-ammunition-in-iraq-and-syria/>.
- Creveld, Martin v. 2004. *Supplying War: Logistics from Wallenstein to Patton*. New York: Cambridge University Press.
- Croicu, Mihai, and Ralph Sundberg. 2016. UCDP GED Codebook version 5.0. N.p.: Department of Peace and Conflict Research, Uppsala University.
<https://ucdp.uu.se/downloads/ged/ucdp-ged-50-codebook.pdf>.
- Department of Defense. 2018. "Counterinsurgency," Joint Publication 3-24. jcs.mil.
https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_24.pdf.
- Encyclopaedia Britannica. 2024. "Peninsular War | Definition, Battles, Dates, Significance, & Facts." Britannica. <https://www.britannica.com/event/Peninsular-War>.
- Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *The American Political Science Review* 97, no. 1 (February): 75-90.
<https://doi.org/10.1017/S0003055403000534>.

- Fettweis, Christopher J. 2000. "Sir Halford Mackinder, Geopolitics, and Policymaking in the 21st Sir Halford Mackinder, Geopolitics, and Policymaking in the 21st Century." *The US Army War College Quarterly: Parameters* The US Army War College Quarterly: *Parameters* 30, no. 2 (Summer): 58-71. 10.55540/0031-1723.1974.
- Foerster, Andreas. 2022. "Reassessing Jomini on Hybrid Warfare and Counterinsurgency." *Small Wars Journal*.
<https://smallwarsjournal.com/jrnl/art/reassessing-jomini-hybrid-warfare-and-counterinsurgency>.
- Gelvin, James L. 2015. *The Arab Uprisings: What Everyone Needs to Know*. New York: Oxford University Press.
- Geofabrik. n.d. "About Page." Geofabrik Download Server. Accessed July 3, 2024.
<https://download.geofabrik.de>.
- Glenn, Cameron, John Caves, Mattisan Rowan, and Garrett Nada. 2019. "Timeline: the Rise, Spread, and Fall of the Islamic State." Wilson Center.
<https://www.wilsoncenter.org/article/timeline-the-rise-spread-and-fall-the-islamic-state>.
- Goraya, G. S. 2019. *The Great Game: Afghanistan, British India and the Hundred Year War for India's Northern Frontier*. Panjab: Independently Published.
- Gordillo, Gastón. 2018. "Terrain as insurgent weapon: An affective geometry of warfare in the mountains of Afghanistan." *Political Geography* 64, no. May (May): 53-62.
<https://doi.org/10.1016/j.polgeo.2018.03.001>.
- Guardado, Jenny, and Steven Pennings. 2020. "The Seasonality of Conflict." World Bank Group, Development Economics, Development Research Group.
<https://documents1.worldbank.org/curated/en/261391598906198836/pdf/The-Seasonality-of-Conflict.pdf>.
- Gubbins, Colin, and Ian Hall. 2019 (revised). *The Art of Guerilla Warfare and The Partisan Leader's Handbook*. 2nd ed. 2 vols. N.p.: Independently Published.
- Guevara, Che. 2013. *Guerrilla Warfare*. Translated by Joseph P. Morray. N.p.: BN Publishing.
- Hahlweg, Werner. 1986. "Clausewitz and Guerrilla Warfare." *Journal of Strategic Studies* 9, no. 2-3 (June): 127-133. <https://doi.org/10.1080/01402398608437262>.
- Hammond, Jesse. 2017. "Maps of mayhem: Strategic location and deadly violence in civil war." *Journal of Peace Research* 55, no. 1 (September): 32-46.
<https://doi.org/10.1177/0022343317702956>.

- Hanna, Andrew. 2019. "Timeline: The Life and Death of Abu Bakr al Baghdadi." Wilson Center. <https://www.wilsoncenter.org/article/timeline-the-life-and-death-abu-bakr-al-baghdadi>.
- Hassan, Hassan. 2018. "How Did ISIS Really Emerge?" *The Atlantic*, November 30, 2018. <https://www.theatlantic.com/ideas/archive/2018/11/isis-origins-anbari-zarqawi/577030/>.
- Heidelberg Institute for International Conflict Research. 2022. "Conflict Barometer 2022." Heidelberg Institute for International Conflict Research. <https://hiik.de/?lang=en>.
- Hendrix, Cullen S. 2011. "Head for the Hills? Rough Terrain, State Capacity, and Civil War Onset." *Civil Wars* 13, no. 4 (December): 345-370. <https://doi.org/10.1080/13698249.2011.629863>.
- Inowlocki, Tania. 2016. *Tracing the Supply of Components Used in Islamic State IEDs*. N.p.: Conflict Armament Research. <https://www.conflictarm.com/reports/tracing-the-supply-of-components-used-in-islamic-state-ieds/>.
- Inowlocki, Tania. 2017. *Weapons of the Islamic State*. N.p.: Conflict Armament Research. <https://www.conflictarm.com/reports/weapons-of-the-islamic-state/>.
- Inowlocki, Tania. 2020. *Procurement Networks Behind Islamic State Improvised Weapon Programmes*. N.p.: Conflict Armament Research. <https://www.conflictarm.com/reports/procurement-networks-behind-islamic-state-improvised-weapon-programmes/>.
- Jawad Al-Tamimi, Aymenn. 2019. "'Stations' of Abu Bakr al-Baghdadi's Life: Translation and Analysis." *Pundicity*, November 7, 2019. <https://www.aymennjawad.org/2019/11/stations-of-abu-bakr-al-baghdadi-life-translation>.
- Jelineo, Jake R. 2012. *Napoleon's Logistics; or How Napoleon Learned to Worry about Supply*. N.p.: Air Command and Staff College, Air University. <https://apps.dtic.mil/sti/tr/pdf/AD1022125.pdf>.
- Kalyvas, Stathis N. 2012. *The Logic of Violence in Civil War*. Boston: Cambridge University Press. <https://doi.org/10.1017/CBO9780511818462>.
- Kamolnick, Paul. 2017. *The Al-Qaeda Organization and the Islamic State Organization: History, Doctrine, Modus Operandi, and U.S. Policy to Degrade and Defeat Terrorism Conducted in the Name of Sunni Islam*. Edited by Army War College (U.S.). Strategic Studies Institute. N.p.: Strategic Studies Institute and U.S. Army War College Press. <https://press.armywarcollege.edu/monographs/3/>.
- Keegan, John. 2002. *Intelligence in War*. New York: Vintage Books.

- Kibris, Arzu. 2021. "The geo-temporal evolution of violence in civil conflicts: A micro analysis of conflict diffusion on a new event data set." *Journal of peace research* 58, no. 5 (March): 885-899. <https://doi.org/10.1177/0022343320978695>.
- Kilani, Feras. 2021. "A Caliph Without a Caliphate: The Biography of ISIS's New Leader." *New Lines Magazine*, April 15, 2021. <https://newlinesmag.com/reportage/a-caliph-without-a-caliphate-the-biography-of-isis-new-leader/>.
- Kilcullen, David. 2008. "Political Maneuver in Counterinsurgency." *Small Wars Journal*. <https://smallwarsjournal.com/blog/political-maneuver-in-counterinsurgency>.
- Knights, Michael. 2018. "The Islamic State Inside Iraq: Losing Power or Preserving Strength?" *CTC Sentinel* 11, no. 11 (December): 1-10. <https://ctc.westpoint.edu/islamic-state-inside-iraq-losing-power-preserving-strength/>.
- Koven, Barnett S. 2017. "The End of Afghanistan's Spring Fighting Seasons and the Demise of the Afghan National Security Forces?" *Small Wars Journal*. <https://smallwarsjournal.com/jrnl/art/the-end-of-afghanistan%E2%80%99s-spring-fighting-seasons-and-the-demise-of-the-afghan-national-secu>.
- Lawrence, Thomas E. 2017. *Seven Pillars of Wisdom*. N.p.: CreateSpace Independent Publishing Platform.
- Levallios, Agnès, and Jean-Claude Cousseran. 2017. "The Financing of the 'Islamic State' in Iraq and Syria (ISIS)," Report. European Parliament. https://www.europarl.europa.eu/RegData/etudes/IDAN/2017/603835/EXPO_IDA%282017%29603835_EN.pdf.
- Lewis, Janet I. 2020. *How Insurgency Begins: Rebel Group Formation in Uganda and Beyond*. Boston: Cambridge University Press. <https://doi.org/10.1017/9781108855969>.
- Mackinder, Halford J. 1904. "The Geographical Pivot of History." *The Geographical Journal* XXIII, no. 4 (April): 421-437.
- Mahan, Alfred T. 1918. *The Influence of Sea Power Upon History, 1660-1783*. Boston: Boston Little, Brown and Company. <https://doi.org/10.1017/CBO9780511783289>.
- Maizland, Lindsay. 2023. "The Taliban in Afghanistan." Council on Foreign Relations, January 19, 2023. <https://www.cfr.org/background/taliban-afghanistan>.
- Maizland, Lindsay, Zachary Laub, and Mohammed Aly. 2016. "The Islamic State." Council on Foreign Relations. <https://www.cfr.org/background/islamic-state>.

- McColl, Robert W. 1969. "The Insurgent State: Territorial Bases Of Revolution." *Annals of the Association of American Geographers* 59, no. 4 (December): 613-631. 10.1111/j.1467-8306.1969.tb01803.x.
- Milton, Daniel. 2016. "Communication Breakdown: Unraveling the Islamic State's Media Efforts," Report. Combating Terrorism Center at West Point. <https://ctc.westpoint.edu/communication-breakdown-unraveling-the-islamic-states-media-efforts/>.
- Moroni Bracamonte, José A., and David E. Spencer. 1995. *Strategy and Tactics of the Salvadoran FMLN Guerrillas: Last Battle of the Cold War, Blueprint for Future Conflicts*. Westport: Bloomsbury Academic. <https://doi.org/10.1215/00182168-76.2.369>.
- Mu'ller-Crepon, Carl, Philipp Hunziker, and Lars-Erik Cederman. 2021. "Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa." *Journal of Conflict Resolution* 65, no. 2-3 (October): 563-590. 10.1177/0022002720963674.
- Nagl, John A. 2005. *Learning to Eat Soup with a Knife: Counterinsurgency Lessons from Malaya and Vietnam*. Westport: University of Chicago Press. <https://doi.org/10.1353/jmh.2003.0205>.
- National Army Museum. 2024. *Peninsular War*, Featured Article. London, United Kingdom: n.p. <https://www.nam.ac.uk/explore/peninsular-war>.
- National Counterterrorism Center. 2022. "Islamic State of Iraq and Ash-Sham (ISIS)." Counterterrorism Guide: Foreign Terrorist Organizations. https://www.dni.gov/nctc/ftos/isis_fto.html.
- O'Loughlin, John, and Frank Witmer. 2011. "The Localized Geographies of Violence in the North Caucasus of Russia, 1999-2007." *Annals of the Association of American Geographers* 101, no. 1 (December): 178-201. <https://doi.org/10.1080/00045608.2010.534713>.
- O'Loughlin, John, Frank Witmer, and Andrew Linke. 2010. "The Afghanistan-Pakistan Wars, 2008-2009: Micro- geographies, Conflict Diffusion, and Clusters of Violence." *Eurasian Geography and Economics* 51, no. 4 (May): 437-471. <https://doi.org/10.2747/1539-7216.51.4.437>.
- Open Street Map. n.d. "About Open Street Map." OpenStreetMap. Accessed July 3, 2024. <https://www.openstreetmap.org/about>.
- Orton, Kyle. 2016. "Governing the Caliphate: Profiles of Islamic State Leaders." Centre for the Response to Radicalization and Terrorism. <https://henryjacksonsociety.org/wp-content/uploads/2016/07/IS-leaders-report.pdf>.

- Otero-Bahamon, Silvia, Simon Uribe, and Isabel Penaranda-Currie. 2022. "Seeing like a Guerrilla: the logic of infrastructure in the building of insurgent orders." *Geoforum* 133 (July): 198-207. <https://doi.org/10.1016/j.geoforum.2021.10.009>.
- Owens, Mackubin T. 2019. "Sanctuary: The Geopolitics of Terrorism and Insurgency." Institute of World Politics. https://www.iwp.edu/wp-content/uploads/2019/05/20180620_SanctuaryTheGeopoliticsofTerrorismandInsurgency.pdf.
- Oxford English Dictionary. 2019. "Definition of Guerrilla." Oxford English Dictionary. https://www.oed.com/dictionary/guerrilla_n?tab=etymology.
- Parker, Ned. 2015. "After Iraqi forces take Tikrit, a wave of looting and lynching." *Reuters*, April 3, 2015. <https://www.reuters.com/investigates/special-report/mideast-crisis-iraq-tikrit/>.
- Payne, Ed. 2015. "Sinjar: Kurds try to retake key Iraqi town from ISIS." *CNN*, November 12, 2015. <https://www.cnn.com/2015/11/11/middleeast/iraq-free-sinjar-isis/index.html>.
- Raleigh, Clionadh, and Havard Hegre. 2009. "Population size, concentration, and civil war. A geographically disaggregated analysis." *Political Geography* 28, no. 4 (May): 224-238. <https://doi.org/10.1016/j.polgeo.2009.05.007>.
- Ramm, Frederik. 2022. "OpenStreetMap Data in Layered GIS-Format Version 0.7.12." Geofabrik. <https://www.geofabrik.de/data/geofabrik-osm-gis-standard-0.7.pdf>.
- Roggio, Bill. 2022. "Islamic State's emir dead after U.S. military raid in Syria." *Long War Journal*, February 3, 2022. <https://www.longwarjournal.org/archives/2022/02/islamic-states-emir-dead-after-u-s-military-raid-in-syria.php>.
- Schutte, Sebastian. 2014. "Geography, outcome, and casualties: a unified model of insurgency." *The Journal of Conflict Resolution* 59, no. 6 (March): 1101-1128. <https://doi.org/10.1177/0022002713520534>.
- Schutte, Sebastian, and Nils B. Weidmann. 2011. "Diffusion patterns of violence in civil wars." *Political Geography* 30, no. 3 (March): 143-152. <https://doi.org/10.1016/j.polgeo.2011.03.005>.
- Shrader, Charles R. 1995. *Communist logistics in the Korean War*. Westport: Bloomsbury Academic.
- Simon-Skjodt Center for the Prevention of Genocide. March. "From a Democratic Uprising to Full-Scale War." United States Holocaust Memorial Museum. <https://www.ushmm.org/genocide-prevention/countries/syria/full-scale-war>.

- Staff Writers. 2016. "A century on: Why Arabs resent Sykes-Picot." *Al Jazeera*, 2016. <https://interactive.aljazeera.com/aje/2016/sykes-picot-100-years-middle-east-map/index.html>.
- Staff Writers. 2019. "The rise and fall of the Islamic State group: The long and short story." *BBC*, March 23, 2019. <https://www.bbc.com/news/world-middle-east-47210891>.
- Staff Writers. 2023. "T.E. Lawrence on guerrilla warfare | Insights & Tactics." Britannica. <https://www.britannica.com/topic/T-E-Lawrence-on-guerrilla-warfare-1984900>.
- Stina, Hogbladh. 2023. *UCDP GED Codebook version 23.1*. N.p.: Department of Peace and Conflict Research, Uppsala University.
- Sundberg, Ralph, and Erik Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset." *Journal of Peace Research* 50, no. 4 (July): 523-532. <https://doi.org/10.1177/0022343313484347>.
- Tollefsen, Andreas, and Halvard Buhaug. 2015. "Insurgency and Inaccessibility." *International Studies Review* 17, no. 1 (March): 6-25. <https://doi.org/10.1111/misr.12202>.
- Townsley, Michael, and Shane D. Johnson. 2008. "Space Time Dynamics of Insurgent Activity in Iraq." *Security Journal* 21, no. 3 (June): 139-146. [10.1057/palgrave.sj.8350090](https://doi.org/10.1057/palgrave.sj.8350090).
- Tse-Tung, Mao. 2017. *On Guerilla Warfare: Mao Tse-Tung On Guerilla Warfare*. Translated by Samuel B. Griffith. Mansfield Centre: Martino Fine Books.
- Turbiville, Graham H. 2005. "Logistic Support and Insurgency." Defense Technical Information Center. <https://apps.dtic.mil/sti/citations/ADA495398>.
- United Nations. 2023. "FAO GIEWS Country Brief on Iraq -." Food and Agriculture Organization of the United Nations. <https://www.fao.org/giews/countrybrief/country.jsp?code=IRQ>.
- United States Marine Corps. 1997. *Warfighting*, Marine Corps Doctrinal Publication 1. N.p.: Department of Defense.
- United States Marine Corps. 2023. *Logistics*, Marine Corps Doctrinal Publication 4. N.p.: Department of Defense.
- Uppsala University. 2024. "About The Uppsala Conflict Data Program." Uppsala Conflict Data Program (UCDP). <https://www.uu.se/en/departement/peace-and-conflict-research/research/ucdp/>.

- U.S. Department of State. 2001. "Glossary." U.S. Department of State Archive. <https://2001-2009.state.gov/s/ct/info/c16718.htm#:~:text=Terrorism%3A%20Premeditated%2C%20politically%20motivated%20violence,As%20per%2022%20USCS%202656f>.
- US Government. 2012. *Guide to the Analysis of Insurgency*.
- U.S. Office of the Director of National Intelligence. 2022. "National Counterterrorism Center." National Counterterrorism Center | FTOs. https://www.dni.gov/nctc/ftos/isis_fto.html.
- Vergun, David. 2022. "Leader of ISIS Dead Following U.S. Raid in Syria." *Defense.gov*, February 3, 2022. <https://www.defense.gov/News/News-Stories/Article/Article/2922796/leader-of-isis-dead-following-us-raid-in-syria/>.
- Walker, Martin. 2013. "Guerilla warfare's epic history." *Wilson Quarterly*. https://www.wilsonquarterly.com/quarterly/_/guerilla-warfares-epic-history#.
- Walter, Barbara F. 2022. *How Civil Wars Start*. New York: Crown.
- Watson, Mason W. 2021. *The Conflict with ISIS: Operation INHERENT RESOLVE, June 2014–January 2020*. N.p.: US Army Center for Military History. <https://history.army.mil/html/books/078/78-2/index.html>.
- Weidmann, Nils B., and Michael D. Ward. 2010. "Predicting Conflict in Space and Time." *The Journal of Conflict Resolution* 54, no. 6 (July): 883-901. <https://doi.org/10.1177/0022002710371669>.
- Weinstein, Jeremy M. 2007. *Inside Rebellion: The Politics of Insurgent Violence*. Boston: Cambridge University Press.
- Weiss, Gordon. 2012. *The Cage: The Fight for Sri Lanka and the Last Days of the Tamil Tigers*. New York: Bellevue Literary Press.
- Wilkins, Frederick. 1954. "Guerrilla Warfare." *U.S. Naval Institute Proceedings* 80, no. 3 (March): 613. <https://www.usni.org/magazines/proceedings/1954/march/guerrilla-warfare>.
- WorldPop. n.d. "About." WorldPop. Accessed July 4, 2024. <https://www.worldpop.org/about/>.
- WorldPop. n.d. "Population Counts." WorldPop :: Population Counts. Accessed July 4, 2024. <https://hub.worldpop.org/geodata/listing?id=29>.
- Wright, Robin, J. M. Berger, William Braniff, Cole Bunzell, Daniel Byman, Jennifer Cafarella, Harleen Gambhir, et al. 2017. "The Jihadi Threat: ISIS, al-Qaeda, and

- Beyond.” U.S. Institute of Peace. <https://www.usip.org/sites/default/files/The-Jihadi-Threat-ISIS-Al-Qaeda-and-Beyond.pdf>.
- Young, PhD, Robert. 2021. “The Guerrilla and the Peninsular War.” *The Saber and Scroll Journal* 10, no. 2 (Winter): 67-80. <https://doi.org/10.18278/sshj.10.2.6>.
- Zhukov, Yuri M. 2012. “Roads and the diffusion of insurgent violence.” *Political Geography* 31, no. 3 (March): 144-156. 10.1016/j.polgeo.2011.12.002.
- Zhukov, Yuri M. 2016. “On the Logistics of Violence.” In *Economic Aspects of Genocides, Other Mass Atrocities, and Their Prevention*, edited by Charles H. Anderton and Jurgen Brauer, 399-424. Oxford: Oxford University Press.
- Zinn, Cherish M. 2016. “Consequences of Iraqi De-Baathification.” Volume IX. [cornell.edu. https://journals.library.cornell.edu/tmpfiles/CIAR_9_2_2.pdf](https://journals.library.cornell.edu/tmpfiles/CIAR_9_2_2.pdf).