Spatial Associations between Land Use and Infectious Disease:
WNV in the United States and Zika in Colombia

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Josh S. Weinstein
Master of Science
George Mason University, 2003
Bachelor of Science
George Mason University, 1998

Director: Timothy F. Leslie, Associate Professor
Department of Geography and GeoInformation Sciences

Fall Semester 2017
George Mason University
Fairfax, Virginia

DEDICATION

To my incredible family in Trish, Ethan, and Hailey, whose encouragement and patience fueled the completion of this research. A special thank you to my committee in Dr. Timothy Leslie, Dr. Cara Frankenfeld, Dr. Arie Croitoru, and Dr. Anthony Stefanidis. Dr. Leslie, in particular, never served answers on a platter, but rather advised in such a way that I discovered answers myself.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Akaike Information Criterion Correlation……………………………………………….. AICC
Bayesian Information Criterion…………………………………………………………… BIC
Centers for Disease Control and Prevention………………………………………………CDC
Database of Global Administrative Areas…………………………………………… GADM
Food and Agriculture Organization…………………………………………………….. FAO
Geographic Information System…………………………………………………………... GIS
Global Administrative Unit Layers……………………………………………………... GAUL
Global Forest Watch………………………………………………………………………...GFW
Global Land Cover SHARE……………………………………………………. GLC-SHARE
Gross Domestic Product…………………………………………………………………... GDP
Health and Human Services……………………………………………………………... HHS
International Geosphere Biosphere Programme……………………………………IGBP
Land Cover Classification System………………………………………………………. LCCS
Land Use / Land Cover…………………………………………………………………... LULC
Local Administrative Unit………………………………………………………………... LAU
Moderate Resolution Imaging Spectrometer……………………………………….MODIS
Modifiable Areal Unit Problem…………………………………………………………MAUP
Normalized Difference Vegetation Index……………………………………………NDVI
National Institute of Allergy and Infectious Diseases……………………………NIAID
National Institutes of Health…………………………………………………………….NIH
Population Estimates Program………………………………………………………… PEP
Principal Component……………………………………………………………………… PC
Principal Component Analysis………………………………………………………... PCA
Research Class………………………………………………………………………………... RC
Research Question………………………………………………………………………….. RQ
Severe Acute Respiratory Syndrome……………………………………………………SARS
Shuttle Radar Topography Mission……………………………………………….. SRTM
Simian Immunodeficiency Virus………………………………………………………... SIV
Statistical Package for the Social Sciences…………………………………………... SPSS
United Nations………………………………………………………………………………...UN
United States Geological Survey…………………………………………………….. USGS
Vector-Borne Infectious Disease………………………………………………………... VBID
World Health Organization…………………………………………………………….. WHO
West Nile Virus………………………………………………………………………………WNV

ABSTRACT

SPATIAL ASSOCIATIONS BETWEEN LAND USE AND INFECTIOUS DISEASE:
WNV IN THE UNITED STATES AND ZIKA IN COLOMBIA

Josh S. Weinstein, PhD

George Mason University, 2017

Dissertation Director: Dr. Timothy F. Leslie

This dissertation seeks to address three research questions through the context of spatial associations between land use / land cover (LULC) and vector-borne infectious disease (VBID). These research questions are: (1) Do spatial associations exist between the linear (edge) density of LULC boundaries and VBID occurrence? (2) Do patterns of spatial associations repeat over time? (3) Do patterns of spatial associations repeat across space?

Understanding how LULC change influences disease emergence informs the prevention and mitigation of local disease outbreaks prior to transmission growth into regional epidemics or global pandemics. Close and frequent human contact with infected arthropod vectors near local-level LULC boundaries drives VBID emergence. Increasingly dense and fragmented LULC boundaries result from human activities in

the expansion of urban, pastoral, and agricultural areas. Fragmentation increases the likelihood of pathogen spillover at local-level LULC boundaries, the human-physical interface. Unmitigated and uncontrolled local spillover events can grow in spatial scale and result in significant social and economic impacts. Measuring the human-physical interface to identify spillover hotspots prior to VBID emergence and increasing levels of disease transmission is paramount to protecting public health.

Methods that measure the human-physical interface influence our ability to identify areas with elevated risk of VBID emergence. Prior research used remote sensing, field research, or literature reviews to identify substantive associations between LULC and VBID emergence. The research within this dissertation focuses on the spatial association between the linear density of LULC boundaries and VBID occurrence through spatial statistical methods, to include Principal Component Analysis and negative binomial regression. Proportion abundance and patch density are supplemental landscape metrics that add context to linear (edge) density. Case studies involve West Nile Virus in the contiguous United States from 2003 through 2014 and Zika in Colombia, South America during 2016. The goal is a method that can make use of land development plans to identify areas that could experience elevated VBID occurrence.

1. INTRODUCTION

Unmitigated and uncontrolled VBID outbreaks result in significant social and economic impacts. In 2003, the global economy incurred an estimated $30-50 billion loss due to a Severe Acute Respiratory Syndrome (SARS) outbreak originating from a live-market in Guangdong province, China (Karesh et al. 2012). While this single SARS outbreak killed 775, malaria kills roughly 2.7 million and infects nearly 400 million people annually (Wilcox and Ellis 2006).

More recently, nearly 50 Central- and South-American countries experienced their first Zika occurrence (WHO 2016). Argentine, Bolivia, Brazil, Ecuador, and Peru have reported an average of 1,246 weekly cases since early 2017 (PAHO WHO 2017) and transmission of the disease is progressing further into temperate zones (WHO 2016). During the first half of 2017, vector-to-human WNV transmission spans 18 U.S. states and severe neuro-invasive diseases, such as meningitis and encephalitis, account for 57 percent of the reported cases (CDC Preliminary 2017).

Close and frequent human contact with infected vectors at local-level LULC boundaries drives VBID emergence (Morse 2004, Morse et al. 2012, Pike et al. 2010, Wilcox and Ellis 2006, Woolhouse et al. 2012). Understanding how and where the human modification of LULC boundaries adversely influences VBID emergence will enable prevention and mitigation before local disease outbreaks grow in spatial scope (McFarlane et al. 2013, Murray and Daszak 2013).

LULC boundaries – the human-physical interface – are fragmenting at an increasing rate as the resource demands of a growing global population drives expansion of urban, pastoral, and agricultural areas (Patz et al. 2008, Walsh et al. 1993, Wilcox and Ellis 2006, Vittor et al. 2006). Agricultural development in parallel with urbanization and population growth represents the most significant reason for LULC change and is the major driver of deforestation and forest fragmentation in the tropics (Patz et al. 2008). In fact, roughly half of all global infectious disease emergence events are associated with changes in LULC for agriculture and food production (McFarlane et al. 2013).

Monitoring local-level emergence through VBID surveillance and predictive analytics is necessary to protect public health. However, the seven leading epidemiology journals evaluated by Auchincloss et al. (2012) cited spatial methodologies in roughly one percent of all articles and the complex spatial association between LULC and VBID is known only for a few diseases (Ezenwa et al. 2007, Jones et al. 2008, Patz et al. 2004). Backed by even less research is the epidemiological impact of LULC fragmentation, which increases the linear (edge) density of LULC boundaries.

This dissertation seeks to address three research questions (RQ) in the context of the spatial association between the linear density of LULC boundaries and West Nile Virus (WNV) in the contiguous United States and Zika in Colombia, South America. Spatial statistical methods aid in the discovery of associations. Proportion abundance and patch density are supplemental landscape metrics that will add context to the spatial associations discovered through the linear density landscape metric. Quantifying such spatial associations sheds light on the degree to which alteration of the land influences the magnitude of VBID occurrence.

Data that spans narrow temporal periods are snapshots that could include uncertainty borne from confounding environmental events. Iteration of regression models over multiple years will mitigate the effect of confounding environmental events, such as El Niño and La Niña (McClintock et al. 2010). Temporal comparisons of the density of LULC boundaries and VBID occurrence for WNV in the contiguous United States over a 12-year period will help identify patterns of spatial associations over time while mitigating the influence of El Niño or La Niña. Comparisons of spatial associations between WNV in the United States and Zika in Colombia, a different study area and VBID, will provide support to or the rejection of patterns across space.

Identifying the magnitude of associations between LULC and VBID will benefit policy makers and spatial epidemiologists alike. McFarlane et al. (2013) and Murray and Daszak (2013) recommend development of a strategy to define high-risk regions so land development plans may be adapted to mitigate future VBID risk. Further, identification of high-risk areas will provide spatial epidemiologists another tool in the allocation of resources. Such an early warning tool could aid in prevention and mitigation, thus dampening the social and economic impacts of VBID.

2. BACKGROUND AND LITERATURE REVIEW

LULC is associated with roughly half of all global infectious disease emergence (McFarlane et al. 2013). The demands of a rapidly growing global population drives LULC change through agricultural development, urbanization, and deforestation. These drivers place humans in close and frequent contact with vectors at local LULC boundaries, increasing the risk of VBID emergence (Jones et al. 2008). Measurement of LULC has enabled the identification, quantification, and mapping of vector range and abundance and the prediction of VBID hotspots (Figure 1).

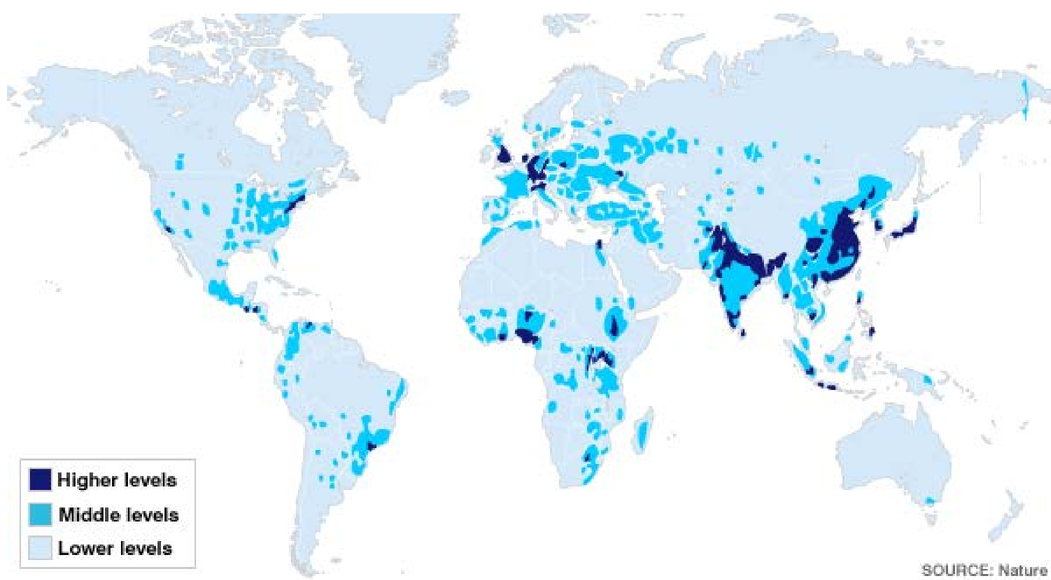


Figure 1. VBID hotspots based on frequency of human-vector contact (Jones et al. 2008).

Vanwambeke et al. (2011), Ezenwa et al. (2007), and Allan et al. (2003) performed field studies that identified strong associations between LULC and increased risk of VBID, such as dengue, WNV, and Lyme respectively. Vector range, prevalence, and abundance discovered using mosquito traps informed the development of vector-to-host ratio maps (Vanwambeke et al. 2011) or the analysis of landscape influence on VBID occurrence (Ezenwa et al. 2007, Allan et al. 2003). Ratmanov et al. (2013) and De La Rocque et al. (2004) exceed the spatial scope and collection frequency of field studies while reducing the time required to collect data by using remotely sensed imagery to identify LULC and VBID associations.

Regardless of research method, an understanding of the complex associations between LULC and VBID exists for only a few diseases (Ezenwa et al. 2007, Jones et al. 2008, Patz et al. 2004). Murray and Daszak (2013), McFarlane et al. (2013), Morse (2012), and Eisenberg et al. (2007) cite the pressing need to increase research into these associations to prevent and mitigate the impact of VBID. This pressing need exists because the "replacement and removal" of natural flora – LULC change – is the most significant influencer of disease emergence (McFarlane at al. 2013). Through LULC change, humans created what Eisenberg at al. (2007) describe as a landscape of human disease and we must now develop early warning methods that examine this landscape to reduce disease outbreaks (Morse 2012).

2.1 Social and Economic Impacts

VBID outbreaks that grow into epidemic or pandemic proportions generate social and economic impacts that resonate globally. More than 2.5 billion humans, nearly 36 percent of the global population, are at risk of dengue, and roughly 390 million people are infected annually (Bhat et al. 2013). Similarly, malaria has annual morbidity and mortality rates of nearly 400 million and 2.7 million respectively (Wilcox and Ellis 2006). The 2003 outbreak of SARS killed 775; a low number relative to malaria, yet panic across financial markets resulted in an estimated $30 to $50 billion hit to the global economy (Karesh et al. 2012).

Social and economic impacts of dengue, malaria, and other VBIDs, compounded by slow government response, can breed regional instability. VBID response by national and international health organizations cannot always diffuse volatile environments in enough time to maintain social and economic stability. As such, the United States National Security Council, Department of Defense, and Department of State cited the need for increased collaboration with the World Health Organization (WHO) and international partners to strengthen detection and response to emerging infectious diseases (Nuzzo and Gronvall 2011).

The 2014 Ebola outbreak in Sierra Leone, Liberia, Guinea, Nigeria, and Senegal exemplifies VBID-borne regional instability. These West African countries, already among the world's poorest yet positioned to leave behind years of civil war and enter Africa's resource extraction economic boom, experienced setbacks in social and economic stability due to increased public fear and reduced levels of tourism, trade, agriculture, and mining (MacDougall and Farge 2014). Starting at local LULC interfaces, unmitigated VBID emergence can produce impacts at various scales.

2.2 Scales of VBID Transmission

Approximately 75 percent of all known infectious diseases are, or once were, transmitted by vectors at the local LULC boundaries – the human-physical interface (Wilcox and Ellis 2006). Understanding interface dynamics is critical to the mitigation of downstream impacts at increased scales of disease transmission. As illustrated in Figure 2, numerous models portray the scales of infectious disease transmission, from local emergence, or spillover, to transmission at the global scale.

Local-level disease transmission begins with disease spillover at the physical interface between humans and pathogenic vectors. An increase in the spatial scope of disease occurrence characterizes each step up the transmission ladder. Continued and unchecked transmission at the local-level can lead to a regional epidemic or global pandemic.

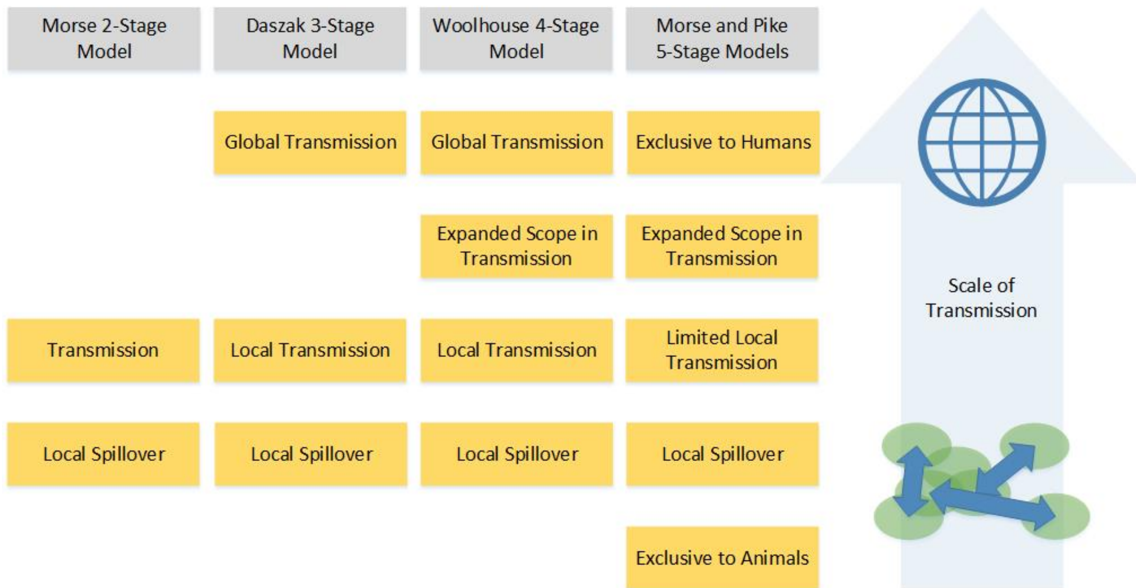


Figure 2. Transmission models share commonalities in local spillover and increasing scales of transmission.

The Morse (2004) two-stage infectious disease transmission model entails spillover, where an infectious disease jumps from pathogenic vector to host at the local interface, followed by transmission within the human host population at an increased spatial scale. The Daszak three-stage model subdivides the post-spillover stage of transmission within the human population into two parts – local area transmission and continued transmission from the local- to the global-level (Bogich et al. 2012). The United States Agency for International Development PREDICT program, designed to detect pathogens at the interface between humans and wildlife, adopted the Daszak three-stage model of local spillover followed by local, then global transmission (USAID 2014).

Woolhouse et al. (2012) adds a fourth stage between local and global transmission, where transmission expands in spatial scope. The Morse et al. (2012) and Pike et al. (2010) five-stage models include the Woolhouse transitional stage that connects local to global transmission, along with a pre-spillover stage where the disease pathogen is found solely within non-human animal populations.

Commonalities between these models include a progressive increase in spatial scales of transmission and initial spillover at the human-physical interface. Developed through academic rigor, these models are high-level representations of complex dynamics between pathogens, vectors, hosts, and the environment.

2.3 VBID Dynamics

Complex dynamics between pathogens, vectors, and hosts at the human-physical interface influence spillover and makes the tasks of targeted VBID prevention and mitigation difficult (Wilcox and Colwell 2005). Arthropod vectors typically become infected by a pathogen while feeding on infected reservoir hosts, such as birds, rodents, and other larger animals. After feeding on a reservoir host, vectors can pass pathogens to susceptible humans. An increase in contact frequency with infected vectors at the human-physical interface accelerates pathogen evolution, which increases infectivity and transmission risk (Wilcox and Ellis 2006). Unpredictable and uneven levels of contact between humans and pathogens contributes to both the dynamism and complexity of forecasting, let alone measuring, pathogen evolution and infectivity, and the association between LULC and VBID.

LULC change further alters the pathogen-vector-host relationship by providing pathogens an opportunity to expand their range and exploit new habitat niches (Murray and Daszak 2013). Malarial outbreaks in the tropics often spike after forest felling for road and ditch construction creates new mosquito breeding habitats. In the northeast United States, the reduction of predator species due to LULC change resulted in a population explosion of tick host species, such as deer and mice, which in turn led to the resurgence in tick-borne Lyme disease (McMichael 2004).

The increase in tsetse fly borne sleeping sickness in East Africa is yet another example of the human influence on vector abundance. Tsetse flies, the sleeping sickness vector, shifted feeding patterns from cattle and wildebeests to humans when the introduction of the rinderpest rapidly depleted the cattle and wildebeest populations. In this example, an invasive species introduced by humans resulted in a change to pathogen dynamics followed by an increase of sleeping sickness occurrences (Karesh et al. 2012).

LULC change and encroachment into vector habitats alters pathogen dynamics, which amplifies the pathogen transmission cycle through an increase in the frequency of human contact with infected vectors (Wilcox and Ellis 2006). Identification of potential disease hotspots through greater understanding of the association between LULC and VBID is critical in the protection of public health.

2.4 LULC and VBID Dynamics

Spatial epidemiologists have identified strong associations between LULC and VBID despite the significant knowledge gaps in our understanding of the complex dynamics at the human-physical interface. Table 1 depicts the major transitions in human interaction with the land and corresponding infectious disease emergence.

Table 1. Major transitions in human interaction with the land increased frequency of contact with pathogenic vectors (McMichael 2004).

Transition	Description	Infectious Disease
8,000 – 3,000 BC	Early settlements	Early versions of cholera, typhus, influenza, smallpox, malaria
1,000 BC – 500 AD	Early Eurasian civilization	Bubonic plague
500 – 1,900s	European exploration	Measles, yellow fever, dengue
Modern day	Globalization & rapid population growth	SARS, Ebola, Zika

The transition from hunting and gathering to agrarian- and pastoral-based societies produced significant and widespread changes at the human-physical interface (McMichael 2004). Such changes to these early settlements resulted in the emergence of cholera, typhus, smallpox, and other infectious diseases (Wilcox and Ellis 2006). Accelerated expansion of our global footprint expressed as rapid and extensive LULC change to meet the demands of unprecedented population growth characterizes the modern transition (McMichael 2004), resulting in the emergence of VBIDs such as Ebola, SARS, and Zika.

McMichael (2004) cited 12 drivers (Table 2) derived from the 2002 Working Group on Land Use Change and Infectious Disease Emergence. The working group, composed of international experts across epidemiologic and environmental disciplines, investigated the association between environmental changes and disease emergence. All emergence drivers are associated with human activities and behaviors, compelled by the need to support a rapidly growing global population.

Table 2. Human behaviors and activities manifest as VBID emergence drivers (McMichael 2004).

- Agricultural Development
- Urbanization
- Deforestation
- Population Movement
- Introduced Species / Pathogens
- Biodiversity Loss
- Habitat Fragmentation
- Water and Air Pollution
- Road Building
- Impact of HIV/AIDS
- Climate Changes
- Hydrological Changes and Dams

VBID transmission drivers entail LULC change to expand settlements and urban areas, farm the land, and extract resources for human consumption (Foley 2005). Grace et al. (2012) associated the human-physical interface, specifically agricultural and pastoral LULC, with roughly 1.7 million annual VBID deaths. Largely overlapping with the 2002 working group findings are the transmission drivers Wilcox and Ellis (2006) associated with specific VBID, as illustrated in Table 3.

Table 3. VBIDs and associated emergence driver (Wilcox and Ellis 2006).

Disease	Emergence Drivers
Dengue	Urbanization
Ebola	Hunting, Logging, Agriculture, and Others
Leishmaniosis	Deforestation, Human Expansion into Forest
Leptospirosis	Watershed Altercation
Lyme Disease	Deforestation, Habitat Fragmentation
Malaria	Deforestation, Human Expansion into Forest
Nipah Virus	Pig and Fruit Production at Forest Border
Rabies	Human Expansion into Forest
SARS	Wildlife Trade, Mixing of Bats and Civet Cats
SIV	Deforestation, Hunting, Settlement Expansion
Yellow Fever	Deforestation, Hunting, Settlement Expansion

Many of these drivers produce LULC fragmentation and/or environmental changes that influence vector, host, and reservoir range and abundance (Patz et al. 2008). Such is the case with the three primary drivers in agricultural development, urbanization, and deforestation (McMichael 2004). While these drivers are highly interrelated and often inseparable in practice, the following sections separate these three primary drivers to aid in the description of their associations with VBID.

2.4.1 Agricultural Development

The daily addition of roughly 275,000 people to the global community created our need to encroach into new landscapes for agricultural development (Lambin et al. 2003), the most frequently cited driver of VBID emergence (Grace et al. 2012, Woolhouse 2011). McFarlane et al. (2013) claim that LULC change for agriculture and food production is responsible for roughly half of all global VBID occurrences.

Agricultural development increases the frequency of contact with vectors at the human-physical interface. The creation of new habitats around irrigation systems and ditches during agricultural development increases vector range and accelerates the mosquito lifecycle. Construction of irrigation systems and ditches around the Aswan Dam in Egypt resulted in an explosion of the mosquito population and a corresponding increase of Bancroftian filariasis occurrence (Patz et al. 2008).

Munga et al. (2006) found that mosquito lifecycles accelerate within agricultural LULC compared to natural habitats, such as swamps and forests. These artificial habitats encourage mosquito breeding and pupation because additional sunlight cast on the agricultural ground surface compared to the prior, naturally vegetated (thus shaded) LULC, translates into higher air and water temperatures, both of which accelerate the mosquito lifecycle.

Patterns in Ebola outbreaks led to the hypothesis that spillover is due to changes in natural fauna associated with agricultural LULC change adjacent to forests. Many Ebola spillover events have occurred along LULC boundaries, where people are in close and frequent contact with pathogenic vectors from forested areas due to agricultural and settlement expansion (Morvan et al. 2000, Patz et al. 2004).

2.4.2 Urbanization

In the context of global population growth of 100 million people annually (King 2004) and growing numbers of megacities in Africa, South America, and Asia (Arinaminpathy et al. 2009), humanity is unwittingly increasing VBID emergence risk. Urbanization and settlement expansion go hand-in-hand with agriculture and food production (Wilcox and Gubler 2005), where humans, animal reservoirs, and vectors co-inhabit urban areas and the surrounding peri-urban, pastoral grassland, and cropland at densities that elevate VBID risk (Arinaminpathy et al. 2009).

Crowded cities and the urban adaptation of the *Aedes aegypti* mosquito is associated with an increase of dengue in Asia and Latin America. Similar to malaria's resurgence in deforested areas, dengue, which is rapidly emerging in impoverished regions of the developing world, transmits via vectors that rely on people to create conditions conducive to vector breeding (Wilcox and Ellis 2006). Such is the case with Zika's primary vectors in the *Aedes aegypti* and *Aedes albopictus* species of mosquito. Marcondes et al. (2016) discovered that urban water storage and residential vessels in Brazil represent the majority of mosquito breeding sites. Mosquitos are now adapted to breed in vessels such as pots and gutters and even refuse such as a bottle cap can store enough water to allow mosquito breeding.

2.4.3 Deforestation

Deforestation that results from urbanization and the expansion of agricultural and pastoral areas fragments LULC, in turn increasing vector range and hastening vector reproduction. Emergence of the first plague-causing pathogens in tropical Asia was a product of such a deforestation and VBID association.

Associations exist between elevated malaria occurrence in many African, South American, and Southeast Asian regions and road and irrigation ditches built during deforestation (Vittor et al. 2006, Walsh et al. 1993). Deforestation-driven road and ditch construction involves the felling of trees. This activity reduces faunal shade and increases the pooling of water, both of which promote mosquito breeding and an acceleration of larval development from one to two weeks to as quickly as 4.5 days (Afrane et al. 2005, de Castro et al. 2006, McMichael 2004).

In addition to the change in natural fauna, deforestation can result in the depletion of predator species. An explosion of small mammal and arthropod vector populations fill these human-induced voids in the natural flora. The 1998 Nipah virus outbreak in Malaysia is an example of predator depletion due to deforestation. A reduction of predator species resulted in greater populations of and interaction between fruit bats and pigs, followed by disease spillover to humans (Pike et al. 2010). While LULC change and VBID emergence both occur at the local-level, impacts of their association increase alongside the scale of transmission.

2.5 VBID Monitoring

Monitoring of local-level VBID emergence to prevent an increase in the scale of transmission via disease surveillance is necessary to protect public health. VBID surveillance is defined by the Centers for Disease Control and Prevention (CDC) as the "ongoing, systematic collection, analysis, interpretation, and dissemination" of public health data with the goal of improving health and reducing morbidity and mortality (Thacker et al. 2012).

VBID surveillance falls into two broad categories in event-based and indicator-based. Traditional, event-based VBID surveillance systems rely on authoritative updates about disease occurrences from official sources, typically clinics or medical laboratories. While highly accurate, these event-based surveillance systems exhibit slow response times. In contrast, indicator-based VBID surveillance systems offer rapid detection of emerging outbreaks, but suffer from a perceived lack of quality due to reliance on unreliable, unvetted, and unofficial data sources.

Laboratory testing provides the accurate, specific, and authoritative VBID and serotype diagnosis that forms the foundation of traditional, event-based surveillance. This authoritative data informs the mitigation, intervention, and response decisions of public health agencies (Feng and Varma 2011).

Calvo-Cano et al. (2014) cite the specificity event-based surveillance afforded medical practitioners investigating the high temperature, lethargy, myalgia, and headache symptoms described by a traveler who spent two weeks in Thailand. The elimination of conditions thought to be responsible for the traveler's ailments, to include various hepatitis serotypes, chikungunya, and dengue was possible through laboratory testing, which diagnosed Leptospira as the culprit.

Authoritative data also affords accurate trend and pattern analysis, as exemplified through a German study that analyzed trends of Campylobacteriosis, the most common gastrointestinal disease. Schielke et al. (2014) analyzed event-based data gathered over multiple years to reveal spikes of Campylobacteriosis cases in January and the summer months. Analysis also revealed higher rates of occurrence in adults in their twenties and children under four years of age, specifically those living in rural areas. This retrospective trend and pattern analysis gave German public health professionals accurate insight into temporal and demographic patterns of Campylobacteriosis emergence, enabling a tailored and targeted response.

While accuracy and specificity of laboratory testing enables tailored and targeted response activities by public health agencies, the retrospective nature of event-based surveillance comes at the sake of the temporal benefits needed address emerging outbreaks. Drawbacks of event-based VBID surveillance also include limited pathogenic scope, spatial scale, and health care participation. Further, disproportionately few VBID surveillance systems surveille developing countries, which are at the greatest risk of disease emergence. The lack of detailed data (disease or otherwise) within developing countries, particularly those in tropical regions, compounds this problem (Wilcox and Ellis 2006).

Sutherst (2004) suggests that infectious diseases endemic to tropical and subtropical regions will expand their latitudinal range in the decades ahead. This regional expansion combined with accelerating LULC change will increase infectious disease risk in temperate zones (Arinaminpathy et al. 2009). However, disease specialization and the limited spatial scope of event-based VBID surveillance prevents systems from surveilling many of the diseases within or on the doorstep of the systems surveillance range (Morse 2012). An assessment of 115 event-based surveillance systems from countries in five continents contends that the fragmented spatial coverage of event-based surveillance is due to the limited spatial range and single-disease focus of most systems (Bravata et al. 2004). Even within the spatial range of a surveillance system, inaccessibility of remote rural areas and the sheer number of inhabitants can limit surveillance effectiveness (Feng and Varma 2011).

The spatial and content quality of the surveillance hinges on the assumption that all actors within the system have the necessary resources to perform surveillance and are working collaboratively. Phalkey et al. (2013) evaluated the Integrated Disease Surveillance system across all 34 districts in Maharashtra, India to discover limited capacity at the lab-, district-, and national-level. Only 53 percent of districts could confirm all priority diseases and all performance factors scored worse at the lab-level than the district-level. Irrespective of the reason, failures in surveillance at the lower levels negatively influences public health decisions of officials at the highest levels.

Similar reporting issues exist within the United States, where each county is responsible for reporting infectious disease occurrences that fall within the parameters of the National Notifiable Disease list through the state to the CDC. Each county handles the reporting of surveillance information differently and not every county or state participates in every surveillance program (Woolhouse 2011). Further, underreporting to the CDC is a major challenge created by inadequate resource levels and policies (Woolhouse 2011). Underreporting could also be a product of the disincentives associated with the potential negative economic impacts an incident report might elicit at the local-level (Bogich et al. 2012).

These capacity and underreporting factors produce inconsistent spatial and content coverage, which adversely influences identification of accurate trends and patterns (Morse 2007). Inference of VBID hotspots, trends, and patterns through measurement of LULC can supplement disease surveillance.

2.6 Measuring LULC and VBID Associations

Evaluation of existing literature, field research, and remote sensing are common methods by which spatial epidemiologists measure associations between LULC and infectious disease. VBIDs are ecological diseases due to vector reliance on particular vegetation, temperature ranges, and other environmental variables (Ceccato et al. 2005). The methods used to measure these variables influence the identification of associations between LULC and VBID.

2.6.1 Review of Existing Research

As the corpus of LULC and environmental data expands, so does the amount of epidemiological research that could leverage the spatial information therein. McFarlane et al. (2013) conducted a systematic review of the association between LULC change and infectious disease emergence within Australia. Research encompassed a review of literature published between 1973 through 2010 using keywords pertaining to emerging infectious disease and Australia. Twenty-two percent of the 90 emerging infectious disease events were associated with LULC change, and the strongest associations involved agricultural LULC change. Further, significant LULC change events were associated with highly clustered infectious disease occurrences.

2.6.2 Field Research

Vanwambeke et al. (2011) conducted field research to discover an association between LULC and dengue risk from the *Aedes albopictus* mosquito vector in select areas within Hawaii. In another field research example, Ezenwa et al. (2007) studied the spatial association between LULC and WNV prevalence in St. Tammany Parish, Louisiana. This research found a strong association between wetland LULC and WNV host abundance and range, suggesting that larger unbroken tracts of wetland would result in greater control of hosts by predators, thus naturally controlling WNV outbreaks. Field research provides robust and invaluable ground-truths of vector habitat characteristics and spatial associations. However, this method does not scale as efficiently or cost-effectively as remote sensing.

2.6.3 Remote Sensing

Remote sensing provides epidemiologists a spectrally fine and spatially broad view of LULC interfaces at ever increasing temporal resolution. Data from satellite, aerial, and ground-based remote sensors has unlocked applications at scales beyond that of localized field research.

Unlike the identification of vectors through field studies, remote sensing infers the existence, range, and abundance of vectors through measurement of elevation and environmental variables such as LULC, vegetation type and health, and land temperature (Kalluri et al. 2007). When validated by or combined with vector habitat characteristics discovered via field research, remotely sensed data can be used to predict disease risk on a broad spatial scale (De La Rocque et al. 2004).

Temperature, humidity, rainfall, and vegetation variables derived from remote sensing aid in the creation of prediction models for malaria based on the preference of many mosquito species to deposit larvae in specific, identifiable LULC classes (Hay et al. 1998). De La Rocque et al. (2004) analyzed tree canopy patterns from high-resolution SPOT imagery to classify vegetation type and evaluated Normalized Difference Vegetation Index (NDVI) to classify vegetative health. Analysis of these countrywide datasets in Burkina Faso identified habitats of six tsetse fly species. This remote sensing method enabled frequent updates of regional-level disease risk maps, a spatiotemporal achievement that leverages findings from and compliments the precision of smaller scale field studies.

2.6.4 LULC Classification

LULC classification from remotely sensed imagery occurs through categorization of physical features and processes based on spectral properties. Similar to the gradient characterized by most climate variables, the physical transition between different LULC classes range from abrupt to gradual. The imagery-derived classification of gradual transitional boundaries between LULC classes results in pixels that contain the spectral signature of more than one land type. Mixed or mosaicked LULC classes mitigate uncertainties in pixel assignment, as attempted by the extensive range of classes within the GlobCover dataset (Table 4).

Table 4. GlobCovers extensive class set accounts for gradual LU transitions compared to GLC-SHARE.

GlobCover	GLC-SHARE
Post-Flooding or Irrigated	Artificial Surfaces
Rain-fed Croplands	Croplands
Mosaic Croplands	Grasslands
Mosaic Forest/Shrubland	Tree-Covered
Mosaic Grassland	Shrub-Covered
Mosaic Vegetation	Herbaceous
Closed to Open Broadleaved	Mangroves
Closed to Open Mixed	Sparse
Closed to Open Shrubland	Bare Soil
Closed to Open Grassland	Snow and Glacier
Closed to Open Vegetation	Inland Water
Closed Broadleaved Deciduous	
Closed Needleleaved Evergreen	
Closed Forest Regularly Flooded	
Closed Semi-Deciduous	
Open Broadleaved Deciduous	
Open Deciduous or Evergreen	
Sparse Vegetation	
Waterlogged Soil	
Artificial Surfaces	
Bare Areas	
Water Bodies	
Permanent Snow and Ice	

A trade-off exists between the addition of classes that depict gradual LULC transitions and the complexity of analysis (Wang and Howarth 1993). Compared to the 11-class GLC-SHARE, GlobCover's 23 classes can increase complexity and bloat compute time. Consolidation of LULC classes considered duplicative within the lens of a research topic reduces the computational and analytic load.

2.6.5 LULC Uncertainty

LULC datasets are the product of human-derived classification schemas and the abstracted definitions therein (Wang and Howarth 1993), factors that introduce uncertainty through ambiguity and vagueness (Leyk et al. 2005). The lack of schema standardization (Buyantuyev and Wu 2007) and the unintentional, yet often unavoidable inclusion of co-occurrences, also contribute to uncertainties in LULC datasets (Altizer et al. 2006, McClintock et al. 2010).

Ambiguity

Classification schemas with slightly different definitions for the same physical feature produce ambiguity when comparing LULC datasets. Slight definition differences exists between the International Geosphere Biosphere Programme (IGBP) and United Nations (UN) LCCS classification schemas. The LCCS criteria for the snow and glacier LULC class includes snow or glacial coverage for at least ten months out of the year (Latham et al. 2014). The IGBP criteria is slightly different in the requirement of snow or glacial coverage throughout the year (FRA 2000). The two-month difference in the LULC definition could result in ambiguity when comparing digital representations of the same physical features from different sources (Leyk et al. 2005). The looser threshold for feature inclusion in the LCCS snow and glacier class will produce a corresponding LULC footprint more spatially extensive relative to the more stringent IGBP threshold. The delta between these two LULC definitions are physical features that fall within different classes depicted in a LULC dataset.

Vagueness

Vagueness is a product of poorly defined classification criteria that create confusion as to which class a pixel belongs (Leyk et al. 2005). The unsupervised Global Forest Watch (GFW) classification algorithm compares the visible and infrared signature of each pixel in an image and the corresponding pixel in subsequent images to assign each pixel a 'forested' or 'not forested' value. Vagueness in the classification criteria and LULC definitions used by GFW could produce uncertainty within a portion of the datasets 143 billion pixels, assuming vague criteria in the definition of mosaic forests that reside in between forested or not-forested LULC (Global Forest Watch 2014). Vagueness is a product of definitions that allow for flexible interpretation. Leyk et al. (2005) suggest mitigation of vagueness via semantic model to compare all definitions.

Estimation of a datasets ambiguity and vagueness should account for the number of contributors listed in the metadata lineage (Leyk et al. 2005). Each additional contributor potentially propagates the influence of different, 'imperfect' definitions or semantics during feature classification. Fewer contributors translates into a greater likelihood that LULC definitions and criteria produce methodological uniformity.

Standardization

Buyantuyev and Wu (2007) express that standardization of LULC classification schemas would enable greater application of analytic methodologies through accurate spatial and temporal comparisons of LULC datasets. The lack of standardization makes apples-to-apples comparison of datasets impossible. Even with datasets from the same data producer, improvements in classification definitions and algorithms result in different pixel values for the same physical feature under the same environmental conditions. The comparison limitations combined with the infrequent creation of LULC datasets adversely influences outcomes when using multiple LULC datasets within a single research project.

Co-Occurrences

LULC, environmental, and climate variables derived from remotely sensed data aid the successful prediction of tick abundance in North America (Ratmanov et al. 2013). While measuring such variables with remote sensors offers many benefits, co-occurrences such as seasonal changes in temperature, rainfall, and resource availability can add uncertainty to the research of LULC and VBID associations. Seasonal changes influence influenza transmission, which is typically more prevalent during winter months, and malaria transmission, which increases during periods of seasonal rain. Co-occurrences such as El Niño and La Niña can result in longer-term deviations of environmental and climate variables (Altizer et al. 2006).

McClintock et al. (2010) suggest repeating research methodologies over time and space will reduce uncertainties borne through seasonal and multi-year co-occurrences. Mitigation of uncertainties through iteration of methodologies over time and space improves reliability when using LULC, environmental, and climate measurements to locate and quantify vector range and abundance, and predict hotspots of human risk for VBID. Be it through evaluation of existing literature, field studies, or remote sensing, spatial epidemiological research is leading to the discovery of associations between LULC and VBID at multiple scales.

2.6.6 Geographic and Thematic Scales

The geographic and thematic scales of analysis influence research findings (Beale et al. 2008), to include the strength of associations discovered between LULC and VBID. Under optimal conditions, research requirements will dictate scale. In practice, theory-based scale criteria dissolves when faced with poor data quality and sparse data availability.

Arsenault et al. (2013) convey that dynamic epidemiological processes collected, analyzed, and communicated at one scale might be insignificantly represented or omitted altogether at another scale. Low-resolution, broad area collection might not capture processes that occur at the local-level while high-resolution, narrow-area collection might contain too much noise to discern regional-scale processes.

Arsenault et al. (2013) applied a set of measurable criteria to select the geographic scale of analysis. Criteria included intra-unit homogeneity, compactness, variation in areal size, and the percentage of areas with sufficient population size. Such criteria led to the identification of municipality as the appropriate scale for their research into Campylobacteriosis within Quebec, Canada. This scale selection methodology proved successful in a developed country with the wealth of population, epidemiologic, and LULC data seldom found in developing countries.

In addition to the geographic scale of analysis, the scale at which data is mapped – thematic scale – requires consideration. Thematic scale determines the amount of information cartographically depicted, serving as a filter that either represents physical features as pixels or ignores physical features smaller than the scale value. As an example, imagery derived maps at a 1:10,000 scale produce significantly denser abstractions of physical features than maps at 1:100,000 scale.

If analysis at the regional, national, or global scale leverages local-level data, aggregation could transform the data into a scale appropriate for a larger area study. However, the modifiable areal unit problem (MAUP) often results from data aggregation (Arsenault et al. 2013). Aggregation of local data to the state- or nation-level can strip associations discoverable at the county-level and produce MAUP-related uncertainty and biases (Beale et al. 2008).

Selection of geographic- and thematic-scales should occur through the lens of the research questions and associated data needs. If fortunate enough to research a topic flush with data at various scales, the selection of scale should compromise between the content richness and noise of local-level data and the generalization and broad spatial coverage of regional-level data.

2.7 Research Opportunity

McFarlane et al. (2013), Murray and Daszak (2013), Morse (2012), and Eisenberg et al. (2007) are among the contingent of spatial epidemiologists whom have called for additional research into the association between LULC and infectious disease to detect potential hotspots of increased human transmission risk.

The social and economic impacts of the 2003 SARS outbreak and the 2014 Ebola outbreak underpins the UN Millennium Development Goal to halt and reverse occurrences of major infectious diseases (United Nations 2014). However, understanding of the complex spatial association between LULC and VBID exists for only a few diseases (Ezenwa et al. 2007, Jones et al. 2008, Patz et al. 2004).

Although the spatial component of infectious disease received more attention during the 2000's, just one percent of all articles within seven leading epidemiology journals evaluated by Auchincloss et al. (2012) cited a spatial methodology, regardless of type. The spatial associations between VBID and LULC fragmentation received even less research attention than the general LULC associations conveyed by Ezenwa et al. (2007), Jones et al. (2008), and Patz et al. (2004).

Fragmentation increases the linear density of LULC boundaries, also known as transitions, ecotones, edges, or interfaces. A Rhode Island field study by Finch et al. (2014) associated the linear density of shrub LULC boundaries with increased densities of tick nymphs and elevated rates of tick related diseases. Lambin et al. (2010) found a strong association between WNV and the density of LULC boundaries produced by fragmentation in southern France. Fragmentation increases the frequency of human contact with vectors, in turn increasing the rate of VBID emergence (McMichael 2004).

Fragmented LULC contains high densities of vectors relative to larger, unbroken blocks of LULC. In Belgium, the probability of tick-borne Lyme disease is greater in areas with a high level of forest and peri-urban LULC fragmentation (Wilcox and Ellis 2006). A southeastern New York field study of a Lyme disease reservoir host in white-footed mice identified similar results in the association between forest fragmentation and Lyme disease whereby the greatest risk of Lyme occurred in settlements adjacent to forest fragments (Allan et al. 2003).

Identifying the point at which the density of LULC boundaries become associated with increased VBID occurrence will benefit policy makers and health professionals alike. McFarlane et al. (2013) and Murray and Daszak (2013) recommend policy makers mitigate VBID risk through strategies that modify land development plans and policies based on identification of high-risk areas. Health care professionals can supplement existing methods to assess the targeted allocation of their overburdened resources. The goal for both user groups is an early warning methodology to aid in targeted VBID prevention and mitigation, thus dampening the social and economic impacts of infectious disease.

3 METHODOLOGY

A disproportionate amount of VBID are transmitted near LULC edges, where the human-physical interface is greatest (Patz et al. 2008, Wilcox and Ellis 2006). This spatial relationship between LULC and VBID is the foundational concept driving the research in this dissertation.

This dissertation seeks to address three research questions through the context of spatial associations between LULC and VBID. These research questions are: (RQ1) Do spatial associations exist between the linear density of LULC boundaries and VBID occurrence; (RQ2) Do patterns of spatial associations repeat over time; and (RQ3) Do patterns of spatial associations repeat across space?

RQ1 examines WNV occurrences in the contiguous United States during 2014. RQ2 expands the temporal breadth of RQ1 through examination of WNV occurrences in the contiguous United States between 2003 through 2014. RQ3 shifts VBID and study area to examine Zika in Colombia, South America during 2016.

The high-level methodology includes constituent steps, such as the collection, preparation, and processing of unstructured, semi-structured, and structured data from official and unofficial sources (described in Section 4). As an example, conversion of temperature and elevation raster files into vector points enables a spatial join with second-order administrative unit polygons to calculate average values per unit.

The high-level methodology applied in part or in full to the three RQs include:

- Consolidation of LULC classes from GlobCover and GLC-SHARE datasets into Research Classes (RC) to reduce compute time and analytic complexity.
- Calculation of landscape metrics in linear density, proportion abundance, and patch density to create LULC independent variables.
- Derivation of Principal Components to reduce the complexity presented through inclusion of multiple RCs in regression analysis.
- Calculation of incidence rate ratio and other negative binomial regression factors to assess the spatial associations between LULC and VBID.

3.1 Consolidation of LULC Classes

Identification of LULC areas occurs through the categorization of remotely sensed physical features and the associated human activities as defined in LULC classification schemas (Hay et al. 1998). However, LULC datasets can contain upwards of 20 classes, many of which are indistinguishable for the purposes of this research. Further, inclusion of the full breadth of LULC classes bloats compute time and adds unnecessary complexity to analysis.

Wang and Howarth (1993) suggest a trade-off evaluation between the number of LULC classes and the complexity of analysis. To this end, LULC classes defined as similar through the lens of this research are consolidated into fewer, often broader RCs. The RCs created for this research include:

- Artificial Surfaces
- Cropland
- Grassland
- Tree-Covered
- Waterlogged
- Bare Soil
- Water Bodies
- Snow and Glacier

Classification schemas define the characteristics of each LULC class and guide the consolidation (Section 4.2) of similar LULC classes into RCs. As an example, the Tree-Covered RC is a consolidation of Closed Broadleaved Deciduous and Closed Needleleaved Evergreen classes from the GlobCover dataset. The resultant eight RC classes will reduce compute time and analytic complexity.

3.1.1 Landscape Metrics

Linear (edge) density serves as the primary landscape metric in the examination of spatial associations between LULC and VBID occurrence. Supplemental landscape metrics in proportion abundance and patch density add context to linear density.

LULC Linear Density

An increase in the linear density of LULC boundaries occurs when a large LULC patch is broken, or fragmented, into smaller patches. Fragmentation increases the density of LULC boundaries within a given area, which increases the amount of human-physical interface and produces a greater likelihood of VIBD emergence compared to larger, unbroken patches (McMichael 2004). Spatial associations exist between fragmentation and emergence of VBIDs, such as Lyme (Wilcox and Ellis 2006, Patz et al. 2004, Allan et al. 2003) and WNV (Lambin et al. 2010).

The calculation of LULC boundary density provides a landscape metric akin to fragmentation, and one in which the spatial unit of the linear density calculation is equal to that of the VBID dataset, the second-order administrative unit. Arsenault et al. (2013) identified municipality as the appropriate scale for VBID research within Quebec, Canada. Similarly, the municipality within Colombia and the county within the United States, both second-order administrative units, are appropriate scales for Zika and WNV research, respectively.

Designation of the Local Administrative Unit (LAU) at the second-order reduces uncertainties caused by the spatial variation between VBID transmission and report locations. Data at a scale more granular than the second-order administrative unit is not possible because such VBID information is not publically available. Further, a more granular research unit increases the likelihood that reports of disease occurrence will fall outside of the unit where transmission took place. Conversely, aggregation of data into administrative units larger than the second-order would introduce the MAUP, whereby spatial detail and insight is lost.

Geographic Information System (GIS) functionality intersects LULC areas where they cross LAU areas. After intersecting LULC and LAU areas, geometry calculations derive LULC boundary length for each RC within each LAU. Dividing boundary length by area determines the linear density (length per area) for each set of RC boundaries. The following equation calculates the linear, or edge, density of LULC boundaries:

$$Linear\ Density = \frac{(\sum_{i=1}^{n} L_i\ )}{A},$$

where L = length of boundary within a LAU, A = LAU area

LULC Proportion Abundance

Proportion abundance is a landscape metric that measures the area of a single RC compared to the area of all RCs in the same second-order administrative unit. Proportion abundance is independent from the two other landscape metrics since patch shape (linear density) and patch count (patch density) do not necessarily correlate to proportion abundance.

The proportion abundance of a given RC will increase as its area increases within an administrative unit. However, the shape and complexity of these RC patches might not result in a corresponding increase in linear density. Simply stated, natural patches with complex boundaries replaced by larger patches with simpler geometric boundaries (e.g. cropland) will increase proportion abundance while decreasing linear density.

LULC Patch Density

Patch density is a landscape metric that measures patch count of a single RC in a unit compared to the total patch count of all RCs in the same unit. Like proportion abundance, this landscape metric is independent from the other landscape metrics. Due to variability in patch shape and area, the patch density landscape metric does not correlate to linear density or proportion abundance.

The patch density of a given RC will increase as the number of its patches increase within an administrative unit. While fragmentation increases the amount of patches within a unit, the shape and complexity of the RC patches within the same unit might not result in a corresponding increase in linear density.

Figure 3 depicts the relationship between the three landscape metrics. The linear density of the gray patches within the first row remain constant, as does proportion abundance in the second row, and patch density in the third row.

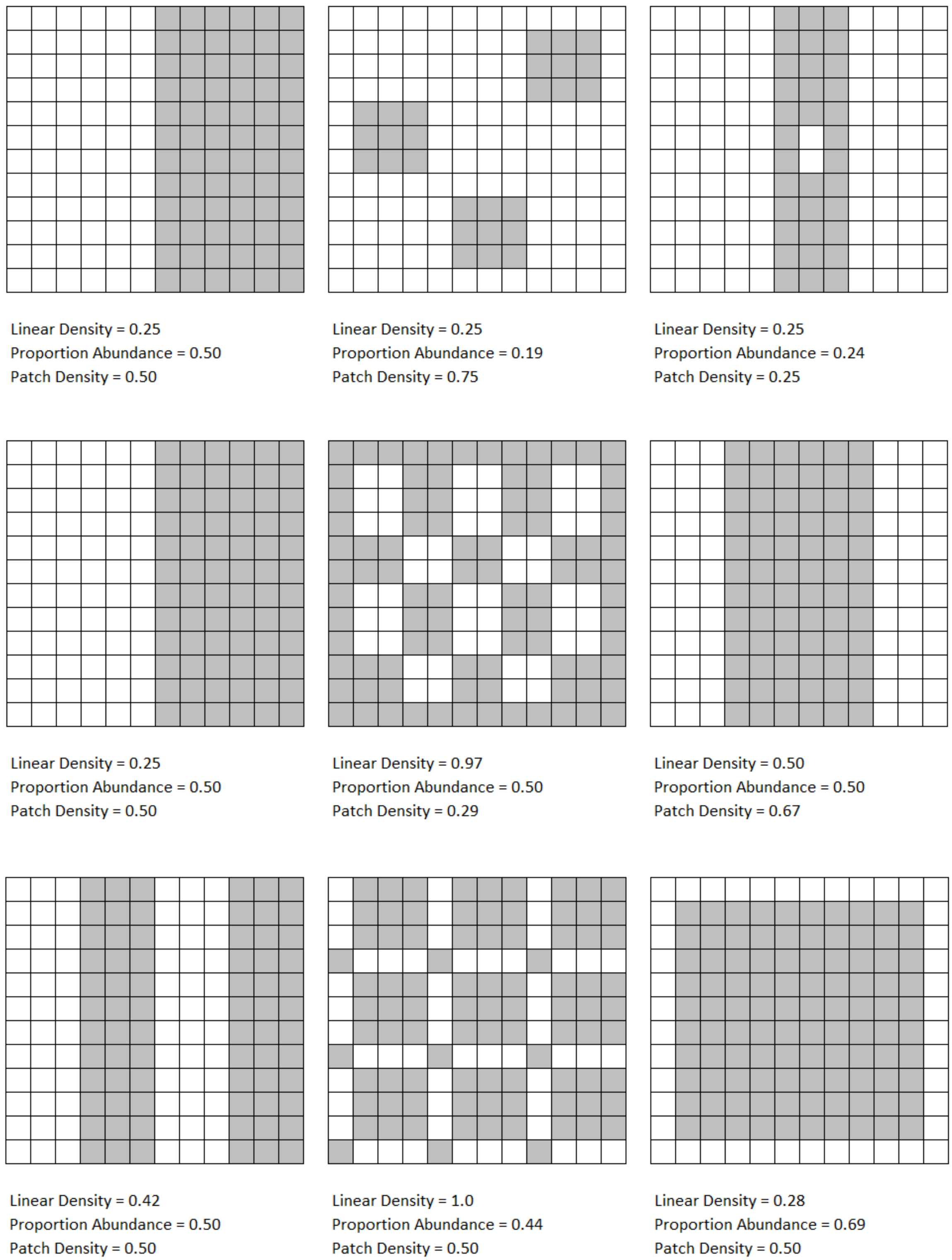


Figure 3. Relationship between linear density, proportion abundance, and patch density.

3.1.2 Principal Component Analysis

Principal Component Analysis (PCA), or dimension reduction, maximizes data variance while reducing the number of original variables into a lesser amount of components. The Principal Component explains the greatest amount of data variance with subsequent components explaining progressively less data variance.

The PCA Correlation Matrix validates selection of the original variables. In addition to the variable correlations within the Correlation Matrix, eigenvalue percent variance and cumulative variance aids downstream analysis by quantifying the variance explained by the original variables. Eigenvalues greater than 0.5 signify that the component's original variables explain a significant amount of data variance.

PCA is performed on targeted RCs using IBMs Statistical Package for the Social Sciences (SPSS) to examine RQ1 and RQ2. RCs used to examine the spatial association between LULC and VBID within the contiguous United States include Cropland, Tree-Covered, Artificial Surface, Water Bodies, and Waterlogged.

Cropland is included in PCA because 86 percent of all 3,107 counties in the contiguous United States have linear density values greater than zero for this RC. Cropland is not just widespread; it is associated with roughly half of all global VBID occurrences (McFarlane et al. 2013). Similarly, Tree-Covered is widespread due to non-zero presence in 74 percent of all counties.

Artificial Surface is included in PCA because this RC represents high densities of humans living adjacent to non-urban RCs. In addition, Arinaminpathy et al. (2009) identified associations between urban areas and the surrounding LULC in peri-urban, grassland, and cropland with elevated VBID risk. Wilcox and Ellis (2006) associated urbanization with specific VBIDs in dengue, SIV, and yellow fever. Water Bodies and Waterlogged are included in PCA because they represent ideal locations for mosquito breeding sites (EPA 2004).

Grassland was initially included in PCA because it accounts for 42.05 percent of all LULC in the contiguous United States. Removal of this RC from PCA and its addition as a stand-alone independent variable was necessary because Grassland is highly associated with VBID regardless of PCA and regression permutation.

3.1.3 Regression Analysis

Negative binomial regression, used across all research questions, reveals spatial associations (or lack thereof) between LULC and VBID. Unlike Ordinary Least Squares, Poisson, and other Generalized Linear Regression methods, negative binomial determines the incidence rate ratio of dependent variables characterized by small numbers.

Although VBID presents a significant and growing threat to public health, occurrence rates are relatively low when compared to the general population. Further, a significant number of Colombian municipalities and US counties are devoid of occurrences. The WNV occurrence data indeed represents small numbers. Based on the available data (2003 through 2014) that met the selection criteria of this research, WNV peaked in the contiguous United States in 2003. During this year, 2,036 out of 3,107 counties, 65.5 percent, did not experience a single reported WNV occurrence. Further, the 9,755 confirmed occurrences out of a total population of 288,208,365 results in an occurrence rate of 0.0034 percent, or 3.4 out of 100,000.

Zika in Colombia during 2016 also represents small numbers. Forty-five percent or 509 out of 1,1,21 municipalities did not experience a single Zika occurrence while the occurrence count of 501,970 in a population of 48,881,635 results in an occurrence rate of 1 percent or 1,027 out of 100,000. Negative binomial appropriately accounts for VBID occurrence characteristics, such as low counts, a significant number of zero values, and dissimilarity between the variance and mean.

The negative binomial regression models use VBID occurrence counts as the dependent variable. Independent variables for the contiguous United States study area include population density, average temperature, Grassland, and the RC-derived Principal Components. Average rainfall, similar to average temperature, is an environmental factor associated with VBID occurrence. However, removal of the rainfall variable was necessary due to collinearity with temperature.

Independent variables for the Colombia study area include population density, average elevation, per capita Gross Domestic Product (GDP), and three RCs in Cropland, Grassland, and Tree-Covered. These three RCs account for more than 98 percent of all LULC within Colombia during 2016. GPD was included to normalize the spatial association for an economic variable.

The primary regression model examines the linear density landscape metric. Supplemental regression models examine the proportion abundance and patch density landscape metrics. These supplemental landscape metrics add context to the linear density landscape metric results.

Output tables from the regression model for RQ1 includes the minimum, maximum, mean, standard deviation, incidence rate ratio (expB), and the 95 percent confidence interval. ExpB, the incidence rate ratio, indicates that a one-unit increase in the independent variable will increase (if expB > 1.0) or decrease (if expB < 1.0) the relative risk of the dependent variable, VBID occurrence, by the expB value.

3.2 Spatial Associations over Time

The methods described in Section 3.1 examined a temporal snapshot of 2014 WNV data in the contiguous United States. Examination of a single years' worth of data could inadvertently introduce uncertainties borne via co-occurrences such as seasonal or long-term confounding environmental events.

Long-term confounding environmental events, such as El Niño and La Niña, result in global climate changes that can last for years. El Niño or La Niña -driven increases or decreases to a regions average temperature and rainfall unevenly influences the range and abundance of VBID hosts and vectors across LULC classes. Due to multi-year aftereffects, a weak La Niña event in 2011-2012 introduces potential uncertainty in the analysis of 2014 WNV data (Null 2015).

McClintock et al. (2010) assert that the comparison of multiple temporal snapshots mitigates the influence of co-occurrences. As a result, negative binomial regression is iterated using WNV data within the contiguous United States for each year between 2003 through 2014. Averaging each variables' incidence ratio derived from negative binomial regression will provide a more reliable and robust analysis of spatial associations between LULC and VBID than any one year can offer.

3.3 Spatial Associations across Space

Comparison of negative binomial regression results for two study areas and infectious diseases will help answer RQ3. This RQ examines spatial associations between LULC and the Zika VBID within Colombia during 2016.

McClintock et al. (2010) not only suggest the repetition of research methodologies over time, they recommend repeating methodologies over space to reduce uncertainty borne through confounding environment variables. The comparison of incidence ratios derived from two geographically disparate study areas in the United States and Colombia and two different diseases in WNV and Zika will advance our understanding of the spatial associations between LULC and VBID.

4 DATA

Data requirements for research into the spatial associations between LULC and VBID includes disease occurrences, land use, administrative units, and environmental and demographic variables. Data source dates are similar (Table 5) to ensure that VBID occurrences temporally align with the population impacted by said VBID. The data requirements and criteria outlined in this section narrows the focus of this research to WNV in the contiguous United States between 2003 through 2014 and Zika in Colombia during 2016.

Table 5. Temporal alignment of data and data sources each study area.

	WNV in the contiguous United States				
	Admin	**Population**	**Temp**	**VBID**	**LULC**
2003	GAUL	US Census	PRISM	CDC/USGS	
2004	GAUL	US Census	PRISM	CDC/USGS	
2005	GAUL	US Census	PRISM	CDC/USGS	GlobCover
2006	GAUL	US Census	PRISM	CDC/USGS	
2007	GAUL	US Census	PRISM	CDC/USGS	
2008	GAUL	US Census	PRISM	CDC/USGS	
2009	GAUL	US Census	PRISM	CDC/USGS	GlobCover
2010	GAUL	US Census	PRISM	CDC/USGS	
2011	GAUL	US Census	PRISM	CDC/USGS	
2012	GAUL	US Census	PRISM	CDC/USGS	
2013	GAUL	US Census	PRISM	CDC/USGS	
2014	GAUL	US Census	PRISM	CDC/USGS	GLC-SHARE

	Zika in Colombia				
	Admin	**Population**	**Elevation**	**VBID**	**LULC**
2016	GADM	Dept. Statistics	SRTM	GitHub	GLC-SHARE*[1]

1. GLC-SHARE produced in 2014

4.1 VBID

Data availability and alignment, spatial resolution, occurrence count and scope, transmission type, and relevancy criteria funneled the breadth of potential VBIDs to the selection of WNV in the contiguous United States between 2003 through 2014 and Zika in Colombia during 2016.

Data Availability and Alignment

This dissertation relies on publically available VBID data. Further, all VBID data must be in temporal alignment with the other data sources to ensure the research accurately captures spatial associations.

Spatial Resolution

The spatial resolution of VBID data must be at or more granular than the second-order administrative unit. Nearly all publically available infectious disease data is at the state or nation level. Similar to Arsenault et al. (2013) identification of municipality as the optimal resolution to research VBID in Quebec, the spatial resolution for this WNV and Zika research is the United States county and Colombian municipality, respectively. VBID data aggregated to the state or national-level would increase spatial uncertainty and decrease analytic accuracy (Arsenault et al. 2013).

Occurrence Count and Scope

VBIDs must represent a significant number of occurrences and broad spatial scope. Schneider et al. (2010) assert that discovery of associations requires a number of observations at least 20 times the number of independent variables. Independent variables include population density, average temperature or elevation, up to three LULC variables, and in the Colombia study area, GDP per capita. The inclusion of up to six independent variables requires at least 120 occurrences per study area.

This minimum occurrence requirement excludes Anthrax, a VBID transmitted by the *Bacillus anthracis* bacterium, as there are only one to two naturally occurring deaths annually within the United States (CDC 2009). While Anthrax and other rare diseases do not meet the occurrence count minimum, 120 occurrences is achievable for a broad range of VBIDs. The fewest WNV occurrences between 2003 through 2014 in the contiguous United States occurred in 2011, with 705 occurrences. Zika in Colombia during 2016 reached 501,970 confirmed occurrences.

Transmission Type

The preferred VBID for study does not pass from human-to-human. The VBID occurrence data assessed for this research does not differentiate between vector-to-human and human-to-human modes of disease transmission. Elimination of contagious diseases from this research is necessary to maintain focus on the human-physical interface and the influence of land use on VBID.

Relevancy

VBIDs should be relevant through the lens of our current public health landscape. The WHO, CDC, Health and Human Services (HHS), National Institute of Health's (NIH) National Institute of Allergy and Infectious Diseases (NIAID), and other national and international organizations fund and resource VBID prevention and mitigation (NIH 2015, NIH News 2009, WHO 2013). Despite the attention priority diseases receive, austere budget environments result in funding levels below what is required for full preparedness, reducing the capacity for proactive VBID mitigation. The International Working Group on Financing Pandemic Preparedness cited the recent Ebola outbreak as an example of the global health community's common practice of reactive response and truncated funding (World Bank 2017).

The need for proactive VBID detection and mitigation in an underfunded environment underscores the relevancy criteria, which focus this dissertations research on VBIDs that produce current social and economic impacts.

4.1.1 West Nile Virus

Over 30 species of mosquitos serve as the WNV vector, carrying the pathogen to humans after feeding on an infected bird. WNV vectors exist across Africa, Europe, Middle East, North America, and West Asia (WHO Media Centre 2011).

WNV mosquito vectors rarely use healthy wetlands as breeding habitats due to the abundance of natural predators. However, degraded wetlands, altered by organic material and contaminant runoff produce algal blooms, which foster mosquito reproduction. Further, mosquito species that serve as WNV vectors, particularly *Culex salinarius* and *Culex tarsalis*, thrive in urban and suburban areas where they use stagnant water as breeding sites (EPA 2004). The greatest concentrations of mosquitos occur in temporary pools of water, such as ditches, tires, gutters, and abandoned swimming pools, which lack the mosquito's natural predators and often contain organic material. Under ideal warm weather conditions, these mosquito-breeding sites can foster reproduction cycles in as quickly as 4.5 days. As such, nearly any warm-weather standing temporary body of water can serve as a productive mosquito-breeding habitat (Gaines 2014).

The first documented case in the Western Hemisphere occurred in 1999. Since the initial outbreak in New York City, WNV has spread across the 48 contiguous states (USGS WNV 2013) and was responsible for nearly 40,000 infections and 2,000 deaths in the United States between 2003 through 2014 (CDC MMWR 2014). Roughly one percent of humans bitten by an infected mosquito will develop severe symptoms, to include encephalitis or meningitis (USGS WNV 2013) and no vaccine or antiviral drug exists to treat a WNV infection. As of October 2017, 7 deaths, 5 cases of paralysis, and 74 cases of encephalitis or meningitis have been attributed to WNV in Los Angeles County (Karlamangla 2017).

The Border Infectious Disease Surveillance program considers WNV a priority disease due to the northern trend of WNV vectors into the United States (Arizona Department of Health Services 2013). The CDC's Division of Vector-Borne Diseases, the US HHS Global Health Strategy, and the WHO consider WNV a priority disease. In addition, the NIH NIAID designated WNV a Category B pathogen, which are relatively easy to transmit and result in moderate morbidity (NIH 2015).

This research utilizes WNV data from the United States Geological Survey (USGS) WNV Human Provisional Data website (USGS 2014), which provides county-level

WNV occurrences annually from 2003 through 2014. The USGS WNV website updates weekly with data from the CDC through state health departments (USGS WNV 2013).

4.1.2 Zika

The *Aedes aegypti* and *Aedes albopictus* species of mosquito are the predominant Zika virus vector. The WHO describes the *Aedes aegypti* mosquito species as opportunistic in its ability to adapt to rapidly changing ecologies, including urbanization. The *Aedes aegypti* mosquito can reproduce in environments ranging from forests to densely populated urban areas (WHO Mosquito Control 2016).

While roughly 80 percent of those infected with the Zika virus will not exhibit symptoms, there is no vaccine or cure for those who do manifest symptoms, such as fever, rash, joint pain, and red eyes. Further, health professionals are finding potential associations between Zika and increased risk of microcephaly, a condition that results in smaller than expected head size and abnormal brain development of a fetus while in utero, along with Guillain-Barré Syndrome, a condition that causes the immune system to attack the nervous system (CDC Zika 2016).

The first Zika case occurred in 1947 within the Zika Forest, Uganda, near the west shore of Lake Victoria where scientists were researching yellow fever in rhesus monkeys. The virus largely remained in Africa with small outbreaks in Asia for decades after the first human infections in Uganda and the United Republic of Tanzania in 1952. In 2007, a major epidemic occurred on the island of Yap in Micronesia, where nearly 75 percent of the population were infected. In 2013, Zika emerged in French Polynesia, sending 28,000 people, 11 percent of the population, to seek medical care. In March 2014, Chile notified the WHO with confirmation of indigenous transmission on Easter Island, with 72 of the cases deemed severe, 40 of whom suffered through Guillain-Barré Syndrome. Since May 2015, when health professionals in Brazil confirmed transmission of Zika, the virus has spread through Central and South America, with local transmission identified in Texas and south Florida (WHO 2017).

Case counts are increasing due to the ongoing nature of the Zika epidemic. As of June 2017, the virus exists in 78 countries (Department of Health, 2017), with 224 locally acquired mosquito transmission cases in the United States, and nearly 37,000 across United States territories (CDC 2017).

The CDC's Division of Vector-Borne Diseases considers Zika a priority disease, resulting in greater levels of prevention and mitigation funding and resources. In April 2016, the Food and Drug Administration added Zika to the list of diseases for priority review to spur the development of a vaccine (CDC Zika 2016).

The GitHub data repository provides publically available Zika data across a breadth of Central and South American countries (Rodriguez et al. 2016). Although contributors to this repository claim the data is not exhaustive or official, it is the most comprehensive publically available and centralized data repository of Zika cases at the second-order administrative unit.

Data availability through GitHub is possible through translation of and information extraction from official Epidemiological Bulletins created by local-level health departments or ministries across Central and South America. The earliest weekly Epidemiological Bulletin is dated late November 2015 (Rodriguez et al. 2016) and the repository maintains currency through the continuous addition of new data.

4.2 Land Use

WNV and Zika occurrences are associated with human-modified LULC classes, specifically urban, agricultural, and pastoral grassland (Kilpatrick 2011), adjacent to natural vegetation (McMichael 2004, Wilcox and Ellis 2006). Due to the infrequency of LULC dataset creation and the corresponding infrequency of source date alignment between all datasets (Table 5), this research must leverage two GlobCover LULC datasets and one GLC-SHARE LULC dataset.

The GlobCover and GLC-SHARE classification schemas and class definitions inform the consolidation of LULC classes to RCs. Table 6 depicts the consolidation of 23 GlobCover LULC classes to eight RCs and the corresponding rationale.

Table 6. Consolidation of 23 GlobCover classes into eight Research Classes.

GlobCover Class	Research Class	Rationale
Artificial Surfaces	Artificial Surfaces	
Post-Flooding or Irrigated Rain-fed Croplands Mosaic Croplands Mosaic Vegetation	Croplands	Land classes that represent human altered areas with the intent of agricultural use.
Mosaic Grassland Closed to Open Grassland Closed to Open Vegetation Closed to Open Shrubland Sparse Vegetation	Grass/Shrub	Land classes that are predominately grassland, shrub, or herbaceous.
Closed to Open Broadleaved Closed to Open Mixed Closed Broadleaved Deciduous Closed Semi-Deciduous Closed Needleleaved Evergreen Mosaic Forest/Shrubland Open Broadleaved Deciduous Open Deciduous or Evergreen	Tree-Covered	Land classes that are predominately deciduous and/or evergreen tree covered.
Closed Forest Regularly Flooded Waterlogged Soil	Waterlogged	Land classes that represent coastal (saline) flooded areas.
Bare Areas	Bare Soil	
Water Bodies	Water Bodies	
Permanent Snow and Ice	Snow and Glacier	

While both GLC-SHARE and GlobCover adhere to the Land Cover Classification System (LCCS), GLC-SHARE is a more streamlined subsection of the full LCCS. GLC-SHARE is a best-of-breed, centralized, mosaicked, and harmonized database of regionally produced LULC datasets, each adhering to different classification schemas. GlobCover 2009, Moderate Resolution Imaging Spectrometer (MODIS) 2010, and Cropland Database 2012 supplement GLC-SHARE in areas where higher resolution authoritative national or regional data is not available. Fitness-for-use for GLC-SHARE sources are determined through imagery and ground-truth comparison with the best available source determined at the pixel level (Latham et al. 2014).

Table 7 depicts the consolidation of 11 GLC-SHARE classes to eight RCs and the corresponding class consolidation rationale, informed by the GLC-SHARE classification schema and class definitions.

Table 7. Consolidation of 11 GLC-SHARE classes into eight Research Classes.

GLC-SHARE Class	Research Class	Rationale
Artificial Surfaces	Artificial Surfaces	
Croplands	Croplands	
Grassland Shrub-Covered Sparse Herbaceous	Grassland	Land classes that are predominately grassland, shrub, or herbaceous.
Tree-Covered	Tree-Covered	
Mangrove	Waterlogged	Land classes that represent coastal (saline) flooded areas.
Bare Soil	Bare Soil	
Water Bodies	Water Bodies	
Snow and Glacier	Snow and Glacier	

4.3 Average Temperature and Elevation

Similar to LULC, air temperature (Ceccato et al. 2005, Hay et al. 1998, Kalluri et al. 2007) and elevation (Kalluri et al. 2007) influence vector range and abundance. Average temperature data from the United States Department of Agriculture funded PRISM Climate Group is publically available through Oregon State University. The most current municipality-level temperature dataset in Colombia is dated 1980. Due to this data vintage, average elevation replaces average temperature for the Colombian environmental variable. Average elevation is derived from the 250-meter Shuttle Topography Radar Mission (STRM) dataset.

4.4 Population

The US Census Bureau Population Estimates Program (PEP), publically available through the Census American Fact Finder website, provides the United States county-level population data for intercensal years. The Census Bureau annually adjusts decennial census data based on birth, death, and migration data to produce PEP data. The inclusion rate is 90.9 percent with a margin of error of 0.2 (Census 2012).

The CityPopulation.com website provides Colombian municipality-level population data. CityPopulation data is derived from the National Department of Statistics, Republic of Colombia (CityPopulation 2017).

4.5 Gross Domestic Product

DANE, the Colombian National Statistics Office, makes municipality-level GDP data publically available. This data, gathered from datlascolombia.com, divided by the municipality-level population counts produces per capita GDP.

4.6 Administrative Units

The UNs Food and Agriculture Organizations (FAO) Global Administrative Unit Layers (GAUL) initiative is the source of second-order administrative units within the contiguous United States study area. GAUL is a global scale administrative unit dataset updated annually by the UN Cartographic Unit (GAUL 2014).

GAUL does not include every municipality listed by the Colombian Statistics Department. As a result, Colombian boundaries are from the Database of

Administrative Areas (GADM). Robert Hijimans produces this dataset with support from colleagues at the University of California, Berkeley Museum of Vertebrate Zoology, the International Rice Research Institute, and the University of California, Davis. GADM is in complete alignment with the Colombian municipality names from the Colombian Statistics Department and the CityPopulation website.

4.7 Variable Visualization

Figures 4 through 10 visualize the spatial variation for the independent variables in each study area. Specifically, population density, average temperature, and each grassland landscape metric in the contiguous United States during 2003 and population density, average elevation, per capita GDP, and each landscape metric for tree-covered, cropland, and grassland in Colombia during 2016.

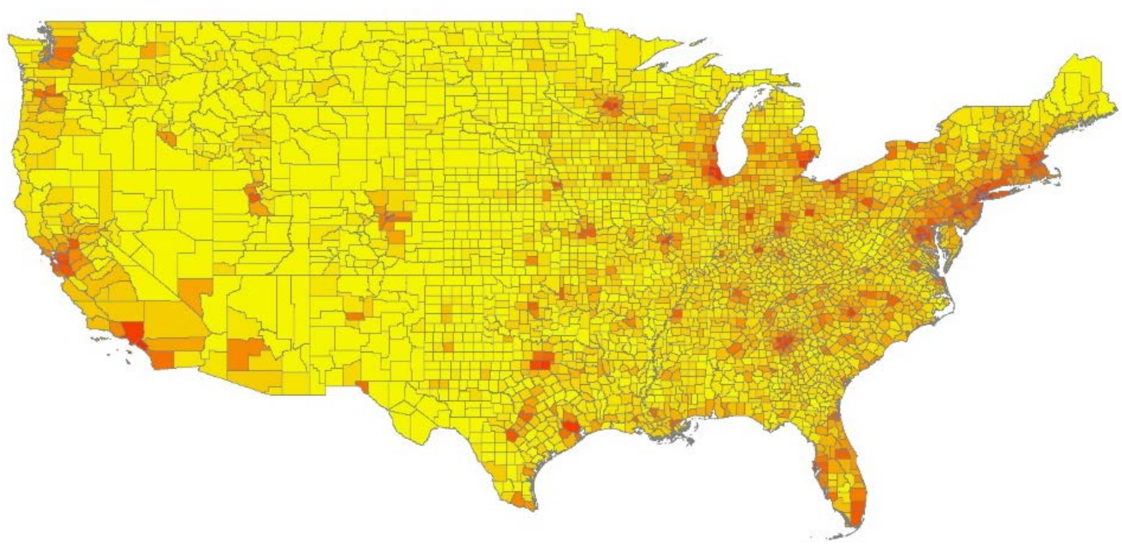


Figure 4. Population Density in the contiguous United States during 2014.

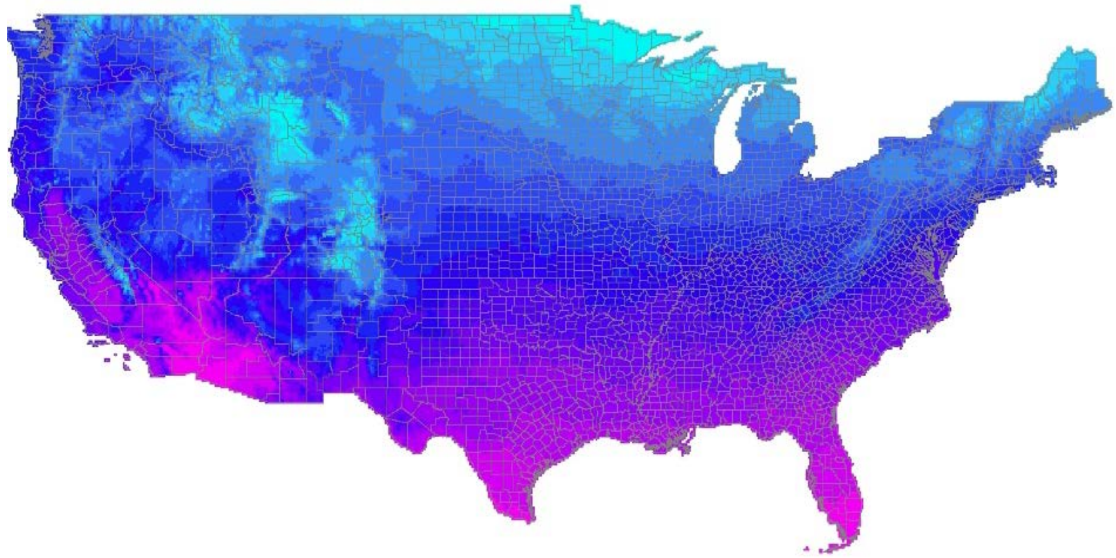


Figure 5. Average Temperature in the contiguous United States during 2014.

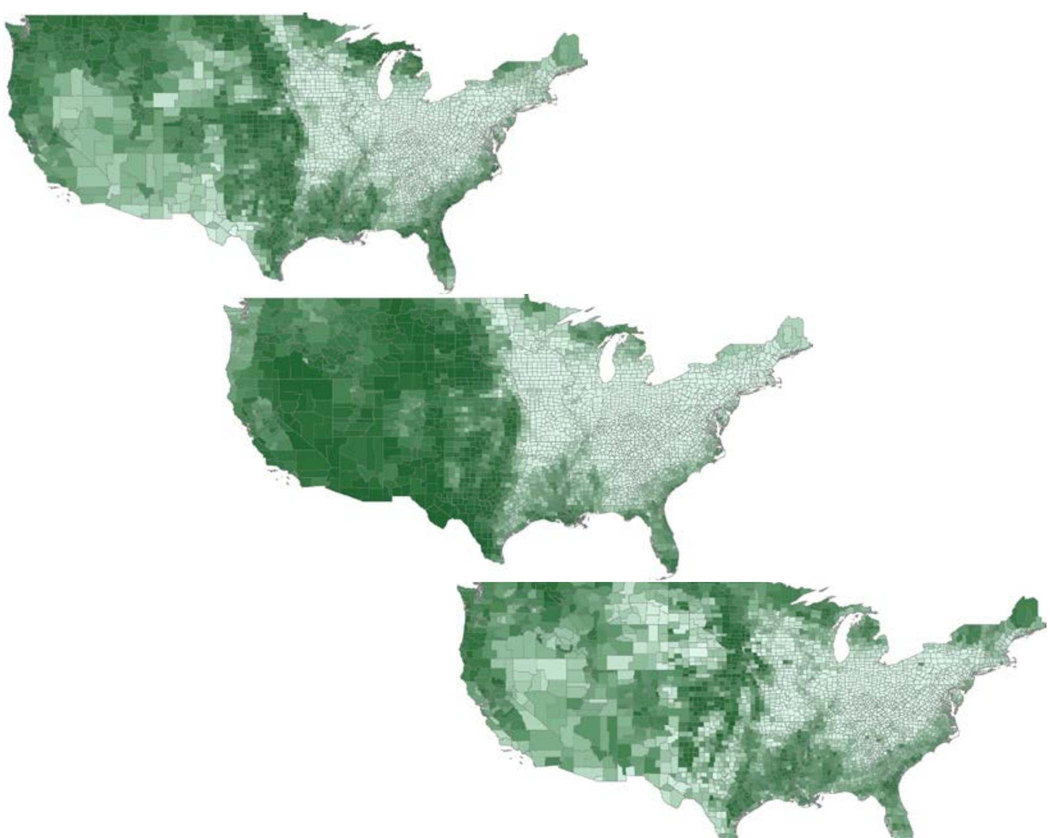


Figure 6. Grassland Linear Density (Top), Proportion Abundance (Middle), and Patch Density (Bottom) in the contiguous United States during 2014.

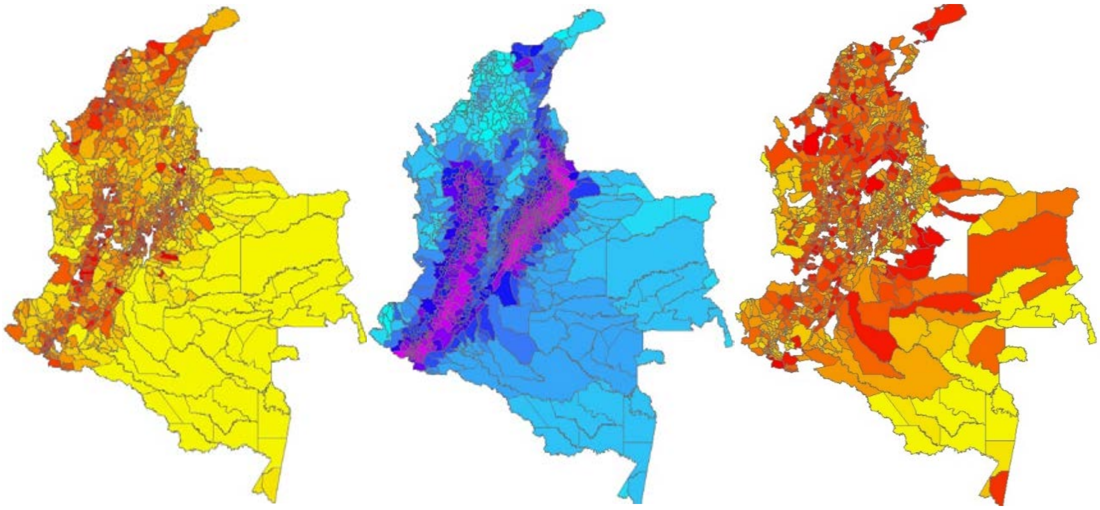


Figure 7. Population Density (Left), Average Elevation (Middle), and Per Capita GDP (Right) in Colombia during 2016.

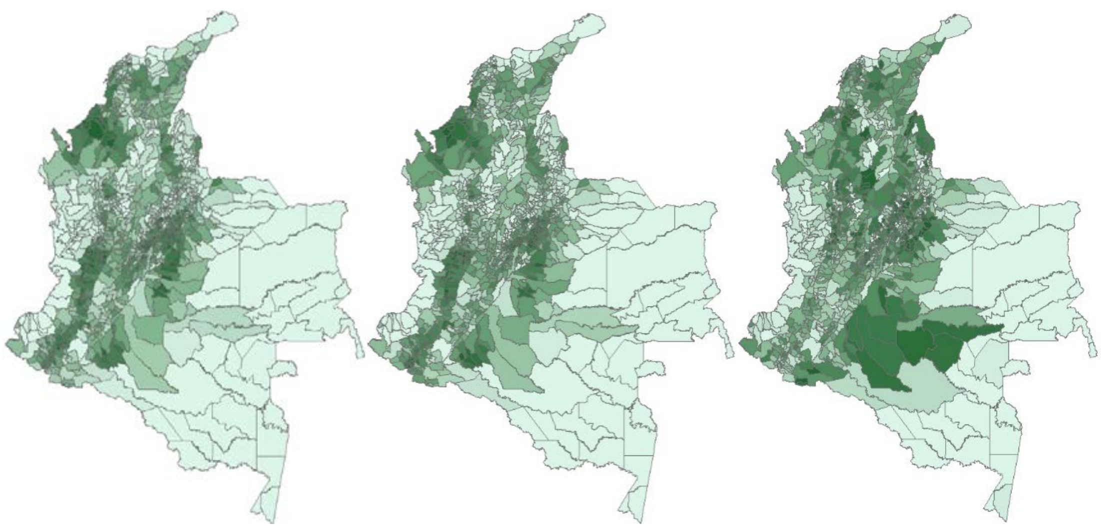


Figure 8. Cropland Linear Density (Left), Proportion Abundance (Middle), and Patch Density (Right) in Colombia during 2016.

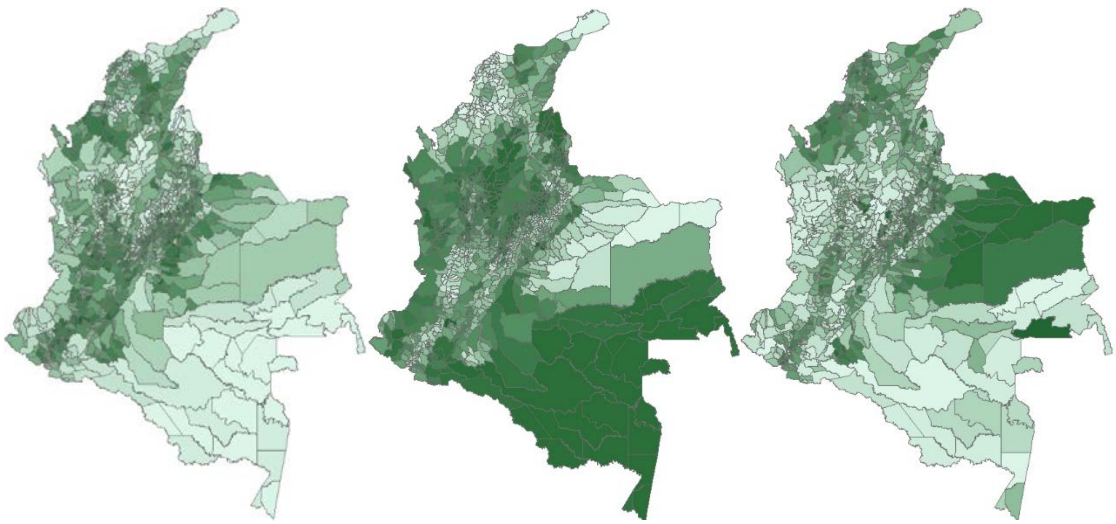


Figure 9. Tree-Covered Linear Density (Left), Proportion Abundance (Middle), and Patch Density (Right) in Colombia during 2016.

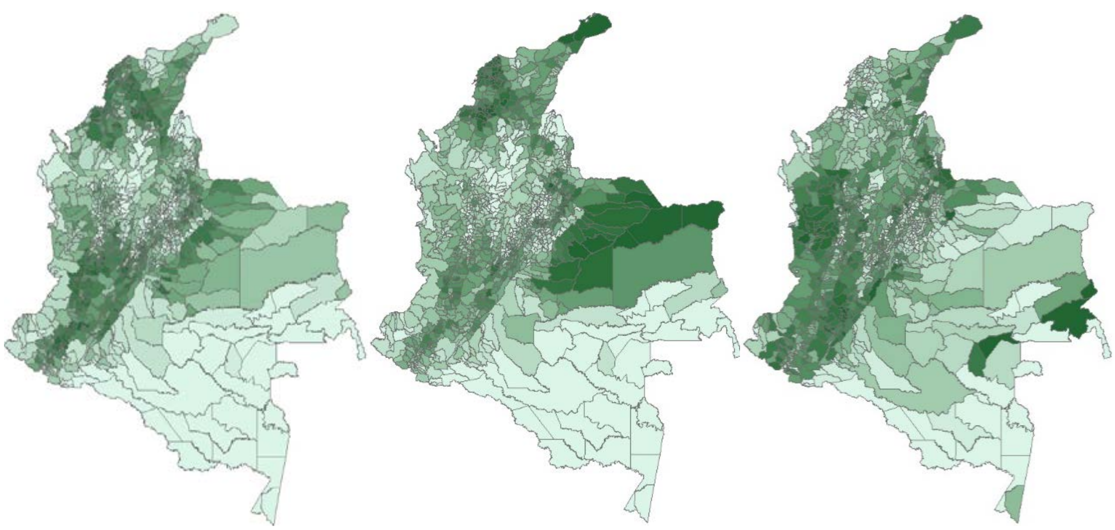


Figure 10. Grassland Linear Density (Left), Proportion Abundance (Middle), and Patch Density (Right) in Colombia during 2016.

Visual analysis of the Colombian variables reveal potential variable correlations. While most correlations (Table 8) are weak to not meaningful, a perfect correlation exists between the cropland proportion abundance and the grassland proportion abundance. An association exists between cropland proportion abundance and cropland linear density while a negative association exists between grassland proportion abundance and tree-covered proportion abundance.

Table 8. Correlation matrix for Colombian independent variables.

	Grass LD	Crop LD	Tree LD	Grass PA	Crop PA	Tree PA	Grass PD	Crop PD	Tree PD	Elevation	Pop Density
Ave GDP	0.04	0.01	0.03	0.11	0.01	(0.08)	(0.04)	(0.03)	0.10	(0.05)	0.00
Pop Den.	0.12	0.00	0.03	0.13	0.47	(0.13)	(0.01)	(0.08)	0.07	0.03	
Elevation	0.08	0.20	0.32	(0.11)	0.03	0.06	0.18	0.08	0.07		
Tree PD	0.09	0.12	0.05	0.37	0.00	(0.46)	(0.45)	(0.27)			
Crop PD	(0.12)	0.31	0.13	(0.11)	(0.03)	0.12	(0.39)				
Grass PD	0.20	(0.14)	0.07	(0.17)	0.01	0.21					
Tree PA	(0.62)	(0.47)	(0.34)	(0.78)	(0.50)						
Crop PA	0.23	0.81	0.45	1.00							
Grass PA	0.58	0.04	0.13								
Tree LD	0.61	0.67									
Crop LD	0.44										

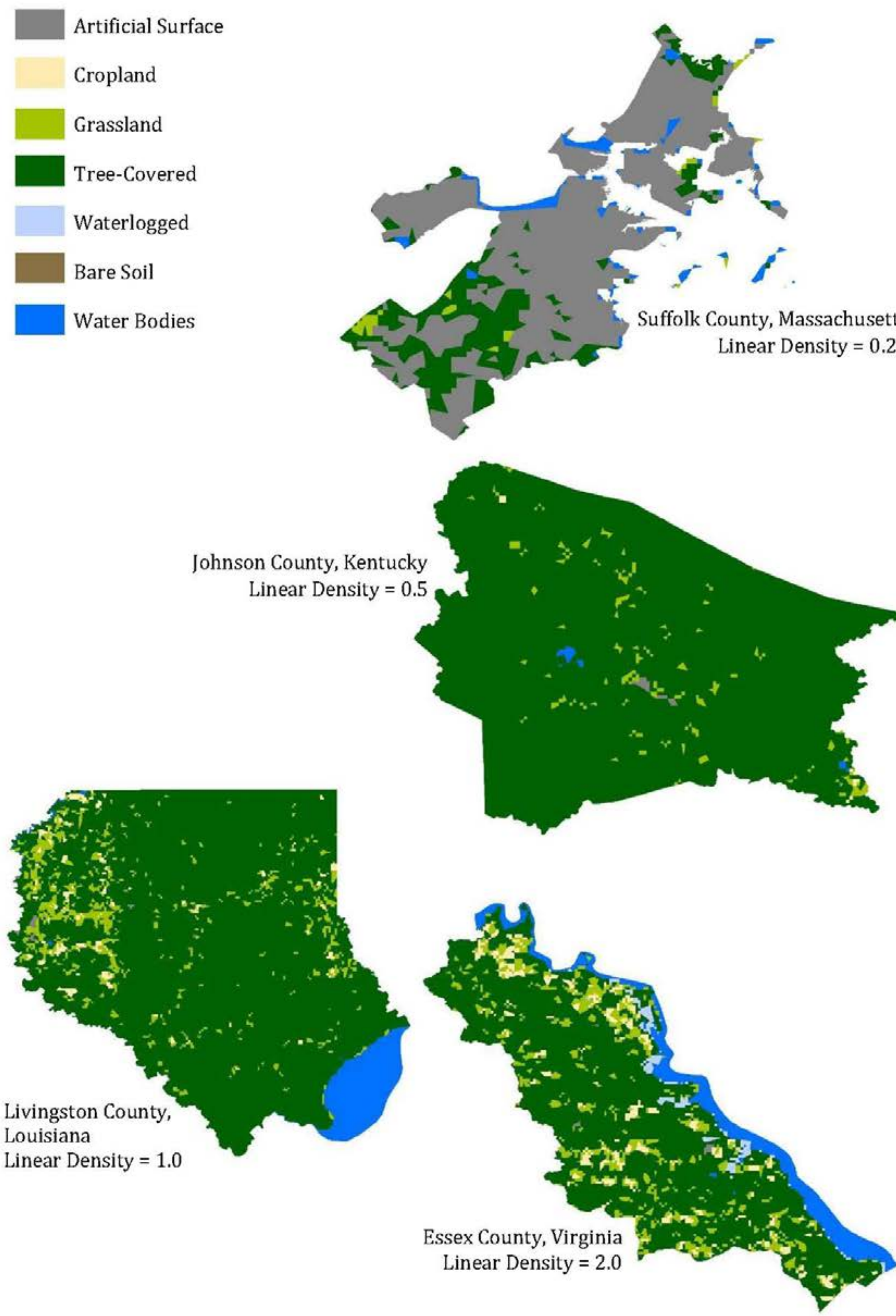
5 RESULTS

The following section contemplates the results through the context of each research question – Do spatial associations exist between the linear density of LULC boundaries and VBID occurrence (RQ1); do these spatial associations repeat over time (RQ2); and do these spatial associations repeat across space (RQ3)?

RQ1 focuses on WNV within the contiguous United States during 2014. RQ2 temporally expands RQ1 to focus on WNV between 2003 through 2014. This broader temporal scope will mitigate the influence of confounding environmental events, an important task since El Niño events occurred in 2003, 2007, and 2011, and La Niña events occurred in 2008, 2011, and 2012 (ESRI 2017). RQ3 shifts geographic study area and VBID type to focus on Zika in Colombia, South America during 2016. Examination of two infectious diseases, each in a different study area, will increase the reliability of analytic outcomes.

RQ1 and RQ2 take advantage of PCA to reduce the number of RC variables into a more manageable amount of Principal Components, reducing the complexity of regression modeling. Removal of Grassland from PCA in RQ1 and RQ2 and inclusion in regression as a stand-alone independent variable was due to its overwhelming spatial significance. To achieve optimal model results, the RCs used to create Principal Components for RQ1 and RQ2 varied across each of the three landscape metrics. However, the RCs used in PCA remained consistent across study years. This PCA and regression method results in a focused analysis of the association between WNV and Grassland in the contiguous United States, with nuanced analysis of the RCs that comprise each Principal Component. RQ3 does not leverage PCA due to the use of fewer RCs. As such, the spatial association between Zika and LULC in Colombia focuses on three primary RCs in Grassland, Cropland, and Tree-Covered.

Research into the spatial associations between land use and infectious disease occurs through the lens of three landscape metrics in linear density, proportion abundance, and patch density. The methodology results described within this section depict a range of landscape metrics values that are often difficult to conceptualize. Figure 11, Figure 12, and Figure 13 visualize incremental increases in the landscape metrics within the contiguous United States during 2003.


Artificial Surface
Cropland
Grassland
Tree-Covered
Waterlogged
Bare Soil
Water Bodies
Suffolk County, Massachusetts
Linear Density = 0.25
Johnson County, Kentucky
Linear Density = 0.5
Livingston County,
Louisiana
Linear Density = 1.0
Essex County, Virginia
Linear Density = 2.0

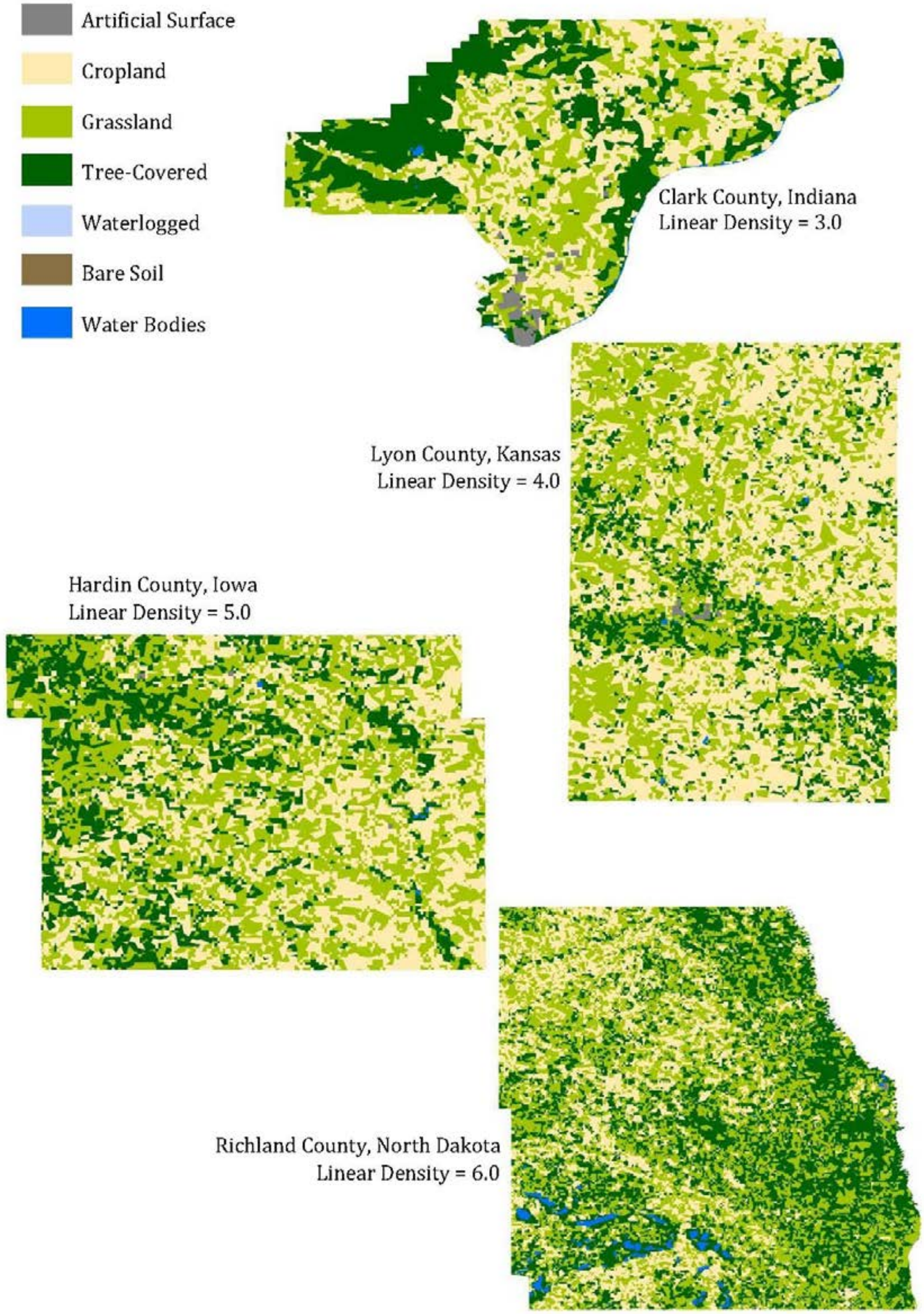
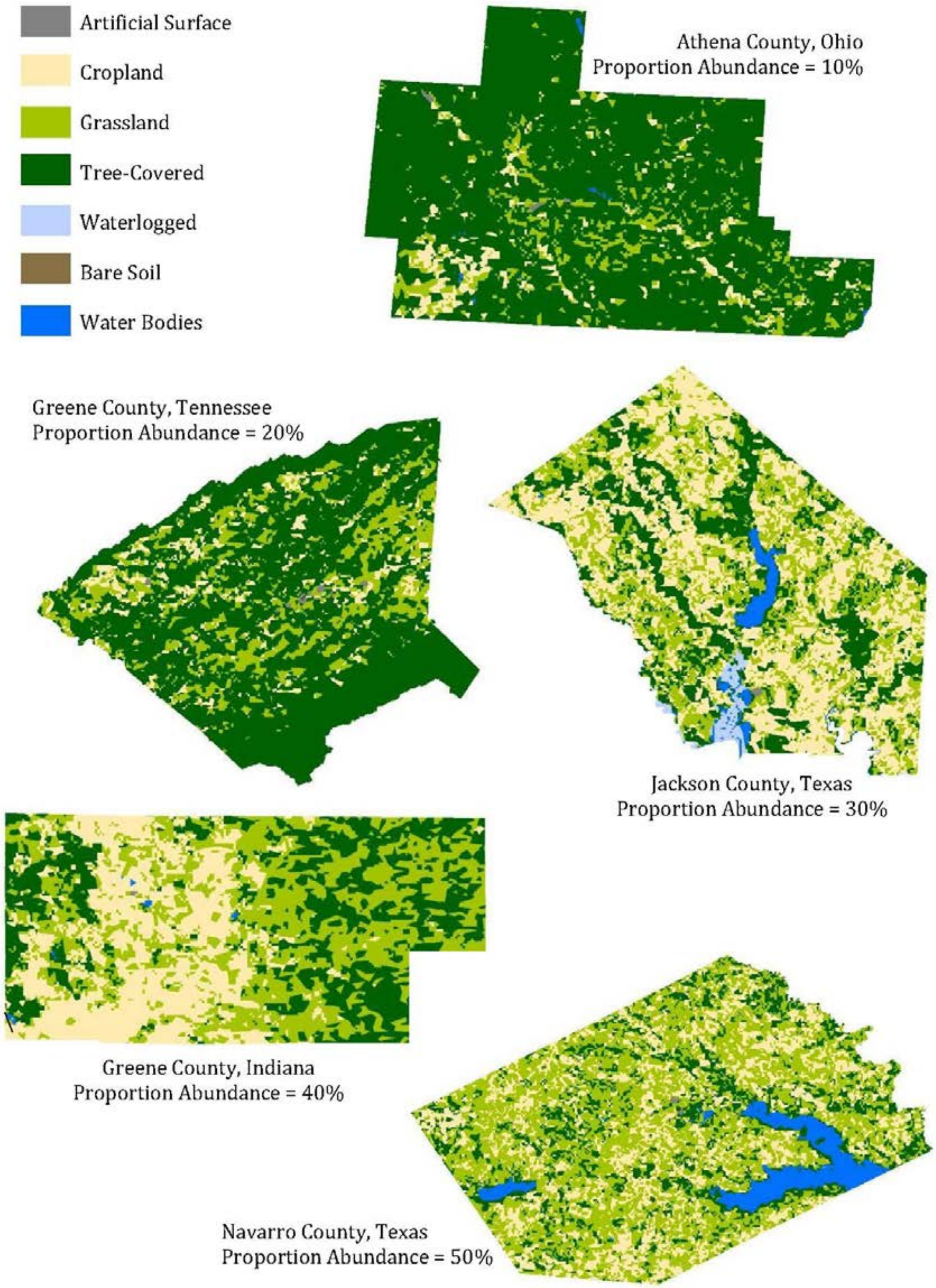


Figure 11. U.S. county examples of incremental increase in Grassland linear density


Artificial Surface
Cropland
Grassland
Tree-Covered
Waterlogged
Bare Soil
Water Bodies
Athena County, Ohio
Proportion Abundance = 10%
Greene County, Tennessee
Proportion Abundance = 20%
Jackson County, Texas
Proportion Abundance = 30%
Greene County, Indiana
Proportion Abundance = 40%
Navarro County, Texas
Proportion Abundance = 50%

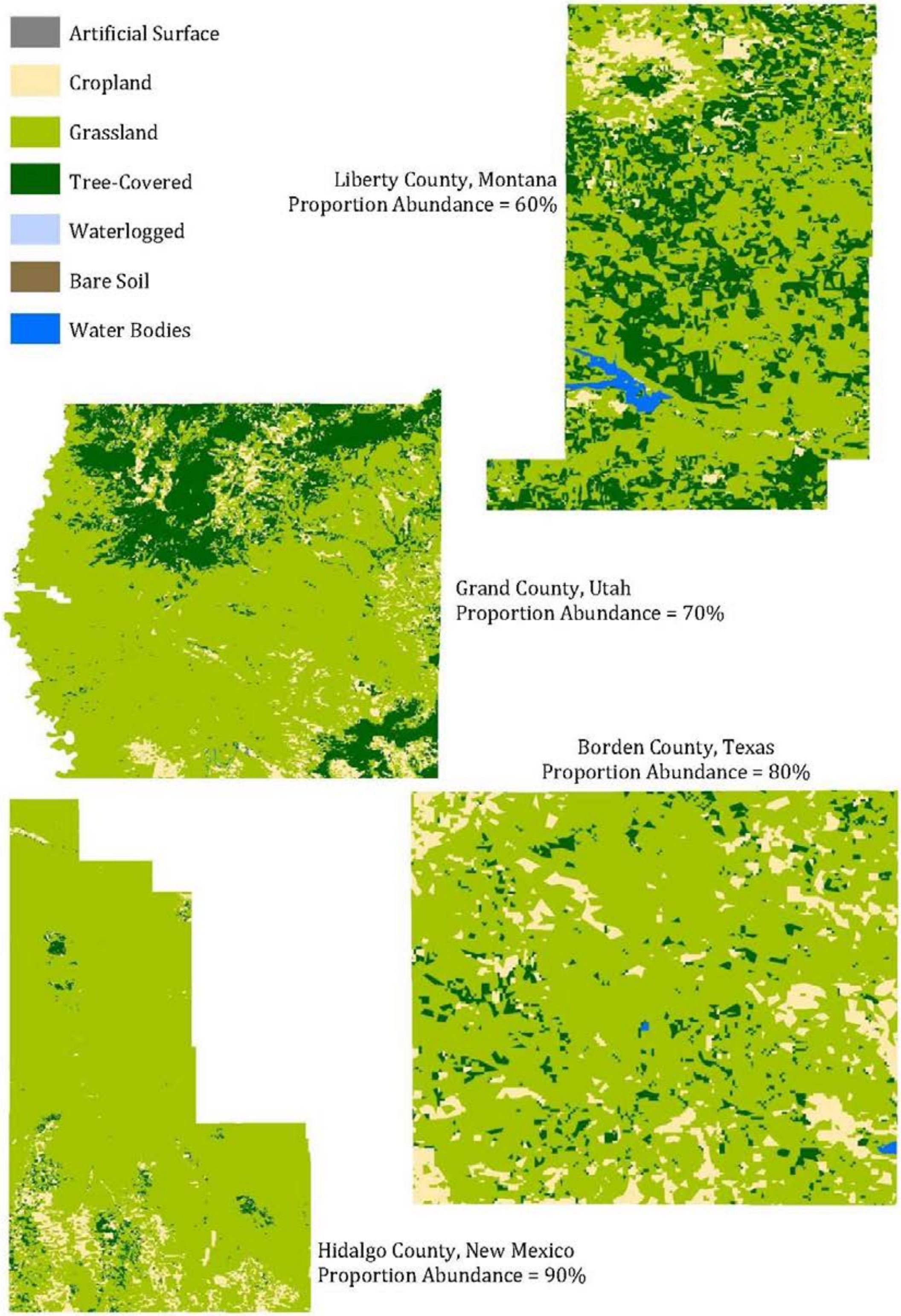
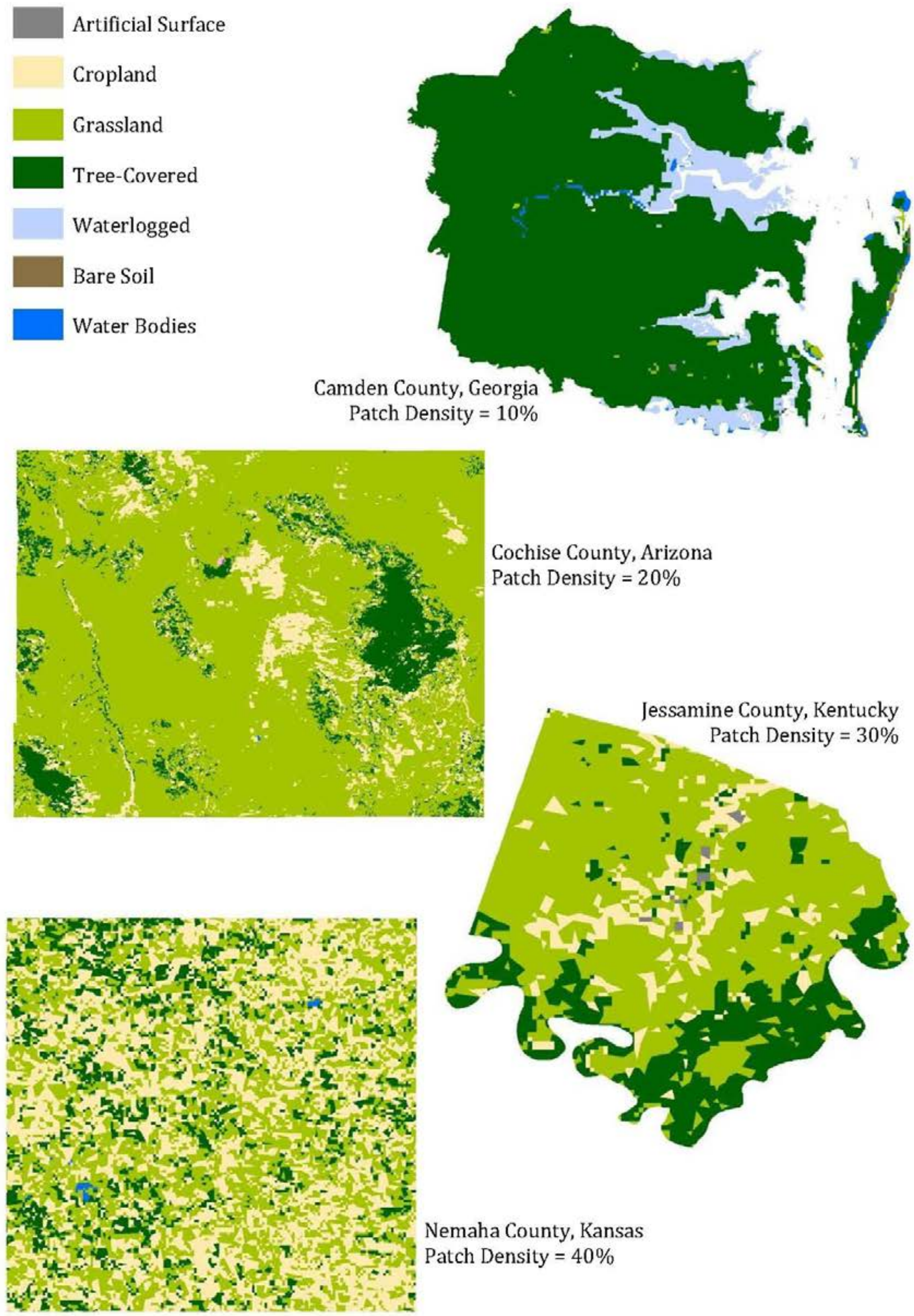


Figure 12. U.S. county examples of incremental increase in Grassland proportion abundance.


Artificial Surface
Cropland
Grassland
Tree-Covered
Waterlogged
Bare Soil
Water Bodies
Camden County, Georgia
Patch Density = 10%
Cochise County, Arizona
Patch Density = 20%
Jessamine County, Kentucky
Patch Density = 30%
Nemaha County, Kansas
Patch Density = 40%

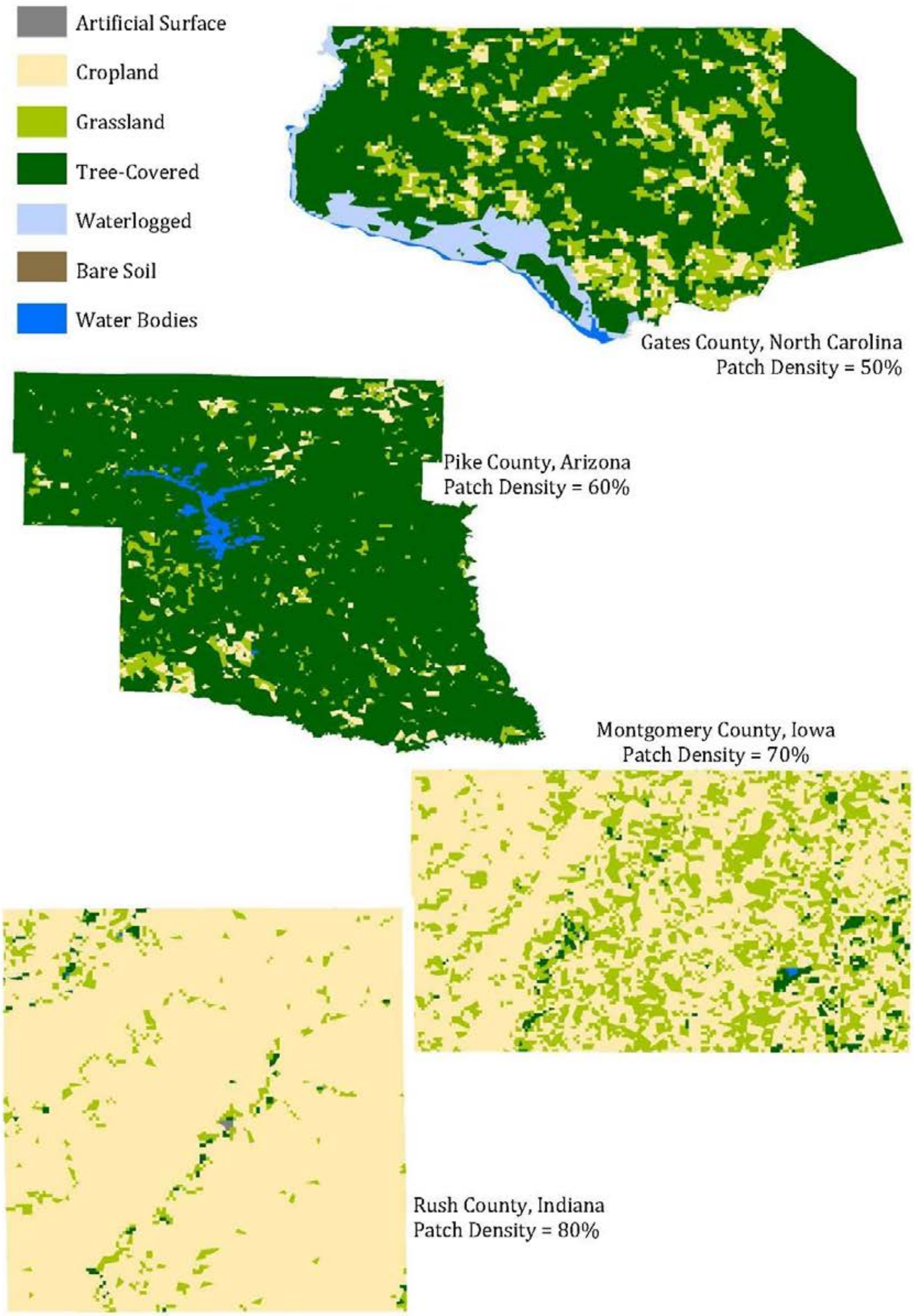


Figure 13. U.S. county examples of incremental increase in Grassland patch density.

5.1 Spatial Associations (RQ1)

Research into the spatial association between LULC and VBID occurrence (RQ1) made use of 2014 WNV data in the contiguous United States. Spatial associations were assessed through the lens of three landscape metrics in linear density, proportion abundance, and patch density. Linear density serves as the primary landscape metric to evaluate the spatial association between LULC and VBID. Supplemental landscape metrics in proportion abundance and patch density add context to the linear density results. The geometric complexity of LULC patches and, to a lesser extent, patch area and count influences linear density. Proportion abundance compares the area of a LULC type to the total area of all LULC. Similarly, patch density compares the patch count of a LULC type to the total patch count.

Linear Density

Cropland, Tree-Covered, Artificial Surface, Waterbodies, and Waterlogged are the RC variables in the PCA dimension reduction process. Originally included in PCA, Grassland was removed and added to regression as a stand-alone independent variable due to the spatial significance of this RC compared to all other RCs (Table 9).

Table 9. Proportion abundance of each Research Class in the contiguous United States during 2014.

Research Class	Proportion Abundance
Grassland	42.05 %
Tree-Covered	27.11 %
Cropland	26.18 %
Artificial Surface	2.51 %
Water Bodies	1.39 %
Bare Soil	0.65 %
Snow and Glacier	0.09 %
Waterlogged	0.02 %

The PCA component matrix (Table 10) depicts strong correlations between all original variables save Waterlogged, resulting in a primary Principal Component that accounts for a large amount of data variance, indicated by values greater than 0.5.

Table 10. PCA component matrix based on the linear density landscape metric.

Principal Component	Cropland	Tree-Covered	Water Bodies	Artificial Surface	Waterlogged
Primary	0.526	0.799	0.537	0.682	0.016
Secondary	(0.599)	(0.248)	0.489	0.362	0.593

The five original RC variables in Cropland, Tree-Covered, Water Bodies, Artificial Surface, and Waterlogged reduce to these two underlying Principal Components, which explain 56.42 percent of the data variance (Table 11).

Table 11. PCA eigenvalues based on the linear density landscape metric.

Principal Component	Eigenvalue	% Variance	Cumulative Variance
Primary	1.679	33.573	33.573
Secondary	1.142	22.847	56.420

VBID occurrence counts serve as the dependent variable in negative binomial regression (Table 12). Independent variables include population density, average temperature, and the linear density of Grassland and the primary and secondary Principal Components. Negative minimum values for the Components are a function of negative variance by one or more of the original variables.

Table 12. Negative binomial regression table based on the linear density landscape metric.

Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
WNV Occurrence (Cases)	0.000	263.000	0.709	7.484	-	-
Ave. Temp (°C)	(1.440)	24.660	11.463	4.800	1.082	1.068, 1.096
Pop Density (People/Sq. Mi.)	0.000	71609.103	261.666	1788.746	1.001	1.001, 1.001
Grassland (Mi/Sq. Mi)	0.000	1.482	0.307	0.346	6.045	4.887, 7.478
Primary PC (Mi/Sq. Mi)	(1.537)	5.790	0.000	1.000	0.487	0.488, 0.529
Secondary PC (Mi/Sq. Mi)	(2.302)	20.747	0.000	1.000	1.583	1.418, 1.767

The decision to use the independent variables listed in Table 12 occurred through multiple model iterations. Model tests prompted the removal of elevation as an independent variable due to collinearity with temperature. Model tests such as Akaike Information Criterion Correlation (AICC) and Bayesian Information Criterion (BIC) validated inclusion or removal of the RCs deemed relevant via prior research. These RCs were modeled individually and in combination with other RCs, the latter to quantify the spatial significance of neighboring (adjacent) RCs.

Eleven of the twelve lowest scoring models based on AICC and BIC values included Grassland, Tree-Covered, Cropland, and Artificial Surface, either individually or in combination with other RCs. The Artificial Surface / Grassland adjacency model scored the lowest AICC and BIC in 11 of the 12 study years for WNV in the contiguous United States. Persistently optimal AICC and BIC scores for the Artificial Surface / Grassland adjacency compared to all other RC adjacency models is not a surprise given the association of urban areas and the surrounding grassland and cropland with elevated VBID risk (Arinaminpathy et al. 2009).

Cropland is present in 86 percent of all (3,107) counties within the contiguous United States, an important measurement since agricultural LULC is responsible for roughly half of all global VBID (McFarlane et al. 2013). Similarly, Tree-Covered is present in 74 percent of contiguous US counties and, when adjacent to Cropland or Artificial Surface, is associated with increased risk of VBID occurrence (Morvan et al. 2000, Patz et al. 2004).

Although the Water Bodies and Waterlogged RCs did not produce AICC and BIC scores as low as models with Cropland, Grassland, Tree-Covered, and Artificial Surface, these two RCs remained in the regression models. The Water Bodies and Waterlogged RCs represent ideal locations for mosquito breeding (EPA 2004). Model tests, adjacency evaluation, and lack of documented associations via prior research eliminated the Bare Soil and Snow/Glacier RCs.

Negative binomial regression reveals that, in 2014, the linear density of Grassland boundaries exhibits a positive association with WNV occurrence in the contiguous United States. Controlling for all other independent variables, each mile per square mile increase in Grassland linear density is associated with a 6.045-fold increase in WNV risk. Waterlogged also displays a positive association with WNV. Each mile per square mile increase in Waterlogged linear density is associated with a 58.3 percent increase in WNV risk. Each one degree Celsius increase in average temperature is associated with an 8.2 percent increase in WNV occurrence while each one-person increase in population density is associated with a 0.1% increase in WNV occurrence.

Proportion Abundance

Proportion abundance measures the area of a single RC compared to the area of all RCs. Appendix I depicts the PCA component matrix, PCA eigenvalue table, and negative binomial output table based on the proportion abundance landscape metric.

The primary Principal Component, largely comprised of the Tree-Covered RC, and the secondary Principal Component, largely comprised of the Artificial Surface RC, account for a combined 54.15 percent of the data variance from the original variables. VBID occurrence counts served as the dependent variable for binomial regression. Population density, average temperature and RCs in Grassland and the Principal Components served as independent variables.

Negative binomial regression reveals that, in 2014, a positive association exists between the proportion abundance of Grassland and WNV occurrence. Controlling for all other independent variables, each one percent increase in Grassland proportion abundance is associated with a 2.4 percent increase in WNV risk. The incidence rate of 1.024 is in stark contrast to the significant association exhibited by Grassland linear density. While the shape (geometry) of Grassland patches influence VBID occurrence, the relative area Grassland occupies has only minor influence.

Each one percent increase in Artificial Surface proportion abundance is associated with a 75.9 percent increase in WNV risk. The average temperature and population density variables repeat their linear density pattern with marginal influence in the proportion abundance model.

Patch Density

Patch density measures patch count of a single RC compared to the total patch count of all RCs. Appendix I depicts the PCA component matrix, PCA eigenvalues, and negative binomial output tables based on patch density. The primary Principal Component, largely comprised of the Tree-Covered RC, and secondary Principal Component, largely comprised of the Water Bodies RC, accounts for a combined 55.66 percent of the data variance from the original variables.

Negative binomial regression reveals that, in 2014, a positive association exists between the patch density of Grassland and WNV occurrence. Controlling for all other independent variables, each one percent increase in Grassland patch density is associated with a 2.9 percent increase in WNV risk. Similar to proportion abundance, the patch density of Grassland marginally influences WNV occurrence.

Landscape Metric Summary

In the contiguous United States during 2014, average temperature and population density display a marginal positive association with WNV occurrence across all landscape metrics. Grassland patch density and proportion abundance exhibit a positive association with WNV occurrence while the Grassland linear density exhibits a positive association. The negative binomial incidence ratios (Table 13) illustrates that relative the risk of WNV occurrence increases significantly in relation to an increase in the linear density of Grassland boundaries.

Table 13. Incidence ratios with WNV (2014) for each independent variable and each landscape metric.

	Linear Density	Proportion Abundance	Patch Density
Ave. Temperature (°C)	1.082	1.068	1.085
Pop. Density (People/Sq. Mi.)	1.001	1.000	1.001
Primary PC	0.487	0.561	0.825
Secondary PC	1.583	1.759	1.307
Grassland	6.045	1.024	1.029

RC variables used in PCA for each landscape metric were adjusted based on model results. Tree-Covered, included in PCA for each landscape metric, exhibits a negative association with WNV occurrence. Waterlogged linear density exhibits a positive association with WNV occurrence, as does Artificial Surface proportion abundance, and Water Body patch density. Stated another way, in 2014, the shape (geometry) of Grassland, the relative area of Artificial Surface, and the quantity of Water Body patches exhibited positive associations with WNV occurrence while the shape (geometry) of Tree-Covered patches exhibited a negative association. This single year analysis represents a snapshot, potentially influenced by multi-year aftereffects of the 2011 El Niño and 2011-2012 La Niña events.

5.2 Spatial Associations over Time (RQ2)

Research into the spatial association between LULC and VBID occurrence over time (RQ2) made use of 2003 through 2014 WNV data in the contiguous United States. Incidence ratios averaged across a longer duration of time provides a robust measure that mitigates uncertainty borne through seasonal and long-term co-occurrences. This temporally focused research question leverages the same independent variables and landscape metrics assigned in RQ1. Similar to RQ1, original variables for PCA include Cropland, Tree-Covered, Artificial Surface, Waterbodies, and Waterlogged.

The twelve-year period between 2003 through 2014 spans three LULC datasets. As depicted in Table 5, GlobCover 2005/2006 is used to assess associations between 2003 through 2007, GlobCover 2009 for associations between 2008 through 2011, and GLC-SHARE 2014 for associations between 2012 through 2014. Slight differences in LULC classification schemas result in analytic outcome differences when shifting from GlobCover 2009 to GLC-SHARE 2014, as evidenced in the PCA Component Matrices for each landscape metric (Table 14, Table 15, and Table 16).

Table 14. PCA Component matrix based on the linear density landscape metric.

LULC Dataset	Year	Principal Component	Cropland	Tree-Covered	Water Bodies	Artificial Surface	Waterlogged
GlobCover 2005/2006	2003 - 2007	Primary	(0.192)	0.284	0.801	0.537	0.724
		Secondary	0.795	0.801	0.064	0.072	(0.227)
GlobCover 2009	2008 - 2011	Primary	0.055	0.456	0.778	0.539	0.678
		Secondary	0.877	0.771	(0.179)	(0.054)	(0.341)
GLC-SHARE 2014	2012 - 2014	Primary	0.536	0.799	0.537	0.682	0.016
		Secondary	(0.599)	(0.248)	0.489	0.362	0.593

Table 15. PCA Component matrix based on the proportion abundance landscape metric.

LULC Dataset	Year	Principal Component	Cropland	Tree-Covered	Water Bodies	Artificial Surface	Waterlogged
GlobCover 2005/2006	2003 - 2007	Primary	(0.955)	0.919	0.259	0.117	0.156
		Secondary	0.085	(0.275)	0.718	0.265	0.751
GlobCover 2009	2008 - 2011	Primary	(0.947)	0.910	0.275	0.124	0.163
		Secondary	0.097	(0.281)	0.716	0.211	0.764
GLC-SHARE 2014	2012 - 2014	Primary	(0.895)	0.834	0.203	0.198	0.014
		Secondary	(0.025)	(0.358)	0.603	0.756	0.259

Table 16. PCA Component matrix based on the patch density landscape metric.

LULC Dataset	Year	Principal Component	Cropland	Tree-Covered	Water Bodies	Artificial Surface	Waterlogged
GlobCover 2005/2006	2003 – 2007	Primary	(0.764)	(0.447)	0.812	0.531	0.596
GlobCover 2009	2008 - 2011	Primary	(0.633)	(0.307)	0.777	0.555	0.637
		Secondary	(0.603)	0.880	(0.162)	(0.080)	0.092
GLC-SHARE 2014	2012 - 2014	Primary	(0.887)	0.597	0.442	0.468	0.115
		Secondary	0.055	(0.684)	0.714	0.196	0.441

Although PCA results are measurably different between the GlobCover and GLC-SHARE datasets, negative binomial regression results are largely consistent. Additionally, for each LULC dataset and landscape metric combination, the first two Principal Components individually account for significant data variance (eigenvalues greater than 0.5) and collectively account for greater than 54 percent of the cumulative variance from all original variables (Table 17).

Table 17. PCA eigenvalues for each LULC dataset and landscape metric combination.

		Linear Density		**Proportion Abundance**		**Patch Density**	
Year	**Principal Component**	**Eigenvalue**	**Cum. Variance**	**Eigenvalue**	**Cum. Variance**	**Eigenvalue**	**Cum. Variance**
2003 - 2007	**Primary**	1.572	31.445	1.862	37.232	2.081	41.610
	Secondary	1.335	58.137	1.233	61.894	0.914	59.898
2008 - 2011	**Primary**	1.566	31.319	1.841	36.813	1.811	36.224
	Secondary	1.516	61.638	1.228	61.377	1.179	59.800
2012 - 2014	**Primary**	1.679	33.573	1.577	31.543	1.570	31.408
	Secondary	1.142	56.420	1.131	54.159	1.213	55.663

Iteration of negative binomial regression for each year between 2003 through 2014 used VBID occurrence counts as the dependent variable and, similar to RQ1, population density, average temperature, Grassland, and the Principal Components as independent variables. Table 18 and Table 19 depict average incidence ratios with WNV for each independent variable and landscape metric combination using the GlobCover and GLC-SHARE LULC datasets, respectively. Appendix II contains the complete regression results for each year and landscape metric combination.

Table 18. Average incidence ratios with WNV for each independent variable and landscape metric using the GlobCover LULC datasets (2003-2011).

	Linear Density	**Proportion Abundance**	**Patch Density**
Ave. Temperature (°C)	1.013	0.995	1.015
Pop. Density (People/Sq. Mi.)	1.001	1.001	1.001
Primary PC	0.907	0.952	0.747
Secondary PC	0.520	1.140	1.466
Grassland	1.749	1.032	0.970

Table 19. Average incidence ratios with WNV for each independent variable and landscape metric using the GLC-SHARE LULC dataset (2012-2014).

	Linear Density	**Proportion Abundance**	**Patch Density**
Ave. Temperature (°C)	1.043	1.031	1.046
Pop. Density (People/Sq. Mi.)	1.001	1.000	1.001
Primary PC	0.653	0.616	0.860
Secondary PC	1.376	1.940	1.199
Grassland	4.712	1.020	1.024

Separation of the annual averages into two tables visualizes the difference in negative binomial results when transitioning from GlobCover (2003 through 2011) to GLC-SHARE (2012 through 2014). Despite the dissimilarities in how the European Space Agency (GlobCover) and the UN FAO (GLC-SHARE) collect and depict LULC, analytic value exists when averaging outcomes from and across both data sources.

Across both sets of averages (2003 through 2011 and 2012 through 2014), regression reveals that the linear density of Grassland boundaries is positively associated with WNV occurrence. Controlling for all other independent variables, each mile per square mile increase in Grassland linear density is associated with a 2.489-fold increase in WNV risk across the entire study timeframe.

The proportion abundance of Grassland reflects a positive association with WNV occurrence. Each one percent increase in Grassland proportion abundance is associated with a 2.9 percent increase in WNV risk. The patch density of Grassland reflects a negative association with WNV occurrence between 2003 through 2011 and a positive association with WNV occurrence between 2012 through 2014. The average incidence ratio of 0.984 reflects a negative association across the entire study timeframe. Each one percent increase in Grassland patch density is associated with a 1.6% decrease in WNV risk. When comparing the incidence ratios for each Grassland landscape metric it becomes evident that the geometric complexity of Grassland patches, rather than their size or count, produces a greater influence on WNV.

Similar to the RQ1 results, average temperature and population density exhibit marginal positive associations with WNV occurrence when averaged from 2003 through 2014. Based on the PCA Component Matrices (Table 14, Table 15, and Table 16), the linear density and patch density of Artificial Surface and Water Bodies exhibit a negative association with WNV occurrence. However, the proportion abundance of Water Bodies is positively associated with WNV occurrence. In the case of Water Bodies, the total area, rather than patch complexity or count, is associated with greater VBID risk. The proportion abundance of Tree-Covered is negatively associated with WNV occurrence while patch density for this RC exhibits a positive association. In the case of Tree-Covered, patch count, rather than area, is associated with greater VBID risk.

Patterns of spatial associations between LULC and VBID repeat over time. The marginal influence of average temperature and population density during 2014 across each landscape metric repeats during the 2003 through 2014 timeframe. Further, Grassland linear density exhibits a positive association during the single year snapshot, averaged within each LU source, and averaged across the entire study.

5.3 Spatial Associations across Space (RQ3)

Research into the spatial association between LULC and VBID across space (RQ3) made use of 2016 Zika data in Colombia. This analysis leverages Zika occurrences as the dependent variable and six independent variables in population density, per capita GDP, average elevation, and the Cropland, Grassland, and Tree-Covered RCs.

There is no need to leverage dimension reduction through PCA since the Cropland, Grassland, and Tree-Covered RCs clearly represent the majority of Colombian LULC. Combined, these three RCs account for more than 98 percent of LULC in Colombia.

For the purpose of study area comparison, Colombia has 36 percent of second-order administrative units found within the contiguous United States – 1,121 municipalities compared to 3,107 counties. Further, Colombia is 8.7 percent of the total area of the contiguous United States. Whereas the maximum annual WNV occurrences in the contiguous United States (between 2003 through 2014) occurred in 2003 with 9,755 occurrences, Colombia experienced 501,970 lab- and clinic-confirmed Zika cases during 2016. Due to the significantly smaller area, yet a Zika occurrence count 52 times greater than that of WNV in the United States, Colombia is a valuable study area for the comparison of spatial associations across space.

Appendix III presents the negative binomial output tables for each landscape metric in the examination of Zika in Colombia during 2016. A positive association exists between the linear density of each RC and Zika occurrence. Controlling for all other independent variables, each mile per square mile increase in the linear density of these RCs is associated with an average 170.6 percent increase in Zika risk. The specific incidence ratio for each RC includes:

- Cropland = 1.083
- Grassland = 1.553
- Tree-Covered = 2.481

The Grassland incidence ratio for the linear density landscape metric exhibits a positive association with VBID in both study areas. While the 2.489 average incidence ratio in the contiguous United States and the 1.553 incidence ratio in Colombia are both positive, the difference is potentially a result of land use composition within each study area. Grassland accounts for 19.21 percent of the total LULC within Colombia, 42.05 percent in the contiguous United States, an important factor since pastoral grassland is associated with increased VBID risk due to the rural co-inhabitance of humans and vectors (Arinaminpathy et al. 2009).

The Tree-Covered incidence ratio for the linear density landscape metric exhibits a negative association for WNV in the contiguous United States and, at 2.481, a positive association for Zika in Colombia. A few factors could contribute to the delta between these incidence ratios, to include differences in tree-covered abundance (27.11 percent in the US, 74.03 percent in Colombia) and vector abundance, along with the infectivity rate differences between WNV and Zika. Morvan et al. (2000) and Patz et al. (2004) assert that frequent contact with pathogenic vectors from forested areas due to settlement expansion increases VBID risk, a factor exacerbated by the significant amount of deforestation driven fragmentation that occurs in Colombia due to settlement expansion (Armenteras et al. 2011).

The proportion abundance of Cropland, Grassland, and Tree-Covered RCs are associated with Zika occurrence. Controlling for all other independent variables, each one percent increase in proportion abundance is associated with an average 5.3 percent increase in Zika risk. Specific incidence ratios include:

- Cropland = 1.040
- Grassland = 1.066
- Tree-Covered = 1.054

A positive association also exists between the patch density of Cropland, Grassland, and Tree-Covered RCs and Zika occurrence. Controlling for all other independent variables, each one percent increase in patch density is associated with an average 2.6 percent increase in Zika risk. The incidence ratio for each RC includes:

- Cropland = 1.026
- Grassland = 1.029
- Tree-Covered = 1.023

Across each Colombian landscape metric, average elevation exhibits an incidence ratio of 0.999. Each meter increase in elevation is associated with a 0.1 percent decrease in WNV risk. With an incidence ratio of 1.000, per capita GDP has no influence on Zika occurrence in Colombia. Similar to Grassland examined in RQ1 and RQ2, the linear density of Grassland displays a positive association with VBID while the proportion abundance and patch density of this RC exhibits a nominal influence. Such patterns support that, to a significant extent, spatial associations between LULC and VBID repeat across space. In addition to these patterns, an interesting association was discovered within the Colombian study area in the marginal influence of LULC proportion abundance and patch density on Zika occurrence.

5.4 Summary of Research Question Results

Patterns exist between Zika in Colombia during 2016 and WNV in the contiguous United States between 2003 through 2014. Across both study areas and VBIDs, the linear density of particular LULC boundaries exhibit positive associations with VBID occurrence, specifically Grassland in the contiguous United States and Grassland and

Tree-Covered in Colombia. The associations between VBID and the proportion abundance and patch density landscape metrics neither supports nor calls to question the linear density results. Further, population density, per capita GDP, temperature, and elevation do not significantly influence VBID occurrence.

The positive spatial association between the density of Grassland boundaries and VBID occurrence exists for each year individually and for the average across all 12-years in the contiguous United States study area. A positive association also exists between the linear density of Grassland and, to a greater extent, Tree-Covered boundaries and Zika despite the RQ3 shift of geographic study area and VBID. This pattern of association between the VBID and the linear density of LULC exists alongside marginal associations exhibited by the proportion abundance and patch density landscape metrics.

Geometric complexity of LULC patches determines linear density. Linear density, however, is driven by human behaviors and activities to meet the demands of a growing population and is expressed through urban morphology and corresponding land development plans and policies. The shape complexity factor plays a more critical role in determining the magnitude of VBID and LULC association than does the relative area or patch count of a single LULC. Urban morphology and land development plans and policies that emphasize simple and compact rather than complex and irregular (Figure 14) patch geometries will reduce linear density and thus reduce the spatial breadth over which humans and vectors come in contact.

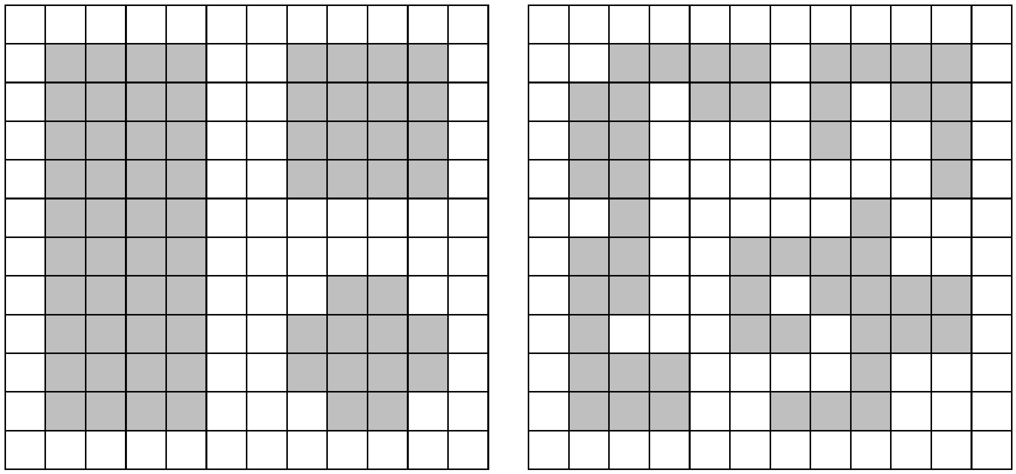


Figure 14. While proportion abundance and patch density remain constant, the simple (left) linear density value of 0.40 is nearly half of the complex (right) linear density value of 0.78.

In addition, a simple and compact geometry increases a patches core area. This boundary related factor, more so than patch area or count, is associated with habitat health, to include the preservation of species at each link of the food chain. While a patch may be large enough to support a given species, it still may not contain a core area large enough to support a diverse range of species (McGarigal 2017). Preservation of species diversity reduces the likelihood that pathogens can exploit vacant ecological niches (McMichael 2004, Murray and Daszak 2013, Pike et al. 2010).

6 CONCLUSION

Research into the spatial relationships between LULC and VBID is limited, though field studies and remote sensing methods have successfully identified associations, predominately within developed countries. In addition to a new methodology, this dissertations broad spatial scope encompasses two countries while the temporal scope spans 12 years, mitigating for confounding environmental variables that could influence smaller scale or temporally limited research. Further, examination of the spatial association between LULC and VBID occurs through the lens of three landscape metrics in linear (edge) density, proportion abundance, and patch density.

This dissertation reveals that spatial associations exist between the linear density of specific LULC boundaries and occurrences of WNV and Zika. Within the contiguous United States, an increase of Grassland linear density exhibits a positive association with WNV occurrence while proportion abundance exhibits a marginal positive association. In Colombia, South America, an increase in Grassland or Tree-Covered linear density exhibits positive associations with Zika occurrence while proportion abundance and patch density of these classes exhibit a marginal positive association.

The significance of LULC boundaries, linear density in particular, on VBID emergence is in line with prior research performed by spatial epidemiologists, to include Morse et al. (2012), Woolhouse et al. (2012), Pike et al. (2010), Wilcox and Ellis (2006), and Morse (2004). However, this dissertation reveals two significant points. First, the cropland landscape metrics displayed a minor positive association with VBID in both study areas, at least when compared to Grassland and Tree-Covered. This marginal association is contrary to the 2002 Working Group (McMichael 2004) designation of agricultural development as the primary driver of VBID emergence. In addition, the significance of LULC boundary linear density on VBID occurrence concurrent with the marginal influence of both proportion abundance and patch density represents an important finding. The landscape factor that influences disease emergence is the amount of specific LULC boundaries (linear density), not the overall area or number of patches.

Patch shape (geometry) determines LULC linear density. Shape and linear density are products of urban morphology and the corresponding land development plans and policies – behaviors and activities that result from our need to support a rapidly growing population. Patch shapes that are simple and compact produce linear density values lower than complex and irregular shapes (Figure 14). As a result, simple and compact patches will reduce VBID occurrences, a potential outcome of the larger core area that characterizes simple and compact patches. Core area is associated with habitat health and preservation of species, including predator species, the lack of which spikes populations of vectors and hosts and provides pathogens niches to exploit. Larger unbroken patches result in greater control of hosts and vectors by predators, thus naturally controlling VBID outbreaks (Ezenwa et al. 2007).

An opportunity exists to mitigate the impact of VBID through changes to land development policies related to zoning and planning. Changes that emphasize the simplification of patch shape (geometry) will reduce LULC linear density within a given area and result in an associated decrease in VBID occurrence. Based on the RQ1 and RQ2 outcomes, such actions applied to Grassland will reduce WNV occurrences within the contiguous United Stated. Similarly, based on the results of RQ3, such actions applied to Grassland or Tree-Covered will reduce Zika occurrence in Colombia. Based on these results, policy/decision makers within local governments can reduce future VBID occurrences through alteration of land development plans. Such proactive avoidance will reduce the social and economic impacts of VBID, along with the burden borne by local health agencies. Consideration of VBID risk based on the impact of land development plans represents a shift in urban morphology practices and the corresponding modification of the environment.

Epidemiologists within the public health community can supplement their prevention and mitigation toolbox through the identification of areas that could experience an increase in VBID occurrences. These potential benefits to policy makers and the health community answers the call for additional research into the association between LULC and infectious disease from spatial epidemiologists such as McFarlane et al. (2013), Murray and Daszak (2013), and Morse (2012).

This dissertation enables research that incorporates additional landscape metrics, areas of interest, and VBIDs. Data availability and quality factors restricted this study to WNV in the United States and Zika in Colombia. Research into the same VBID across two disparate study areas would provide a more robust comparison that improves the assessment of spatial association patterns across space. This dissertation includes the uncertainty inherent in the comparison of a single year snapshot from one geographic region (Colombia – 2016) to 12 years from a different geographic region (US – 2003-2014). Research that utilizes data that spans multiple years across multiple study areas can improve the assessment of spatial association patterns over time. In addition, research outcomes could be further refined through incorporation of additional landscape measurement methods, such as perimeter-area methods to assess fractal dimension, core area methods to assess the core to edge ratio, and contrast methods to assess the magnitude of difference in LULC along patch edges (McGarigal 2017).

The use of second-order administrative units to evaluate LULC and VBID associations offers benefits and drawbacks. Spatial granularity reveals associations hidden at state, regional, or national scales (Arsenault et al. 2013). Conversely, uncertainty arises when VBID transmission occurs outside of the unit where the patient receives health care. Similar to the transmission/report uncertainty, the location of labs and clinics in relation to the patient seeking a diagnosis introduces uncertainty. Factors such as drive time, convenience, and insurance could influence a patient's decision to seek medical attention outside of the VBID transmission unit (Schuurman et al. 2006). However, VBID occurrence data that includes the probable transmission location would allow for an increase in the spatial granularity of analysis, perhaps at the sub-county level. While the aforementioned uncertainties and limitations hinder the accurate identification disease hotspots at the sub-county level, the evaluation of land development plans using the methodology described within this dissertation will reveal counties that could experience greater risk of VBID.

7 APPENDIX I

Table 20, Table 21, and Table 22 present the PCA component matrix, PCA eigenvalues, and negative binomial output tables based on proportion abundance.

Table 20. PCA Component matrix based on the proportion abundance landscape metric.

Principal Component	Cropland	Tree-Covered	Water Bodies	Artificial Surface	Waterlogged
1	(0.985)	0.835	0.203	0.194	0.014
2	(0.026)	(0.355)	(0.603)	0.757	0.261

Table 21. PCA eigenvalues based on the proportion abundance landscape metric.

Principal Component	Eigenvalue	% Variance	Cumulative Variance
1	1.577	31.537	31.537
2	1.131	22.616	54.153
3	0.993	19.868	74.021
4	0.910	18.209	92.231
5	0.388	7.769	100.000

Table 22. Negative binomial regression table based on the proportion abundance landscape metric.

Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
WNV Occurrence (Cases)	0.000	263.000	0.709	7.488	-	-
Ave. Temperature (°C)	(1.440)	24.660	11.466	4.799	1.068	1.053, 1.083
Pop Density (People/Sq. Mi.)	0.000	71609.103	257.239	1782.790	1.000	1.000, 1.001
Grassland (Percent)	0.000	99.943	21.142	30.272	1.024	1.022, 1.026
Primary PC (Percent)	(1.741)	2.367	0.000	1.000	0.561	0.513, 0.614
Secondary PC (Percent)	(1.115)	13.065	0.000	1.000	1.759	1.584, 1.952

Table 23, Table 24, and Table 25 display the PCA component matrix, PCA eigenvalues, and negative binomial outputs based on patch density.

Table 23. PCA Component matrix based on the patch density landscape metric.

Principal Component	Cropland	Tree-Covered	Water Bodies	Artificial Surface	Waterlogged
1	(0.887)	0.597	0.442	0.468	0.155
2	0.055	(0.684)	0.714	0.196	0.441

Table 24. PCA eigenvalues based on the patch density landscape metric.

Principal Component	Eigenvalue	% Variance	Cumulative Variance
1	1.570	31.408	31.408
2	1.213	24.255	55.663
3	0.971	19.413	75.076
4	0.874	17.473	92.549
5	0.373	7.451	100.00

Table 25. Negative binomial regression table based on the patch density landscape metric.

Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
WNV Occurrence (Cases)	0.000	263.000	0.709	7.484	-	-
Ave. Temperature (°C)	(1.440)	24.660	11.464	4.800	1.085	1.071, 1.098
Pop Density (People/Sq. Mi.)	0.000	71609.103	261.666	1788.746	1.001	1.001, 1.001
Grassland (Percent)	0.000	98.214	22.987	22.210	1.029	1.026, 1.032
Primary PC (Percent)	(2.487)	3.194	0.000	1.000	0.825	0.762, 0.894
Secondary PC (Percent)	(2.353)	14.340	0.000	1.000	1.307	1.200, 1.423

8 APPENDIX II

Table 26. Negative binomial regression output table based on the linear density landscape metric for RQ2, WNV in the contiguous United States during 2003 through 2014.

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2003	**WNV Occurrence** (Cases)	0.000	546.000	3.140	18.395	-	-
	Ave. Temperature (°C)	(1.390)	24.560	12.082	4.586	0.881	0.872, 0.889
	Pop Density (People/Sq. Mi.)	0.106	68,365.602	242.297	1,690.168	1.001	1.001, 1.001
	Grassland (Mi/Sq. Mi)	0.000	6.523	2.396	1.469	2.282	2.103, 2.476
	Primary PC (Mi/Sq. Mi)	(1.093)	11.367	0.000	1.000	0.886	0.835, 0.941
	Secondary PC (Mi/Sq. Mi)	(3.409)	2.635	0.000	1.000	0.337	0.300, 0.378
2004	**WNV Occurrence** (Cases)	0.000	355.000	0.804	9.885	-	-
	Ave. Temperature (°C)	(2.120)	24.130	12.290	4.658	1.096	1.082, 1.110
	Pop Density (People/Sq. Mi.)	0.081	68,706.652	243.446	1,691.856	1.001	1.001, 1.001
	Grassland (Mi/Sq. Mi)	0.000	6.523	2.396	1.469	1.400	1.245, 1.574
	Primary PC (Mi/Sq. Mi)	(1.093)	11.367	0.000	1.000	0.729	0.673, 0.789
	Secondary PC (Mi/Sq. Mi)	(3.410)	2.635	0.000	1.000	0.541	0.456, 0.642

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2005	**WNV Occurrence** (Cases)	0.000	163.000	0.956	5.794	-	-
	Ave. Temperature (°C)	(1.450)	24.110	12.616	4.519	1.043	1.031, 1.056
	Pop Density (People/Sq. Mi.)	0.103	68,865.339	244.566	1,691.283	1.001	1.001, 1.001
	Grassland (Mi/Sq. Mi)	0.000	6.523	2.396	1.469	2.158	1.962, 2.375
	Primary PC (Mi/Sq. Mi)	(1.093)	11.367	0.000	1.000	0.950	0.887, 1.016
	Secondary PC (Mi/Sq. Mi)	(3.410)	2.635	0.000	1.000	0.448	0.389, 0.515
2006	**WNV Occurrence** (Cases)	0.000	251.000	1.346	8.115	-	-
	Ave. Temperature (°C)	(1.560)	24.190	12.990	4.493	0.937	0.928, 0.947
	Pop Density (People/Sq. Mi.)	0.111	69,066.565	245.808	1,692.597	1.001	1.001, 1.001
	Grassland (Mi/Sq. Mi)	0.000	6.523	2.396	1.469	1.865	1.692, 2.056
	Primary PC (Mi/Sq. Mi)	(1.093)	11.367	0.000	1.000	0.944	0.886, 1.005
	Secondary PC (Mi/Sq. Mi)	(3.410)	2.635	0.000	1.000	0.460	0.399, 0.529
2007	**WNV Occurrence** (Cases)	0.000	140.000	1.116	5.708	-	-
	Ave. Temperature (°C)	(1.220)	24.550	12.644	4.604	0.934	0.925, 0.943
	Pop Density (People/Sq. Mi.)	0.117	69,207.965	247.475	1,697.527	1.001	1.000, 1.001
	Grassland (Mi/Sq. Mi)	0.000	6.523	2.396	1.469	1.688	1.533, 1.858
	Primary PC (Mi/Sq. Mi)	(1.093)	11.367	0.000	1.000	0.924	0.863, 0.990
	Secondary PC (Mi/Sq. Mi)	(3.410)	2.635	0.000	1.000	0.605	0.525, 0.697
2008	**WNV Occurrence** (Cases)	0.000	156.000	0.426	3.974	-	-
	Ave. Temperature (°C)	(2.190)	24.230	11.803	4.871	1.074	1.058, 1.090
	Pop Density (People/Sq. Mi.)	0.090	69,453.917	249.498	1,707.513	1.001	1.001, 1.001
	Grassland (Mi/Sq. Mi)	0.000	6.133	1.944	1.506	1.741	1.581, 1.917
	Primary PC (Mi/Sq. Mi)	(1.059)	10.949	0.000	1.000	0.789	0.728, 0.854
	Secondary PC (Mi/Sq. Mi)	(5.413)	3.302	0.000	1.000	0.520	0.446, 0.607

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2009	**WNV Occurrence** (Cases)	0.000	27.000	0.228	1.423	-	-
	Ave. Temperature (°C)	(2.000)	24.290	11.812	4.873	1.021	1.004, 1.039
	Pop Density (People/Sq. Mi.)	0.114	69,296.761	251.741	1,715.247	1.000	1.000, 1.000
	Grassland (Mi/Sq. Mi)	0.000	6.133	1.944	1.506	1.831	1.633, 2.054
	Primary PC (Mi/Sq. Mi)	(1.059)	10.949	0.000	1.000	0.735	0.667, 0.809
	Secondary PC (Mi/Sq. Mi)	(5.413)	3.302	0.000	1.000	0.414	0.342, 0.502
2010	**WNV Occurrence** (Cases)	0.000	115.000	0.332	2.766	-	-
	Ave. Temperature (°C)	(1.510)	23.420	12.284	4.422	1.056	1.037, 1.075
	Pop Density (People/Sq. Mi.)	0.000	69,518.337	253.601	1,726.330	1.001	1.000, 1.001
	Grassland (Mi/Sq. Mi)	0.000	6.133	1.944	1.506	1.579	1.426, 1.748
	Primary PC (Mi/Sq. Mi)	(1.059)	10.949	0.000	1.000	1.023	0.944, 1.109
	Secondary PC (Mi/Sq. Mi)	(5.413)	3.302	0.000	1.000	0.551	0.471, 0.646
2011	**WNV Occurrence** (Cases)	0.000	58.000	0.228	1.772	-	-
	Ave. Temperature (°C)	(2.030)	24.570	12.537	4.911	1.071	1.050, 1.092
	Pop Density (People/Sq. Mi.)	0.000	70,460.700	255.875	1,748.637	1.001	1.000, 1.001
	Grassland (Mi/Sq. Mi)	0.000	6.133	1.944	1.506	1.194	1.064, 1.340
	Primary PC (Mi/Sq. Mi)	(1.059)	10.949	0.000	1.000	1.180	1.081, 1.289
	Secondary PC (Mi/Sq. Mi)	(5.413)	3.302	0.000	1.000	0.803	0.676, 0.953
2012	**WNV Occurrence** (Cases)	0.000	396.000	1.760	11.480	-	-
	Ave. Temperature (°C)	(0.280)	24.720	13.590	4.467	1.090	1.078, 1.102
	Pop Density (People/Sq. Mi.)	0.000	71,124.639	258.099	1,767.009	1.001	1.001, 1.001
	Grassland (Mi/Sq. Mi)	0.000	1.482	0.309	0.346	3.766	3.244, 4.371
	Primary PC (Mi/Sq. Mi)	(1.537)	5.790	0.000	1.000	0.893	0.846, 0.943
	Secondary PC (Mi/Sq. Mi)	(2.302)	20.747	0.000	1.000	1.230	1.141, 1.327

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2013	**WNV Occurrence** (Cases)	0.000	158.000	0.788	4.290	-	-
	Ave. Temperature (°C)	(1.650)	24.780	11.802	4.764	0.957	0.946, 0.968
	Pop Density (People/Sq. Mi.)	0.000	71,422.538	260.097	1,779.854	1.001	1.000, 1.001
	Grassland (Mi/Sq. Mi)	0.000	1.482	0.307	0.346	4.324	3.633, 5.147
	Primary PC (Mi/Sq. Mi)	(1.537)	5.790	0.000	1.000	0.578	0.540, 0.619
	Secondary PC (Mi/Sq. Mi)	(2.302)	20.747	0.000	1.000	1.315	1.203, 1.438
2014	**WNV Occurrence** (Cases)	0.000	263.000	0.709	7.484	-	-
	Ave. Temperature (°C)	(1.440)	24.660	11.463	4.800	1.082	1.068, 1.096
	Pop Density (People/Sq. Mi.)	0.000	71,609.103	261.666	1,788.746	1.001	1.001, 1.001
	Grassland (Mi/Sq. Mi)	0.000	1.482	0.307	0.346	6.045	4.887, 7.478
	Primary PC (Mi/Sq. Mi)	(1.537)	5.790	0.000	1.000	0.487	0.448, 0.529
	Secondary PC (Mi/Sq. Mi)	(2.302)	20.747	0.000	1.000	1.583	1.418, 1.767

Table 27. Negative binomial regression output table based on the proportion abundance landscape metric for RQ2, WNV in the contiguous United States during 2003 through 2014.

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2003	**WNV Occurrence** (Cases)	0.000	546.000	3.137	18.401	-	-
	Ave. Temperature (°C)	(1.390)	24.560	12.077	4.583	0.876	0.868, 0.884
	Pop Density (People/Sq. Mi.)	0.106	68,365.602	242.322	1,690.707	1.000	1.000, 1.001
	Grassland (Percent)	0.000	97.575	25.039	18.141	1.033	1.030, 1.036
	Primary PC (Percent)	(2.617)	2.690	0.000	1.000	0.723	0.687, 0.761
	Secondary PC (Percent)	(0.846)	12.917	0.000	1.000	1.231	1.165, 1.301
2004	**WNV Occurrence** (Cases)	0.000	306.000	0.688	7.571	-	-
	Ave. Temperature (°C)	(2.120)	24.130	12.286	4.656	1.039	1.025, 1.053
	Pop Density (People/Sq. Mi.)	0.081	68,706.652	243.468	1,692.396	1.001	1.001, 1.001
	Grassland (Percent)	0.000	97.575	25.039	18.141	1.045	1.041, 1.048
	Primary PC (Percent)	(2.617)	2.690	0.000	1.000	1.274	1.166, 1.393
	Secondary PC (Percent)	(0.846)	12.917	0.000	1.000	1.080	1.020, 1.144
2005	**WNV Occurrence** (Cases)	0.000	163.000	0.929	5.630	-	-
	Ave. Temperature (°C)	(1.450)	24.110	12.612	4.517	1.013	1.000, 1.025
	Pop Density (People/Sq. Mi.)	0.103	68,865.339	244.584	1,691.822	1.001	1.001, 1.001
	Grassland (Percent)	0.000	97.575	25.039	18.141	1.027	1.024, 1.031
	Primary PC (Percent)	(2.617)	2.690	0.000	1.000	0.731	0.683, 0.782
	Secondary PC (Percent)	(0.846)	12.917	0.000	1.000	1.162	1.095, 1.234
2006	**WNV Occurrence** (Cases)	0.000	251.000	1.322	8.007	-	-
	Ave. Temperature (°C)	(1.560)	24.190	12.986	4.492	0.933	0.924, 0.943
	Pop Density (People/Sq. Mi.)	0.111	69,066.565	245.823	1,693.137	1.001	1.000, 1.001
	Grassland (Percent)	0.000	97.575	25.039	18.141	1.029	1.025, 1.032
	Primary PC (Percent)	(2.617)	2.690	0.000	1.000	0.738	0.696, 0.783
	Secondary PC (Percent)	(0.846)	12.917	0.000	1.000	1.343	1.269, 1.421

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2007	**WNV Occurrence** (Cases)	0.000	140.000	1.140	5.578	-	-
	Ave. Temperature (°C)	(1.220)	24.550	12.639	4.602	0.923	0.914, 0.933
	Pop Density (People/Sq. Mi.)	0.117	69,207.965	247.488	1,698.068	1.000	1.000, 1.001
	Grassland (Percent)	0.000	97.575	25.039	18.141	1.040	1.036, 1.044
	Primary PC (Percent)	(2.617)	2.690	0.000	1.000	1.028	0.959, 1.101
	Secondary PC (Percent)	(0.846)	12.917	0.000	1.000	1.130	1.064, 1.201
2008	**WNV Occurrence** (Cases)	0.000	156.000	0.425	3.975	-	-
	Ave. Temperature (°C)	(2.190)	24.230	11.801	4.871	1.048	1.033, 1.063
	Pop Density (People/Sq. Mi.)	0.090	69,453.917	249.562	1,707.784	1.001	1.001, 1.001
	Grassland (Percent)	0.000	98.300	21.045	21.004	1.033	1.030, 1.037
	Primary PC (Percent)	(2.507)	2.824	0.000	1.000	1.044	0.945, 1.155
	Secondary PC (Percent)	(0.797)	13.218	0.000	1.000	1.062	0.988, 1.142
2009	**WNV Occurrence** (Cases)	0.000	27.000	0.228	1.413	-	-
	Ave. Temperature (°C)	(2.000)	24.290	11.810	4.873	1.026	1.009, 1.044
	Pop Density (People/Sq. Mi.)	0.114	69,296.761	251.805	1,715.520	1.000	1.000, 1.000
	Grassland (Percent)	0.000	98.300	21.045	21.004	1.029	1.024, 1.033
	Primary PC (Percent)	(2.507)	2.824	0.000	1.000	0.893	0.803, 0.994
	Secondary PC (Percent)	(0.797)	13.218	0.000	1.000	1.062	0.965, 1.169
2010	**WNV Occurrence** (Cases)	0.000	115.000	0.319	2.763	-	-
	Ave. Temperature (°C)	(1.510)	23.420	12.283	4.423	1.037	1.019, 1.055
	Pop Density (People/Sq. Mi.)	0.000	69,518.337	253.665	1,726.605	1.001	1.001, 1.001
	Grassland (Percent)	0.000	98.300	21.045	21.004	1.030	1.026, 1.034
	Primary PC (Percent)	(2.507)	2.824	0.000	1.000	0.886	0.796, 0.986
	Secondary PC (Percent)	(0.797)	13.218	0.000	1.000	1.161	1.077, 1.253

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2011	**WNV Occurrence** (Cases)	0.000	58.000	0.227	1.772	-	-
	Ave. Temperature (°C)	(2.030)	24.570	12.536	4.912	1.062	1.043, 1.081
	Pop Density (People/Sq. Mi.)	0.000	70,460.700	255.940	1,748.915	1.001	1.001, 1.001
	Grassland (Percent)	0.000	98.300	21.045	21.004	1.022	1.018, 1.027
	Primary PC (Percent)	(2.507)	2.824	0.000	1.000	1.255	1.103, 1.428
	Secondary PC (Percent)	(0.797)	13.218	0.000	1.000	1.033	0.951, 1.122
2012	**WNV Occurrence** (Cases)	0.000	396.000	1.760	11.480	-	-
	Ave. Temperature (°C)	(0.280)	24.720	13.590	4.467	1.092	1.079, 1.104
	Pop Density (People/Sq. Mi.)	0.000	71,124.639	258.099	1,767.009	1.000	1.000, 1.000
	Grassland (Percent)	0.000	99.943	21.115	30.262	1.011	1.010, 1.013
	Primary PC (Percent)	(1.743)	2.371	0.000	1.000	0.725	0.683, 0.770
	Secondary PC (Percent)	(1.122)	13.019	0.000	1.000	2.301	2.117, 2.501
2013	**WNV Occurrence** (Cases)	0.000	158.000	0.788	4.290	-	-
	Ave. Temperature (°C)	(1.650)	24.780	11.802	4.764	0.932	0.921, 0.944
	Pop Density (People/Sq. Mi.)	0.000	71,422.538	260.097	1,779.854	1.000	1.000, 1.000
	Grassland (Percent)	0.000	99.943	21.115	30.262	1.024	1.022, 1.026
	Primary PC (Percent)	(1.743)	2.371	0.000	1.000	0.559	0.519, 0.602
	Secondary PC (Percent)	(1.122)	13.019	0.000	1.000	1.740	1.583, 1.912
2014	**WNV Occurrence** (Cases)	0.000	263.000	0.709	7.484	-	-
	Ave. Temperature (°C)	(1.440)	24.660	11.463	4.800	1.068	1.054, 1.083
	Pop Density (People/Sq. Mi.)	0.000	71,609.103	261.666	1,788.746	1.000	1.000, 1.000
	Grassland (Percent)	0.000	99.943	21.115	30.262	1.024	1.022, 1.027
	Primary PC (Percent)	(1.743)	2.371	0.000	1.000	0.563	0.514, 0.616
	Secondary PC (Percent)	(1.122)	13.019	0.000	1.000	1.779	1.603, 1.975

Table 28. Negative binomial regression output table based on the patch density landscape metric for RQ2, WNV in the contiguous United States during 2003 through 2014.

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2003	**WNV Occurrence** (Cases)	0.000	546.000	3.140	18.395	-	-
	Ave. Temperature (°C)	(1.390)	24.560	12.082	4.586	0.887	0.878, 0.895
	Pop Density (People/Sq. Mi.)	0.106	68,365.602	242.297	1,690.168	1.000	1.000, 1.000
	Grassland (Percent)	0.000	95.735	48.984	15.533	0.968	0.965, 0.971
	Primary PC (Percent)	(1.555)	6.253	0.000	1.000	0.701	0.661, 0.744
2004	**WNV Occurrence** (Cases)	0.000	355.000	0.804	9.885	-	-
	Ave. Temperature (°C)	(2.120)	24.130	12.290	4.658	1.089	1.076, 1.102
	Pop Density (People/Sq. Mi.)	0.081	68,706.652	243.446	1,691.856	1.001	1.001, 1.001
	Grassland (Percent)	0.000	95.735	48.984	15.533	0.949	0.944, 0.953
	Primary PC (Percent)	(1.555)	6.253	0.000	1.000	0.616	0.597, 0.669
2005	**WNV Occurrence** (Cases)	0.000	163.000	0.956	5.794	-	-
	Ave. Temperature (°C)	(1.450)	24.110	12.616	4.519	1.019	1.007, 1.030
	Pop Density (People/Sq. Mi.)	0.103	68,865.339	244.566	1,691.283	1.001	1.001, 1.001
	Grassland (Percent)	0.000	95.735	48.984	15.533	0.969	0.965, 0.973
	Primary PC (Percent)	(1.555)	6.253	0.000	1.000	0.661	0.615, 0.710
2006	**WNV Occurrence** (Cases)	0.000	251.000	1.346	8.115	-	-
	Ave. Temperature (°C)	(1.560)	24.190	12.990	4.493	0.948	0.939, 0.958
	Pop Density (People/Sq. Mi.)	0.111	69,066.565	245.808	1,692.597	1.000	1.000, 1.001
	Grassland (Percent)	0.000	95.735	48.984	15.533	0.965	0.962, 0.969
	Primary PC (Percent)	(1.555)	6.253	0.000	1.000	0.702	0.656, 0.750
2007	**WNV Occurrence** (Cases)	0.000	140.000	1.166	5.708	-	-
	Ave. Temperature (°C)	(1.220)	24.550	12.644	4.604	0.940	0.930, 0.949
	Pop Density (People/Sq. Mi.)	0.117	69,207.965	247.475	1,697.527	1.000	1.000, 1.001
	Grassland (Percent)	0.000	95.735	48.984	15.533	0.960	0.956, 0.964
	Primary PC (Percent)	(1.555)	6.253	0.000	1.000	0.626	0.581, 0.675

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2008	**WNV Occurrence** (Cases)	0.000	156.000	0.426	3.974	-	-
	Ave. Temperature (°C)	(2.190)	24.230	11.803	4.871	1.074	1.059, 1.089
	Pop Density (People/Sq. Mi.)	0.090	69,453.917	249.498	1,707.513	1.001	1.001, 1.001
	Grassland (Percent)	0.000	96.679	44.462	15.387	0.976	0.970, 0.981
	Primary PC (Percent)	(1.593)	6.477	0.000	1.000	0.690	0.635, 0.751
	Secondary PC (Percent)	(3.034)	4.321	0.000	1.000	1.371	1.269, 1.483
2009	**WNV Occurrence** (Cases)	0.000	27.000	0.228	1.423	-	-
	Ave. Temperature (°C)	(2.000)	24.290	11.812	4.873	1.044	1.027, 1.062
	Pop Density (People/Sq. Mi.)	0.114	69,296.761	251.741	1,715.247	1.000	1.000, 1.000
	Grassland (Percent)	0.000	96.679	44.462	15.387	0.987	0.980, 0.993
	Primary PC (Percent)	(1.593)	6.477	0.000	1.000	0.776	0.696, 0.864
	Secondary PC (Percent)	(3.034)	4.321	0.000	1.000	1.742	1.595, 1.903
2010	**WNV Occurrence** (Cases)	0.000	115.000	0.322	2.766	-	-
	Ave. Temperature (°C)	(1.510)	23.420	12.284	4.422	1.060	1.042, 1.078
	Pop Density (People/Sq. Mi.)	0.000	69,518.337	253.601	1,726.330	1.001	1.000, 1.001
	Grassland (Percent)	0.000	96.679	44.462	15.387	0.971	0.965, 0.977
	Primary PC (Percent)	(1.593)	6.477	0.000	1.000	0.910	0.844, 0.980
	Secondary PC (Percent)	(3.034)	4.321	0.000	1.000	1.575	1.447, 1.714
2011	**WNV Occurrence** (Cases)	0.000	58.000	0.228	1.772	-	-
	Ave. Temperature (°C)	(2.030)	24.570	12.537	4.911	1.071	1.052, 1.090
	Pop Density (People/Sq. Mi.)	0.000	70,460.700	255.875	1,748.637	1.001	1.000, 1.001
	Grassland (Percent)	0.000	96.679	44.462	15.387	0.986	0.980, 0.993
	Primary PC (Percent)	(1.593)	6.477	0.000	1.000	1.042	0.965, 1.125
	Secondary PC (Percent)	(3.034)	4.321	0.000	1.000	1.174	1.064, 1.295

Year	Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
2012	**WNV Occurrence.** (Cases)	0.000	396.000	1.760	11.480	-	-
	Ave. Temperature (°C)	(0.280)	24.720	13.590	4.467	1.101	1.089, 1.113
	Pop Density (People/Sq. Mi.)	0.000	71,124.639	258.099	1,767.009	1.001	1.001, 1.001
	Grassland (Percent)	0.000	98.214	22.987	22.210	1.022	1.019, 1.024
	Primary PC (Percent)	(2.487)	3.194	0.000	1.000	0.992	0.935, 1.053
	Secondary PC (Percent)	(2.353)	14.340	0.000	1.000	1.066	1.001, 1.134
2013	**WNV Occurrence** (Cases)	0.000	158.000	0.788	4.290	-	-
	Ave. Temperature (°C)	(1.650)	24.780	11.802	4.764	0.952	0.942, 0.963
	Pop Density (People/Sq. Mi.)	0.000	71,422.538	260.097	1,779.854	1.001	1.001, 1.001
	Grassland (Percent)	0.000	98.214	22.987	22.210	1.022	1.019, 1.025
	Primary PC (Percent)	(2.487)	3.194	0.000	1.000	0.764	0.717, 0.814
	Secondary PC (Percent)	(2.353)	14.340	0.000	1.000	1.225	1.141, 1.315
2014	**WNV Occurrence** (Cases)	0.000	263.000	0.709	7.484	-	-
	Ave. Temperature (°C)	(1.440)	24.660	11.463	4.800	1.085	1.071, 1.098
	Pop Density (People/Sq. Mi.)	0.000	71,609.103	261.666	1,788.746	1.001	1.001, 1.001
	Grassland (Percent)	0.000	98.214	22.987	22.210	1.029	1.026, 1.032
	Primary PC (Percent)	(2.487)	3.194	0.000	1.000	0.825	0.762, 0.894
	Secondary PC (Percent)	(2.353)	14.340	0.000	1.000	1.307	1.200, 1.423

9 APPENDIX III

Table 29, Table 30, and Table 31 present additional negative binomial outputs for each landscape metric in the examination of Zika in Colombia during 2016.

Table 29. Negative binomial regression table based on the linear density landscape metric.

Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
Zika Occurrence (Cases)	0.000	72,900.000	447.788	2,982.191	-	-
Population Density (People/Sq. Mi.)	0.234	31,754.989	402.433	1,715.979	1.000	1.000, 1.000
Per Capita GDP (Annual Income/Person)	984.425	889,550.410	12,033.233	32,643.741	1.000	1.000, 1.000
Average Elevation (Meters)	2.990	6,167.250	1,326.178	1,006.648	0.999	0.999, 0.999
Cropland (Mi/Sq. Mi)	0.000	1.523	0.334	0.311	1.083	1.071, 1.095
Grassland (Mi/Sq. Mi)	0.000	1.419	0.398	0.292	1.553	1.534, 1.571
Tree-Covered (Mi/Sq. Mi)	0.000	1.835	0.489	2.750	2.481	2.442, 2.521

Table 30. Negative binomial regression table based on the proportion abundance landscape metric.

Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
Zika Occurrence (Cases)	0.000	72,900.000	447.788	2,982.191	-	-
Population Density (People/Sq. Mi.)	0.234	31,754.989	402.433	1,715.979	1.001	1.001, 1.002
Per Capita GDP (Annual Income/Person)	984.425	889,550.410	12,033.233	32,643.741	1.000	1.000, 1.000
Average Elevation (Meters)	2.990	6,167.250	1,326.178	1,006.648	0.999	0.999, 0.999
Cropland (Percent)	0.000	93.553	12.735	17.719	1.040	1.030, 1.049
Grassland (Percent)	0.000	100.000	21.356	23.994	1.066	1.055, 1.076
Tree-Covered (Percent)	0.000	100.00	63.614	29.698	1.054	1.045, 1.063

Table 31. Negative binomial regression table based on the patch density landscape metric.

Variable	Minimum	Maximum	Mean	Standard Deviation	Exp(B)	95% Confidence Interval
Zika Occurrence (Cases)	0.000	72,900.000	447.788	2,982.191	-	-
Population Density (People/Sq. Mi.)	0.234	31,754.989	402.433	1,715.979	1.000	1.000, 1.000
Per Capita GDP (Annual Income/Person)	0.815	159,753.000	582.509	5,126.290	1.000	1.000, 1.000
Average Elevation (Meters)	2.990	6,167.250	1,326.178	1,006.648	0.999	0.999, 0.999
Cropland (Percent)	0.000	88.889	30.043	19.441	1.026	1.025, 1.026
Grassland (Percent)	0.000	100.000	38.223	20.086	1.029	1.029, 1.029
Tree-Covered (Percent)	0.000	100.000	22.986	18.395	1.023	1.022, 1.023

REFERENCES

Afrane, Y. A.; Lawson, B. W.; Githeko, A. K.; Yan, G. (2005): Effects of microclimatic changes caused by land use and land cover on duration of gonotrophic cycles of Anopheles gambiae (Diptera: Culicidae) in western Kenya highlands. *Journal of medical entomology*, *42*(6), 974-980.

Allan, Brian F.; Keesing, Felicia; Ostfeld, Richard S. (2003): Effect of Forest Fragmentation on Lyme Disease Risk. In *Conservation Biology* 17 (1), pp. 267–272.

Altizer, Sonia; Dobson, Andrew; Hosseini, Parviez; Hudson, Peter; Pascual, Mercedes; Rohani, Pejman (2006): Seasonality and the dynamics of infectious diseases. In *Ecology Letters* 9 (4), pp. 467–484. DOI: 10.1111/j.1461-0248.2005.00879.x.

Arinaminpathy, N.; McLean, A. R.; Godfray, H.C.J. (2009): Future UK land use policy and the risk of infectious disease in humans, livestock and wild animals. In *Land Use Policy* 26, pp. S124. DOI: 10.1016/j.landusepol.2009.08.026.

Arizona Department of Health Services (2013): Border Infectious Disease Surveillance. Retrieved from http://www.azdhs.gov/diro/borderhealth/bids/vector-borne.htm

Armenteras, D., Rodríguez, N., Retana, J., & Morales, M. (2011): Understanding deforestation in montane and lowland forests of the Colombian Andes. *Regional Environmental Change*, *11*(3), 693-705.

Arsenault, Julie; Michel, Pascal; Berke, Olaf; Ravel, André; Gosselin, Pierre (2013): How to choose geographical units in ecological studies. Proposal and application to campylobacteriosis. In *Spatial and Spatio-temporal Epidemiology* 7, pp. 11–24. DOI: 10.1016/j.sste.2013.04.004.

Auchincloss, Amy H.; Gebreab, Samson Y.; Mair, Christina; Diez Roux, Ana V. (2012): A Review of Spatial Methods in Epidemiology, 2000–2010. In *Annual Review of Public Health)* 33 (1), pp. 107–122. DOI: 10.1146/annurev-publhealth-031811-124655.

Beale, Linda; Abellan, Juan Jose; Hodgson, Susan; Jarup, Lars (2008): Methodologic Issues and Approaches to Spatial Epidemiology. In *Environmental Health Perspectives* 116 (8), pp. 1105–1110. DOI: 10.1289/ehp.10816.

Bhatt, Samir; Gething, Peter W.; Brady, Oliver J.; Messina, Jane P.; Farlow, Andrew W.; Moyes, Catherine L. et al. (2013): The global distribution and burden of dengue. In *Nature* 496 (7446), pp. 504–507. DOI: 10.1038/nature12060.

Bogich, Tiffany L.; Chumara, Rumi; Scales, David; Chan, Emily; Pinhero, Laura C.; Chmura, Aleksei A. et al. (2012): Preventing Pandemics Via International Development. A Systems Approach. In *PLOS Medicine* 9 (12), pp. 1–4. DOI: 10.1371/journal.pmed.1001354.

Bontemps, Sophie; Defourny, Pierre; van Bogaert, Eric; Arino, Olivier; Kalogirou, Vasileios; Perez, Jose Ramos (2011): GLOBCOVER 2009. Products Description and Validation Report. UCLouvain & ESA Team. Retrieved from http://due.esrin.esa.int/globcover/LandCover2009/GLOBCOVER2009_Validation_Report_2.2.pdf

Bravata, DM; McDonald, KM; Smith, WM; Rydzak, C; Szeto, H; Buckeridge, DL, et al. (2004): Systematic Review: Surveillance Systems for Early Detection of Bioterrorism-Related Diseases. In *Annals of Intern Medicine* 140, pp. 910-922. DOI: 10.7326/0003-4819-140-11-200406010-00013.

Brown, Carolyn. "Zika Virus Outbreaks in Asia and South America." *CMAJ : Canadian Medical Association Journal* 188.2 (2016): E34. *PMC*. Web. 12 June 2017.

Buyantuyev, Alexander; Wu, Jianguo (2007): Effects of thematic resolution on landscape pattern analysis. In *Landscape Ecology* 22 (1), pp. 7–13. DOI: 10.1007/s10980-006-9010-5.

Calvo-Cano, A.; Aldasoro, E.; Ramiraz, M. F.; Martinez, M. J.; Requena-Mendez, A.; Gascon, J. (2014): Two Cases of Laboratory-Confirmed Leptospirosis in Travelers Returning to Spain from Thailand, September 2013. In *Eurosurveillance* 19 (2), pp. 1–3. DOI: 10.4269/ajtmh.12-0551.

CDC (2009): *Anthrax*. Retrieved from http://www.cdc.gov/nczved/divisions/dfbmd/diseases/anthrax/technical.html

CDC (2017): Zika cases in the United States. Retrieved from https://www.cdc.gov/zika/reporting/case-counts.html

CDC Preliminary (2017): Preliminary Maps & Data for 2017. Retrieved from https://www.cdc.gov/westnile/statsmaps/preliminarymapsdata2017/index.html

CDC Zika (2016): Zika. Retrieved from https://www.cdc.gov/zika/

CDC MMWR (2014): *West Nile Virus and Other Arboviral Diseases – United States, 2013.* Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6324a1.htm

Ceccato, P.; Conner, S. J.; Jeanne, I.; Thomson, M. C. (2005): Application of Geographical Information Systems and Remote Sensing Technologies for Assessing and Monitoring Malaria Risk. In *Parassitologia* 47, pp. 81–96.

Census (2012): Methodology for the Intercensal Population and Housing Unit Estimates: 2000 to 2010. Retrieved from http://www.census.gov/popest/methodology/2000-2010_Intercensal_Estimates_Methodology.pdf

CityPopulation (2017): *Colombia*. Retrieved from https://www.citypopulation.de/php/colombia-admin.php

DANE (2017): *Colombian GDP*: Retrieved from http://datlascolombia.com/#/downloads?locale=en-col

de Castro, M. C., Monte-Mór, R. L., Sawyer, D. O., & Singer, B. H. (2006): Malaria risk on the Amazon frontier. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(7), 2452-2457.

De La Rocque, S.; Michel, V.; Plazanet, D.; Pin, R. (2004): Remote sensing and epidemiology. Examples of applications for two vector-borne diseases. In *Comparative Immunology, Microbiology and Infectious Diseases* 27 (5), pp. 331–341. DOI: 10.1016/j.cimid.2004.03.003.

Department of Health (2017): List of Zika virus infected countries. Retrieved from http://www.health.gov.au/internet/main/publishing.nsf/content/ohp-zika-countries.htm

Eisenberg, Joseph N.S.; Desai, Manish A.; Levy, Karen; Bates, Sarah J.; Liang, Song; Naumoff, Kyra; Scott, James C. (2007): Environmental Determinants of Infectious Disease: A Framework for Tracking Causal Links and Guiding Public Health Research. In *Environ Health Perspect* 115 (8), pp. 1216–1223. DOI: 10.1289/ehp.9806.

EPA (2004): Wetlands & West Nile Virus. Retrieved from http://www.epa.gov/owow/wetlands/pdf/WestNile.pdf

ESRI (2017): El Nino Southern Oscillation (ENSO). Retrieved from https://www.esrl.noaa.gov/psd/enso/past_events.html

European Space Agency (2006): GlobCover 2005/2006. Retrieved from http://due.esrin.esa.int/page_globcover.php

European Space Agency (2010): GlobCover 2009. Retrieved from http://due.esrin.esa.int/page_globcover.php

Ezenwa, Vanessa O.; Milheim, Lesley E.; Coffey, Michelle F.; Godsey, Marvin S.; King, Raymond J.; Guptill, Stephen C. (2007): Land Cover Variation and West Nile Virus Prevalence. : Patterns, Processes, and Implications for Disease Control. In *Vector-Borne and Zoonotic Diseases* 7 (2), pp. 173–180. DOI: 10.1089/vbz.2006.0584.

Fagan, William F.; Cantrell, Robert Stephen; Cosner, Chris (1999): How Habitat Edges Change Species Interactions. In *The American Naturalist* 153 (2), pp. 162–182.

Feng, Z.; Li, W.; Varma, J. K. (2011): Gaps Remain In China's Ability To Detect Emerging Infectious Diseases Despite Advances Since The Onset Of SARS And Avian Flu. In *Health Affairs* 30 (1), pp. 127–135. DOI: 10.1377/hlthaff.2010.0606.

Finch C, Al-Damluji MS, Krause PJ, Niccolai L, Steeves T, O'Keefe CF, et al. (2014): Integrated Assessment of Behavioral and Environmental Risk Factors for Lyme Disease Infection on Block Island, Rhode Island. PLoS ONE 9(1): e84758. doi:10.1371/journal.pone.0084758.

Foley, J. A. (2005): Global Consequences of Land Use. In *Science* 309 (5734), pp. 570–574. DOI: 10.1126/science.1111772.

FRA (2000): Forest Cover Mapping and Monitoring with NOAA-AVHRR and other Course Spatial Resolution Sensors. Food and Agriculture Organization of the United Nations (Forest Resources Assessment (FRA) Working Paper, 29). Retrieved from http://www.fao.org/docrep/007/ae161e/ae161e00.htm

Gaines, David N. (2014): Mosquito Breeding Habitats. Virginia Department of Health. Retrieved from http://www.vdh.virginia.gov/lhd/CentralShenandoah/EH/WNV/mosquito_breeding_habitats.htm

GAUL (2014). Global Administrative Unit Layers (GAUL). FAO GEONETWORK. Rome, Italy: FAO, created March 1, 2017. Retrieved from http://www.fao.org/geonetwork/srv/en/metadata.show?id=12691

Global Administrative Areas (2015). GADM database of Global Administrative Areas, version 2.8. [online] URL: www.gadm.org.

Global Forest Watch (2014): Data Sources. World Resources Institute. Retrieved from http://www.globalforestwatch.org/sources/forest_change

Grace, D.; Mutua, F.; Ochungo, P.; Kruska, R.; Jones, K.; Brierley, L. et al. (2012): Mapping of poverty and likely zoonoses hotspots. 2nd ed. Retrieved from http://hdl.handle.net/10568/21161

Hay, S. I.; Snow, R. W.; Rogers, D. J. (1998): From Predicting Mosquito Habitat to Malaria Seasons Using Remotely Sensed Data. Practice, Problems and Perspectives. In *Parasitology Today* 14 (8), pp. 306–313. DOI: 10.1016/S0169-4758(98)01285-X.

HHS (2011): The Global Health Strategy of the U.S. Department of Health and Human Services. Retrieved from http://www.globalhealth.gov/pdfs/Global%20Health%20Strategy.pdf

Jones, Kate E.; Patel, Nikkita G.; Levy, Marc A.; Storeygard, Adam; Balk, Deborah; Gittleman, John L.; Daszak, Peter (2008): Global trends in emerging infectious diseases. In *Nature* 451 (7181), pp. 990–993. DOI: 10.1038/nature06536.

Kalluri, Satya; Gilruth, Peter; Rogers, David; Szczur, Martha (2007): Surveillance of Arthropod Vector-Borne Infectious Diseases Using Remote Sensing Techniques. A Review. In *PLoS Pathogens* 3 (10), pp. e116. DOI: 10.1371/journal.ppat.0030116.

Karesh, William B.; Dobson, Andy; Lloyd-Smith, James O.; Lubroth, Juan; Dixon, Matthew A.; Bennett, Malcolm et al. (2012): Ecology of zoonoses. Natural and unnatural histories. In *The Lancet* 380 (9857), pp. 1936–1945. DOI: 10.1016/S0140-6736(12)61678-X.

Karlamangla, Souyma (2017): Death toll from West Nile climbs to 7 in L.A. County. In *Los Angeles Times*. Retrieved from http://www.latimes.com/local/california/la-me-ln-west-nile-deaths-20170923-story.html

Kilpatrick, A. Marm (2011): Globalization, Land Use, and the Invasion of West Nile Virus. In *Science* 334 (6054), pp. 323–327. DOI: 10.1126/science.1201010.

King, Lonnie J. (Ed.) (2004): Emerging and Re-Emerging Zoonotic Diseases. Challenges and Opportunities. 72nd Annual General Session of the International Committee of the Office International des Epizooties (OIE). Paris, France, May 23-28. World Organization for Animal Health. Retrieved from http://www.oie.int/doc/ged/D695.PDF

Lambin, Eric F.; Geist, Helmut J.; Lepers, Erika (2003): Dynamics of Land-Use and Land-Cover Change in Tropical Regions. In *Annual Review of Environmental Resources.* 28 (1), pp. 205–241. DOI: 10.1146/annurev.energy.28.050302.105459.

Lambin, Eric F.; Tran, Annelise; Vanwambeke, Sophie O.; Linard, Catherine; Soti, Valérie (2010): Pathogenic landscapes: Interactions between land, people, disease vectors, and their animal hosts. In *International Journal of Health Geographics* 9 (1), p. 54. DOI: 10.1186/1476-072X-9-54.

Latham, John; Cumani, Renato; Rosati, Ilaria; Bloise, Mario (2014): Global Land Cover SHARE. Food and Agriculture Organization of the United Nations. Retrieved from http://www.glcn.org/downs/prj/glcshare/GLC_SHARE_beta_v1.0_2014.pdf

Leyk, Stefan; Boesch, Ruedi; Weibel, Robert (2005): A Conceptual Framework for Uncertainty Investigation in Map-based Land Cover Change Modelling. In *Transactions in GIS* 9 (3), pp. 291–322. DOI: 10.1111/j.1467-9671.2005.00220.x.

MacDougall, Clair; Farge, Emma (2014): Ebola sucks life from W. African states weakened by war, poverty. MONROVIA/DAKAR. Reuters. Retrieved from http://www.reuters.com/article/2014/08/06/health-ebola-economies-idUSL6N0QC45F20140806, checked on 9/16/2014

Marcondes, C. B., & Ximenes, M. D. F. F. D. (2016). Zika virus in Brazil and the danger of infestation by Aedes (Stegomyia) mosquitoes. *Revista da Sociedade Brasileira de Medicina Tropical*, *49*(1), 4-10.

McClintock, Brett T.; Nichols, James D.; Bailey, Larissa L.; MacKenzie, Darryl I.; Kendall, William. L.; Franklin, Alan B. (2010): Seeking a second opinion: uncertainty in disease ecology. In *Ecology Letters* 13 (6), pp. 659–674. DOI: 10.1111/j.1461-0248.2010.01472.x.

McFarlane, Rosemary; Sleigh, Adrian; McMichael, Anthony (2013): Land-Use Change and Emerging Infectious Disease on an Island Continent. In *International Journal of Environmental Research and Public Health* 10 (7), pp. 2699–2719. DOI: 10.3390/ijerph10072699.

McGarigal, K. (2017). Landscape metrics for categorical map Patterns. Retrieved from http://www.umass.edu/landeco/teaching/landscape_ecology/schedule/chapter9_metrics.pdf

McMichael, A. J. (2004): Environmental and social influences on emerging infectious diseases: past, present and future. In *Philosophical Transactions of the Royal Society B: Biological Sciences* 359 (1447), pp. 1049–1058. DOI: 10.1098/rstb.2004.1480.

Morse, Stephen S. (2004): Factors and determinants of disease emergence. In *Revue scientifique et technique* 23 (2), pp. 443–451. Retrieved from http://www.researchgate.net/publication/8029656_Factors_and_determinants_of_disease_emergence/file/60b7d515c2163c37e0.pdf

Morse, Stephen S. (2007): Global Infectious Disease Surveillance And Health Intelligence. In *Health Affairs* 26 (4), pp. 1069–1077. DOI: 10.1377/hlthaff.26.4.1069.

Morse, Stephen S. (2012): Public Health Surveillance and Infectious Disease Detection. In *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science* 10 (1), pp. 6–16. DOI: 10.1089/bsp.2011.0088.

Morse, Stephen S.; Mazet, Jonna A. K.; Woolhouse, Mark; Parrish, Colin R.; Carroll, Dennis; Karesh, William B. et al. (2012): Prediction and prevention of the next pandemic zoonosis. In *The Lancet* 380 (9857), pp. 1956–1965. DOI: 10.1016/S0140-6736(12)61684-5.

Morvan, J., Nakoune, E., Deubel, V., & Colyn, M. (2000): Ebola virus and forest ecosystem. *Bulletin de la Société de Pathologie Exotique*, *93*(3), 172-175.

Munga, Stephen; Minakawa, Noboru; Zhou, Guofa; Mushinzimana, Emmanuel; Barrack, Okeyo-Owuor J.; Githeko, Andrew K.; Yan, Guiyun (2006): Association Between Land Cover and Habitat Productivity of Malaria Vectors in Western Kenyan Highlands. In *American Journal of Tropical Medicine and Hygiene* 74 (1), pp. 69–75.

Murray, Kris A.; Daszak, Peter (2013): Human ecology in pathogenic landscapes. Two hypotheses on how land use change drives viral emergence. In *Current Opinion in Virology* 3 (1), pp. 79–83. DOI: 10.1016/j.coviro.2013.01.006.

NIAID (2014): *Biodefense and Emerging Infectious Diseases.* Retrieved from http://www.niaid.nih.gov/topics/BiodefenseRelated/Biodefense/Pages/CatA.aspx

NIH (2015): Biodefense and Emerging Infectious Diseases. Retrieved from http://www.niaid.nih.gov/topics/BiodefenseRelated/Biodefense/Pages/CatA.aspx

NIH News (2009): NIAID Renews Funding for National Emerging Infectious Diseases Research Network. Retrieved from http://www.nih.gov/news/health/jun2009/niaid-08.htm

Null, Jan (2015): El Niño and La Niña Years and Intensities; Based on Oceanic Niño Index (ONI). Retrieved from http://ggweather.com/enso/oni.htm

Nuzzo, Jennifer B.; Gronvall, Gigi Kwik (2011): Global Health Security. Closing the Gaps in Responding to Infectious Disease Emergencies. In *Global Health Governance* IV (2), pp. 1–15.

PAHO WHO (2017): Regional Zika Epidemiological Update (Americas) May 25, 2017. Retrieved from http://www.paho.org/hq/index.php?option=com_content&id=11599&Itemid=41691

Patz, Jonathan A.; Daszak, Peter; Tabor, Gary M.; Aguirre, A. Alonso; Pearl, Mary; Epstein, Jon et al. (2004): Unhealthy Landscapes: Policy Recommendations on Land Use Change and Infectious Disease Emergence. In *Environ Health Perspect* 112 (10), pp. 1092–1098. DOI: 10.1289/ehp.6877.

Patz, Jonathan A.; Olson, Sarah H.; Uejio, Christopher K.; Gibbs, Holly K. (2008): Disease Emergence from Global Climate and Land Use Change. In *Medical Clinics of North America* 92 (6), pp. 1473–1491. DOI: 10.1016/j.mcna.2008.07.007.

Phalkey, Revati K.; Shulka, Sharvari; Shardul, Savita; Ashtekar, Nutan; Awate, Pradip (2013): Assessment of the core and support functions of the Integrated Disease Surveillance system in Maharashtra, India. In *BMC Public Health* 13. DOI: 10.1186/1471-2458-13-575.

Pike, Brian L.; Saylors, Karen E.; Fair, Joseph N.; LeBreton, Matthew; Tamoufe, Ubald; Djoko, Cyrille F. et al. (2010): The Origin and Prevention of Pandemics. In *Clinical Infectious Diseases* 50 (12), pp. 1636–1640. DOI: 10.1086/652860.

PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created March 1, 2017.

Ratmanov, P.; Mediannikov, O.; Raoult, D. (2013): Vectorborne diseases in West Africa. Geographic distribution and geospatial characteristics. In *Transactions of the Royal Society of Tropical Medicine and Hygiene* 107 (5), pp. 273–284. DOI: 10.1093/trstmh/trt020.

Rodriguez, Dania M.; Johansson, Michael A.; Moiradillon2; Mier-y-Teran-Romero, Luis; YoJimboDurant; Mietchen, Daniel (2016). zika: September 7, 2016 [Colombia]. Zenodo. http://doi.org/10.5281/zenodo.61753. Retrieved from https://github.com/cdcepi/zika

Schielke, Anika; Rosner, Bettina M.; Stark, Klaus (2014): Epidemiology of campylobacteriosis in Germany. Insights from 10 years of surveillance. In *BMC Infectious Diseases* 14, pp. 1–8. DOI: 10.1186/1471-2334-14-30.

Schneider, A.; Hommel, G.; & Blettner, M. (2010): Linear Regression Analysis: Part 14 of a Series on Evaluation of Scientific Publications. *Deutsches Ärzteblatt International*, *107*(44), 776–782. doi:10.3238/arztebl.2010.0776

Schuurman, Nadine; Fielder, Robert S.; Grzybowski, Stephan CW; Grund, Darrin (2006): Defining rational hospital catchments for non-urban areas based on travel-time. *International Journal of Health Geographics*, *5*(1), 43.

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proceedings of the National Academy of Sciences, 109(40), 16083-16088.

Sutherst, R. W. (2004): Global Change and Human Vulnerability to Vector-Borne Diseases. In *Clinical Microbiology Reviews* 17 (1), pp. 136–173. DOI: 10.1128/CMR.17.1.136-173.2004.

Thacker, Stephen B.; Qualters, Judith R.; Lee, Lisa M. (2012): Public Health Surveillance in the United States. Evolution and Challenges. CDC's Vision for Public Health Surveillance in the 21st Century. In *Morbidity and Mortality Weekly Report* 61, pp. 3–9. Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/su6103a2.htm

Theobald, D. (2005): Landscape patterns of exurban growth in the USA from 1980 to 2020. *Ecology and Society*, *10*(1).

United Nations (2014): Millennium Development Goals and Beyond 2015. United Nations. Retrieved from http://www.un.org/millenniumgoals/aids.shtml

United States Census Bureau / Population Estimates Program. "Annual Estimates of the Resident Population for Counties" April 1, 2010 to July 1, 2014," http://www.census.gov/popest/data/counties/totals/2014/CO-EST2014-01.html, created March 1, 2017.

United States Census Bureau / Population Estimates Program. "Intercensal Estimates of the Resident Population for Counties: April 1, 2009 to July 1, 2010," http://www.census.gov/popest/data/intercensal/county/county2010.html, created March 1, 2017.

USAID (2014): PREDICT. United States Agency for International Development. Retrieved from http://avianflu.aed.org/eptprogram/

USGS (2014): West Nile Virus Provisional 2014 Data. Retrieved from http://diseasemaps.usgs.gov/mapviewer/

USGS WNV (2013): West Nile Virus Background. Retrieved from http://diseasemaps.usgs.gov/wnv_background.html

Vanwambeke, Sophie O.; Bennett, Shannon N.; Kapan, Durrell D. (2011): Spatially disaggregated disease transmission risk: land cover, land use and risk of dengue transmission on the island of Oahu. In *Tropical Medicine & International Health* 16 (2), pp. 174–185. DOI: 10.1111/j.1365-3156.2010.02671.x.

Vittor, A. Y., Gilman, R. H., Tielsch, J., Glass, G., Shields, T. I. M., Lozano, W. S., & Patz, J. A. (2006): The effect of deforestation on the human-biting rate of Anopheles darlingi, the primary vector of falciparum malaria in the Peruvian Amazon. *The American Journal of Tropical Medicine and Hygiene*, *74*(1), 3-11.

Walsh, J. F., Molyneux, D. H., & Birley, M. H. (1993): Deforestation: effects on vector-borne disease. *Parasitology*, *106*(S1), S55-S75.

Wang, Minhua; Howarth, Philip J. (1993): Modeling errors in remote sensing image classification. In *Remote Sensing of Environment* 45 (3), pp. 261–271. DOI: 10.1016/0034-4257(93)90109-B.

Wessels, K.J.; De Fries, R.S.; Dempewolf, J.; Anderson, L.O.; Hansen, A.J.; Powell, S.L.; Moran, E.F. (2004): Mapping regional land cover with MODIS data for biological conservation: Examples from the Greater Yellowstone Ecosystem, USA and Pará State, Brazil. In *Remote Sensing of Environment* 92 (1), pp. 67–83. DOI: 10.1016/j.rse.2004.05.002.

WHO (2013): Research priorities for the environment, agriculture and infectious diseases of poverty. *World Health Organization technical report series*, (976), i.

WHO (2016): Zika situation report 16 June 2016. Retrieved from http://www.who.int/emergencies/zika-virus/situation-report/16-june-2016/en/

WHO (2017): Emergencies – The History of Zika Virus. Retrieved from http://www.who.int/emergencies/zika-virus/history/en/

WHO Media Centre (2011): *West Nile Virus*. Retrieved from http://www.who.int/mediacentre/factsheets/fs354/en/

WHO Mosquito Control (2016): Mosquito Control. Retrieved from http://www.who.int/emergencies/zika-virus/articles/mosquito-control/en/

WHO Zika Virus (2016): Zika virus. Retrieved from http://www.who.int/mediacentre/factsheets/zika/en/

Wilcox, Bruce A,; Colwell, Rita R. (2005): Emerging and Reemerging Infectious Diseases: Biocomplexity as an Interdisciplinary Paradigm. In *EcoHealth* 2 (4), pp. 244–257. DOI: 10.1007/s10393-005-8961-3.

Wilcox, Bruce A.; Ellis, B. (2006): Forests and emerging infectious diseases of humans. In *Unasylva* 57 (224), pp. 11–18. Retrieved from http://www.hawaii.edu/publichealth/ecohealth/si/course-eids/readings/Wilcox_Ellis_2006.pdf

Wilcox, Bruce A.; Gubler, Duane J. (2005): Disease Ecology and the Global Emergence of Zoonotic Pathogens. In *Environmental Health and Preventative Medicine* 10, pp. 263–272.

Woolhouse, M. (2011): How to make predictions about future infectious disease risks. In *Philosophical Transactions of the Royal Society B: Biological Sciences* 366 (1573), pp. 2045–2054. DOI: 10.1098/rstb.2010.0387.

Woolhouse, M.; Scott, F.; Hudson, Z.; Howey, R.; Chase-Topping, M. (2012): Human viruses. Discovery and emergence. In *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1604), pp. 2864–2871. DOI: 10.1098/rstb.2011.0354.

World Bank. (2017): *From panic and neglect to investing in health security: financing pandemic preparedness at a national level*. Washington, D.C.: World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/979591495652724770/From-panic-and-neglect-to-investing-in-health-security-financing-pandemic-preparedness-at-a-national-level

Wu, Jiangou; Jenerette, G. Darrel; Buyantuyev, Alexander; Redman, Charles L. (2011): Quantifying spatiotemporal patterns of urbanization: The case of the two fastest growing metropolitan regions in the United States. In *Ecological Complexity* 8, pp. 1-8.

BIOGRAPHY

Josh Weinstein earned a B.S. in Earth Systems Sciences with a minor in Geology from George Mason University in 1998 and a M.S. in Geographic and Cartographic Sciences from George Mason University in 2003. He earned the Project Management Institute's Project Management Professional accreditation in 2006 and their Program Management Professional accreditation in 2016. He earned the Certified Mapping Scientist accreditation from the American Society for Photogrammetry and Remote Sensing in 2006. Josh is an alumnus of the National Outdoor Leadership School and served as President of the Lee Roadies Club of Toastmasters International. He served on the GMU Consortium for Northern Virginia Workforce Development of Geospatial Professionals and mentored high school students during multiple iterations of the NoVA Geospatial Leaders Workshop. Josh currently serves as President of the Bull Run Mountain Civic Association.

Josh Weinstein has worked alongside national security clients since 1997. He believes in building a strong team with a laser focus on solving his client's most difficult challenges, a tenet that earned him the SAIC ASPIRE award for Excellence in Program Performance. He earned the SAIC ASPIRE award for Excellence in Business Development for leading the Geospatial Technologies Operation's Emerging Markets Division. His efforts spearheading the creation and implementation of a formal career progression roadmap for geospatial professionals resulted in selection as the Operations representative to the SAIC Corporate IDIQ/GWAC Council and Business Unit representative to the SAIC Corporate Program Management Council. Josh currently serves as a Senior Program Manager in support of growth across the ABI/Multi-INT portfolio for Novetta's Information Exploitation Division.