2021

Spatializing Coupled Human and Natural System (CHANS)

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

SPATIALIZING COUPLED HUMAN AND NATURAL SYSTEM (CHANS)

by

YAXIONG MA

B.Sc., China University of Geosciences, 2013
M.A., Boston University, 2015

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2021

Approved by

First Reader _____

Sucharita Gopal, Ph.D.
Professor of Earth and Environment

Second Reader _____

Les Kaufman, Ph.D.
Professor of Biology

Third Reader _____

Nathan G. Phillips, Ph.D.
Professor of Earth and Environment

Fourth Reader _____

Kevin P. Gallagher, Ph.D.
Professor of Global Development Policy

ACKNOWLEDGMENTS

I want to extend my deepest gratitude to the people who have supported and encouraged me throughout the development of my dissertation. First and foremost, I want to thank my major advisor Suchi Gopal, who has guided me since I started my master's program at Boston University, for her motivation, patience, and guidance for my research. I appreciate the advice you have given me, both professionally and personally. It has been a pleasure and honor to work with you for all these years. Many thanks to my committee members: Les Kaufman, Nathan Phillips, Kevin Gallagher, and Magaly Koch. I feel fortunate for the opportunity to have worked with you during my doctoral program. I have learned so much from all of you. I also want to thank Lawrence Were and Julio Castrillon for the collaborations that extended my knowledge in health care and mathematics.

Many thanks to all the professors that I have learned from over my master's and doctoral programs. I would also like to thank my friends and fellow graduate students for their inspiration and support. Special thanks to Yingtong Zhang, Xiaojing Tang, Chi Chen, and Shixiong Wang. Our daily group lunch is very rewarding and the highlight of my days. Thanks to my lab mate, Jessica Wright, it's wonderful to work with you, and I wish you the best. Many thanks to my friends in CAS334 for have been there since the beginning: Ilyun Koh, Minkyu Moon, Taejin Park, Amani Al-Abri, and Arnold Fernandes.

Thanks to all the staff in Earth and Environment for your help. Thanks to Initiative on Cities (Katharine Lusk and Conor Leblanc) for helping to organize the

workshop. I want to like to thank members of the Global Development Policy Center (Kevin Gallagher, Rebecca Ray, Xinyue Ma, Hongbo Yang, and Blake Simmons) for your support and advice.

Last and most importantly, I want to thank my family, partner, and friends for their interest, support, and encouragement along my journey abroad. This degree would not be possible without you all.

SPATIALIZING COUPLED HUMAN AND NATURAL SYSTEM (CHANS)

YAXIONG MA

Boston University Graduate School of Arts and Sciences, 2021

Major Professor: Sucharita Gopal, Professor of Earth and Environment

ABSTRACT

Human sustainability is one of the most pressing issues of the 21st century. Coupled Human and Natural Systems (CHANS) offers a useful framework to focus on understanding the complex process and pattern that characterizes the dynamical interactions between human and natural systems. This dissertation research integrates the geospatial analysis into the CHANS framework from three perspectives: temporal, spatial, and organizational coupling.

Using the temporal coupling aspect, we monitor the risk of deforestation and biodiversity threats from energy investments in Southeast Asia. We assess the energy investment evaluate changes to forest morphology and the risk to biodiversity. In terms of land cover change, we find that hydroelectric power plants tend to have more extensive biodiversity impacts than coal-fired plants, which are usually built within proximity to major population centers.

Next, we explore spatial coupling by examining the spatial heterogeneity and homogeneity in home prices across Massachusetts, using Geographically Weighted Regression models with natural and socio-demographic variables. We discovered models

that utilized spatial heterogeneity perform better. However, statistical tests of significance are required to determine the model specification to avoid over-fitting.

In the fourth chapter, we examined a critical refugium for endangered fish species in East Africa by mapping the organizational dynamics of aquatic vegetation on Lake Kyoga, Uganda. A CHANS organizational coupling involving the natural infrastructure of aquatic vegetation and fishes can adversely impact endangered species and the surrounding human communities. Floating aquatic vegetation can protect the native fishes from predation by Nile Perch by creating hypoxic barriers between water bodies. We developed an algorithm to locate and identify various types of aquatic vegetation. Profiles of lakes are created to examine the spatiotemporal dynamics of refugia. The results are valuable in shaping strategies to conserve both fish species and human livelihoods.

The fifth chapter explores emerging technologies, Virtual Reality, in communicating the complex CHANS coupling of green (trees) and gray infrastructure (gas pipelines). This chapter demonstrates the building of 3D urban landscapes from remote sensing data and the emerging use of VR to communicate, educate and empower the stakeholders on sustainability issues related to aging natural gas infrastructure and resulting methane emissions.

This dissertation research aims to build a set of methodologies based on extensive geospatial data, spatially explicit models, and tools essential for operationalizing and monitoring CHANS in studies ranging from local to regional scales. Each application builds or revises a new model or algorithm to address a real-world CHANS problem.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 – Introduction

Human sustainability is one of the most pressing issues of the 21st century. Humans alter the ecological, hydrological, and thermal characteristics of their environment through deforestation, urbanization, transportation, modifications to the landform and vegetation cover, and many other infrastructure activities, resulting in an inextricable coupling of human and natural systems. Coupled Human and Natural Systems (CHANS) offers a useful framework to focus on understanding the complex process and pattern that characterizes the two-way dynamical interactions between human and natural systems. These are highly dependent on the spatial and temporal scale. This dissertation research integrates the geospatial analysis components into the CHANS framework from three perspectives: temporal, spatial, and organizational coupling.

Using the temporal coupling aspect, we monitor the risk of deforestation and biodiversity threats that play out over time from energy investments related to coal and hydroelectricity in Southeast Asia. We assess the energy investment in terms of 3 years prior and post-construction to evaluate the risk to biodiversity. We find that hydroelectric power plants (though much lower in carbon emissions than the burning of coal or other fossil fuels) tend to have more extensive biodiversity impacts than coal-fired plants. We used forest morphology metrics to examine the nature and magnitude of these impacts. We found that there are statistical differences in energy investments with versus without Chinese financing. China invests more in coal power plants in this region, where the power plants are usually built within proximity to major population centers. As such,

these power plants have a lower impact on total forest cover but must contend with a high ratio of forest edges and bridges (bridges connect fragments or edges with core forest).

Next, we explore spatial coupling by examining the spatial heterogeneity as well as spatial homogeneity in determining home prices across the state of Massachusetts, using Geographically Weighted Regression models. There are ten independent variables in this model, describing both natural and socio-demographic variables. We examined the impact of each determinant on the home value. For example, unprotected forest or available green space acts as a non-stationary determinant. It adds a positive value, resulting in an increase in home prices in the eastern part of the state, around suburban towns of the greater Boston area. In contrast, it adds a negative value in the western part of the state, resulting in decreased home prices.

In the third chapter, we examined a critical refugium for endangered fish species in East Africa by mapping the organizational dynamics of aquatic vegetation on Lake Kyoga, Uganda, and its associated satellite lakes. We did this using a long-time series of Landsat images. Societies that reside in the Lake Kyoga region have traditionally fished in and farmed beside the waters of this lake. A CHANS organizational coupling involving the natural infrastructure of aquatic vegetation and fishes can result in adverse impacts both to endangered species and to the surrounding human communities. Floating aquatic vegetation (e.g., papyrus, hippo grass, water hyacinth, water lettuce) can protect the native fishes from predation by Nile Perch by creating hypoxic barriers between water bodies difficult for the perch to cross. This work makes it easier to quantify and track shifts in floating vegetation over space and time. We developed an algorithm to

identify the location of aquatic vegetation as well as differentiate floating versus emergent vegetation. We also created profiles of satellite lakes that are refugia for endangered fish species to examine the spatiotemporal dynamics of these small lakes and changes over time in suitable fish habitats. The results are valuable in shaping strategies to conserve both fish species and human livelihoods in this region.

The fourth chapter explores emerging technologies, Virtual Reality, in communicating the complex CHANS coupling of green (trees) and gray infrastructure (gas pipelines). This chapter demonstrates the building of 3D urban landscapes from remote sensing data and the emerging use of VR to communicate, educate and empower the stakeholders on sustainability issues related to gas leaks. Thus, VR is bridging a broken gap in knowledge by enabling the stakeholders to see the invisible infrastructure and better understand the CHANS urban infrastructure.

This dissertation research aims to build a set of methodologies based on extensive geospatial data, spatially explicit models, and tools essential for operationalizing and monitoring CHANS in studies ranging from local to regional scales. Each application builds or revises a new model or algorithm to address a real-world CHANS problem. The purpose of each application is to find. sustainable solutions and provide decision and policy insights to stakeholders.

CHAPTER 2 – A Big Data Spatial Analytical Framework for Deforestation and Biodiversity Risks of Power Generation Projects in Southeast Asia

Ecosystem destruction and biodiversity loss due to human activities are widespread and now extremely rapid. They are rated among the top global risks both in terms of likelihood and in terms of impact. When considering development projects, thorough environmental impact evaluation requires good analysis of local ecological data, among other information besides social and economic impacts. Understanding deforestation and attendant biodiversity risks from existing projects is essential for conservation purposes and informing future project planning. Multiple geospatial data sources related to location, type, and the characteristics of energy investments and landcover data from various sources are used to integrate, analyze, and characterize the deforestation and biodiversity impact of energy investments in this study. Spatial morphological pattern analysis is also used on forest cover data to examine the relationships between forest core and edge effects on biodiversity. The analysis enables us to address conservation questions relevant to developmental finance in energy in different forest structures. We found that there are statistical differences in energy investments with and without Chinese development finance. The locational siting of investments and the capacity differences in energy production drive the spatial pattern in biodiversity risks.

2.1 Introduction

Southeast Asia is one of the regions with the world's highest species richness and endemism (Mittermeier et al. 1999, Duckworth et al. 2012). Its tropical ecosystems- freshwater, terrestrial, and marine- are characterized by rich biodiversity and complex biotic interactions among their species. Unfortunately, today this region exhibits the world's highest vertebrate extinction rates, primarily due to severe habitat loss (Sodhi et al. 2010). The countries in this region have seen deforestation rates that are unchecked and higher than in any other major tropical region globally, with forest cover has decreased anywhere from 20% to 70% in the past four decades. Southeast Asia is expected to lose three-quarters of its original forests by 2100 and up to 42% of its biodiversity (Sodhi et al. 2004, Estoque et al. 2019), a massive rate of species decline and extinction (Schipper et al. 2008). Human activities such as clear-cutting, illegal logging (Schipper et al. 2008), agriculture, and, most importantly, infrastructure projects, including energy (Drollette 2013), are the underlying drivers of this loss. In addition, there are many other drivers of biodiversity loss, including the harvest of wild species for luxury food, medicine, tonics, horns, and other trophy parts, and captive animals (Uhm 2016).

Prior studies have studied the culprits and mechanisms of deforestation and biodiversity loss, focusing on habitat fragmentation (Hughes 2018), increased roadkill (Coffin 2007), increased wildlife trafficking (Bush et al. 2014), eroding indigenous populations (AIIB 2016), and biosecurity concerns from the introduction of invasive species (Liu et al. 2019). Deforestation often begins with the construction of road

networks and fences that have negative direct and indirect impacts on biodiversity (Forman and Alexander 1998, Forman et al. 2002, Ree et al. 2015, Borda-de-Água et al. 2017). Indirect impacts from illegal logging, poaching, mining, and urban encroachment often accompany or result from road construction (Wilkie et al. 2000, Laurance et al. 2008, 2009, Ali et al. 2015, Bebbington et al. 2018). Any existing road in a forested area leads to further deforestation; about 5% of deforestation occurs within a five-kilometer buffer of a road network or a one-kilometer buffer of a navigable river (Benítez-López et al. 2010, Barber et al. 2014). The greatest threat to wild orangutan populations in Borneo and Sumatra is the loss of their habitat to road development, which exacerbates agricultural expansion, illegal timber harvesting, mining, and human encroachment. As connectivity within and among remaining forest patches is compromised, the negative impacts to biodiversity accelerate (Laurance et al. 2009). The infrastructure impacts on existing biodiversity and forests in the area can further weaken ecosystems and increase anthropogenic carbon emissions (Zadek 2019). Loss of biodiversity also disrupts ecosystem processes, and ecosystem service flows that benefit human well-being (Turner 1996, Sodhi et al. 2010).

The trade-offs between biodiversity conservation and human development activities (Adams et al. 2004, Redford and Fearn 2007, Reed et al. 2016) can be analyzed at multiple spatial scales from both conservation and sustainable development perspectives (ideally, they should be part of a single, systemic perspective). At the landscape scale, forest loss and fragmentation severely affect biodiversity (Morris 2010, Alroy 2017). Deforestation creates patterns of remnant forest fragments and ecologically

degraded forest edges. Species richness decreases in fragmented forests because remnants may be too small for species to persist, too exposed to edge effects, or too isolated to be colonized from other remnants (Gillespie et al. 2012). Edge effects and spill-over of species from secondary habitats (Vetter et al. 2011) impacts species composition in fragmented forests. Thus, there is a growing need to understand the impact of particular spatial patterns of deforestation and degradation, as these areas may be open to strategic modification even if the overall amount of deforestation is challenging to influence. The application of spatial ecology can also set a strategy for reforestation. All this should be incorporated into relevant environmental impact assessments, environmental education programs, and management plans to strengthen habitat protection and reforestation activities and other efforts to preserve and protect natural habitats and biodiversity (Quintero 2007, Buntaine 2011, Buchanan et al. 2016).

2.1.2 Development finance and energy investment

Development banks and development finance institutions (DFIs) serve as important channels for infrastructure finance and have adopted safeguards systems to minimize and manage the environmental and social risks associated with their projects (Morgado and Taşkın 2019). The banks have to address Sustainable Development Goals (S.D.G.s), laid out by the United Nations in 2015. Different development banks also apply differing levels of safeguards. Development projects have to comply with various national and international biodiversity regulations and protect the vulnerable and indigenous groups' rights and sustainable harvesting of forest products (Vardon et al. 2016). However, the effectiveness of these environmental regulations in different

countries is open to debate. We need timely and accurate spatial data to analyze the impacts of development projects on biodiversity loss.

In 2013, China introduced the Belt and Road Initiative (B.R.I.) to strengthen infrastructure, trade, and investment links between China and the participating countries. The B.R.I. has reached a vast portion of the world, particularly China's neighbors in Southeast Asia. All eleven countries in Southeast Asia - Brunei, Cambodia, East Timor, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam have signed B.R.I. cooperation agreements with China. China's Policy Banks have emerged as significant players in investing in development projects in Southeast Asia. Figure A1 shows international development finance commitments in the region between 2013 and 2019 (a period when data is available for all institutions listed). Figure A1 also indicates that in this period, the energy sector received a significant share of China's development finance in Southeast Asia (specifically, the power sector). Note that this investment amount is greater than the combined amount of all the other major development banks active in the region (Buchanan et al. 2016, Gallagher 2019, A.D.B. n.d., Japan International Cooperation Agency n.d., KfW n.d., World Bank n.d.). The two Chinese development banks that engage in international businesses are the China Development Bank (C.D.B.) and the Export-Import Bank of China (CHEXIM). The total amount of development finance committed by the two banks over this period has exceeded all other major development financiers in the region, including the World Bank (W.B.), the Asian Development Bank (A.D.B.), and the Japan International Cooperation Agency (JICA). However, compared with the multilateral development banks and some

western national development institutions, the Chinese development banks have lax internal environmental standards and principles that defer local regulations. While the trend in the multilateral development banks safeguards is moving toward a country systems approach, local regulations and local execution can be insufficient. Thus, China's development finance in the energy sector in Southeast Asia is critical in examining changes in biodiversity.

Many parts of the B.R.I. intersect terrestrial and marine biodiversity hotspots, wilderness areas, and critical conservation areas in Southeast Asia, including primary dipterocarp forests and the Coral Triangle (Lechner et al. 2018). Recent studies have investigated deforestation from the B.R.I. and its impact on biodiversity (Hughes 2019), species interactions, and ecosystem processes in tropical forests (Schleuning et al. 2011, Liu et al. 2019).

Although prior studies have considered how electricity generation impacts other aspects of the environment, few have looked specifically at how much land different energy sources require (Stevens et al. 2017). Electricity production in the U.S. shows that coal, natural gas, and nuclear power all feature physical footprints of about 4.85 Hectares per megawatt installed capacity. Solar and wind power respectively take up 17.6 and 28.6 hectares per megawatt installed capacity, while hydroelectricity generated by large dams has a significant footprint of 127.56 hectares per megawatt installed capacity (Stevens et al. 2017). Land cover changes associated with power projects include the construction of the plant, resource production, and the inundation of forested valleys by reservoirs. Transmission and transportation can also lead to deforestation, fragmentation of habitats

and migration routes, as well as pollution that could harm the ecosystems. The development of hydroelectric dams in the Mekong River basin illustrates the trade-offs and impacts of energy development on biodiversity (Ascensão et al. 2018, Shi and Yao, 2019). Hydropower produces relatively cheap and sustainable electricity for the lifetime of a dam (which is often not very long), flood and drought protection (Hecht 2017), and new economic and employment opportunities (Sabo et al. 2017). However, dam construction impacts water flow and fish migrations (Wild et al. 2016) and ultimately can cause the loss of traditional livelihoods and a significant reduction in the seasonal monsoon-driven flooding that had long brought nutrient-rich sediments out onto the floodplain. The dams also impede the downstream flow of these sediments (Baird 2016, Middleton et al. 2019). Such energy projects spotlight the need for strengthened project evaluation and monitoring and reforestation and conservation plans in any future infrastructure development through ecosystem trade-off analysis.

2.1.3 Framework

Our paper examines the increasing deforestation and threat to biodiversity fueled by power generation projects across the region by China and other countries. We adopt a "big data" spatial analytical framework to catalog deforestation and biodiversity changes in different ecoregions that have potentially resulted from power generation projects financed by China and other (including domestic) entities. The study utilizes a time series of forest cover and loss data (2000-2018) derived from Landsat to map the baseline information of forests, biodiversity, and energy investment locations to analyze forested areas bisected by the B.R.I. corridors. Forest cover data are analyzed four years prior and

as of the commission of the project. In contrast to prior studies, our study does not treat all forests as equivalent. We characterized the spatial structure of forest cover distribution using distinctions such as core, edges, and bridges. The spatial data and analysis enable us to investigate the morphological patterns of deforestation impact and specific biodiversity risks of power generation projects with different financing sources, energy sources, characteristics, and locations. Specifically, this paper attempts to:

- Based on forest cover and loss data from 2000-2018, inspect if deforestation patterns four years prior and as of the commission of energy projects are different and if the potential impacts of projects financed by Chinese development banks differ from non-Chinese financing.
- Investigate the biodiversity risks using IUCN criteria within critical buffer distances of power generation investments with and without Chinese development finance. Are there statistical differences?
- Examine habitat criticality and forest cover changes near different types of power plants (coal or hydro) with and without Chinese development finance in Southeast Asia before and after the project's commission to estimate if there are significant differences in biodiversity risks and deforestation impacts. Do the biodiversity impacts differ based on energy sources? How can our analysis inform policy on the ground?
- Using a forest spatial pattern methodology, we identify core, periphery, and other spatial fragmentation metrics to analyze the relationships between forest core and edge effects in more detail for energy investments in Southeast Asia.

Our project offers metrics and a replicable analytical framework to identify critical areas where forest and habitat preservation should be most heavily emphasized in Southeast Asia and other regions in the world. In particular, it demonstrates the merit of looking at deforestation in ways that dig deeper into its particulate causes and their specific impacts on forest structure and ecological integrity.

2.2 Materials and methods

2.2.1 Data

This study's spatial domain covers eight countries in Southeast Asia - Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Thailand, and Vietnam, all of which are also signatories of B.R.I. with China and are fit for comparison. Our methodology requires a selection of biodiversity factors, including habitats, forest boundaries and corridors, locations of power generation facilities, and a list of endangered and threatened species differentiated by ecoregions. China's Overseas Development Finance Database (Ray et al. 2021) provides projects' financial information financed by Chinese development banks. Figure 2.1 shows the data and methodology framework of the research.

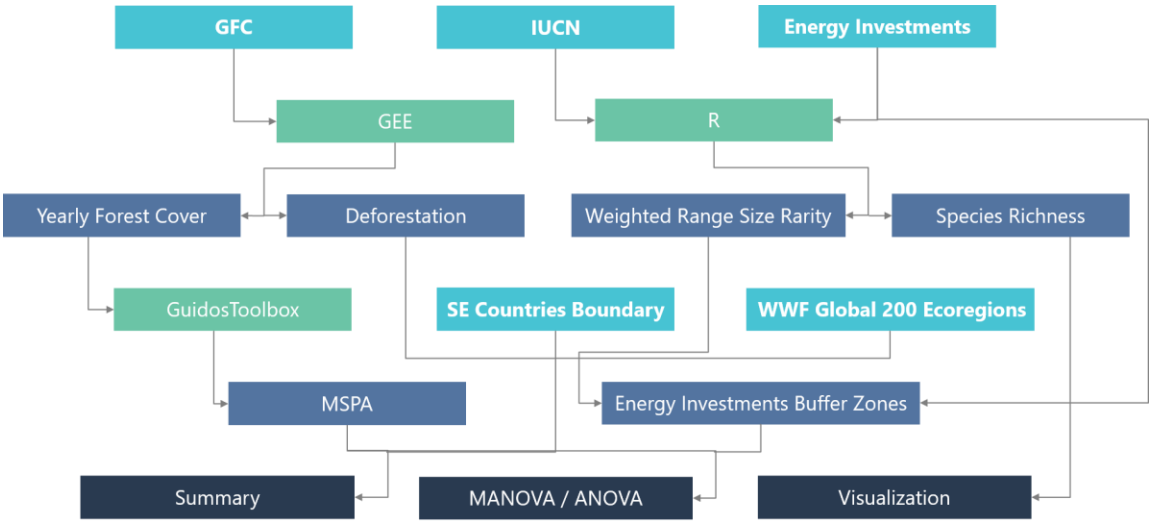

Figure 2.1. Research data and framework

2.2.1.1 Forest cover

Infrastructure constructions are likely to erode forest cover and further exacerbate impacts on biodiversity. We perform a time-series analysis characterizing forest extent and change in Southeast Asia based on a published data product, Global Forest Change v1.6 (Hansen et al. 2013), processed using Google Earth Engine. In this product, the variables are defined as follows:

- Trees are defined as vegetation taller than 5 m in height and are expressed as a percentage per output grid cell using the base year 2000 (denoted as ‘2000 Percent Tree Cover'). Tree cover makes up 25% or more at the Landsat pixel scale (30-m × 30-m spatial resolution) represents forest cover in this scheme.
- Forest Cover Loss is defined as a stand-replacement disturbance or a change from a forest to a non-forest state during 2000–2018.
- Forest Loss Year is a disaggregation of total 'Forest Loss' to annual time scales.

Based on forest cover data, it is possible to describe spatial patterns of forests in conjunction with non-forests, including fragmentations occurring within the forest and along its exterior boundaries. These fragmentations have differential impacts on related landscape properties, especially structural connectivity, that can be extracted and evaluated. Figure 2.2 shows the deforestation loss in different countries. Every country in this region has seen deforestation, from 2000 to 2018, the most severe being in Indonesia and Malaysia, 17% and 28% of their total tree cover respectively. The least amount of deforestation is in Brunei (5.2%), mainly attributable to its leadership, and East Timor (3.8%) in the south, which is dry and has little or no wet forest.

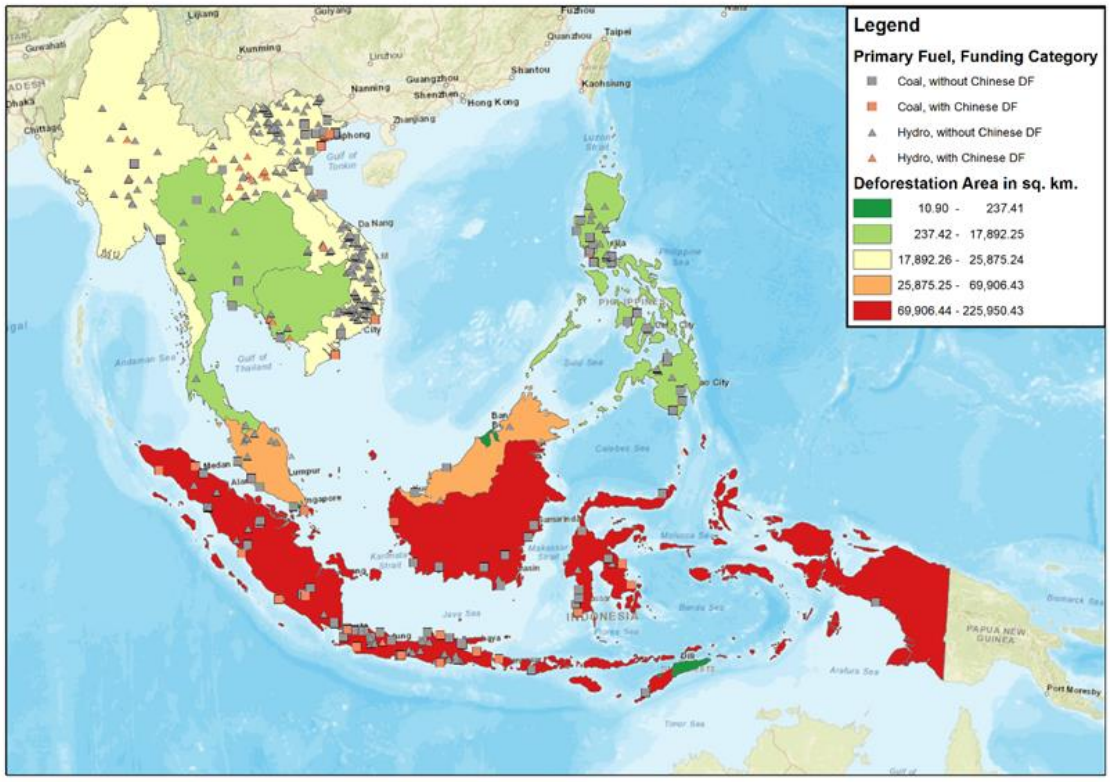

Figure 2.2 The map shows deforestation between 2000-2018 where red indicates the highest area of deforestation.

2.2.1.2 W.W.F.'s ecoregions

It is vital to interpret biodiversity loss in the context of ecoregions to account for the magnitude and nature of loss correctly. World Wildlife Fund (W.W.F.) defines an ecoregion as a "relatively large unit of land or water containing a characteristic set of natural communities that share a large majority of their species dynamics, and environmental conditions" (Olson and Dinerstein 2002). Each ecoregion has its distinct species, whose populations in an area of interest can be critically endangered or endangered, vulnerable, or relatively intact. We consider twelve ecoregions (Figure A2) in Southeast Asia, identified by W.W.F. Ecoregions vary significantly in population density, type and amount of vegetation cover, biodiversity, and development projects. Thus, it is possible to compare and contrast relatively undisturbed from more disturbed areas. For example, the Cardamom Mountains, a part of the regional Indo-Burma Hotspot, represent a rainforest ecoregion. It is (or was) home to several large mammals, including the Asian elephant (Elephas maximus) and tiger (Panthera tigris, now extirpated). Its remote location, protected status, and investment in enforcement had until recently prevented deforestation, which has now seen severe degradation.

In contrast, the peat swamp and heath forests, characterized by sandy soils, have no large mammals. The peat forest called kerangas forest on the islands of Indonesia (Borneo, Belitung, and Bangka) is home to some iconic specialist plants such as carnivorous Nepenthes pitcher plants and myrmecophiles (or ant plants). These peat forests are impacted by two forces, rapid clearing for human settlements and large-scale development projects, resulting in a highly vulnerable ecoregion (Dohong et al. 2017).

Each ecoregion in Southeast Asia, characterized by a different mix of plants and animals, is experiencing different threat levels due to human settlement expansion and logging. Hence, the increasing threats fueled by energy investments across the region by China and other countries further exacerbate biodiversity in Southeast Asia.

The W.W.F. has produced a global ecoregion map of the world of 867 ecoregions. In the present study, the Global Forest Change (G.F.C.) is overlaid on the ecoregion map to characterize deforestation and morphological spatial pattern trends in Southeast Asia.

2.2.1.3 The IUCN Red List of Threatened Species

The world's most complete inventory of the global conservation status of amphibians, fish, birds, and mammals is the IUCN Red List of Threatened Species (also known as the IUCN Red List or Red Data List). This global analysis of all species' status and distributions is based on the best available data and expert analysis for each species. (Lamoreux et al. 2003, Rodrigues et al. 2006). The IUCN Red List classifies species into nine groups, set through criteria such as the rate of decline, population size, area of geographic distribution, and degree of population and fragmentation. The IUCN criteria recognize significant differences between species and the circumstances leading to their extinction risk. This methodology categorizes each species' threat level based on the available data and the specified quantitative thresholds. IUCN's goal is to provide information and analyses on the status, trends, and threats to species to inform and catalyze biodiversity conservation action (Mace et al. 2008). The IUCN list can inform global and regional biodiversity targets, aid conservation planning, evaluate conservation actions, and inform legislative frameworks to protect species (Hoffmann et al. 2010).

IUCN Red List criteria cover three threatened categories -- Critically Endangered, Endangered, and Vulnerable. Species are allocated to one of the three types of extinction risk based on population size, range area, and decline rate. This allocation can aid in studies of biodiversity. There is also a probability of extinction in the wild associated with these three categories. The Red List has become an increasingly powerful tool for global conservation, management, monitoring, and decision making.

2.2.1.4 Power generation project locations and Chinese development finance

The paper compared and combined the Global Power Plant Database published by the World Resources Institute (Byers et al. 2018), the WEPP (World Electric Power Plant) database (Platts 2018), and power generation project information in China's Overseas Development Finance database (Ray et al. 2021). According to the data compiled, as of 2018, Southeast Asia has approximately 184.3GW of installed power generation capacity. About 59.3% of the current installed capacity is commissioned after 2000. However, the energy mix has remained almost unchanged. Coal, natural gas, and hydropower are the dominant power generation sources in the region (Figure 2.3). China Development Bank (CDB) and the Export-Import Bank of China (CHEXIM) are the two major Chinese state-owned development- and policy-oriented banks that engage in overseas development finance (DF). Among the 109.3 GW power generation capacity added between 2000 and 2018 in Southeast Asia, more than 19% have some Chinese development finance involvement. According to China’s Overseas Development Finance database, which is the best available database on the overseas lending of China Development Bank and Export-Import Bank of China compiled from public sources,

during this period in Southeast Asia, the Chinese development banks only financed coal and hydropower in the power sector. More than 36.9% of the coal power generation capacity and over 10.2% of the hydropower added over this period in the region have the Chinese development banks' financial participation. Thus, to make a fair comparison, this study only examines coal and hydropower projects in the region that are the foci of the Chinese development finance.

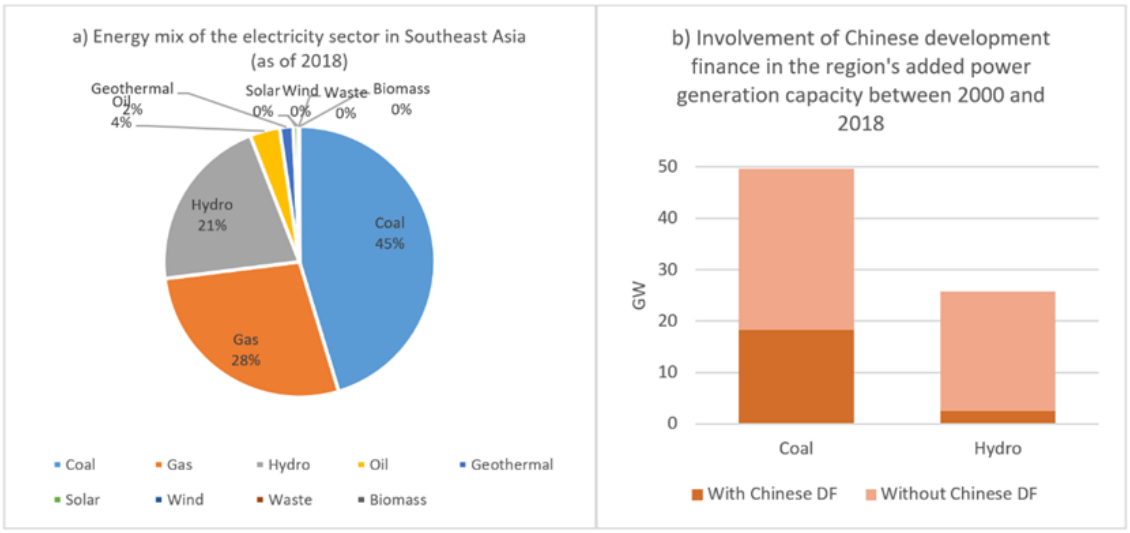

Figure 2.3. Energy mix and Chinese development finance involvement of the electricity sector in Southeast Asia, compiled from the Global Power Plant Database, World Resources Institute; the Platts World Electric Power Plant Database, and China's Overseas Development Finance database. Note: b) does not include projects whose commission year is not available. The capacity with Chinese development finance does not include project expansion on older projects commissioned before 2000.

The spatial database also includes relevant spatial data related to forest cover and biodiversity for assessing the impact of the energy investments. There are 398 coal or hydro energy investments without Chinese development finance and 51 with Chinese development finance in this database. The two sets of investments in coal and hydro are shown in Table A1.

2.2.2 Methods

To visualize Southeast Asia's biodiversity, we created a species richness map using IUCN Red List species geographic ranges at a raster resolution of the GFC dataset. IUCN has assessed terrestrial taxonomic groups and prepared polygon maps. This study's first step rasterized the vector IUCN ranges, to sum up across all species, resulting in a species richness count, representing the number of species potentially occurring in each pixel. Figure 2.4 shows the IUNC's species richness in Southeast Asia. Each panel uses a range of colors from red to blue, indicating high to low species richness. Panel 3a displays amphibians, while 6b - 6e shows reptiles, mammals, freshwater groups, and marine fish. Each panel shows a different pattern of richness. Mammals (6c) exhibit the highest species richness in peninsular Malaysia, Borneo, Sumatra, northern Vietnam, and eastern Myanmar. Reptiles are similar but are limited to only peninsular Malaysia, Borneo, and Sumatra. Panel 6e shows sharks, rays, and chimeras with the highest concentration near Taiwan.

2.2.2.1 Species richness & weighted range size rarity

Many studies aimed at identifying high-priority areas have utilized species richness to highlight biodiversity hotspots. As high/moderate-resolution land cover and climate data are more readily available from remote sensing, some studies have suggested rasterizing IUCN geographic ranges with high-resolution cell size. However, doing so would overestimate the actual species richness value due to scale mismatch (Hurlbert and Jetz 2007). Instead of using species richness as a metric for biodiversity, we calculated species weighted range size rarity for each energy investment buffer zone to address the

scaling issue. We calculated a weighted range size rarity for buffer zones in R (R Core Team, 2020) based on species range maps of IUCN, which is defined as:

$$wrsi_i = \sum_j w_j q_{ij}$$

where w_j is the weight assigned to species j in the prioritization and q_{ij} is the fraction of species j's range falling within the buffer zone of energy investment i (Moilanen *et al.* 2014). This measure lowers the contribution of wide-ranging species to overall species richness and highlights areas with a relatively high proportion of narrow-range species.

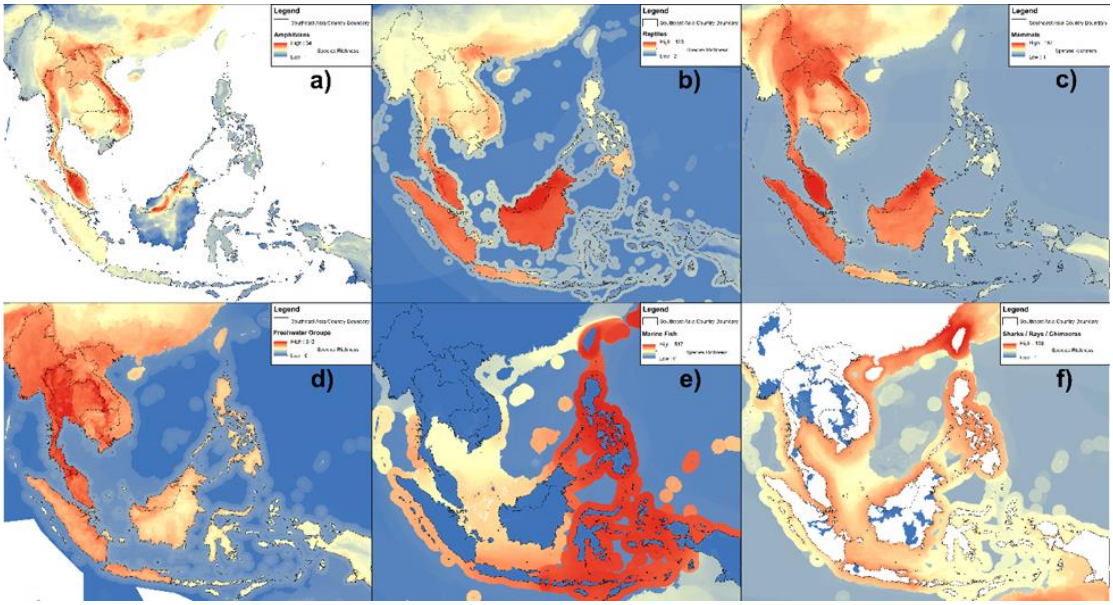

Figure 2.4. Species richness of a. amphibians; b. reptiles; c. mammals; d. freshwater groups; e. marine fish; f. sharks, rays, and chimaeras.

Our choice of weights assigned to species is guided by the severity level in the IUCN categories. We assign the following weights: critically endangered = 8; endangered = 6; vulnerable = 4; near Threatened = 2; least concern = 1; data deficient = 2 (Montesino Pouzols et al. 2014)

A high value of weighted range size rarity within the buffer zone indicates that this buffer zone can have one of the following:

- High species count overall.
- High species count with greater weight.
- Species with narrow geographic ranges.
- Any of the combination above.

2.2.2.2 Deforestation estimation 2000-2018 in Southeast Asia

Deforestation is estimated using Hansen's Global Forest Change. Google Earth Engine is utilized to estimate forest loss every observation year in each Southeast Asia country in the WWF ecoregions and energy investment buffer zones. The year 2000 is selected as our baseline to estimate subsequent yearly forest loss. We also looked at the forest cover changes for each energy investment four years before the commission year. A four-year window was chosen because it is the average time from construction to a typical power plant commission.

2.2.2.3 Geocoding and buffering power generation projects

The impact of power generation projects could reach beyond the site of the power plant. For thermal power facilities, transporting fuels and electricity transmission are the main factors that cause landuse change; for hydropower, wind, and solar power, the generation facilities are much more land intensive (Stevens et al. 2017). Meanwhile, indirect impacts such as other road constructions, land clearing for agriculture, human migration, increased poaching, etc., might result from the construction of power facilities.

As comprehensive location data is only available for power generation facilities, this study only compares deforestation trends near power plants, using spatial buffer areas for accounting for transportation impacts of fuel and electricity transmission.

Power plants have two labels - type of fuel source and whether they involve finance from Chinese development banks. They are geocoded as point information on the map, and each location is buffered at 1, 2, and 10km distances to estimate biodiversity species richness that faces potential impact and deforestation (Benítez-López et al. 2010). World bank and available local guidelines require buffer distances of within 1km. We expand the minimum buffer to 1km for safe estimates. However, past studies find significant differences in disturbance sensitivity between species groups and specific cases. Bird populations seem to be affected within a few hundred meters from infrastructure. In contrast, a reduction in mammal populations has been found at distances of a few hundred meters up to several kilometers from infrastructure (McLellan and Shackleton 1989, Cameron et al. 1992, Ortega and Capen 1999, Nellemann et al. 2003). Therefore, larger buffer distances, which may better reflect the impact on different species and footprints of transportation, transmission, or other indirect impacts related to the power generation project, are also used in this study (Benítez-López et al. 2010).

The section below elaborates the study's specific methodology. On deforestation trends, we analyzed deforestation in buffer areas of projects commissioned between 2000 and 2018, four years before the commission of each investment. This scope limitation narrows the number of projects studied in this paper down to 272 projects without Chinese development finance and 30 with Chinese development finance. Past studies

found that mammal and bird population densities significantly declined with their proximity to infrastructure on biodiversity risks. Bird populations were more impacted at a shorter distance to infrastructure than mammal populations. We categorize three levels of risk based on buffer distances around each energy investment. Each level is selected based on sensitivity to impact. Level 1 impact is defined with a buffer distance of 1 km, considering the IUCN categories in the immediate vicinity; Level 2 impact is a distance of 2 km, while Level 3 impact is a buffer of 10 km. Each IUCN level is examined in each risk buffer, providing insights on conservation risk priorities resulting in a total of 18 combinations (3 risk categories, 6 species groups). We then examine those combinations in each category using multivariate ANOVA (MANOVA) and ANOVA tests. We used MANOVA to extend ANOVA's capabilities to assess multiple dependent variables simultaneously. MANOVA tests helped differentiate the group significance of energy investments based on five variables, weighted range size rarity, power capacity, and changes in deforestation area, core, and bridge areas during the four-year period. ANOVA enabled us to test the significance of one variable at a time. We used the R (R Core Team, 2020) to derive MANOVA and ANOVA results.

2.2.2.4 Morphological analysis

We estimated spatial forest pattern and landscape connectivity using a morphological spatial pattern analysis (MSPA) toolbox (GuidosToolbox 2.6 version 4) (Vogt and Riitters 2017). The Global Forest Change was preprocessed into binary foreground/background image (forest/non-forest) for each year from 2000 to 2018. We selected a threshold of 900-m to define intact forest patches, forest corridors, and islet

forest-isolated patches. Due to Southeast Asia's geographic extent, the yearly tree cover image was recoded and split into tiles with dimensions of 9984 pixels (multiple of shardSize, an export parameter of GEE). This tiling process allows for parallel processing and overcomes computing memory constraints (10,000 pixels). The MPSA explicitly addresses the forest's morphology and segments the image into seven distinct forest landscape elements. Core forest refers to interior areas of forest located at a distance of (≥900 m) from the nearest forest edge, while edge forest refers to external forest perimeter of (<900 m). The 900 m edge distance threshold is the midrange indicator of the potential edge effect on forest patches (Alamgir et al. 2019). Next, a loop refers to a corridor connecting the same core area on either side, while a bridge refers to a corridor connecting different sections of a core-forest patch. Islet refers to disjoint and small locations of intact forest in contrast to a perforation that characterizes an internal object perimeter (Vogt and Riitters 2017, Alamgir et al. 2019). We deploy the MSPA application to examine the deforestation process and use the segmentation methodology to characterize differential impacts in China and other countries' energy investments.

Workflow using the GFC dataset:

- Using the Tree Cover of 2000 as a baseline, we can determine which forest cover pixels are lost per year based on a simple differencing. Each loss is denoted by the specific loss year.
- Power plant buffer zones, country boundaries, and ecoregion boundaries were used to generate deforestation areas (for the boundaries mentioned above).

- The remaining forest cover each year (determined by Tree Cover 2000 and loss year), were used as input for GuidosToolbox for MSPA analysis.
- The MSPA analysis results were tabulated with the Power plant buffer zones, country boundary, and ecoregion boundary.

2.2.2.2 Ecoregions

We used the WWF Global 200 ecoregion product and filtered Southeast Asia ecoregions. Figure A3 shows each ecoregion's distribution based on country boundaries. Indonesia has the largest number of ecoregions. The two largest ecoregions in terms of area are Borneo Lowland and Montane Forests and Sumatran Islands Lowland and Montane Forests. An ecoregion can be distributed across one or multiple countries. For instance, Naga-Manapuri-Chin Hills Moist forest is found mainly in Myanmar, while Indo-China dry forest is located over multiple countries, Vietnam, Thailand, Laos, and Cambodia.

2.3 Results

2.3.1 Characterizing deforestation trends in Southeast Asia based on MSPA

First, we estimate the deforestation areas in countries during this period based on a time series of global forest change, assessing deforestation impacts in forest patch morphology. Most lowland Southeast Asia was historically covered in hyperdiverse mixed dipterocarp forest of various types, with higher altitudes under a lower stature but still diverse montane forest. Nearly all of the original forest cover is already gone

(Grantham et al. 2020), with bridges and loops the dominant deforestation morphological structures. Bridges connecting different core areas (shown in red) are usually somewhat degraded and strongly modified by edge effects. A single core forest patch can be broken up into smaller areas, connected by loops (shown in yellow). Other morphological structures such as perforations or islets are often not visible at this scale due to their small sizes. The net result is a degraded forest in which only minimal areas are free of edge effects from some or all directions. We have yet to examine the impact of this pattern of erosion on tree species loss, but it could have a major impact on the current forest structure and its future trajectory. For example, tree species that require deep shade to germinate and mature are likely to be at higher risk than sun-tolerant forms. These tree species composition shifts will likely have knock-on consequences for animal communities as well as structurally dependent plant life forms such as lianas and epiphytes, with multiplicative impacts on all other taxa. Figure 2.5 shows the MSPA analysis for Southeast Asia in 2000 and 2018 with an Edge distance of 900 m. Figure A4 shows that deforestation is consistently increasing within the region, particularly in Cambodia, Indonesia, Vietnam, and Laos, as viewed from the perspective of a rapidly expanding matrix (green line) of non-forested landuse. Core forest loss, shown in pink, has been phenomenal, particularly in Indonesia in the last two decades. There are very few intact core areas (shown in green) left in Indonesia during this time. The MSPA analysis shows other countries, including Cambodia, Laos, Malaysia, and Vietnam, as having experienced core and edge forest losses over the last decade. These are of varying degrees but uniformly alarming with respect to the implications for biodiversity loss.

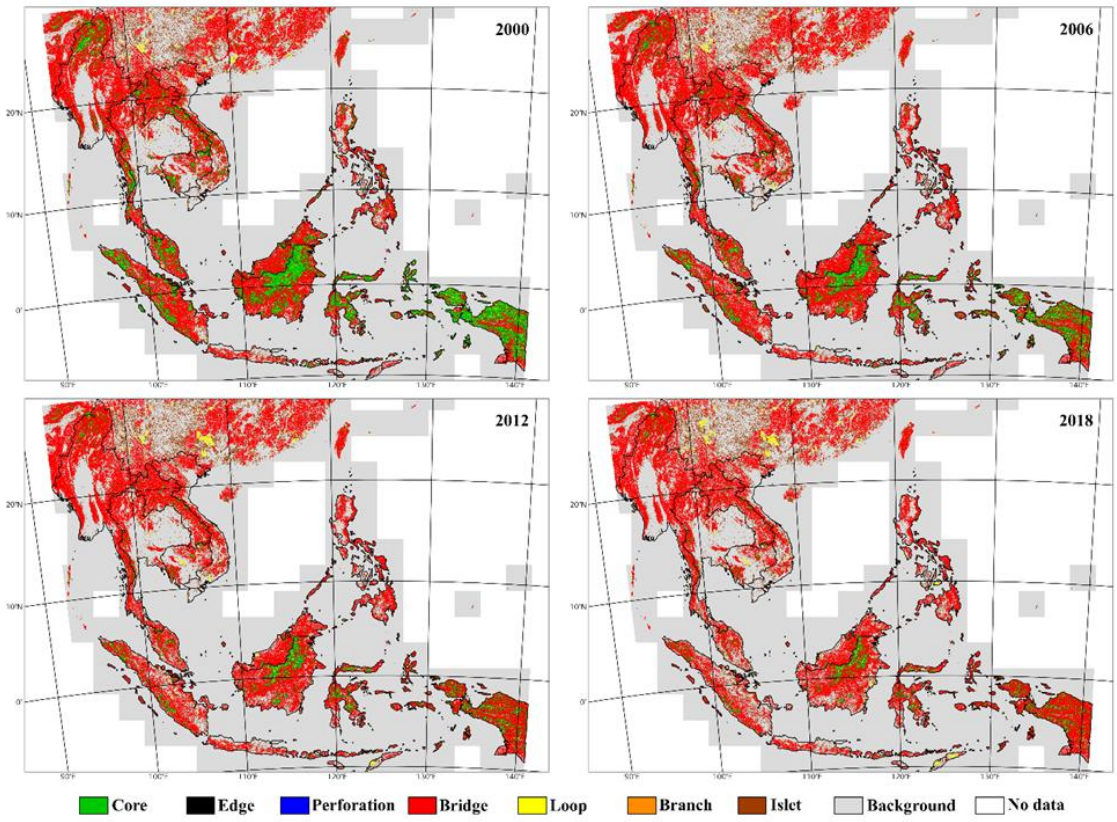

Figure 2.5. MSPA for 2000, 2006, 2012, and 2018 with edge distance of 900m; green represents core forest area (interior intact forest) while red represents the bridges connecting different core areas, often fragmented or deforested areas.

Next, we examine the types of significant differences in the morphological patterns of deforestation in various ecoregions. The core forest in Sulawesi moist forest and Sumatran lowland and montane forest has been roughly halved, dropping from around 50,000 km2 to less than 25,000 km2 in the last two decades. Even more dramatic is the core forest loss in the Borneo lowland and montane forest ecoregion, which have dropped from 190,000 to 50,000 km2. Cardamom mountain moist forest has similarly lost core forest from 100,000 to less than 50,000 km2 within the last two decades- this despite periods of relative stasis when conservation measures were episodically effective.

The patch morphology called “bridge” is of critical importance as wildlife highways, preserving the value of diminishing core forest patches by connecting them and maintaining gene flow and population viability. Here changes in forest bridge habitat are shown in purple. In some countries (Myanmar, Cambodia, Indonesia, Laos, Malaysia, Philippines), forest bridges increase at first and then decrease, indicating forest encroachment, followed by development and rising or total deforestation. The feature called "edge" is too small to be displayed in these graphs but plays a vital role in the forest disturbance and recovery regimes. Such morphological changes in deforestation have impacts on animal ranges and habitats. Orangutans once lived in large contiguous areas of intact rainforest, both in Borneo and Sumatra, but are now increasingly found at the edges of agricultural (oil palm) plantations and forest fragments, placing them at considerable risk. Another iconic species is the loris (Loris tardigradus), living in human-modified landscapes in Java, needing to use bridges to move from one patch of forest to another. Increasing human expansion (Birot et al. 2020) results in forest fragmentation, leading to an ever-greater reliance on bridges connecting population fragments of forest-dependent species. The morphological change in the forest may have altered the loris’ habitat, and now the loris is on the IUCN critically endangered list.

Figure A5. shows the MSPA results by ecoregion from 2000 to 2018, profiling temporospatial change in forest patch morphology. The Cardamom Mountains rain forests ecoregion (top right first row) spans 44,288.8 km2 of rain forest plus some drier dipterocarp forest, a part of the Indo-Burma global biodiversity hotspot. It contains or until recently had included many rare, threatened, and endemic species, including the

Asian elephant (Elephas maximus), tiger (Panthera tigris), and clouded leopard (Pardofelis nebulosa). It is also home to 450 birds and reptiles, such as the Cardamom Mountains wolf snake (Lycodon cardamomensis) and the Cardamom Mountains bent-toed gecko (Cyrtodactylus cardamomensis). Waterways in the Cardamoms are some of the last redoubts of many critically endangered freshwater species, including the Asian arowana (Scleropages spp.) and Siamese crocodile (Crocodylus siamensis). Its remote location and rugged terrain offered some protection in prior decades, but now this region is experiencing rapid and rampant deforestation. Its core forest has decreased from 9,500 to around 2,500 km2. Another ecoregion, the Borneo lowland, and montane forests have experienced a loss in core forests from close 170,000 to less than 52,000 km2 in this period (Cushman et al. 2017). This core forest loss has impacted several endangered and iconic species, including the Bornean orangutan, twelve other primate species, the Bornean bearded pig, and the Bornean yellow muntjac deer (Cheyne et al. 2016, Alamgir et al. 2019). While these species vary in their degree of dependence on intact primary forest, its rapid loss and the drivers responsible for this loss pose an imminent existential risk to them all.

2.3.2 Differences between power plants with and without Chinese development finance in Southeast Asia

We examine the differences between the impacts of power plants with and without Chinese development finance. The number of coal power plants with Chinese development finance is slightly higher (19 vs. 11), while for hydro investments, Chinese investments are lower (63 vs. 209). Chinese coal energy investments have greater power

generation capacity on average (661 MW) than their counterparts (501MW). Similarly, Chinese hydro investments have a greater average power generation, around 234 MW, compared with the hydro plants without Chinese development finance (average around 110 MW). Chinese invested power plants consistently have the same power generation capacity. However, it is worth noting that power plants (coal or hydro) with the largest capacity are non-Chinese investments. The maximum capacity of the Chinese invested coal and hydro energy power plants in Southeast Asia is 1244 MW and 790 MW; corresponding non-Chinese invested coal and hydro energy project capacities in Southeast Asia are 4180 MW and 2400 MW, respectively. Figure 2.6 shows the total power plant capacity in MW and GWh power production of power projects with Chinese investment (in red) and projects without Chinese investments (blue). The spread in the range of plant capacity of the latter group is visible.

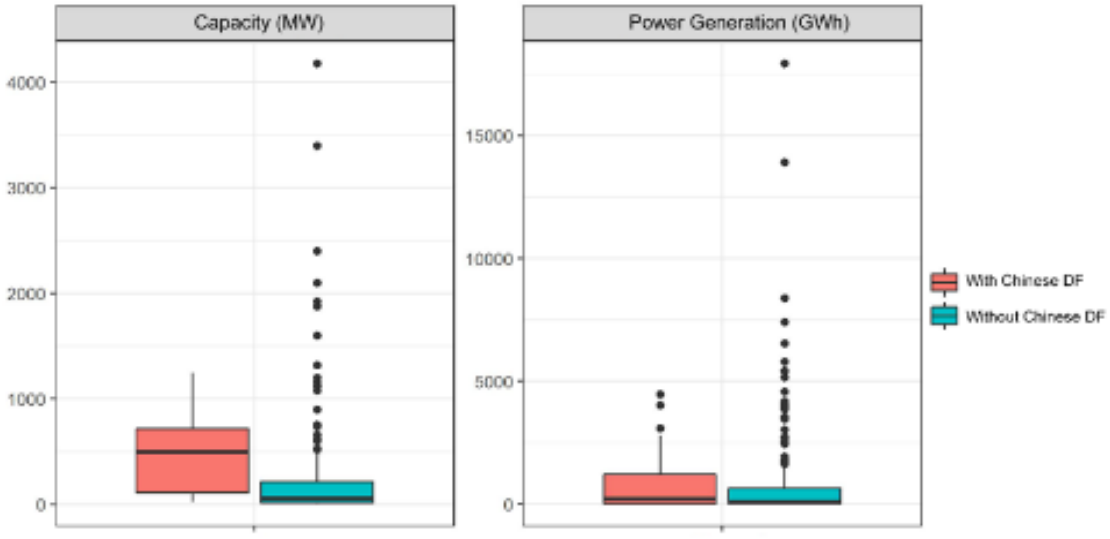

Figure 2.6. Power plant capacities of energy investments with and without Chinese development finance.

2.3.2.1 Significance testing of differences in the biodiversity impact of power plant projects at different buffer distances

We conducted MANOVA analyses to test if the impact of power plants with and without Chinese development finance within critical buffer distances are significantly different, with variables representing power plant characteristics and biodiversity metrics. Table 1 shows differences in statistical significance between the two investments in terms of buffer distances and species. We find that all three IUCN categories differ significantly in Chinese vs. non-Chinese financing at various buffer distances. This signifies differences between Chinese and non-Chinese energy investments, probably attributable to differences in investment strategy and location.

Table 2.1. Statistical significance testing at various buffer intervals between powerplants with and without Chinese development finance. *p<0.1; **p<0.05; ***p<0.01

Buffer	Species	MANOVA	Weighted rarity	Deforested area	Core area change	Bridge area change	Capacity
1 km	Amphibians	0.010*	0.196	0.445	0.692	0.785	4.94E-04***
	Freshwater groups	0.013*	0.540				
	Mammals	0.002**	0.027*				
	Marine Fish	0.004**	0.028*				
	Reptiles	0.013*	0.556				
	Sharks/Rays/ Chimaeras	3.85E-05***	6.79E-06***				
2 km	Amphibians	0.016*	0.231	0.974	0.454	0.744	
	Freshwater groups	0.018*	0.514				
	Mammals	0.002**	0.035*				
	Marine Fish	0.005**	0.015*				
	Reptiles	0.017*	0.485				
	Sharks/Rays/ Chimaeras	3.36E-06***	1.59E-07***				
10 km	Amphibians	0.003**	0.233	0.044*	0.800	0.093	
	Freshwater groups	0.004**	0.373				
	Mammals	0.003**	0.214				
	Marine Fish	0.001***	0.003**				
	Reptiles	0.001**	0.066				
	Sharks/Rays/ Chimaeras	5.09E-10***	7.22E-12***				

ANOVA tests for marine fish like vertebrates are significant at all buffer distances. The number of species in the buffer zone is also higher for power plants with Chinese development finance (shown in Figure 2.7) This is probably because Chinese invested power plants are closer to the coast compared with their counterparts.

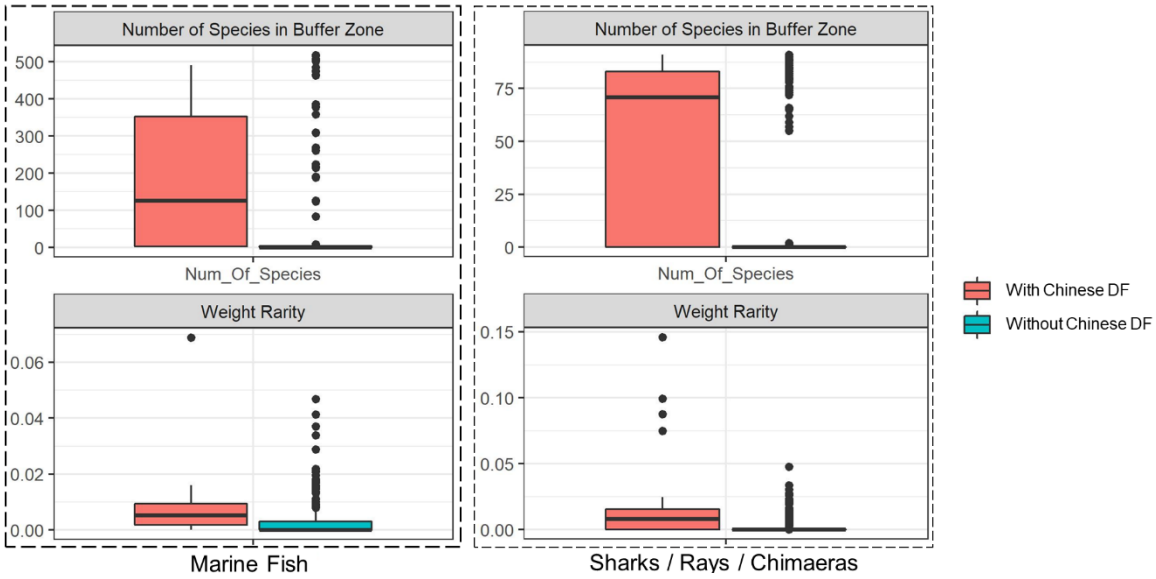

Figure 2.7. Biodiversity statistics of Marine Fish and Sharks / Rays / Chimaeras.

The weighted rarity for mammals in 1 km and 2 km buffer zones is significant, mainly because of one Chinese-funded power plant with a relatively high weighted range size rarity, the highest point in Figure A6. For example, PLTU Tanjung Kasam in Indonesia is a coal plant with a power capacity of 110MW. Although the 2 km buffer zone of this plant overlaps with 49 mammal species, the weights based on IUCN categories are high, resulting in a biased estimate for weighted rarity for mammals resulting from Chinese coal investment. In the following Table A2, the total weighted rarity of species is around 86, where the three endangered and one critically endangered makes 26 of the 86 in the index. PLTU Tanjung Kasam is the highest point in Figure A6,

and its weighted range size rarity value is more than twice as much as the second highest power plant in the lower right panel.

2.3.2.2. Significance testing of differences in the deforestation impacts of power plant with and without Chinese development finance

We use ANOVA tests to find significant discrepancies between the two groups of investments four years before and after the commission year. Figure A7 displays the statistical differences in the amount of forested area impacted by Chinese and non-Chinese investments at a 10km buffer distance. We compare the top left panel in Figure A7 with the lower right panel. The range in the amount of forest area before and during the commission years is higher for Chinese than non-Chinese power investments. Also, both groups show deforestation during the four-year period. The rate of deforestation and the area of deforestation are higher for energy investments without Chinese development finance. The proportion of hydro power plants is likely higher in the non-Chinese investment group. Since Chinese invested coal power plants are closer to the coast and population center, the available forest is lower in their buffer zones

2.3.3 Differences between coal and hydro power plants in Southeast Asia

We run the similar MANOVA and ANOVA test on coal and hydro power plants to determine if the corresponding differences emerge as the previous tests. We found no significant differences in deforested areas during the four years before commissioning. However, there are critical differences between coal and hydro energy power plants in the changes occurring in the core and bridge area in the same period. For hydroelectric

investments, the morphological differences (Figure 2.8) are noticeable in larger buffer zones (10 km). In general, flooding areas associated with a hydroelectric reservoir results in the loss of core forest in a dam's buffer zones. For example, Nam Ngum 2 Dam, located on the Nam Ngum River, one of the major tributaries to the Mekong, is a Chinese financed hydroelectric project in Laos. The forest cover decreased by 7.3 square kilometers within the 10 km buffer area. Interestingly, the loss of bridge forest in buffer zones of hydroelectric and coal energy investments is similar, but some hydroelectric investment buffer zones lead to an extensive bridge area. Increasing forest area may be related to tree plantations, including non-native species such as rubber, eucalyptus, and acacia (Holt et al. 2016, Ingalls et al. 2018).

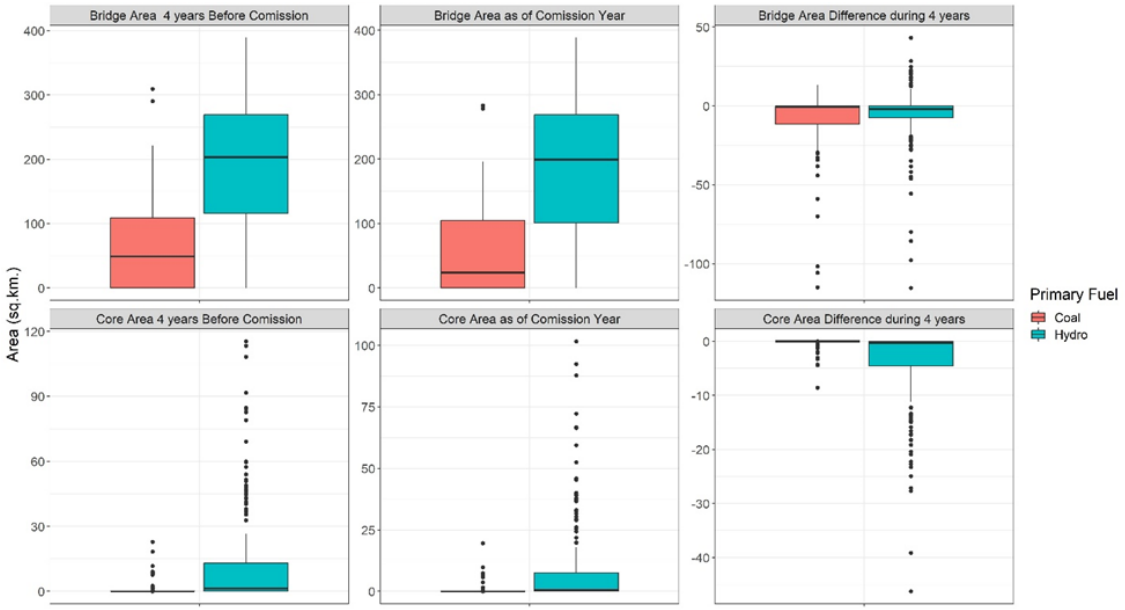

Figure 2.8. Bridge area and core area summaries in 10 km buffer showing four years before commission years, at commission years, and changes during the construction of coal and hydro power plants.

These patterns need to be understood based on the differences between coal and hydropower projects. Coal power plants generally have greater power capacity. However,

four of the top five power plants in power generation in Southeast Asia are hydroelectric. Some morphological differences in the energy investments may stem from the variation in coal and hydro power plants' site selection factors. Coal power plants are located in relatively densely populated areas with high end-use demand nearby. Since large hydroelectric energy investments are constrained by proximity to rivers, they tend to be situated in forested or unpopulated areas that would “only” directly impact the indigenous population.

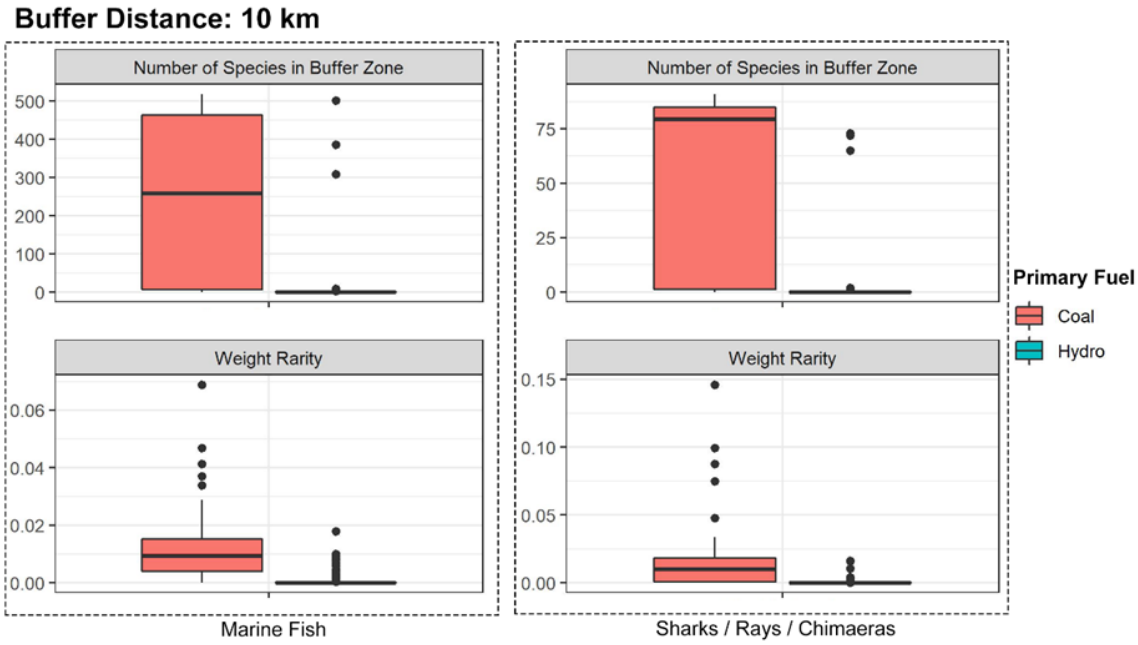

Figure 2.9. Biodiversity statistics of Marine Fish and Sharks / Rays / Chimaeras for coal and hydro power plants.

There are significant statistical differences in weighted range size rarity for amphibians as well as for marine fishes, sharks, rays, and chimaeras. The weighted range size rarity is much higher in the buffer zones of coal power plants than for hydro plants (Figure 2.9). That is perhaps due to coal power plants’ location near major population centers and major cities in Southeast Asia, mainly near the coasts.

Table 2.2. Statistical significance testing at various buffer intervals between coal and hydro power plants. *p<0.1; **p<0.05; ***p<0.01

Buffer	Species	MANOVA	Weighted rarity	Deforested area	Core area change	Bridge area change	Capacity
1 km	Amphibians	2.47E-11***	0.025**	0.588	0.465	0.989	9.22E-14***
	Freshwater groups	1.38E-11***	0.145				
	Mammals	2.03E-11***	0.358				
	Marine Fish	9.47E-25***	6.02E-19***				
	Reptiles	5.10E-11***	0.319				
	Sharks/Rays/ Chimaeras	5.56E-16***	2.61E-11***				
2 km	Amphibians	5.83E-12***	0.029**	0.372	0.090*	0.618	
	Freshwater groups	3.35E-12***	0.189				
	Mammals	5.84E-12***	0.548				
	Marine Fish	4.06E-29***	4.50E-24***				
	Reptiles	9.43E-12***	0.248				
	Sharks/Rays/ Chimaeras	3.01E-19***	9.25E-15***				
10 km	Amphibians	1.58E-14***	0.029**	0.225*	1.49E-05***	0.021**	
	Freshwater groups	1.52E-14***	0.434				
	Mammals	4.11E-14***	0.939				
	Marine Fish	9.47E-38***	3.02E-31***				
	Reptiles	9.05E-15***	0.003***				
	Sharks/Rays/ Chimaeras	4.41E-15***	6.78E-19***				

2.4. Discussion and conclusion

Our study in Southeast Asia provides a big data spatial framework to analyze the impact of energy investments on deforestation and biodiversity throughout Southeast Asia. We use the Global Forest Change - an extensive remote sensing dataset to estimate deforestation at 30 m resolution enabling us to examine yearly deforestation trends in every location. We further enriched our analysis by capturing the forest's morphological structure through MSPA analysis. We utilized the IUCN data to analyze further biodiversity risk, a proxy for the complex biotic interactions among component species in Southeast Asia's twelve ecosystems. Our study is one of the few studies to analyze deforestation highlighting the forest's morphological structure, a key to understanding the impact of biodiversity loss in each ecoregion. Our study metrics would be relevant to developing robust, smart conservation plans in each country as they grapple with energy development and biodiversity trade-offs, the focus in a future study.

The path to sustainable development in Southeast Asia emphasizes energy investments. China's energy investments in Southeast Asia are compared with other countries' investments to understand and compare the risk to biodiversity in both coal and hydroelectricity. We analyzed the energy investments in terms of specific locations, power production, and impacts at various buffer distances. There are statistical differences in energy investments with and without Chinese development finance, mainly driven by the capacity differences and the locations of the two investment groups. We show the differential impact of forest core and edge relationships of each power plant

type on weighted species richness and species richness, based on statistical significance tests.

The rapid rate of tropical forest loss poses such a great and immediate threat to global biodiversity that most authors have gone for the throat and looked exclusively at total forest area lost. Operationally, however, the morphology of forest loss is of enormous biological importance and is also the scale at which policy modifications could make a big difference in species conservation, even independent of deforestation rates per se. The total forest area remaining tells only half the story: if core forest vanishes and corridor habitats are amputated, the days are numbered for all obligately forest-interior species. Prior studies in biodiversity and energy development rarely consider either the extraordinary diversity of forest types in Southeast Asia or their spatial disposition and morphology. A lack of accurate representation of ecosystem processes or proxies for them at a regional scale is the primary source of uncertainty in these studies. Our approach identifies the deforestation and biodiversity risk at a regional ecosystem scale. We are able to differentiate biodiversity impacts in Borneo montane rain forest from peat swamp ecosystems in Borneo using existing data sources for species richness and weighted rarity of species. While there is ubiquitous globally relevant earth observation data, challenges remain in translating the resulting analysis into biodiversity strategies on a regional scale. In this paper, we analyzed the morphological structure of the forest to further differentiate the impacts of the loss of core and edge forests in each ecoregion. The loss of core forests is critical in many regions. The apparent regeneration or restoration of corridor and bridge forests can look good superficially, but if the trees

involved are exotic species, this can result in the substitution of weedy species for a large fraction of the native biota.

The tropical forest estate of Southeast Asia is very close to ruin. However, the remaining pockets and protected areas still provide a basis for maintaining much of the indigenous biodiversity and for restoring viable tracts of forest habitat that could then be self-sustaining, along with all of their contained wildlife, once again. For this to happen, in addition to a development plan, each country and every major development initiative- particularly the Belt and Road Initiative- must also have a rigorous conservation plan. First, they must consider alternative development scenarios so that initial impacts can be minimized. Second, they must include an aggressive restoration program and timeline complete with the full extent of both remote and on-the-ground monitoring required for accountability and adaptive management in the development process. We propose that deforestation and afforestation morphology analysis, in conjunction with regular on-the-ground biodiversity survey work, can be a valuable tool in providing the much-needed accountability to provide metrics in assessing such trade-offs.

Funding: This research was funded by support from the Climate and Land Use Alliance (G-1812-55950) and David and Lucile Packard Foundation (2020-68743) to the Global Development Policy Center at Boston University.

Appendix A contains additional figures, and additional tables.

CHAPTER 3 – Geographically Weighted Regression Models in Estimating Median Home Prices in Towns of Massachusetts Based on an Urban Sustainability Framework

Housing is a key component of urban sustainability. The objective of this study was to assess the significance of key spatial determinants of median home price in towns in Massachusetts that impact sustainable growth. Our analysis investigates the presence or absence of spatial non-stationarity in the relationship between sustainable growth, measured in terms of the relationship between home values and various parameters including the amount of unprotected forest land, residential land, unemployment, education, vehicle ownership, accessibility to commuter rail stations, school district performance, and senior population. We use the standard geographically weighted regression (GWR) and Mixed GWR models to analyze the effects of spatial non-stationarity. Mixed GWR performed better than GWR in terms of Akaike Information Criterion (AIC) values. Our findings highlight the nature and spatial extent of the non-stationary vs. stationary qualities of key environmental and social determinants of median home price. Understanding the key determinants of housing values, such as valuation of green spaces, public school performance metrics, and proximity to public transport, enable towns to use different strategies of sustainable urban planning, while understanding urban housing determinants—such as unemployment and senior population—can help modify urban sustainable housing policies.

3.1 Introduction

3.1.1 Background

More than 83 percent of the US population now lives in cities, up thirty percent from 50 years ago. By 2050, the US urban population is projected to increase to more than 90 percent (of 423 million in 2050). Urbanization is transforming farmland, wetlands, forests, and other natural ecosystems into urban landscapes at an unprecedented rate resulting in urban sprawl. Urban landscape patterns and dynamics are the physical manifestation of complex interactions between environment, society, and economy (Alberti, 1999; Grimm et al., 2000; Ostrom, 2009; Pickett et al., 1997). Thus, urban areas are highly relevant, if not central, to any discussion on sustainable development.

The central goal of urban sustainability is efficient use of natural resources within a city region, while simultaneously improving its livability, through social amenities, economic opportunity, and health (Newman & Kenworthy, 1998). We offer a conceptual framework to understanding urban sustainability through the lens of urban housing that is at the intersection of economic, ecological and social dimensions. Housing plays a pivotal role in determining the financial (economic) security and well-being of individuals, neighborhoods, and cities. Owning a home is part of the “American Dream”. Research on housing has focused on determining home prices, traditionally modeled in economics using hedonic models (Shiller, 2007). Some factors that are typically included in hedonic price analyses are socio-economic and demographic factors such as housing age/tenure, vacancy rates, racial and ethnic demographics (Harris, 1999). Boston metro has been well studied in the context of hedonic pricing models for the last 3 decades. An early paper

(Case & Mayer, 1996) studied patterns of home price appreciation in the Boston area from 1982 to 1994 and showed that changes in the cross-sectional pattern of home prices are related to differences in manufacturing employment, demographics, new construction, proximity to the downtown, and to aggregate school enrollments. Hedonic pricing models use parcel level housing data to estimate model parameters. Social factors including income, population density, race, and education impact urban housing and school districts. This has led to a renewed interest in social justice addressing accessibility to parks (Wolch et al., 2014), food (Alkon & Agyeman, 2011), and environmental health (Cutts et al., 2009) in the context of urban sustainability. Ecological studies in the urban sustainability context have examined challenges and opportunities in addressing imbalances between green and built landscapes (Wu et al., 2014) and more recently in mitigating urban carbon footprint (Wang et al., 2015). Urban sustainability research traditionally has addressed only a single ecological issue in individual neighborhoods (Wachsmuth et al., 2016). We need to expand this research to include regional or state level scales in order to address sustainability in a larger context.

In the last two decades, spatial econometric studies have been introduced that demonstrate the fundamental role of spatial autocorrelation and spatial heterogeneity in the analysis of housing in varied regional contexts (Anselin, 2013; A. C. Goodman & Thibodeau, 2003; LeSage, 2008). Spatial models use a variety of spatial scales, ranging from parcel level data to census town level data, to address different questions in relation to urban patterns and dynamics. Prior studies have focused either on individual housing units (Bitter et al., 2007; Helbich et al., 2014; Huang et al., 2010), or slightly larger scale,

i.e., districts or metro areas (Collins & Woodcock, 1996; Hasse & Lathrop, 2003; Paulsen, 2013). At a regional or continental scale, remote sensing instruments are the most viable option for data-driven mapping, monitoring, and assessment of the urban systems (Jeon et al., 2014; Kennedy et al., 2010; Maliene & Malys, 2009). Given the ready availability of remote sensing data, many studies have incorporated the ecosystem service valuation of green space (forest, woodland, or parks) into pricing models (Wolch et al., 2014), thus opening new avenues of research in urban sustainability in the coming decades.

The objective of this study was to assess the significance of key economic, environmental and social determinants of median home prices in Massachusetts towns using spatial models at the level of the state, including many urban agglomerations. We used a sustainability lens in setting our theoretical framework and incorporated the effects of environmental, economic and social factors over a time period. In a sustainable development context, the dynamics of housing growth need to be examined in order to formulate relevant policy or development decisions. Real estate market dynamics are inseparable from the sustainability concept as cities are designing green buildings, public housing, and affordable housing (Maliene & Malys, 2009). This analysis uses a select set of determinants based on prior empirical research. Population density (Case & Shiller, 2003) is often correlated to income levels; wealthier affluent communities are characterized by higher value homes. Proximity to highways or public transport impacts home values (Higgins & Kanaroglou, 2016). Residential development is also constrained by specific town conservation laws and regulations, therefore woodlands and forests not

protected are open to development (Wolch et al., 2014). Economic factors, such as unemployment, impacts home prices especially in periods of economic downturn (C. J. Goodman & Mance, 2011). Social factors such as age demographics (Myers & Ryu, 2008), public schools (Gibbons & Machin, 2008) and education (Aughinbaugh, 2013) may also impact home prices. The benefit of investigating the change in housing prices on a town scale is that demographic and socio-economic data are available over multiple years of observation. The data can be used to investigate the changing pattern of spatial non-stationarity that can be modeled and mapped using GIS. Such a mapping shows the non-stationarity in urban housing and sustainability.

The main contribution of our approach is the introduction of determinants related to urban sustainability at a town scale rather than the traditional hedonic home pricing related to individual houses. Our approach incorporates a plethora of data including remote sensing, census, and other published state reports. This “big data” approach to sustainability science is needed to develop public policies on housing, conservation, transportation and employment. We believe that our approach provides a broader perspective on urban home values and enables the characterization of patterns of spatial non-stationarity across Massachusetts. There are policy implications related to development, demographics, housing and transportation. What forest or woodland to cut depends on value placed on open lands in various towns across the state? Suburban towns north and west of Boston place a higher premium on forest compared with rural western Massachusetts. Senior populations living in suburban Boston homes have seen considerable increase in their home prices and seniors choose to continue to live in these

towns with greater access to senior services (free rides to malls or hospitals). A town may adopt stricter regulations to protect unprotected forests open from development, or to improve the quality of education by imposing more taxes on its citizens, knowing that school performance positively impacts home prices in these towns.

The structure of the paper is as follows: the next section examines the modeling framework and the long history of hedonic and GWR modeling. Section 2 outlines the data sources and discusses the methodology including spatial modeling considerations. Section 3 provides the results of analysis relating to spatio-temporal patterns of median home prices derived from the ordinary least squares (OLS), and two types of GWR models. We highlight the differences in the degree of spatial non-stationarity in the determinants of the median home price in Massachusetts, as well as characterize the temporal differences in median home prices in the period of bust and boom. Section 4 provides conclusions related to the theory and practice of GWR in this field.

3.1.2 Modeling framework

From the methodological point of view, the hedonic price function f, typically describes the property price P as a function of three categories of independent variables: structural, locational, and environmental characteristics. Traditional hedonic approaches adopt a model structure which reduces heteroscedasticity and nonlinearity to produce a single solution for the intercept term, along with the coefficients that determine the significance of independent locational, structural, or environmental characteristics, and the overall model's goodness-of-fit. Hedonic modeling has been executed at a variety of spatial scales ranging from a block or neighborhood (Bitter et al., 2007; Helbich et al.,

2014; Huang et al., 2010), to the district or metro scale (Hasse & Lathrop, 2003; Paulsen, 2013). However, these models cannot account for spatial autocorrelation resulting from spatially correlated omitted variables or spatial externalities and spatial heterogeneity (A. C. Goodman & Thibodeau, 2003).

Spatial autocorrelation indicates that homes in a neighborhood tend to be more similar. Real estate companies (including the popular Zillow), in effect, use spatial autocorrelation to determine the price of a home at a certain location based on the prices of nearby (similar) homes. Furthermore, many homes in a neighborhood tend to be built around the same time, and proximity to both positive and negative externalities has similar effects on the market values of nearby properties (Fotheringham et al., 2015). However, spatial heterogeneity or spatial non-stationarity results when the relationship between two or more variables determining the median home price is not constant across space, resulting in locally varying submarkets. Ignoring spatial non-stationarity leads to misspecification in the model including missing local effects that can have profound implications for understanding the temporal and spatial relationship between housing prices, location and housing attributes (Huang et al., 2010; Wu et al., 2014).

The spatial modeling literature provides a variety of local and global models to deal with spatial dependence (Anselin, 1998; Pace et al., 2009), as well as models that explicitly incorporate spatial heterogeneity such as the geographically weighted regression (GWR) methodology (Fotheringham et al., 2002). GWR calibrates a series of local regression models separately at each location and offers the ability to map local estimates of the intercept, variable coefficients, and other regression diagnostics and a

check for spatial variations in the relationships between dependent and independent variables at each location (Bitter et al., 2007; Yu et al., 2007). The keys to GWR modeling are the spatial kernel functions, with fixed or varying bandwidth that impact the shape and size of local neighborhood at each location (e.g., a circular neighborhood of fixed radius or a fixed number of neighbors for each location), and the weighting functions that determine the significance of neighbors with fixed or varying bandwidth. The definition of spatial neighborhood is a critical consideration in the analysis. GWR has been applied in the contexts of determining home prices in many cities (Crespo et al., 2007; Demšar et al., 2008; Helbich et al., 2014; Lu, Charlton, et al., 2014; Saphores & Li, 2012).

The basic GWR model assumes the same degree of spatial smoothness for each coefficient, which may not hold true in all contexts. GWR therefore overfits the data and produces a bias. Hence the basic GWR has undergone the following significant revisions: First, traditional GWR models define distances as straight line or Euclidean, while more recent modifications of the distance function adopt non-Euclidean distance metrics (Lu, Charlton, et al., 2014) to improve the model fit. Second, traditional GWR models use a fixed bandwidth for all variables to estimate the spatial relationship between variables while a revised GWR can use a flexible bandwidth (Yang, 2014) to estimate spatially varying relationships at various geographical scales within one model. The resulting model estimates coefficient surfaces that may vary at different spatial scales for different variables leading to better model fits. Third, Wheeler (Wheeler, 2007, 2010) proposed regularized GWR models, by combining ridge and/or lasso regression with GWR that

have shown robustness in addressing the multicollinearity problem. Fourth, there has been a focus on diagnostics to check the model fit such as cross-validation (CV) score (Bowman, 1984) to derive an optimal kernel bandwidth for GWR regression to reduce model bias. Another measure is the Akaike Information Criterion (AIC) (Akaike, 1974) that is traditionally used to account for model parsimony dealing with the trade-off between prediction accuracy and complexity. In GWR, a corrected version of the AIC is used that accounts for sample size (Hurvich et al., 1998) and entails fitting bandwidths with different penalty functions. Fifth, the incorporation of temporal non-stationarity into GWR model is providing more insights on market trends and depreciation of home prices through time (Huang et al., 2010). Finally, not all variables in GWR models exhibit non-stationarity in all contexts; hence assuming that all of the independent variables in the GWR exert a spatial influence on the dependent variable can lead to biased estimations (Wei & Qi, 2012). For example, real estate markets may be economically connected through common federal policies such as governmental subsidies while some price-determining effects vary across space resulting in spatial heterogeneity. It would be wrong to assume that both factors exert a spatial influence. Research addressing this issue has led to the formulation of the Mixed Geographically Weighted Regression model (MGWR) that incorporates both linear regression and the GWR (Fotheringham et al., 2002). MGWR is a regression model in which the first step involves differentiating non-stationary and stationary variables by explicitly testing for spatial variability. This testing results in some independent variable coefficients being held constant, which are considered global parameters, while some others spatially vary, denoted as local

parameters (Brunsdon et al., 2000). The second step in MGWR involves in mapping spatial variability of local parameters while global parameters have no spatial variability. Thus, MGWR differs from GWR, where all independent variables are assumed to have spatial variability resulting in different spatial distributions of local parameters.

3.2 Materials and methods

3.2.1 Study area

Massachusetts is located in the northeastern United States and has an area of 27,340 km2, 25.7% of which is water and 61% is covered by forest. It is the 7th-smallest state in the United States and accounts for around 2.75% of the total GDP. The state has a population of 6.8123 M (2016) of which 4.7 M live in the Boston Metro region. Boston is the 10 largest metropolitan area in the US. The population of the state is mostly urban (83%). Figure 3.1 shows the location of Boston and other towns mentioned throughout this paper. Metro Boston (#7 on Figure 3.1) is located on the eastern coast of the state.

The GWR analysis is conducted at the town and city scale. There are 351 towns and cities in Massachusetts. Each town's data across the time period was selected for census years 2000 and 2010 (Bureau, n.d.-b) and ACS data for years 2009, 2011–2013 (Bureau, n.d.-a). We only included 336 towns and cities due to data availability. Our data includes the Great Recession, a highly influential period for housing prices in the US economy. The period of the Great Recession, from 2008 to 2009, was characterized by loss of wealth, reduction in consumer spending, and massive job loss that resulted in a decline of home prices (Elsby et al., 2010; C. J. Goodman & Mance, 2011).

3.2.2 Data sources—socio-economic and environmental variables

Our data was collected from various sources including the census, remote sensing (land cover), transportation, tax department, and labor statistics. We describe our data shown in Table 3.1. Our dependent variable is the Median Home Prices, recorded from the census (2000 and 2010) and ACS (2009, 2011, 2012, and 2013). Our choice of time periods was dictated by the availability of relevant data. We differentiated our temporal analysis into two sets based on the source of data—decennial census collected in 2000 and 2010, and ACS data from 2009 to 2013. (Note ACS data for 2008 is unavailable). The American Community Survey (ACS) data are estimates and therefore cannot be compared with census data directly. We used the ACS data since it represents the period of Great Recession starting in 2009. We used inflation-adjusted median home prices in order to compare today's real estate prices to their historical norm.

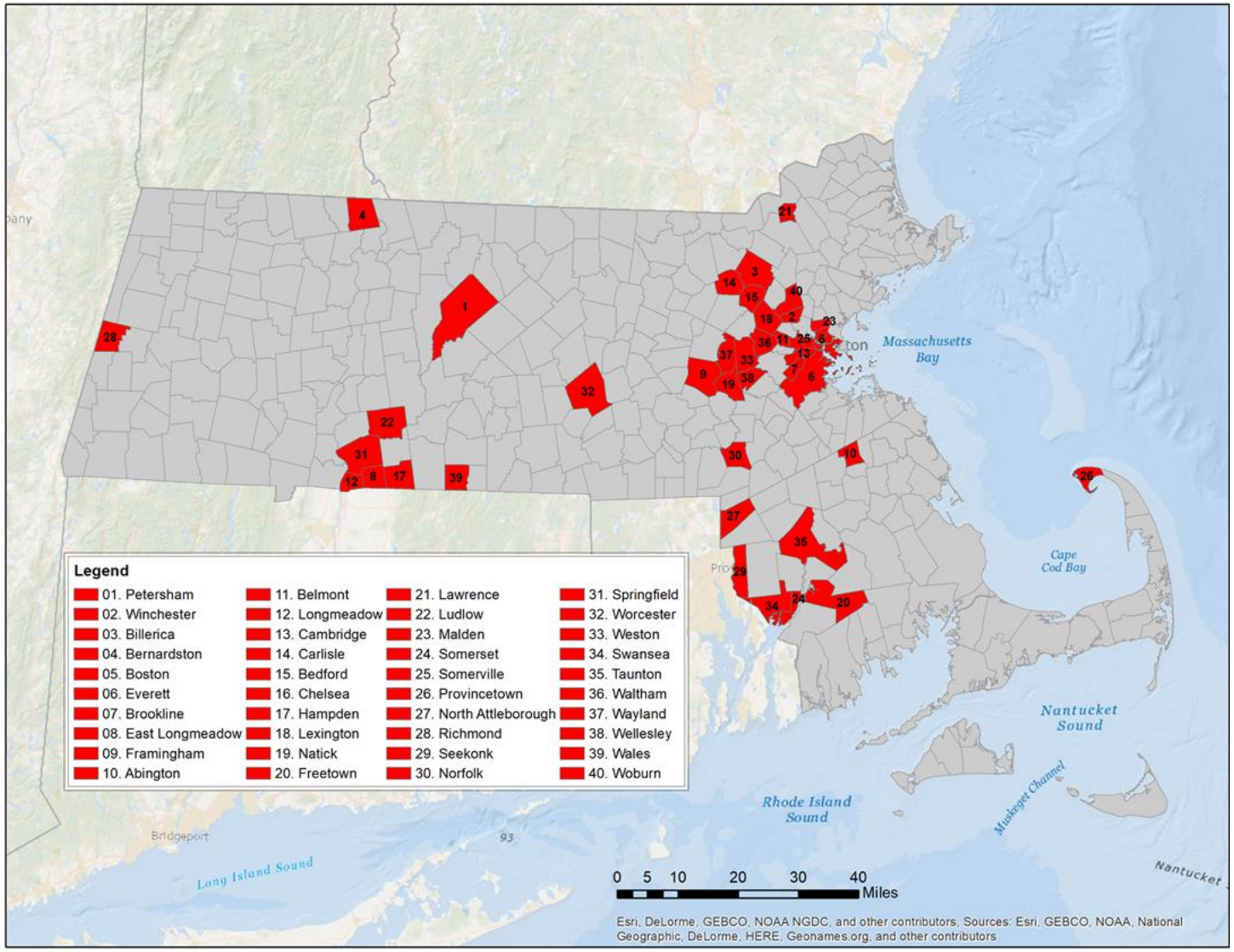

Figure 3.1. Location of Boston and other towns in Massachusetts mentioned in the paper.

The independent variables used in our models were selected based on past theoretical and empirical works examining their relevance in estimating home values. Population Density and Unemployment Rate have an impact on home prices (Case & Shiller, 2003; Rogers & Winkler, 2013; Shiller, 2006). The impact of unemployment on housing during the Great Recession (Byun, 2010; Higgins & Kanaroglou, 2016) shows that employment was key for recovery in some metro areas of the US. Unemployment was a significant factor in the wake of the housing collapse in 2008 in Massachusetts and impacted home prices (Holly et al., 2010). Therefore, we assume that unemployment is likely to impact the housing prices in our study context. The impact of education (college degree) and home price highlight that since 1988, those who completed college owned homes at higher rates in the US than those with no college education (Aughinbaugh, 2013). Sellers of existing homes provide a major share of the annual supply of homes sold in the US; home sales are driven by the aging of the population since seniors are net home sellers (Myers & Ryu, 2008). Older homeowners have emerged as a dominant segment of the housing market following the housing collapse in 2008. The homeownership rate for Americans aged 65 and over has remained at 80 percent while dropping for every other age group. Seniors typically have less mortgage debt than younger homeowners; they typically downsize and sell, or increasingly some stay at the same home (*Projections and Implications for Housing a Growing Population*, 2016). Seniors play a significant role in housing dynamics and hence are included in our analysis.

Residential property tax rates vary across the state related to home values and the mix of residential and commercial holdings in each town. Suburban towns such as

Weston and Wellesley have some of the lowest residential tax rates in the state while Longmeadow in western Massachusetts, with few business establishments, relies heavily on residential property taxes to fund town services. Residential taxes may therefore reflect the economic structure of each town and is included in our analysis. Prior studies identified and measured land value uplift (LVU) resulting from rapid transit and other mass transit (Higgins & Kanaroglou, 2016; Mulley, 2014). Thus, we have included distances from town centroids to commuter rail stations as an independent variable.

Vehicle ownership is derived by normalizing the census variable called vehicle availability for each town by the population of each town. This data may provide information on town's travel accessibility and modal choices. More vehicle ownership may imply more travel by car while less vehicle ownership may imply the use of public transport, walking or biking (Rodrigue et al., 2016). Vehicle ownership has direct implications in urban sustainability studies. The city of Boston is planning to go carbon neutral by 2050 and is seeking solutions for reducing air pollution caused by automobile transportation.

Variables	Description	Source
Median Home Price	Median home value in thousand dollars (adjusted for inflation)	Census, ACS
Population Density	Population density (number of people per hectare)	Census, ACS
Unprotected Forest	Percent coverage of unprotected forest in each town	Landsat
Unemployment Rate	Percent of unemployed people in each town	Mass. Labor and Workforce Development
Residential Area	Percent coverage of residential areas	Landsat
Vehicle ownership	Number of vehicles per capita	Census, ACS
Higher Education	Percent of people have bachelor's or higher degree above the age of 25	Census, ACS
Senior Population	Percent of senior population (over age 65)	Census, ACS
Distance to Commuter Rail Sta.	Distance from town centroid to nearest Commuter Rail Station	MassGIS, MBTA
Residential Property Tax	Amount per $1000 assessed home price	Mass. Department of Revenue
Composite Performance Index	Students' performance on Mathematics	Mass. Department of Elementary and Secondary Education

Table 3.1. Variables used in the study.

There is strong evidence to suggest that school quality substantially affects home prices in the US (Brasington & Haurin, 2006; Clapp et al., 2008). Earlier studies focused on the relationship between home prices and the quality of local education, using public school expenditures per pupil as the key school variable. However, more recent research highlights that the measure of K-12 student achievement is a more appropriate variable in home value estimations. We use the Composite Performance Index (CPI) scores of school districts in our analysis. The CPI is a measure of the extent to which students are progressing toward proficiency (a CPI of 100) in ELA (English Language Arts) and mathematics on the state's MCAS (Massachusetts Comprehensive Assessment System) (*2010 Glossary of AYP Reporting Terms*, 2010). CPIs are generated separately for ELA and mathematics at all levels in Massachusetts. For this study, we considered CPI for mathematics performance in public school district. In order to estimate the MCAS score for each individual town, we used Area Interpolation Tool in ArcGIS. First, we created an interpolated surface based on CPI from all public school districts. Then we applied Massachusetts Town and City boundary shapefile to estimate the CPI for each district. The State implemented and collected CPI data starting 2003. Hence, we substituted the CPI score of 2003 for 2000 since we did not have data for 2000. Better public schools' CPI (MCAS) generally correlates with a higher home price (Gibbons & Machin, 2008), exploited by realtors (such as Zillow) in selling homes.

In our study, we obtained the land cover classes designated as residential (low and high) and forest from the Mass Audubon's publication called Losing Ground (Lautzenheiser et al., 2014), derived from processing and classifying Landsat time series

of Landsat-5, 7, and 8 data for this time period. Two thematic classes called low and high residential classes described in the Losing Ground report were summed up to define the residential class in our study. A recent study by Cunningham et al., (Cunningham et al., 2015) used Landsat archives to examine the change from undeveloped (forest) to developed land-use during the real estate bubble (2000–2006) and the subsequent bust (2006–2013) in Massachusetts. The results in this paper further highlight the significance of the land cover change during this period.

According to US Forest Service, Massachusetts had an estimated 3.0 million acres of forest land in 2014. About 61% of the land area of Massachusetts meets the Forest Inventory and Analysis (FIA) definition of forestland (Butler, 2016). Forests are not evenly distributed across the state, and are largely influenced by development patterns. The lowest occurrences of forestland are seen in areas surrounding Metro Boston, Springfield, and Worcester, as well as along the coast and the major transportation corridors. Unprotected forest is an important determinant of home price and hence we created this new class of forest that is at the greatest risk for development. We derived this class based on two different data sources—Losing Ground (Lautzenheiser et al., 2014) report and MassGIS data, as follows: First, forest areas were extracted from the Losing Ground dataset for the entire study period. Second, MassGIS dataset called Protected and Recreational Open Space layer was extracted; this class includes recreation, conservation, surface water supply protection areas and wellhead protection areas, or scenic sections (*MassGIS (Bureau of Geographic Information) | Mass.Gov*, n.d.). Third, we created a new class called unprotected forest by differencing layers

created in Step 1 and 2 to derive forest that is at risk for development. This thematic class provides us the areas at risk for conversion from forest to some form of development in the towns of Massachusetts.

3.2.3 Spatial model considerations

In this study, we use conventional ordinary least squares (OLS) as a benchmark, GWR and MGWR to describe spatial heterogeneity in housing across the towns in the state of Massachusetts.

The first step in the analysis was to use a traditional OLS of the form:

$$y_i = a_0 + \sum_{j=1}^{p} a_j x_{ij} + e_i \, , i = 1,2,\ldots,n$$

where y_i is home price of each town i at a specified time of the study and x_{ij} is a row vector of explanatory variables for town i, a_j is a column vector of regression coefficients, and e_i is the random error for town i. The first element of the equation is the intercept. The initial model considered the relationship of the median price of a home in each town using independent variables—Population Density, Unprotected Forest, Unemployment rate, Percent Residential Area, Vehicle Ownership, Higher Education, Senior Population, distance to the nearest Commuter Rail station, Residential Property Taxes and CPI (Composite Performance Index) for each town. OLS results were interpreted based on an assessment of multicollinearity, adjusted R^2 and Akaike's information criterion (AIC) (Akaike, 1974). The variance inflation factor (VIF) statistic, which measures redundancy among explanatory variables, was used to assess multicollinearity. Explanatory variables with large VIF values—above a threshold of

7.5—were considered to be multicollinear. This process ensured that the model became unbiased.

The next step in the analysis was to use GWR that explicitly incorporates the spatial structure of the variables into the estimation of the regression and shows how those estimates vary across space. We have selected an adaptive kernel whose bandwidth was found by minimizing the AIC value. The bandwidth is a count of the number of nearest observations to be included under the kernel. Preference is given to lower values of AIC since they indicate a closer fit to the data.

$$y_i = a_0(u_i, v_i) + \sum_{j=1}^{p} a_j(u_i, v_i)x_{ij} + e_i\,, i = 1,2,\ldots,n$$

y_i is the dependent variable at location i

x_{ij} is the jth independent variable at location i

p is the number of independent variables

We explore the spatial variability of relationships between median home prices and the explanatory variables by mapping GWR coefficients and local R^2 values. We also performed an F3 test to probe whether the GWR estimates are a significant improvement on the conventional globally estimated OLS (Leung et al., 2000). The Akaike Information Criterion (AIC) is used in this study as a test diagnostic to select flexible bandwidth b (Gollini et al., 2015; Lu, Harris, et al., 2014) from:

$$AICc(b) = 2n\,ln(\hat{\sigma}) + n\,ln(2\pi) + n\{\frac{n + tr(S)}{n - 2 - tr(S)}\}$$

where n is the local sample size (according to bandwidth); $\hat{\sigma}$ is the estimated standard deviation of the error term; and $tr(S)$ represents the trace of the hat matrix S. The hat matrix denotes the projection matrix from the observed y to the fitted values.

As highlighted before, GWR is not always appropriate if some of the variables do not exhibit spatial non-stationarity and can be held constant. We used a Mixed Geographically Weighted Regression model (MGWR) after testing for spatial variability of all variables (Mei et al., 2006). In MGWR, some contributing factors that have no spatial variability will generate a global parameter, while others with spatial variability will produce a local parameter. The MGWR is defined as:

$$y_i = a_0(u_i, v_i) + \sum_{j=1}^{q} a_j x_{ij} + \sum_{j=q+1}^{p} a_j(u_i, v_i) x_{ij} + e_i\,, i = 1,2,\ldots,n$$

a_j are global coefficients

We used a Monte Carlo approach to test for significant (spatial) variation in each regression coefficient of the basic GWR against a series of randomized data sets (Lu, Harris, et al., 2014). If the true variance of the coefficient did not fall in the top 5% tail of the ranked results, the corresponding variable was treated as a global variable in the specification of MGWR.

3.3 Results

We discuss three sets of results. We first examine the spatio-temporal patterns of median home prices derived by comparing and validating the results of the OLS and GWR models. Our second set of results show the application of MGWR to highlight the

differences in the degree of spatial non-stationarity in the determinants of the median home price in towns of Massachusetts, while the third set of results describes the urban sustainability from the perspective of economic, social and ecological determinants. The OLS results are presented in Tables 3.2–3.4; GWR results are presented in Tables B1–B6 and Table 3.5; and MGWR are presented in Tables B7-B12 and Table 3.6. Summary results for each model are presented in Tables 3.4 (OLS), 3.5 (Basic GWR) and 3.6 (MGWR).

3.3.1 Spatial-temporal patterns of home prices

We first explore and model the spatiotemporal variability of median home prices and associated determinants in the state of Massachusetts by benchmarking the performance of the global regression model (OLS) with its GWR counterpart with the same set of variables. Table 3.2 shows the results from the OLS regression for the median home prices using decennial census data from 2000 and 2010, while Table 3.3 shows similar results for the ACS years—2009, 2011, 2012, and 2013. Table 3.4 shows overall summary results for the OLS model.

Median Home Price										
	OLS Model (2000)					OLS Model (2010)				
Variables	**Coeff.**	**t-Value**	***p*-Value**	**Sig.**	**VIF**	**Coeff.**	**t-Value**	***p*-Value**	**Sig.**	**VIF**
Intercept	289.69	3.53	4.77×10^{-4}	***	-	252.91	1.66	0.099	.	-
Population Density	0.79	0.93	0.353	-	4.01	0.04	0.04	0.971	-	3.40
Unprotected Forest	30.66	0.73	0.468	-	4.00	−52.01	−0.91	0.363	-	3.94
Unemployment Rate	−7.31	−1.94	0.053	.	1.53	−8.69	−3.28	1.15×10^{-3}	**	1.43
Residential Area	−63.12	−1.35	0.179	-	6.61	−65.83	−1.10	0.272	-	5.95
Vehicle ownership	−390.28	−5.13	5.11×10^{-7}	***	2.34	−427.09	−5.91	8.45×10^{-9}	***	2.09
Higher Education	1634.22	14.75	$<2 \times 10^{-16}$	***	2.26	1527.87	10.50	$<2 \times 10^{-16}$	***	1.98
Senior Population	162.68	1.36	0.176	-	1.90	335.00	2.20	0.028	*	2.09
Dist. to Stations	−0.60	−4.019	7.26×10^{-5}	***	2.01	−0.95	−4.495	9.69×10^{-6}	***	2.26
Property Tax	−7.61	−5.64	3.80×10^{-8}	***	1.35	−13.03	−6.52	2.73×10^{-10}	***	1.29
CPI	1.95	3.21	1.48×10^{-3}	**	2.24	5.14	3.56	4.28×10^{-4}	***	2.44

Table 3.2. Ordinary least squares (OLS) results for home prices in MA towns 2000 and 2010. (Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.'). The coefficients, intercept and 10 independent variables along with t-values, *p*-values, significance, VIF values are listed. VIF is a metric for testing multicollinearity (threshold 7.5).

Median Home Price										
	OLS Model (2009)					OLS Model (2011)				
Variables	**Coeff.**	**t-value**	***p*-Value**	**Sig.**	**VIF**	**Coeff.**	**t-Value**	***p*-Value**	**Sig.**	**VIF**
Intercept	129.38	0.93	0.351	-	-	220.08	1.544	0.124	-	-
Population Density	0.69	0.64	0.520	-	3.54	1.27	1.335	0.183	-	3.51
Unprotected Forest	16.39	0.30	0.767	-	3.73	5.81	0.119	0.905	-	3.65
Unemployment Rate	−7.47	−2.86	4.54×10^{-3}	**	1.35	−9.60	−4.006	7.65×10^{-5}	***	1.59
Residential Area	−17.39	−0.30	0.766	-	5.72	−70.84	−1.351	0.177	-	5.75
Vehicle ownership	−394.14	−4.82	2.21×10^{-6}	***	2.25	−363.62	−4.911	1.44×10^{-6}	***	2.31
Higher Education	1520.59	10.85	< 2e−16	***	1.88	1326.71	9.754	$< 2 \times 10^{-16}$	***	2.21
Senior Population	237.89	1.54	0.125	-	2.07	355.34	2.527	0.012	*	2.14
Dist. to Stations	−0.90	−4.29	2.33×10^{-5}	***	2.20	−1.06	−5.696	2.75×10^{-8}	***	2.20
Property Tax	−13.68	−6.31	9.17×10^{-10}	***	1.35	−10.27	−5.779	1.76×10^{-8}	***	1.41
CPI	5.68	4.27	2.57×10^{-5}	***	2.30	4.38	3.238	1.33×10^{-3}	**	2.67
	OLS Model (2012)					OLS Model (2013)				
Variables	**Coeff.**	**t-Value**	***p*-Value**	**Sig.**	**VIF**	**Coeff.**	**t-Value**	***p*-Value**	**Sig.**	**VIF**
Intercept	198.77	1.66	0.097	.	-	101.55	0.75	0.457	-	-
Population Density	1.32	1.55	0.122	-	3.37	0.98	1.15	0.251	-	3.33
Unprotected Forest	9.66	0.21	0.831	-	3.63	0.71	0.01	0.988	-	3.67
Unemployment Rate	−10.95	−4.99	9.66×10^{-7}	***	1.55	−10.19	−4.85	1.88×10^{-6}	***	1.57
Residential Area	−85.82	−1.76	0.079	.	5.73	−77.12	−1.58	0.116	-	5.68
Vehicle ownership	−391.33	−5.81	1.50×10^{-8}	***	2.35	−449.96	−6.45	4.04×10^{-10}	***	2.51
Higher Education	1234.99	10.05	$< 2 \times 10^{-16}$	***	1.99	1157.21	9.32	$< 2 \times 10^{-16}$	***	2.05
Senior Population	403.83	3.12	1.98×10^{-3}	**	2.11	487.12	3.87	1.31×10^{-4}	***	2.06
Dist. to Stations	−1.02	−6.09	3.23×10^{-9}	***	2.08	−1.00	−6.05	4.05×10^{-9}	***	2.01
Property Tax	−10.31	−6.60	1.64×10^{-10}	***	1.39	−9.88	−6.46	3.81×10^{-10}	***	1.41
CPI	4.94	4.53	8.45×10^{-6}	***	2.09	6.24	4.84	2.05×10^{-6}	***	2.26

Table 3.3. OLS results for home prices in MA towns 2009, 2011–2013. (Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.'). The coefficients, intercept and 10 independent variables along with t-values, *p*-values, significance, VIF values are listed. VIF is a metric for testing multicollinearity (threshold 7.5).

Of the covariates, only population density, unprotected forest, and residential area are not significant (at significance level 0.05) in the model in any year of observation, while others including unemployment, vehicles owned, residential taxes, educated population above 25 years, senior population, or distance to commuter rail stations, were statistically significant in all years. Unemployment rate was consistently significant from 2009 to 2013 as seen in other studies. Suburban towns exemplify this pattern the best indicating that an increase in unemployment causes a drop in median home prices in suburban towns such as Natick, Framingham, Wayland, Wellesley and Cambridge. Senior population impacted median home prices more significantly in 2013 (at significance level 0.001) and to a lesser degree from 2010 to 2012 (at significance level 0.05). Perhaps this segment of population held on to their homes that depreciated in value (in 2008) and sold their home starting 2010. The ownership of homes by seniors and its impacts on median home prices deserve further empirical scrutiny. Coefficients, associated with unemployment rate and residential taxes, are negative indicating that a decrease in home price is associated with an increase in both unemployment and residential taxes. Unprotected forest may not have overall significance in the state, but it may have more significance in the eastern part of the state around Boston where development of housing in unprotected forest has resulted in urban sprawl (Wolch et al., 2014) and the building of “McMansions”. Such non-stationarity patterns in this determinant have to be explored using GWR since traditional OLS is unable to account for spatial heterogeneity in local submarkets.

Figure 3.2 shows the OLS residuals in the median price of homes in various towns in different time periods. In general, OLS underestimates median home price in eastern Massachusetts, which is more populated with greater housing density, and overestimates the median home price in the rural western part of the state. This pattern is consistent over the entire time period. The ratio of unprotected forest to residential area in 2013 is computed to highlight such price variations. Eastern towns around Boston have ratios ranging from 0 to 0.626, while in western and central Massachusetts, the ratios range from 5.82 to 75. Higher ranges of ratio in the west suggest that more expansive forest is not influencing median home prices in these towns, while the opposite may be true in the eastern towns. Therefore, unprotected forest may be characterized by non-stationarity and needs to be explored using GWR.

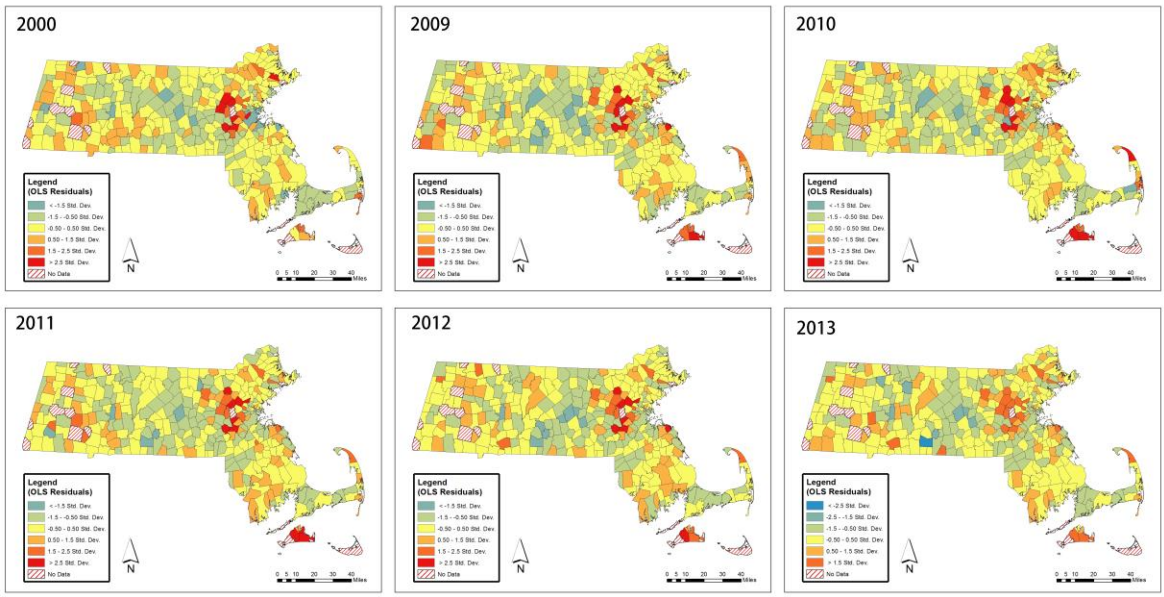

Figure 3.2. OLS residuals of 2000 and 2009–2013.

Decennial Census Years			ACS Years			
Year	**2000**	**2010**	**2009**	**2011**	**2012**	**2013**
RSS	1,609,745	2,946,544	2,918,609	2,332,915	1,993,794	2,033,210
AIC	3825.92	4029.05	4025.85	3950.59	3897.81	3904.38
Adjusted R^2	0.71	0.73	0.64	0.65	0.67	0.66

Table 3.4. OLS diagnostics for home prices in MA towns. RSS, AIC and adjusted R^2, for decennial years (2000 and 2010), ACS years (2009, 2011–2013).

Table 3.4 shows the overall OLS results. The coefficient of determination (R^2) across the time period ranges between 0.71 (2000) and 0.73 (2010) in the census years and from 0.64 (2009) to 0.65 (2011), 0.67 (2012), and 0.66 (2013). The standard error for the intercept is not significant in any year, except 2000. We expect that median home prices exhibit spatial heterogeneity and vary in the state from east to west reflecting proximity to Boston. As highlighted before, most of western Massachusetts is rural with large areas of forest. Hence unprotected forest may not impact overall home prices in the west but may have an impact for residential development in the east.

GWR results are presented in Tables B1–B6; each table pertains to one year of observation and shows the medians and ranges in the values of each coefficient across all towns (Columns 1–3). These coefficient numbers, in general, have large ranges. Hence, the results are next summarized using percentages of coefficient estimates that were positive and negative (columns 4 and 5) for each variable across all towns. GWR results for census years are shown in Tables B1–B2 along with *p*-value (F3 test) and significance. Population density is not a significant factor in both census years. The public-school CPI scores are less significant in all years. The remaining determining variables are highly significant across both census years.

The ACS survey years are shown in Tables B3–B6. Examining the impact of determinants, CPI score is significant in all ACS years (at significance level 0.001).

Population density is only significant in recent years 2012–2013, probably driven by increasing urban growth in towns such as Everett, Lawrence, Malden, Arlington and Brookline (see Figure 3.1). Overall population increased the most in Boston, Cambridge, Somerville, Chelsea and Brookline 2011–2013. Other factors are significant across all years. To summarize, GWR results highlight that most determinants are characterized by non-stationarity leading to spatial variation in median home prices.

We next interpret columns 4 and 5 in Table B6 for 2013. Most towns (65.77%) seem to value unprotected forest (positive coefficient) while some towns (34.2%) in the rural west (negative coefficient) may not. Senior population accounts for 91.96% in the positive coefficient column, indicating that most senior citizens may have higher incomes, and better homes (longer tenure of ownership). The coefficient of this determinant is negative only in 8.04% of towns. This pattern is exemplified in towns outside of Boston Metro, such as Wales, Richmond, and Bernardston, where the total population of the town may be decreasing due to migration of all segments of population except seniors. Overall, the results across census and ACS years, show that unemployment rate, vehicles ownership, residential property taxes, residential area and distance to commuter rail stations have a larger percentage of towns with negative coefficients (column 5 in all tables). The model predicts a negative relationship between these determinants and median home prices across most towns. On the other hand, CPI, educated population above 25 years, population density, unprotected forest and senior population have a range between 60% and 100% of towns with positive coefficients (column 4 in all tables). The model predicts a positive relationship between these

determinants and median home prices across most towns. The model fit in terms of positive and negative coefficient signs is consistent across most years highlighting the strong role of the selected determinants in determining the median home prices in the state.

Figure B1 maps the estimated coefficients of senior population in 2000, and 2009 to 2013 which indicate that proximity to Boston plays a strong role in differentiating eastern (colored red) and western towns (blue). Figure B2 shows coefficients of unprotected forest class for the time period are highly significant in towns northwest of Boston including Lexington, Winchester, Woburn, and Belmont. Finally, the R^2 values are examined to measure the fit of GWR in each town displayed in Figure 3.3 for year 2013. Metro Boston has good GWR R^2 values. The towns of Bernardston, East Longmeadow, Hampden, Longmeadow, Ludlow, and Springfield have higher R^2 values. These towns are located in the central and western regions of the state. Bernardston (Route 91 south of NH Border) had the highest GWR R^2 value. In contrast, Seekonk, Somerset, Swansea, Taunton, and Freetown had lower GWR R^2 values. These towns are all located south of Boston closer to Providence RI and therefore may be impacted by the determinants in Rhode Island. Table 3.5 shows the overall results in terms of bandwidth, RSS, AIC and Adjusted R^2. The R^2 value was the highest in 2000 and averages around 0.80 over other years. To summarize, most determinants in our study are characterized by spatial non-stationarity.

Decennial Census Years			ACS Years			
Year	**2000**	**2010**	**2009**	**2011**	**2012**	**2013**
Bandwidth	84	98	98	91	77	90
RSS	531,883.1	111,8276	1,166,110	951,077.8	633,734.2	786,608.1
AIC	3684.43	3888.25	3904.78	3858.53	3778.99	3798.11
Adjusted R^2	0.86	0.81	0.80	0.79	0.84	0.81

Table 3.5. Basic geographically weighted regression (GWR) diagnostics for home prices in MA towns. Bandwidth, RSS, AIC and adjusted R^2, for decennial years (2000 and 2010), ACS years (2009, 2011–2013).

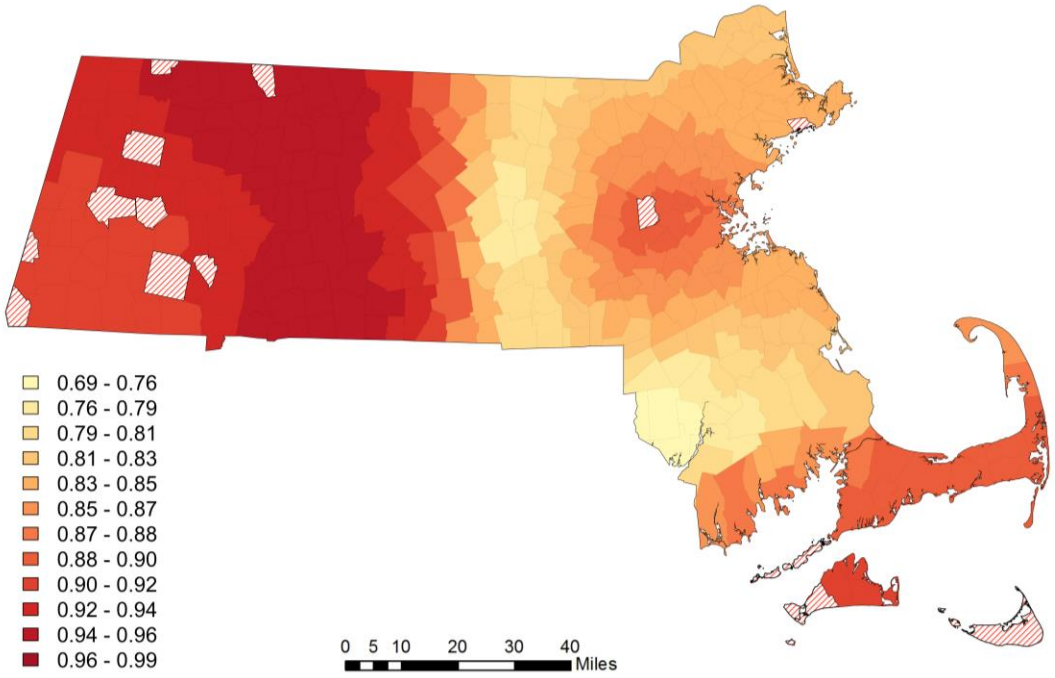

Figure 3.3. Local R^2 of Basic GWR for 2013

3.3.2 Are all determinants of median home price non-stationary?

We next examine the question of non-stationarity of determinants using MGWR that differentiates the spatially non-stationary from stationary determinants. Results are shown in the Tables B7–B12 where variables with *p*-value (Monte Carlo simulations) in column 7 greater than 0.05 are spatially stationary and should be treated as fixed global variables (listed in the lower section of the table). Vehicle ownership and residential taxes should be treated as fixed global variables in most years. The importance of vehicle

ownership is uniform across the state, as are town taxes. These should be treated as fixed global variables, while MCAS CPI scores were spatially non-stationary (local) in 2009 and 2011. Population density presents contrasting results in GWR and MGWR analysis. It is not a fixed global variable in MGWR suggesting spatial non-stationarity; population density was not significant in the GWR analysis until 2012–2013, coinciding with the increase in population in metro Boston area in this time period.

A closer examination of unprotected forest indicates that towns such as Belmont, Billerica, Carlisle, Lexington and Burlington, within the Greater Boston region, consistently perform highest on this coefficient indicating the greatest potential for residential development in these towns, if town regulations permit. On the other hand, towns such Provincetown (on the Cape), North Attenborough (on the intersection of Routes 95 and 295), and Norfolk rank the lowest for at least two years on this coefficient, indicating they have less potential for development.

Figure B3 displays unemployment coefficient maps of the GWR and MGWR that differentiate the spatial non-stationarity in this determinant around Worcester (west of Boston) and suburbs south of Boston. Similarly, results are shown for the display of senior population coefficient maps of the GWR and MGWR in Figure B4. Differences emerge in the western suburbs of Boston. Both determinants are non-stationary using GWR and MGWR but are producing different spatial patterns of coefficients around Boston. They produce similar patterns of coefficients in western part of the state.

MGWR results presented in Table 3.6 show better results compared with GWR results shown in Table B7 across both census and ACS years. As shown in Figure B5,

GWR and MGWR have similar AIC in 2009, 2010, but display some differences in the remaining years. MGWR produces consistently the lowest AIC values for all years. GWR is the second-best performer using the AIC measure while OLS is the worst performer. This result highlights the spatial non-stationary of some factors—Population Density, Unemployment Rate, Residential Area, Vehicle Ownership, Senior Population, Distance to Commuter Rail Stations and Property Tax—that influence median home prices in Massachusetts towns. The highest difference between two model results occurs in 2013 indicating that MGWR is a better choice for this year since the nature of spatial non-stationarity of some determinants has changed.

Decennial Census Years			ACS Years			
Year	2000	2010	2009	2011	2012	2013
Bandwidth	84	98	98	84	77	90
RSS	635,797	1,402,213	1,479,242	1,188,077	944,864	922,170
AIC	3666	3883	3905	3840	3765	3778

Table 3.6. Mixed GWR diagnostics for home prices in MA towns. Bandwidth, RSS and AIC for decennial years (2000 and 2010), ACS years (2009, 2011–2013).

3.3.3 Impact of housing on urban sustainability

Economically, our study captures the Great Recession period of bust and its aftermath based on the ACS data 2009, as well as census 2010 and recovery in 2011–2013 ACS data. Determinants such as unemployment rate show differences from 2009 to 2013, while unprotected forest is significant beginning in 2009, suggesting that housing demand is impacted forests in the period of boom that followed the economic recovery. Population density is significant in recent years 2012–2013 and was likely driven by increasing employment in Metro Boston. Our spatial models capture the economic boom and bust in the various towns from 2009 to 2013 displaying the linkages between

economic, social, and ecological factors in urban environments during critical periods. Addressing real estate expansion during periods of boom, one has to consider ecological and social impacts, while during periods of bust, urban planning has to consider the social implications of unemployment and drop in housing value. Social factors such as unemployment impacted certain towns in Metro Boston (such as Natick, Framingham, Wayland, Wellesley, and Somerville) during this time period. This perspective on social determinants may inform policy makers about provisioning social services and employment opportunities in these towns. Ecological factors related to unprotected forest are also significant in Metro Boston towns (such as Bedford, Lexington, Burlington, Woburn, and Waltham), suggesting these towns should balance residential expansion and forest cover in the future.

3.4 Discussion

Our data driven approach to model housing in this study calls for the integration of remote sensing, socio-economic, town and other data. The study covers 336 towns and cities in the state of Massachusetts for the period 2000–2013 that was characterized by periods of boom and bust in the US economy following the housing market collapse in 2008 (Elsby et al., 2010). Our modeling approach of GWR and MGWR, enables us to analyze how spatial non-stationarity of environmental, economic and social determinants lead to changes in median home prices. The basic OLS model appears least suited in modeling median home prices, while both GWR and MGWR models using adaptive bandwidths perform better.

The analysis presented here can help us address a series of questions focused on social determinants: What is the relationship between an educated workforce and home prices? Boston and Cambridge are home to sectors that attract educated workforce with high-wage jobs. Hence, people with higher education and income move closer to the suburban towns around Boston. This finding is shown in our analysis with marked differences between 2009 and 2013. What is the impact of senior citizens on home prices? Our GWR and MGWR suggests that senior citizens living in close vicinity to metro Boston exerted an influence on median home prices in 2013 and not in the period 2009–2012. Perhaps this segment of population held on to their homes that depreciated in value (in 2008) and sold their home starting 2010. The ownership of homes by seniors and its impacts on home prices deserve further empirical scrutiny. We used public school CPI scores as a determinant on home value. We found that these scores were significant using GWR model but were fixed as global variables using MGWR in 2000, 2010, 2012, and 2013. For future studies, we could incorporate other test scores (such as English Language Arts) to obtain a more nuanced understanding of the correlation between test performance across school districts in each town and their corresponding home prices. What are the implications of spatial non-stationarity on environmental determinants? Unprotected forest cover in the eastern part of the state is more valuable than in the rural western Massachusetts. This may give towns an opportunity to build resilient communities that can preserve unprotected forest or mobilize citizen activists to transform the unprotected forest into protected forest. This investigation has implications in urban planning as it can guide us in conservation of key species. Work presented here

can be extended to address reducing ecological footprints, reducing carbon emissions through reduction in deforestation, and conservation in towns that could benefit from taking actions to be resilient in the future. How does unemployment impact home prices? Unemployment rate is significant in determining home prices in all years using the GWR model. Suburban towns exemplify this pattern, indicating that an increase in unemployment causes a drop in home prices. The model predicts a negative relationship between these determinants and home prices across most towns. Property tax is another significant variable in our GWR analysis in determining home prices.

Housing impacts household behavior, policy, and the environment and therefore directly relates to urban sustainability (Saiz, 2010). Housing construction includes land, energy, and materials. Housing is directly connected to transportation and other externalities since people need to move back and forth from their homes for employment, recreation and other activities (Mills, 1967; Muth, 1969). Given the complex connectivity between housing, transportation, and other economic sectors, sustainability in housing has typically been examined as a function of resource and location efficiency (Koebel et al., 2015). Home prices are related to environmental determinants, as described in our paper and elsewhere (Rauterkus et al., 2010; Tsatsaronis & Zhu, 2004; Tu & Eppli, 2001), that will be transformed given the sustainability focus of many cities, including Boston. Attributes such as green homes and solar panels may increasingly impact future home prices. Many new housing units are built with a focus on energy efficiency. Modern urban architecture emphasizes eco-friendly, or “green” homes that use materials and building methods with less energy requirements as well as result in reduced energy

bills for the homeowner. The decreasing price in solar panels and their increasing availability will impact home prices in the future. Accessibility to public transport is relevant in this paper today but in the future, accessibility to electric car charging stations may become equally important. We discussed the nature and valuation of open space that is relevant in any discussion of urban sustainability as cities lower energy consumption, reduce emissions, and promote healthier lifestyles. Our approach can incorporate and estimate the non-stationarity of these sustainability driven factors that may impact future home prices. Thus, home prices can inform us of the changing public attitudes towards sustainability in cities across the US in the coming decades.

3.5 Conclusions

We proposed an urban sustainability framework centered on housing that could address issues involving economic, social and ecological dimensions. We used GWR model to examine spatial non-stationarity of key economic, ecological and social dimensions that are easily available at a larger spatial scale. Traditional approaches both in economics and in spatial econometrics involve hedonic and spatial models that estimate prices of individual homes using block level census data or town parcel data. Such models do not account for spatial non-stationarity at a regional or state scale. Our approach is to address urban sustainability at a broader spatial scale to encapsulate the entire state. We show that a data driven approach to modeling home prices at this broad scale calls for integration of geospatial data from a number of sources, including remote sensing, and census data to address multiple facets of sustainability. The understanding of

key determinants to housing enables towns to take different strategies to sustainability such as valuation of green spaces and proximity to public transport, while unemployment rates and public-school performance can help shape urban housing policies. We hope to incorporate other variables such as green homes, residential solar panels, charging stations, air quality and other components to make our work relevant for assessing future urban sustainability.

Acknowledgements: This research is based upon work supported by the National Science Foundation under Grant No. CNH #1617053, CNH-S and Boston University Initiative on Cities, 2017.

CHAPTER 4 – Mapping the Dynamics of Aquatic Vegetation in Lake Kyoga and Its Linkages to Satellite Lakes

This study focuses on the dynamics of aquatic vegetation in the water bodies of Lake Kyoga in East Africa. Lake Kyoga is one of the outflows of Lake Victoria, which is the second-largest freshwater lake in the world. As a result, the inflow from Lake Victoria strongly influences Lake Kyoga's water balance. The fluctuation in the water level would cut off the linkages from the satellite lakes to the main lake. Furthermore, aquatic vegetation usually covers these linkages, creating a refugia of a hypoxic environment that protects native fish species from Nile perch, a near-top-level predator introduced in the 1950s. Societies that reside in the Lake Kyoga region have traditionally fished in and farmed beside the waters of this lake. Any changes in the CHANS organizational coupling between aquatic vegetation and ecosystems of the lakes can adversely impact both the endangered species and the surrounding human communities that depend on fisheries and other ecosystem services provided by the lake.

This research aims to develop a new and improved algorithm to map the spatial distribution and dynamics of floating and emergent aquatic vegetation. The study utilizes a time series of 440 Landsat images dating from 1986-2020. A series of water and vegetation indices are designed to characterize and map the maximum extent of water bodies and identify aquatic vegetation. First, two types of water masks are derived using a majority rule - a separate water mask for each image and a composite water mask of the region over the study period. Second, the difference between the two masks is then used to delineate the potential location of macrophytes over the image. Third, an algorithm is

developed to separate the floating vegetation from emergent vegetation; this algorithm uses Landsat spectral bands and two additional spatial and temporal metrics that considerably improve classification accuracy. The algorithm and data products developed in this research are valuable in shaping strategies to conserve fish species and human livelihoods in this region. The goal of this research is to design a decision support system focused on environmental risk and resource management related to fisheries in the Lake Victoria basin.

4.1 Introduction

The Lake Victoria Basin (LVB) is transforming with rapid increases in human population, climate change and geopolitical tensions for access to natural resources, including water rights and fisheries. Lake Victoria (Figure 4.1) is the world's second largest freshwater lake by surface area bordered by Kenya, Tanzania, and Uganda. The LVB is part of the East African Rift valley lakes containing about 25% of the Earth's unfrozen surface fresh water and 10% of the world's fish species. About 40 million people live in this region, with an annual growth rate of 3.5 percent (compared with the global population's annual growth rate of 1.05%). LVB is facing environmental and social challenges. Hence the CHANS framework applied in this context enables us to model the intrinsic coupling between the natural environment of the basin and its surrounding ecosystem and the population living around the basin supported by ecosystem services including fisheries, forests, wetlands, and rangelands.

Anticipated changes in climate pose a great risk to human wellbeing and natural resource flows in LVB that imperil the character and tradition of the system via reduced

food security due to reductions in fish yield (the principal protein source for people); reduced access to safe drinking water; and the irreversible loss of native species that underlie the region's exceptional natural capital (Cohen et al., 1996). LVB is at the epicenter of unique biodiversity, harboring more than 350 species of cichlid species. Although cichlids are small fish, they were a major food resource for 60 million people in the three countries surrounding the lake. The value of the catch is estimated to be more than US$ 550 Million (Outa et al., 2020). The problems facing LVB are significant reduction and risk of extinction of some fish species, including the endemic tilapiine species (Oreochromis esculentus and Oreochromis variabilis), catfishes (Xenoclarias eupogon), haplochromines and cyprinids (Labeo victorianus and Barbus altinialis) in the lake. (Achieng et al., 2020; Njiru et al., 2018; Tungaraza et al., 2012; Yunana et al., 2017). The reduced biodiversity, as well as extinctions within the lake, has been attributed to the introduction of alien fish species, e.g., Nile perch (Lates niloticus) and Nile tilapia (Oreochromis niloticus) into the lake resulting in habitat loss and cultural eutrophication (Marshall, 2018; Yongo et al., 2018). If left unchecked, these changes will have devastating effects on the lake's resources, as well as on people living around the lake and beyond who depend on the lake fisheries for their livelihood. Additionally, climate fluctuations impact the distribution of Nile Perch, which is favored over cichlids resulting in a complete change in the lake ecosystem. (Marshall, 2018; Yongo et al., 2018). Thus, this region is a hotspot of global biodiversity risk stemming from anthropogenic activity as well as climate change.

The focus of this study is to examine one facet of the CHANS related to the spatial dynamics of aquatic vegetation in a satellite lake called Lake Kyoga in LVB. Areas of Lake Kyoga are covered by aquatic vegetation, a distinct type of photosynthetic vegetation that grows periodically or permanently in water bodies (Twongo et al., 1995). This vegetation serves a critical ecological and biological role in inland water bodies, maintaining water quality, regulating oxygen levels, providing food and shelter for organisms and wildlife. Aquatic vegetation is characterized by rapid biomass accumulation and a faster growth rate (Gunnarsson & Petersen, 2007). Areas of Lake Kyoga less than 3 meters deep are completely covered by water lilies from the lake shores to the open water zone, while much of the swampy shoreline is covered with papyrus and the invasive water hyacinth (Eichhornia crassipes (Mort.) Solms (Pontederiaceae)). Aquatic vegetation colonized the habitat and can be divided into the following groups: emergent - rooted in the sediment with foliage extending into the air (papyrus, Cyperus papyrus), floating - plants that float on or under the water surface with or without roots in the sediments (water lilies, Nymphaea spp, and water hyacinth, Eichhornia crassipes), and submerged -plants that grow submerged with roots in the sediment (elodea, Elodea canadensis) (Thomaz et al., 2008).

Satellite remote sensing is the only feasible and economic means of continuously monitoring Earth's surface at multiple spatial and temporal scales for the identification and change detection of land cover. Prior to satellite remote sensing, traditional in situ measurements and aerial images were utilized to identify and map aquatic vegetation (Marshall et al., 1994, Welch et al, 1994). Current remote sensing technologies offer

better alternatives ranging from moderate-resolution imaging spectroradiometer (MODIS) satellite images with coarse resolution but higher frequency (Fusilli et al., 2013; Zhang et al., 2016) to Landsat thematic mapper (Donchyts et al., 2016), enhanced thematic mapper plus (ETM+); as well as QuickBird, IKONOS, and PlanetScope high-resolution images resulting in better mapping and change detection (Gabr et al., 2020). PlanetScope provides images of the entire land surface of the Earth every day (a daily collection capacity of 200 million km^2/day) at approximately 3 meters per pixel resolution. It offers limited but higher resolution data for validation and ground-truthing. Both data streams can be used to monitor changes in the aquatic vegetation of the lake basin.

Rivera et al. (2013) used a simple band ratio and NDVI as input datasets for regression trees to monitor aquatic vegetation cover. Villa et al. (2015) used a rule-based classification scheme for mapping different aquatic vegetation types based on vegetation indices. Both studies showed decent accuracy in extracting floating vegetation from Landsat Images but less than satisfactory results in classifying submerged vegetation and separating emergent vegetation from terrestrial vegetation. This is similar to previous studies on aquatic vegetation based on remote sensing data (Albright & Ode, 2011; Birk & Ecke, 2014; Jiang et al., 2012; Zhao et al., 2012). Pixel-based classification has relatively less accuracy for certain aquatic vegetation classes due to sensor resolution limitation and mixed pixels, where vegetation classes are confused with one another. Another study used UAV (Unmanned aerial vehicle) to collect high resolution, sub 1 meter, images for visual interpretation and manual mapping to monitor aquatic plant

species and achieved an accuracy of over 90% (Husson et al., 2014). Studies on aquatic vegetation monitoring via remote sensing mostly rely on in situ sample collection, which is labor and time consuming and often covers a small section of the study area (Cavalli et al., 2009; Damtew et al., 2021; Thamaga & Dube, 2018). UAVs are cost-efficient; however, the data collected are rarely consistent in terms of quality and temporal resolution compared to satellite observations.

This research aims to derive thematic products, delineating floating vegetation and emergent vegetation maps from multi-temporal medium-resolution Landsat Images with the aid of additional spatial and temporal metrics. This research addresses the design and derivation of the thematic products. Ultimately, the derived data products would capture the ecosystem dynamics to develop a decision support system dedicated to environmental resource management, including threats to the native fish population and fisheries.

4.2 Materials and methods

4.2.1 Study area

Lake Victoria is the largest freshwater lake in Africa and the second largest in the world. Lake Kyoga, an outflow of Lake Victoria, is located north of Lake Victoria in central Uganda. As a result, Lake Kyoga's water balance is strongly influenced by the inflow from Lake Victoria. The fluctuation in the water level would cut off the linkages from the satellite lakes to the main lake. Furthermore, the linkages are usually covered by aquatic vegetation (i.e., papyrus swampland and water hyacinth), which would create a

refugium of hypoxic environments that protect native fish species from Nile perch, a near-top-level predator introduced in LVB. Nile perch and four tilapiine species were introduced in the late 1950s and early 1960s (Mbabazi et al., 2004) in LVB to improve the stocks of the declining fishery but have led to a massive decline of native species. Haplochromine species and trophic groups have been depleted from this lake as well as the other lakes in the LVB. As stocks of introduced species increased, those of the native species declined rapidly or disappeared altogether. The contribution of haplochromines dropped from about 80% of the fish biomass in Lake Victoria in the 1970s to less than 1% in the 1980s (Ogutu-Ohwayo, 1990).

Lake Kyoga (south) and two connected lakes, Kwania and Kojweri (north) in Uganda, are the focus of the study. We use the term Lake Kyoga to describe all three connected lakes. Victoria Nile is the only outflow of the entire LVB that passes through Lake Kyoga. The lake has a depth ranging from less than 3 meters to about 5.7 meters. Lake Kyoga has 14 fish species (Mbabazi et al., 2004), with an additional 37 species in its satellite lakes (Ogutu-Ohwayo, 1990). In the Kyoga satellite lakes, harbor haplochromine species and trophic groups (not found elsewhere) are separated from the main lake by papyrus swamps; swamps serve are effective barriers as fish movement across the swamps is limited to hypoxia‐tolerant species (and not the Nile perch). Thus, satellite lakes of the LVB are potential refugia for the native fish species, enhancing trophic diversity and serving as a significant fishery resource. Therefore, analyzing vegetation dynamics of the satellite lakes, based on new metrics derived from remote

sensing time series data, may offer new critical insights. The fringe of the shoreline between land and water is occupied by papyrus and swampland.

Swamps separate the Kyoga satellite lakes from the main lake; there was no introduction of exotic species in most of the satellite lakes. These lakes share water via the papyrus swamps, but fish movement across the swamps is limited to hypoxia‐tolerant species (Chapman et al., 1996). Nile perch requires high concentrations of oxygen, while some haplochromines can tolerate a low oxygen environment for extended periods. The swamps are thus effective barriers to the dispersal of Nile perch (Kaufman et al., 1997). These satellite lakes harbor some haplochromine species and trophic groups depleted from the Kyoga and Victoria main lakes. Hence satellite lakes have been characterized as possible refugia for the Lake Victoria fish species and trophic diversity as well as a significant fishery resource. Apart from overexploitation and invasive species, major systematic drivers of loss of aquatic species include land cover and land-use change (LUCC). Figure 4.1 shows the location of Kyoga in LVB and false-color Landsat images of Lake Kyoga in Panel a (on the right side), Lake Nawampasa in panel c, and some satellite lakes are shown in panel d.

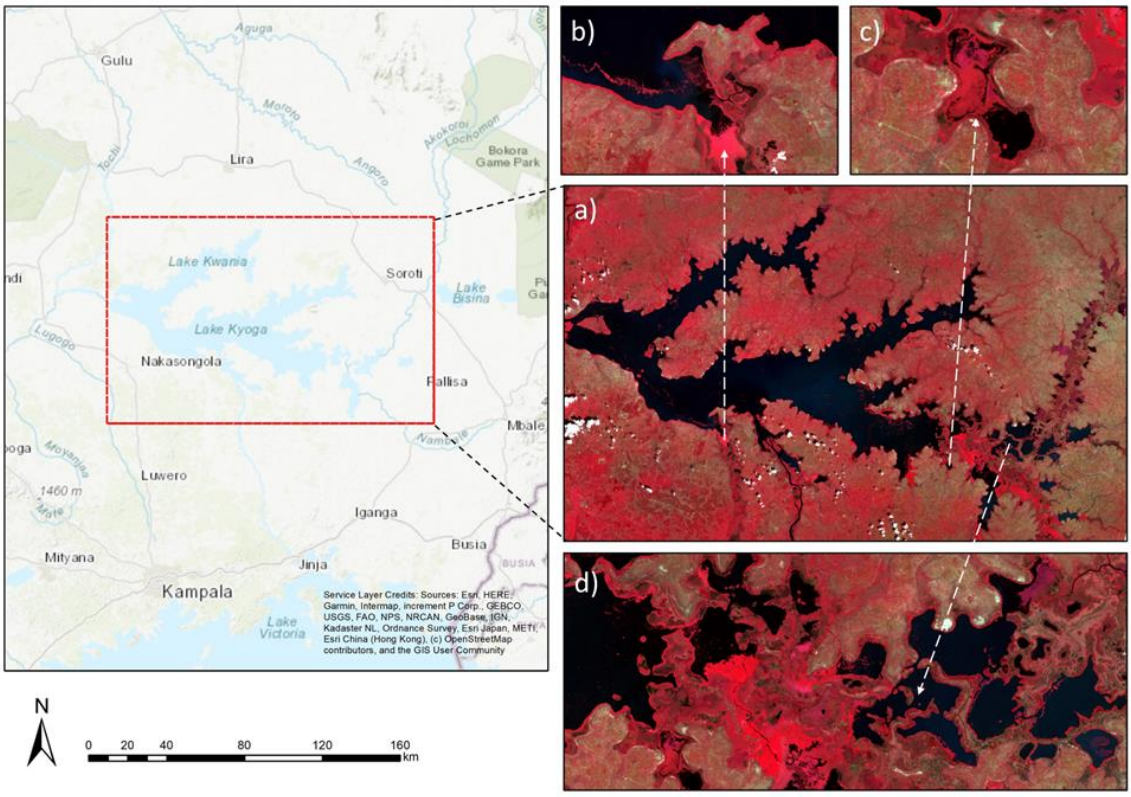

Figure 4.1. Location and false-color Landsat images (December 14, 2016) of Lake Kyoga. a) Lake Kyoga. b) Aquatic vegetation near Victoria Nile. c) Lake Nawampasa. d) Satellite lakes

4.2.2 Data

This research utilized imagery from different satellites and spatial resolutions. They are Landsat, from January 1986 to April 2020, at 30 meters resolution, and PlanetScope archived from June and December during the period 2015-2020, at 4.77 meters resolution.

4.2.2.1 Landsat

Landsat is the source of the data for change detection. Landsat scene WRS-2 127-043 covers almost the entire Lake Kyoga. A total of 449 scenes from 127-043 from

Landsat 4 and 5 TM, Landsat 7 ETM+ and Landsat 8 OLI images are included in this analysis, starting from January 1986 to April 2020. We used an atmospherically corrected surface reflectance data product, where cloud and cloud shadow pixels were masked using the Pixel Quality Assurance (*pixel_qa*) band, provided with Landsat Surface Reflectance products (*Landsat Collection 1 Level-1 Quality Assessment Band, n.d.*).

Early data observations were limited; 2 to 4 years of data gaps existed before 1999. Only 20 images were available until 1995, followed by a period of four years with no images. Starting from 1999 and onwards, the observation density (the number of images per year) increased due to Long-Term Acquisition Plans and the launch of Landsat 7 (Pickens et al., 2020). The study included early observations since they can still provide invaluable insights regarding the dynamics of aquatic vegetation, which is closely linked to the eutrophication of water bodies.

The primary constraints in mapping aquatic vegetation are the difficulties differentiating them from terrestrial classes based on spectral bands alone in a single image. It is challenging to separate marginal vegetation on the shores from aquatic vegetation using just one optical remote sensing image, especially when dense. However, because of the movement and life cycle of aquatic vegetation, it is possible to map aquatic vegetation when there are enough observations. With enough images, we can delineate the maximum extent of the lake by combining clear water pixels from all available images. To address this constraint, we first identify the spatial location and the dynamics of movement of aquatic vegetation captured by the time series of imagery. Second, we identify the type of vegetation- floating and emergent. While emergent

vegetation is in the edge or margins of the lake, floating vegetation can be located in the middle of the lake. Further, spatial and temporal metrics can facilitate identifying the location and type of aquatic vegetation, floating and emergent.

Due to the difficulty of separating the terrestrial vegetation from aquatic vegetation based on a single image, we used all Landsat images available for the study period to map the maximum extent of Lake Kyoga. In addition, we also used all images to analyze the change in aquatic vegetation coverage, semi-terrestrial vegetation, and open water.

4.2.2.2 NICFI imagery

The second data source for this research consists of PlanetScope's tropical mosaic, high-resolution images from Norway's International Climate and Forests Initiative (O'Shea, 2021) satellite program. This data enables us to identify the type of aquatic vegetation for ground-truthing. The PlanetScope Surface Reflectance Mosaics are optimized for scientific analysis, display, and visual interpretation (*NICFI DATA Program*, 2021). The mosaic covers global tropic regions with sub 5 meters per pixel resolution per band (visible bands and near-infrared bands). In terms of temporal resolution, images are available every June and December between 2015 and August 2020, monthly after September 2020.

We used tropical mosaic images consistently available in June and December as ground truth for training and reference data collection. The mosaics are assembled from individual date images. The exact dates of acquisition of the PlanetScope images are unclear, making it difficult to determine the lapse of time between ground truth and

Landsat images. However, we have confidence in using PlanetScope's high-resolution data for ground-truthing since the growth of aquatic vegetation in a short amount of time (number of days) is not expected to create a significant difference in type and coverage of vegetation.

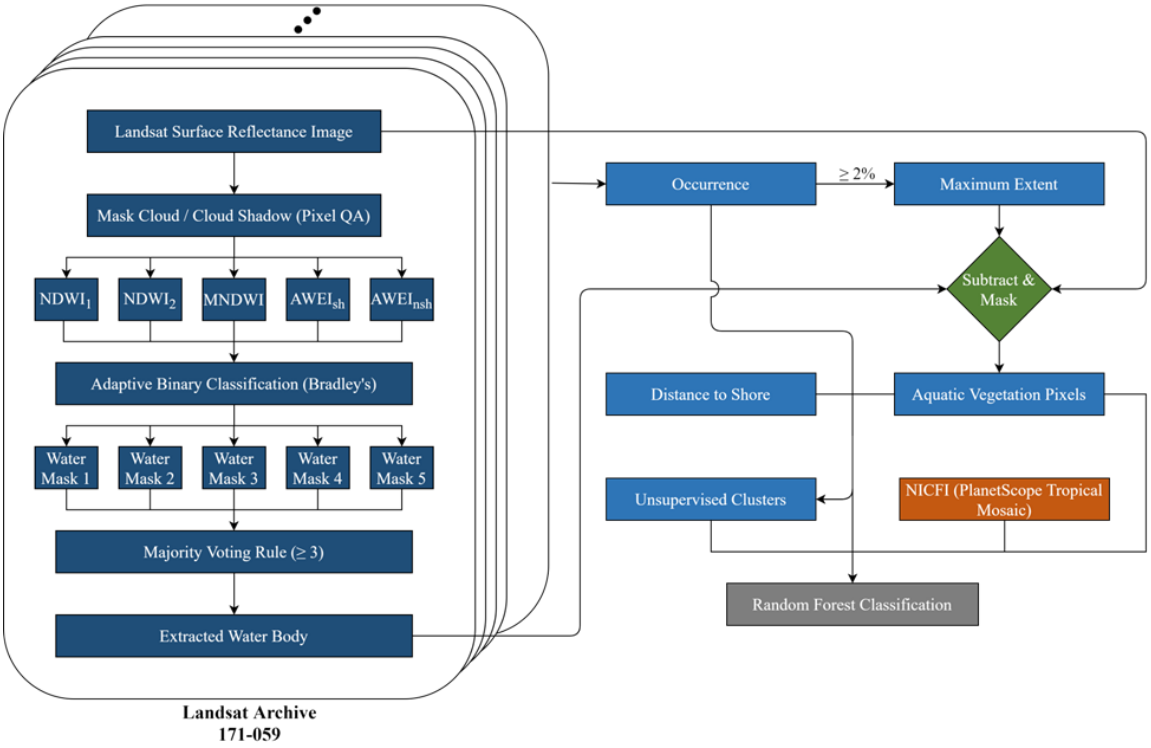

Figure 4.2 Methodological flowchart showing steps in the extraction of water bodies (left) and classification of aquatic vegetation (right).

4.2.3 Methods

Figure 4.2 shows the methodology used in this study. The left panel shows the processing of Landsat data to derive the water body, and the right panel shows the steps involved in locating and classifying aquatic vegetation. The data of the Landsat archive in the left panel of Figure 4.2 shows steps in processing, including cloud masking, derived water indices, adaptive binary classification, water masks, majority voting, and finally,

extracted water body. The final product, combined extracted water body across the time period of observations, serves as the input to the right panel focused classifying aquatic vegetation. Details are described below.

4.2.3.1 Water indices

The first step in our study is the identification of the water body. When and where water is found on the earth's surface is highly significant since it influences the climate system, species distribution, food, and water security of regions and countries. Global to regional surface water dynamics has been recorded using satellite imagery. Prior research has focused on the development of various water indices to highlight water bodies in remote sensing images across the globe (Donchyts et al., 2016; Hui et al., 2008; Li et al., 2016; Shen & Li, 2010). The simplest measure is the Normalized Difference Water Index (NDWI) that was introduced to delineate open water features using the green band and a single near-infrared band (McFEETERS, 1996), where all positive values would be labeled as water, and negative values would be labeled as non-water. The next measure is NDWI (Rogers & Kearney, 2004), which uses the red and short-wave infrared bands for water extraction. However, NDWI was found to misclassify built-up surfaces as water when a threshold value of 0 is applied (Xu, 2006). This limitation was overcome in the modified NDWI (MNDWI) method using the green and short-wave infrared bands (Xu, 2006). Thus, NDWI has become one of the most widely used water indices for surface water mapping (Feyisa et al., 2014).

The disadvantage of using two-band ratio water indices for surface water mapping is that the accuracy suffers when the albedo of background landcovers is low, such as

cloud shadows, mountain valleys, asphalt surfaces, etc. (Feyisa et al., 2014). Automated Water Extraction Index (AWEI), a multiband water index, was introduced to improve the accuracy of surface water mapping by suppressing classification noise from dark surfaces and shadows (Feyisa et al., 2014). There are two formulas for the AWEI; while $AWEI_{nsh}$ is optimized for images where shadows and dark surfaces are not a significant concern, $AWEI_{sh}$ is designed for effectively eliminating shadows and dark surfaces from classification noise. However, $AWEI_{sh}$ may misclassify land cover with high albedo as water (Feyisa et al., 2014). Thus, in this study, both versions of AWEI are used as water indices to overcome the limitations of the individual index.

Index	Formula	References
$NDWI_1$	$(\rho_{Green} - \rho_{NIR})/((\rho_{Green} + \rho_{NIR})$	McFeeters, 1996
$NDWI_2$	$(\rho_{Red} - \rho_{SWIR1})/(\rho_{Red} + \rho_{SWIR1})$	Rogers and Kearney, 2004
MNDWI	$(\rho_{Green} - \rho_{SWIR1})/((\rho_{Green} + \rho_{SWIR1})$	Xu, 2006
$AWEI_{sh}$	$\rho_{Blue} + 2.5 * \rho_{Green} - 1.5 * (\rho_{NIR} + \rho_{SWIR1}) - 0.25 * \rho_{SWIR2}$	Feyisa et al., 2014
$AWEI_{nsh}$	$4 * (\rho_{Green} - \rho_{SWIR1}) - (0.25 * \rho_{NIR} + 2.75 * \rho_{SWIR2})$	

Table 4.1. Water indices used in this study (band designation according to Landsat satellites)

In this study, as shown in Table 4.1, two versions of NDWI and MNDWI are used as the three water indices. Table 4.1 shows each index's formulae used in this research and lists prior research that utilized each index.

4.2.3.2 Waterbody delineation

The single band thresholding method is widely used to separate surface water from the background because of its ease of use, computational efficiency, and greater

precision in specific conditions (Hui et al., 2008). However, water bodies are seldom clear. Variations in water turbidity, depth, chlorophyll, etc., make it challenging to use spectral indices with a single threshold which is often manually determined, to delineate water bodies from other land covers. Otsu thresholding (Otsu, 1979) is deployed to perform automatic image thresholding based on a histogram of spectral index values located within the edge buffer zones of water bodies. A threshold value is detected that best separates the two groups, water, and non-water. A global threshold per image is still not satisfactory in a highly variable water color condition, as it fails to differentiate water and non-water.

We adopted the Bradley method (Bradley & Roth, 2007), an adaptive binary classification that binarizes the image using a locally adaptive threshold. It computes a threshold per pixel using the local mean intensity within the window of the pixel (that could be ⅛ of the image). We used this method on the five versions of water indices (See Table 4.1) for each Landsat image to extract foreground features (water bodies) from the background.

Even improved water indices cannot accurately highlight water bodies in areas with low-albedo surface backgrounds. Water indices have their strengths and weaknesses in addressing water features and suppressing non-water features. To reduce noise, we utilized a majority rule: for a pixel to be labeled as water, in one Landsat image, at least three binary classification results out of the five water indices (described in Table 4.1) have to agree on the label. This step results in the derivation of the water mask for the given Landsat scene. Finally, the pixels with an NDVI value greater than 0.3 are removed

from the water mask to eliminate the pixels where some dark vegetation is misclassified as water.

A recent product called the JRC Global Surface Water provides monthly water history (extent), maximum extent, and occurrence. The JRC maps are generated using Landsat 5, 7, and 8 satellites, where the pixels were classified using an expert system (Pekel et al., 2016). This product provides the location and temporal distribution of global surface water from 1984 to 2020. A metric called monthly water history is estimated by combining water body masks for each month over the time period. Each month may have more than one Landsat image. Thus, the JRC Global Surface Water product is useful for deriving monthly water history by combining water body masks for each month over the time period. However, JRC Global Surface Water product cannot extract aquatic vegetation extent for each individual Landsat image due to the aggregation over the month. Thus, JRC Global Surface Water is used to validate our water extent product.

4.2.3.3 Characterizing the maximum extent of Lake Kyoga and potential locations of aquatic vegetation

Unlike open water, extracting the actual extent of water bodies from a single image is not trivial due to the dynamic nature of aquatic vegetation. However, the movement and life cycle of aquatic vegetation facilitate determining the maximum extent by combining all extracted water bodies from each Landsat image. We propose three mapped products, Aggregated Water Occurrence (AWO), Aggregated Maximum Extent (AME), and Aquatic Vegetation Extent using the following steps:

1) Cell-wise sum of water bodies. For each water body layer, we estimate the pixel value. We give a value of 1 for water and 0 for non-water. In this step, we sum up or aggregate all water pixels.
2) Aggregated Water Occurrence (AWO). AWO is the frequency with which water was present in each pixel covering the study period of Landsat observations. It is the ratio between the number of times the pixel is classified as water and the number of valid observations (not counting cloud, cloud shadow, and Landsat & ETM+ SLC-off gap).

$$Occurrence = \frac{\#\ of\ water\ presence}{\#\ of\ valid\ observations - (\#\ of\ cloud + \#\ of\ cloud\ shadow)}$$

3) Aggregated Maximum extent (AME). Areas with occurrences greater than 2% (to reduce noise) are established as the maximum extent of Lake Kyoga.
4) Aquatic Vegetation Extent. The potential location of aquatic vegetation for a specific image is estimated from the difference between the maximum extent (of water) image and the water mask of the image (see Figure 4.3).

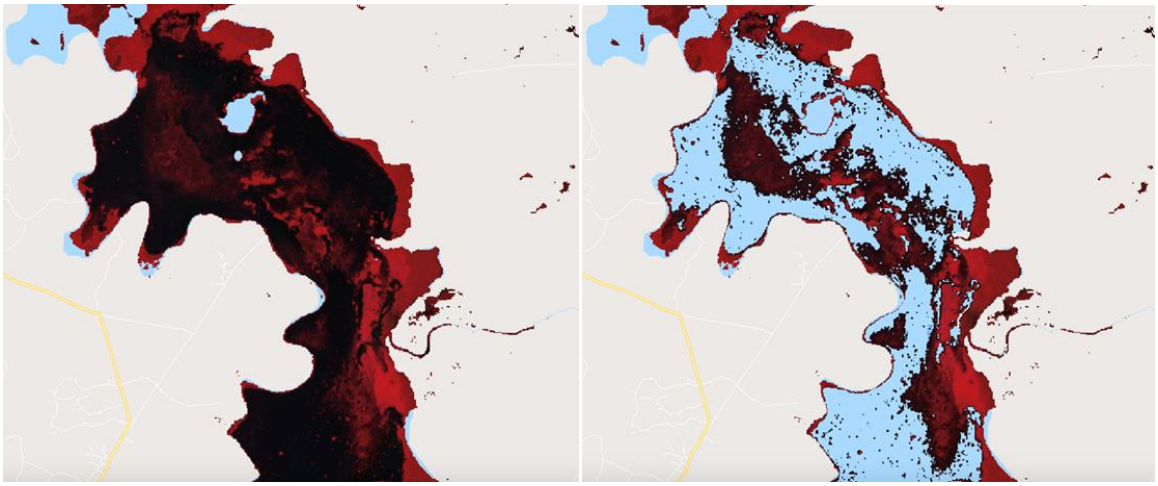

Figure 4.3. Left: shows Landsat 8 OLI image (March 23, 2018) masked by maximum extent. Right: Same Landsat 8 OLI image masked by the difference mask. (of the AME and water mask of Landsat image on March 23, 2018)

4.2.3.4 Adding spatial and temporal metrics to spectral bands

Similar to terrestrial vegetation, aquatic vegetation has the same spectral characteristics, high reflectance in infra-red wavelength and low in red and blue wavelengths (Cavalli et al., 2009). At the micro-level, the leaf density, orientation, and distribution are important. Plants with broad leaves and dense canopy reflect more light. At the macro level, not only the biomass and density influence the spectral signal, but also the water in the background. Because water absorbs electromagnetic radiation strongly in the visible and shortwave infrared spectrum, reflectance measurement of floating aquatic vegetation is highly variable when mixed with or submerged in water (Silva et al., 2008). In addition, the bottom of water bodies affects the absorption and scattering of light, which also needs to be considered when interpreting optical satellite images of aquatic vegetation in shallow water (Silva et al., 2008).

In order to improve the accuracy of aquatic vegetation classification, occurrence and distance to the shoreline (see Figure 4.2 right panel) were added to the spectral bands of Landsat Images:

- Occurrence is the frequency with which water was present. It is the intermediate step when deriving the maximum extent of Lake Kyoga. Thus, occurrence acts as a temporary measure in our estimation algorithm.
- Distance to the banks of the lake is a spatial measure of how far the aquatic vegetation is to the boundary of the maximum extent of Lake Kyoga. It was

calculated on the Google Earth Engine using the cumulative cost function (*Cumulative Cost Mapping | Google Earth Engine, n.d.*).

Aquatic plants may be free-floating or rooted in the bottom of the sediments emerging from the water. As a result, floating vegetation pixels should have relatively high occurrence and are near or in the open lake waters, at a greater distance from the bank of the lake. Emergent rooted vegetation is often located in shallow littoral waters with low occurrence.

4.2.3.5 Aquatic vegetation using unsupervised and supervised classification

A combination of unsupervised and supervised image classification was performed. Earth Engine's K-means (unsupervised) clustering (*WekaKMeans | Google Earth Engine, n.d.*) was applied on potential aquatic vegetation pixels of images of June and December covering the period, 2015 to 2019. A total of 40 clusters for each image are derived from the unsupervised clustering. An expert interpreted and merged similar clusters using PlanetScope Surface Reflectance Mosaics. This data served as the ground truth for cluster labeling of the pixels. Finally, the expert (LK) labeled seven clusters. There are three dominant aquatic vegetation classes - high density floating vegetation, low density floating vegetation, and emergent vegetation as shown in Figure 4.4. Additionally, water, semi-terrestrial vegetation, mudflat, and cloud, were included as target classes as well. To summarize, labeling of unsupervised clusters, NICFI PlanetScope Tropic mosaics, and expert labeling resulted in the identification of seven classes. The above steps created training data for random forest supervised classification of images.

Figure 4.4 Papyrus (left) and water hyacinth (right) in Lake Victoria (Photo credit: Ann Weru/IRIN and Shutterstock)

The random forest classification was performed using the *randomForest* package in R (Cutler & Wiener, 2018). The input data consists of spectral bands from Landsat images, occurrences, and distance to the shoreline. Subsequently, the random forest classifier was calibrated using the *tuneRF* function in the *randomForest* package to look for optimal *ntree* and *mtry* parameters. Finally, the classification map and area estimation of each class distribution were estimated.

A total of 500 stratified probability samples were drawn from the 1986 - 2020 aquatic vegetation maps by stratifying along the 3 major classes: water (n=109); Low density floating vegetation (n=70); High density floating vegetation (n=102); Emergent vegetation (n=117) and 4 other target classes: Semi-terrestrial vegetation (n=24); Mudflat (n=11); Cloud (n=7). Each 30m pixel sample was analyzed using Landsat images, PlanetScope Tropic mosaics and Google Earth images. An error matrix was constructed from the reference sample points to assess user and producer's accuracy.

4.3 Results

The analysis enables us to map the extent of open water in the lakes. This water resource is the most accessible to human populations living around the lakes, and provides wide-ranging ecosystem services, including fisheries. The proposed AWO and AME is validated against another method called JRC Global Surface Water product. Next, dynamics of the open water and aquatic vegetation for the main and satellite lakes are compared to understand the trend through time. Finally, we provide results in classification accuracy of floating vegetation with and without consideration of the distance to shoreline and occurrence.

4.3.1 Temporal dynamics and maximum extent of water bodies - AWO and AME

Our first set of mapped products, Aggregated Water Occurrence (AWO) and Aggregated Maximum extent (AME) account for temporal dynamics and the maximum extent of water bodies. We compare AWO and AME mapped products with the JRC Global Surface Water product (2016) to validate our approach. Figure 4.5 left panel highlights differences in percentiles between the proposed AWO and the JRC Global Surface Water occurrence (2016). The 1% and 99% percentile differences map to -0.273 and 0.085 respectively (negative value shows JRC product has larger occurrence). The most significant differences are 1 percentile (shown in yellow). The highest differences between the two are found at the edge of the lake and satellite lakes. Closer to the edge or in satellite lakes, JRC occurrence has a higher occurrence value. Overall, there is very little difference between AWO and JRC occurrence, especially in the middle of the lake, where the differences range between -0.022 to 0.007.

The right panel of Figure 4.5 looks at the maximum extent of water bodies using the two methods; the JRC Maximum extent (green) is very similar to the AME (blue). While JRC Global Surface Water's maximum extent estimated area is 3650.8437 km2, the AME estimated area is 3569.30 km2. The area of overlap between the two is 3511.75 km2, which is 96.2% of JRC's maximum extent and 98.4% of the AME. This analysis shows that our proposed methods, AWO and AME in particular, are valid in capturing the maximum extent of water in the context of aquatic vegetation. As previously highlighted, JRC cannot be used for mapping aquatic vegetation due to its composite nature.

In summary, both AWO and AME are underestimating water areas since our overall objective is to identify aquatic vegetation. Overestimating water will remove the aquatic vegetation pixels.

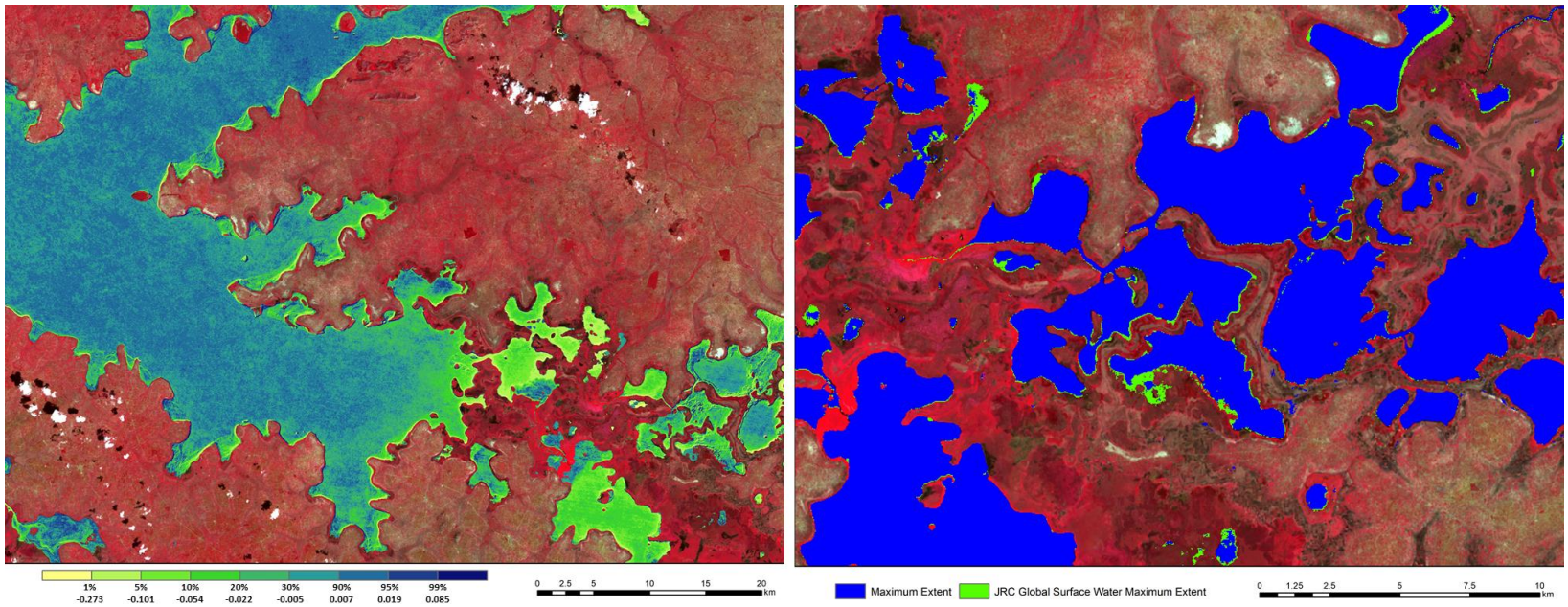

Figure 4.5. Left: Difference between proposed Aggregated Water Occurrence (AWO) and JRC Global Surface Water occurrence is shown as percentiles. Right: Overlay of Proposed Aggregated Maximum Extent (AME) and JRC Global Surface Water maximum extent.

4.3.2 Dynamics of open water in the main and satellite lakes over the entire period

Landsat images are often contaminated with clouds and hence we cannot directly estimate the area of open water; instead, we need to take a ratio estimate of water area to the area of the pixel clear of clouds, yielding an open water ratio for every image. This metric allows us to compare open water dynamics of the main and satellite lakes over the entire period of observation, yielding insights on climate signals and rainfall. Figure 4.6 plots the entire time series of the open water ratio over time that can be summarized as follows:

- The open water ratio is consistent over the entire study period. The open water ratio is around the lower 80% range and then elevated to around 90% during 1999 and 2006 and then a drop to around 80%. This trend agrees with the G-REALM radar altimetry data (*G-REALM - Kyoga, n.d.*), providing us greater confidence in our approach.
- There are high fluctuations in the open water ratio, for example, September 2001, November 2004, and August 2018. These trends are mainly observed in high cloud cover images and indicate that there could be a bias in the estimated open water ratios since the number of valid pixels is reduced in that image.

To summarize, the proposed open water ratio corresponds well with the G-REALM radar water level altimetry data.

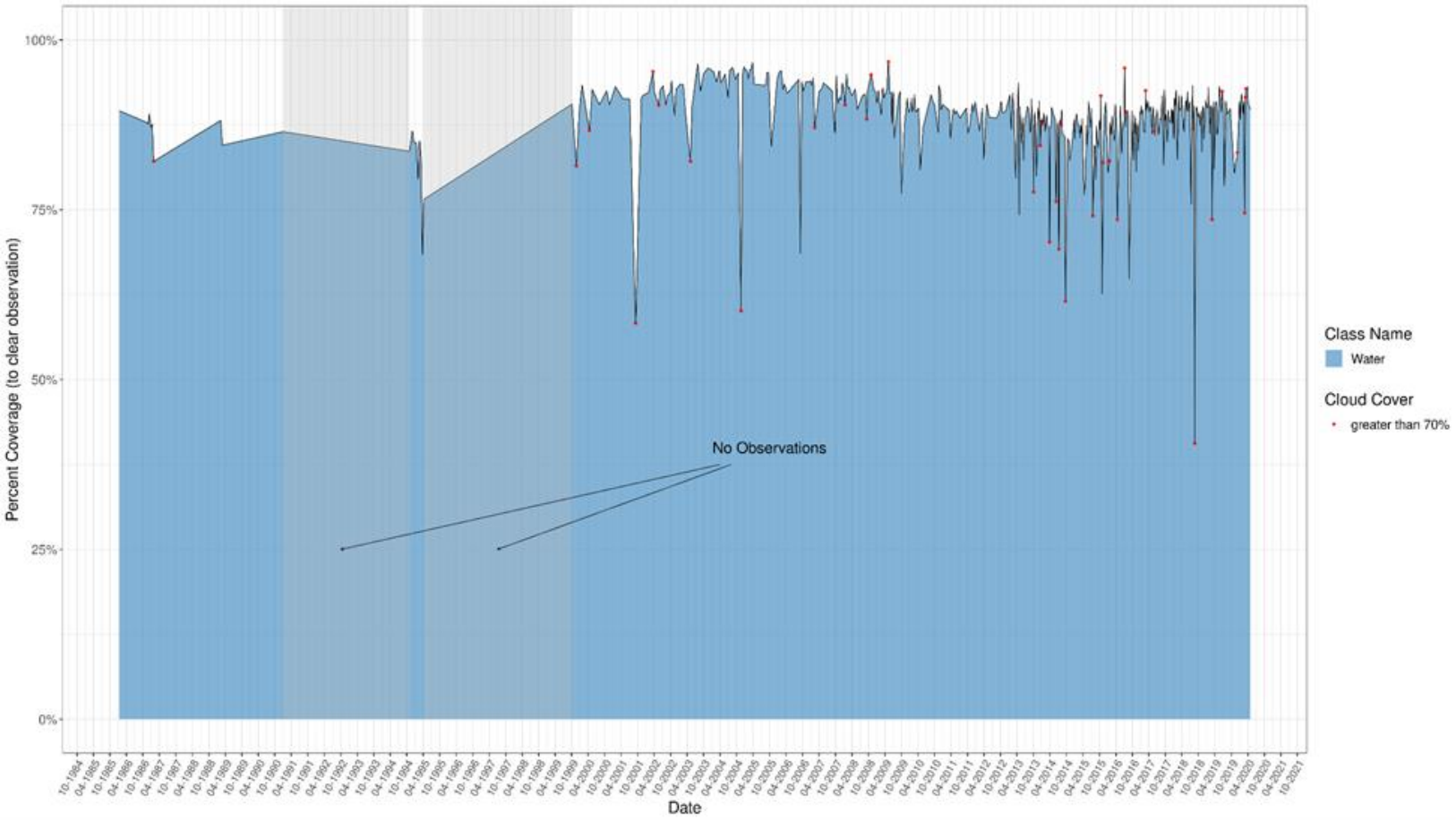

Figure 4.6. Open water ratio (vs. valid observation) for each Landsat image. Images with high cloud coverage (over 70%) are marked in red.

4.3.3 Spatio-temporal changes in aquatic vegetation

Aquatic vegetation is classified using the methodology described in Section 4.2.3.5. Aquatic vegetation includes three types - high density, low density, and emergent vegetation. The vegetation coverages are estimated as ratios of vegetation to the area of the pixel clear of clouds in each image (similar to open water ratios in Section 4.3.2). Areas of vegetation are high during the mid-1990s with a peak value of around 22% recorded in March 1995 (1995-03-24). Vegetation consists of three types - high density (9.6%) and low density (8.3%) floating vegetation, and emergent (6.2%) vegetation around March 24, 1995. Before 1999, the aquatic vegetation ratio was high, probably due to low water level, exposing the shallow lake bottom. However, from 1997 to 2000, the water level increased by around 2 meters, submerged the exposed section of the Lakebed (*G-REALM - Kyoga, n.d.*), resulting in a reduction of aquatic vegetation as shown in Figure 4.7.

- Before 1999, there were two major data gaps, Jan 1991 to Oct 1994, and Apr 1995 to Oct 1999. Missing data during the nineties makes it impossible to interpret what happened to patches of floating vegetation subsequently.
- Starting from 1999, images are more frequent, there are some large spikes in terms of aquatic vegetation coverage. Some are skewed by cloud coverage.

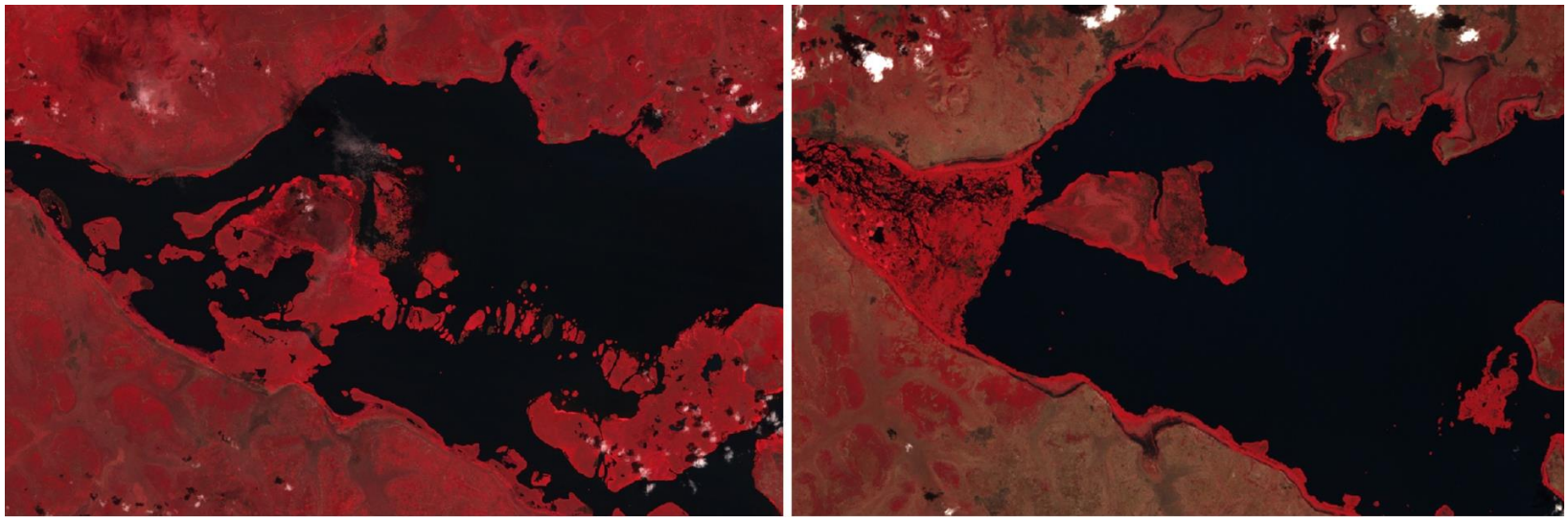

Figure 4.7. Left: False color Landsat 5 TM (11/10/1986); Right: Landsat 7 ETM+ (01/25/2000)

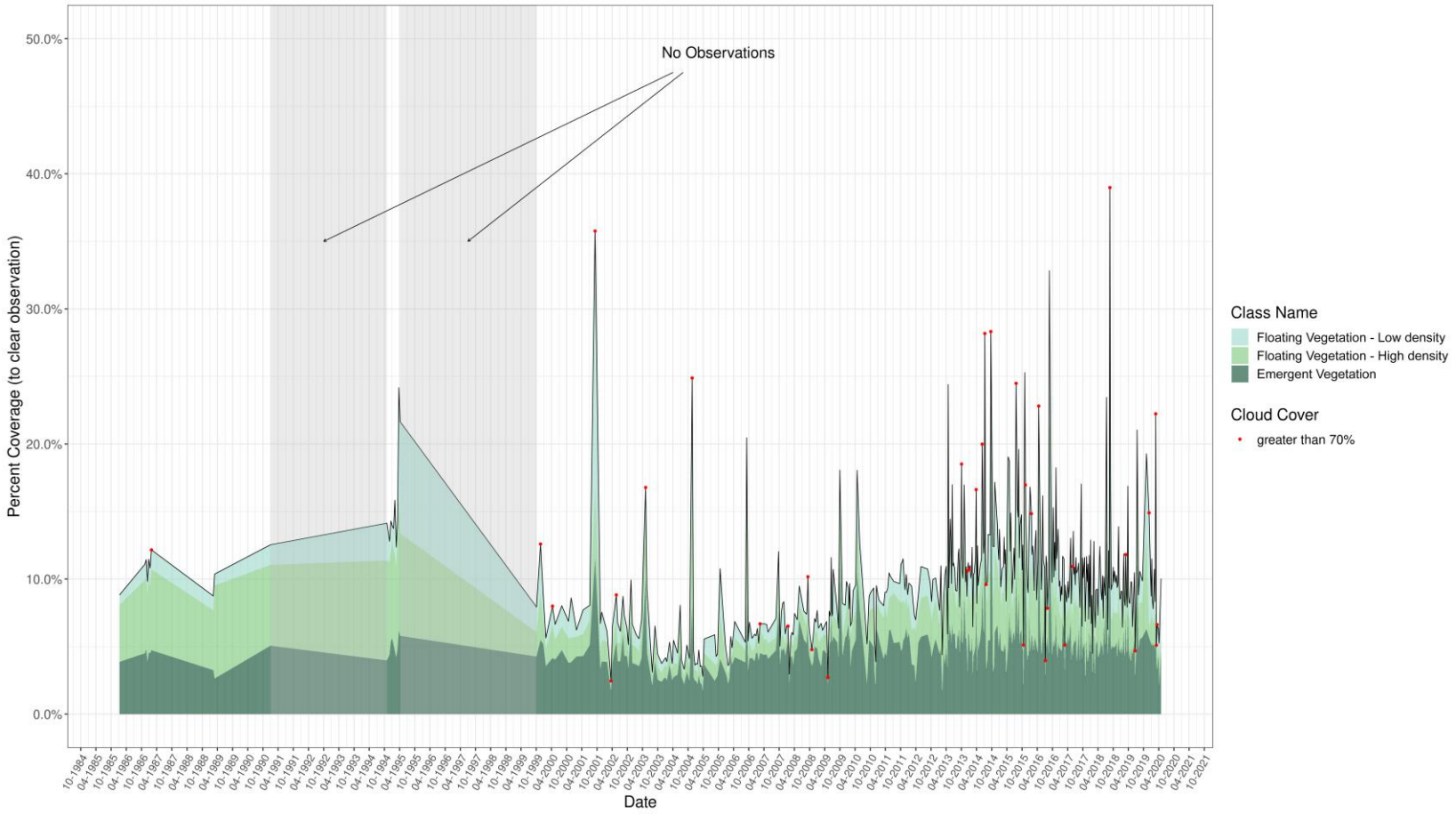

Figure 4.8. Aquatic vegetation ratio (vs. valid observation) for each Landsat image. Images with high cloud coverage (over 70%) are marked in red.

In general, there is an inverse relationship between the aquatic vegetation ratio and open water ratios. As water areas decrease, aquatic vegetation tends to increase. Figure 4.8 shows the aquatic vegetation ratio. The ratio ranges from 10% to 20% before 1999. Then it dropped to 5% to 10% from 2000 to 2011. In the recent decade, the ratio slowly increased to around 12 % with occasional spikes to mid-20%. Further, high-density floating vegetation area was prevalent before 1999 and is less in the last two decades. Emergent vegetation dynamics were stable in the last two decades.

3.3.4 Dynamics of the satellite lakes: the case of Lake Nawampasa

The satellite lake called Nawampasa (Figure 4.1c), south-east of Lake Kyoga, is examined in detail to characterize vegetation dynamics. Figure 4.9 shows 3 aquatic vegetation classes in four time periods (01/10/1986, 12/24/1999, 02/14/2016, and 07/10/2017) (left panel) and plots of AWO and distance to the shoreline (right). Lake Nawampasa had some high-density floating vegetation in the southwest and southern regions in 1986 (top left image). In 1999, the second time period (top right image), vegetation increased by a small amount in the north; in 2016 (lower left image), it increased in the south and central regions, and finally in 2017 (lower right image), there is smaller vegetation cover. Four corresponding scatter plots on the right panel show similar trends. The horizontal axes of each scatter plot show distance to the shore (M) while the vertical shows the percent occurrence of 3 types of vegetation. Emergent is very close to the shoreline in all four plots characterized by low occurrence. The plots also show that the swath of both low and high density is higher in 2016 compared with 2017. To summarize, lake Nawampasa had no floating vegetation in 1986, a slight

increase in 1999, the greatest increase in 2016, with a marked decrease in 2017. This pattern of progressive eutrophication needs further study and field validation.

The study of floating vegetation is crucial for conservation of endangered fish, because submerged aquatic vegetation is a source of food and shelter. However, once the surfaces of the satellite lakes are covered by floating vegetation, it hinders the development of submerged vegetation by denying sunlight and oxygen. Too much floating vegetation could harm the main habitat of the endangered fish and other ones needed for fishery. The total amount of floating vegetation is influenced by many factors, one of which is nutrient washing into the lake. The whole Lake Kyoga area is heavily farmed, and the soil is nutrient rich due to volcanic ash. On the other hand, the biomass and spatial distribution of aquatic vegetation is heavily influenced by water level in shallow lakes. Both factors are directly related to rainfall which is predicted to be more variable manifested through climate change. Such a trend has already been observed in recent years, water level in Lake Kyoga has increased around 3 meters since early 2020. Future flood events could become more frequent and would have a more serious impact on the shallow lake ecosystem.

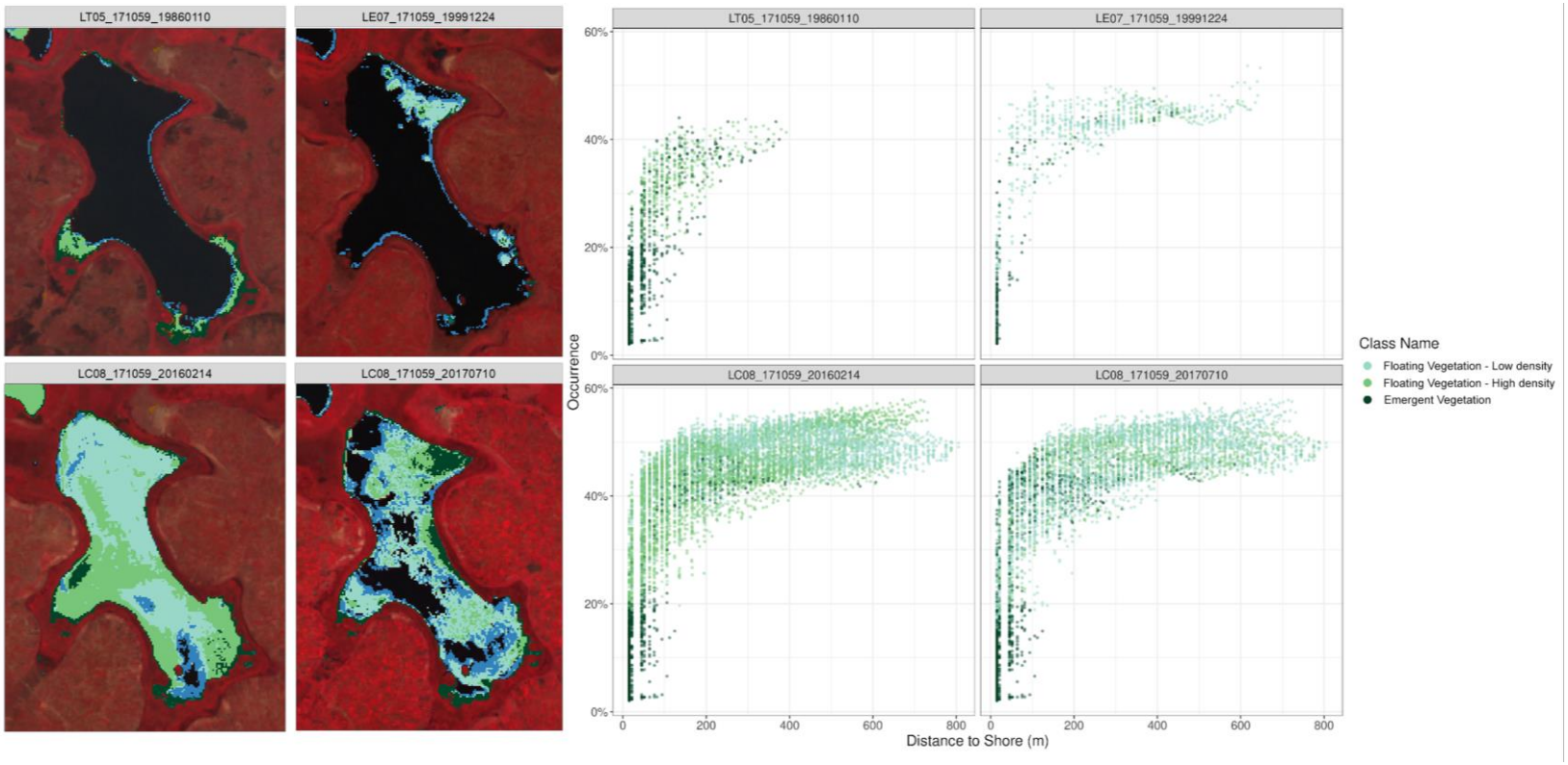

Figure 4.9. Left: classification results on Lake Nawampasa. Right: Scatter plot showing occurrence vs. distance to shoreline of classified pixels.

4.3.5 Improvement in accuracy with spatial and temporal data

In this analysis, we estimated the accuracy of different land cover classes based on spectral bands using a random forest classifier. These results are shown in Table 4.2. The overall accuracy is 60.2%. The best user's accuracy is seen in water followed by emergent vegetation; low-density floating vegetation has lower accuracy, indicating confusion with water. The producer's accuracy is higher for emergent vegetation compared with low and high-density floating vegetation. This classification suggests that spectral bands alone are not sufficient for classifying aquatic vegetation.

Table 4.3 shows the estimated accuracy of different land cover classes based on spectral bands, and two additional factors, distance to shoreline and AWO. The random forest classifier results are shown in Table 4.3. The overall accuracy is 86.8%, a considerable (26.6%) improvement over the accuracy shown in Table 4.2. The best user's accuracy is seen in high density floating vegetation class while emergent vegetation has lower accuracy. The opposite is true for producer's accuracy where emergent vegetation is more accurate than high density floating vegetation. Low density floating vegetation has the lowest accuracy since it is often confused with water.

The two accuracy assessment results highlight the following:

- The accuracy assessment based on spectral bands, AWO, and distance to shore produced higher overall accuracy (86.8%) compared with spectral bands alone (60.2%). Producer's and user's accuracy are also higher using 3 measures (spectral bands, AWO, and distance to shore) compared with spectral band

classification. Thus, adding spatial and temporal measures greatly improves the overall classification of aquatic vegetation.

- Because water is distinctive spectrally, it has the highest producer's and user's accuracy.
- Cloud has the lowest producer's and user's accuracy due to a variety of reasons. Note that most of the cloudy pixels were masked/removed from the Landsat image using CFmask. But the remaining cloudy pixels may be missed by CFmask, usually on the edge of the cloud where some cloudy pixels still remain. In addition, the presence of some cirrus clouds or low cumulus clouds causes confusion. The cloud and cloud shadow could hide subtle spectral differences of different land covers.
- The producer's accuracy of semi-terrestrial vegetation is also low, because they share a similar spectral signature with emergent vegetation, as well as have low occurrence and are located close to the shoreline.

	Cloud	Emergent Vegetation	Floating - HD	Floating - LD	Mudflat	Semi-terrestrial Vegetation	Water	Total	User's Accuracy
Cloud	4	1	0	0	1	0	3	9	44.4%
Emergent Vegetation	1	108	46	9	0	11	1	176	61.4%
Floating - HD	0	26	49	4	0	5	0	84	58.3%
Floating - LD	5	6	5	36	0	2	24	78	46.2%
Mudflat	0	2	1	0	7	4	0	14	50.0%
Semi-terrestrial Vegetation	1	0	16	1	1	15	0	34	44.1%
Water	4	3	1	15	0	0	82	105	78.1%
Total	15	146	118	65	9	37	110	500	
Area estimates [ha]	355,966	4,089,576	3,365,740	1,697,703	183,536	891,528	2,983,004		
95% CI [ha]	172,058	407,749	461,717	355,674	94,090	268,270	314,533		
Producer's Accuracy	26.7%	73.9%	41.5%	55.4%	77.8%	40.5%	74.5%		
Overall Accuracy	60.2%								

Table 4.2. Accuracy assessment of aquatic vegetation using only spectral bands,

	Cloud	Emergent Vegetation	Floating - HD	Floating - LD	Mudflat	Semi-terrestrial Vegetation	Water	Total	User's Accuracy
Cloud	6	0	0	0	0	0	1	7	85.7%
Emergent Vegetation	2	145	13	4	0	8	0	172	84.3%
Floating - HD	0	5	100	1	0	0	0	106	94.3%
Floating - LD	6	0	1	53	0	2	8	70	75.7%
Mudflat	0	0	0	0	8	3	0	11	72.7%
Semi-terrestrial Vegetation	0	1	2	1	0	21	0	25	84.0%
Water	1	0	0	6	1	0	101	109	92.7%
Total	15	151	116	65	9	34	110	500	
Area estimates [ha]	389,448	4,218,744	3,069,918	1,764,575	238,672	915,575	2,988,739		
95% CI [ha]	163,526	290,576	244,082	265,187	96,362	211,226	209,194		
Producer's Accuracy	40%	96.0%	86.2%	81.5%	88.8%	61.7%	91.8%		
Overall Accuracy	86.8%								

Table 4.3. Accuracy assessment of aquatic vegetation using spectral bands with distance to shoreline and AWO

4.4 Discussion

Remote sensing methods have been widely used for monitoring aquatic vegetation over the in-situ field survey, as remote sensing provides images with consistent and repeatable data which reduces the cost and makes the measurement across time comparable. Prior research focused on the aquatic vegetation classification based on the spectral differences of different vegetation types, and spectral libraries of signatures collected from field samples. However, limited radiometric and spatial resolution satellite sensors pose great challenges for accurately distinguishing certain aquatic vegetation types from terrestrial vegetation from individual images. This study demonstrated that a long time series of remote sensing data is useful for mapping the spatial distribution and dynamics of aquatic vegetation, addressing the gaps in the earlier data.

In this study, water bodies are identified using five versions of the water spectral index; each index consists of various combinations of spectral bands to suppress noise and non-water features. In many water indices, the threshold that separates water and non-water features is non-stable, changing from image to image. Ostu thresholding provides an adequate solution to determine the optimal global threshold per image. But in complex scenes, global thresholding is less effective. The Bradley method, a moving window approach, was applied to determine the threshold on a pixel level. To further improve accuracy, a majority voting approach was applied to produce the final water body masks.

Deriving the water masks, two data products were proposed: aggregated water occurrence and aggregated maximum extent. Comparison with the JRC Global Surface

Water showed almost identical results in the open lake region and minor underestimation in edges and satellite lakes. This is mainly due to the algorithm being tuned conservatively to avoid removing potential aquatic vegetation pixels.

Different types of aquatic vegetation possess distinct spatial and temporal characteristics. Distance to shoreline and AWO that were added to spectral signals from Landsat images, which greatly improved the classification accuracy. Drastic changes in water level can negatively impact the accuracy of aquatic vegetation identification when the distance to shoreline and AWO are added. During the early period of the observations, the water level of the Lake region was 1032 MSL (mean sea level) in 1994 which increased to 1035 MSL at the beginning of 2000. A patch of land in the western region close to Victoria Nile of Lake Kyoga was submerged. However, the same patch of land never re-emerged when the water level dropped back to 1032 MSL in 2010. Therefore, prior to 2000, the vegetation at that location was misclassified as high-density floating vegetation due to its low occurrence and high distance to the shoreline. Changes in land use could also negatively affect the classification results, as the AWO and AME metrics were generated using water history from the past 30 years. In the later years of observation, nearshore areas of the lake with agricultural landuse may be misidentified as emergent vegetation. To reduce the limitations, a more granular control can be implemented on what images are used to construct the AWO and AWE. For example, to reduce the influence of change in water level on misclassification is to use images with similar water levels, which can be measured through satellite radar altimetry data.

Another issue is cloud cover, which influences the ratio of aquatic vegetation area and open water, and therefore affects the classification accuracy. When the number of clear pixels is low, (or cloud cover is high), it can often result in bias particularly if the clear pixels are clustered mainly on land or water. Also, the cloud mask algorithm used in this study has limited ability to detect thin cirrus, low cumulus clouds, and cloud edges, which obscures spectral signature for certain classes. Data fusion using images from multiple types of sensors and resolutions could increase the data availability. Influences from clouds can be reduced from fusing data from optical sensors and SAR sensors since clouds are transparent to radar waves. High temporal resolution sensors can be fused with Landsat images through mathematical models to create near real time tracking of aquatic vegetation dynamics.

4.5 Conclusion

A novel algorithm is proposed in this study to classify and map the spatial distribution and dynamics of aquatic vegetation in the Lake Kyoga region in Uganda using Landsat imagery. The mapped products include Aggregated Water Occurrence (AWO), Aggregated Maximum Extent (AME), and Aquatic Vegetation Extent (AQE). Water occurrence is a measure of the number of times the pixel is classified as water from 1986 to 2020. Lake Kyoga's maximum water extent is identified using a combination of water indices ranging from the Normalized Difference Water Index (NDWI) to modified NDWI (MNDWI) and Automated Water Extraction Index (AWEI). The disadvantage of using two-band ratio water indices for surface water of NDWI

mapping is overcome using AWEI, a multiband water index. AWO and AME enable us to identify aquatic vegetation and separate aquatic vegetation pixels from lake water pixels. The classification procedure requires class labels and training data generated in this study using an unsupervised SOM clustering algorithm. An expert identified the clusters and labeled them aided by high-resolution PlanetScope imagery. The experts selected seven meaningful clusters out of 40 clusters generated by the algorithm, which was then used to create training data for a supervised decision tree classifier. Thus, the study provides a valuable methodology for generating class labels and training data in the classification of aquatic vegetation when such data is sparse and unavailable.

Two additional spatial and temporal variables better classify aquatic vegetation in Lake Kyoga in the random forest classification procedure. The resulting classification maps have better overall accuracy and vegetation class accuracies than the basic spectral imagery, indicating that distance to shore and temporal dynamics (captured in AWO) produced higher overall accuracy and class accuracies. Floating and emergent aquatic vegetation are clearly differentiated using additional variables. Hence spectral bands alone cannot differentiate aquatic vegetation. We envision that our algorithms could monitor the aquatic dynamics in lakes in other regions where their environments are similar to the Lake Kyoga region.

CHAPTER 5 – Seeing the Invisible: From Imagined to Virtual Urban Landscapes

Urban ecosystems consist of infrastructure features working together to provide services for inhabitants. Infrastructure functions akin to an ecosystem, having dynamic relationships and interdependencies. However, with age, urban infrastructure can deteriorate and stop functioning. Additional pressures on infrastructure include urbanizing populations and a changing climate that exposes vulnerabilities. To manage the urban infrastructure ecosystem in a modernizing world, urban planners need to integrate a coordinated management plan for these co-located and dependent infrastructure features. To implement such a management practice, an improved method for communicating how these infrastructure features interact is needed. This study aims to define urban infrastructure as a system, identify the systematic barriers preventing implementation of a more coordinated management model, and develop a virtual reality tool to provide visualization of the spatial system dynamics of urban infrastructure. Data was collected from a stakeholder workshop that highlighted a lack of appreciation for the system dynamics of urban infrastructure. An urban ecology VR model was created to highlight the interconnectedness of infrastructure features. VR proved to be useful for communicating spatial information to urban stakeholders about the complexities of infrastructure ecology and the interactions between infrastructure features.

5.1 Introduction

5.1.1 Urban ecology

Cities function akin to ecosystems, consisting of complex features and systems that are interconnected and dependent on one another. These urban ecosystems are fragile and face many challenges. As the population in many areas of the world continues to grow and urbanize, cities are forced to adapt and, as a result, the functioning urban ecosystem becomes stressed while trying to supply services to more people (Colding & Barthel, 2017). Additionally, the urban ecosystem is threatened by a changing climate and extreme weather events - from flooding and land subsidence in New Orleans (Qiang, 2019), to wildfires destroying areas on the west coast of the United States (Schweizer, Cisneros, Traina, Ghezzehei, & Shaw, 2017), environmental hazards test urban ecosystems worldwide with increasing frequency and extremity (Salas & Yepes, 2018). The combination of environmental threats and an ever-growing population has put unprecedented stress on aging urban ecosystems, exposing vulnerabilities and posing a risk of collapse. Improving these ecosystems and increasing the resiliency of infrastructure systems is going to be crucial for cities moving into the future.

5.1.2 Infrastructure systems

Urban infrastructure system (UIS) is a term that will be referred to throughout this paper. The UIS is defined as the dynamically interrelated pieces of individual infrastructure, both above and below the ground, that make cities function. The UIS can change as a whole in response to a shift in one individual feature (Pandit, Lu, &

Crittenden, 2015). An UIS is expansive and is maintained by a variety of stakeholders, including local governance, municipal and public facilities, municipal utilities, and engineers (Ferrer, Thomé, & Scavarda, 2018). In the UIS framework, it is important to understand that a change or failure in one infrastructure feature can cause a ripple effect throughout an urban environment. As stated in Upadhyaya, Biswas, and Tam (2014): "There are multiple and layered negative effects on societal health and well-being when infrastructure systems break down and are unable to adapt to sudden increased demands … Unsustainable and inadequate infrastructure can fail causing stress on resources and endangering public health."

There have been countless incidents where a piece of infrastructure fails and causes damage and inconvenience to large urban populations. Power outages, flooding, and major repair projects are just a few examples of the inconveniences and dangers that occur when the UIS is disrupted (Upadhyaya et al., 2014). Unfortunately, in many municipalities care for infrastructure as separate features and approach management with a narrow technological approach rather than holistically addressing the entire system (Pandit et al., 2015). This separation in infrastructure management is demonstrated in the separation of management between above and belowground infrastructure features. Seldom do the stakeholders at the parks and recreation department interact with the water and sewage workers on coordinating repairs. This siloed approach has created an urban infrastructure management system in which there is little professional and/or public understanding of how these two infrastructure environments interact as a system (Nelson, 2016).

To understand how infrastructure below the ground affects infrastructure above the ground (Rinaldi, Peerenboom, & Kelly, 2001), it is important to distinguish between the components of above and belowground infrastructure. Aboveground infrastructure encompasses the infrastructure citizens walk on, live in, and ride on. Roads, sidewalks, buildings, parking lots, green spaces, and public transportation comprise the above ground infrastructure ecosystem (Andersson et al., 2014). Belowground infrastructure is less conspicuous but is equally, if not more, important to care for in order to reach resilient urbanization goals (Ferrer et al., 2018). Below city streets there is a complex system of utilities, transportation, biomass, and structures that enable urban areas to run. Gas lines, water pipes, sewage, stormwater management, electricity, and cable provide services to urban citizens upon which they rely (Guneralp et al., 2015). Subway lines, tunnels, and skyscrapers' massive foundations also have to find a niche underground to provide ease of movement and support for people living above the ground (Sun & Cui, 2018). In addition to these human requirements, the natural world is a strong competitor below the ground. Microbial communities and root structures compete for space in this highly disturbed environment (Mullaney, Lucke, & Trueman, 2015).

Above and belowground infrastructure features are separate “adaptive entities” that interact and relate to one another in complex ways (Pandit et al., 2015). Unfortunately, infrastructure management in many urban areas is focused on the individual utility. Shifting this focus from the current short-term ad hoc repairs to a comprehensive integrative repair plan will be necessary for cities to be more sustainable and resilient (Derrible, 2017). To make this transition possible, an increased

understanding of the relationship between above and belowground infrastructure will be crucial for all infrastructure stakeholders to promote a healthy urban ecosystem.

In order to efficiently plan for cities of the future that are resilient in a changing climate and an urbanizing world, coordinated infrastructure management of both above and belowground utilities will be necessary. For this study, we focus on the City of Boston, MA and other Massachusetts municipalities that are currently tackling aging belowground infrastructure (Hendrick, Ackley, Sanaie-Movahed, Tang, & Phillips, 2016). Our team hosted a workshop for stakeholders involved in all realms of urban infrastructure management to sit down together and discuss the systematic and foundational barriers that exist for implementing a more coordinated infrastructure management approach. From the conversations,our team uncovered a need for a tool that could not only help stakeholders visualize spatial information but highlight the interconnectedness of various infrastructure features. Virtual Reality (VR) became a clear choice for communicating the complexity of interrelated spatial data to the stakeholders and so our team created an immersive VR tool to demonstrate the UIS.

5.2 Data and methodology

5.2.1 Virtual reality for urban planning (virtual landscapes)

An emerging tool with exciting and growing application in urban planning is Virtual Reality (VR) (Kersten, Deggim, Tschirschwitz, Lindstaedt, & Hinrichsen, 2018). VR is an immersive tool that allows a user to experience and “reproduce a realistic… detailed and accurate visual and audio model as similar as possible” to the real world in

the comfort of their own office or home (Echevarria Sanchez, Van Renterghem, Sun, De Coensel, & Botteldooren, 2017). VR models create an environment that stakeholders can enter, providing a "common language" for them to use and relate to while making planning decisions (Lovett, Appleton, Warren-Kretzschmar, & Von Haaren, 2015). A VR environment is immersive; creating a "multisensory' visualization… [that] track[s] user movements [and] show[s] a virtual environment wherever the user is looking" (Berger & Bill, 2019). Additional benefits of a VR model is that this environment can be created in an office building, not requiring stakeholders to travel to a location to visualize infrastructure like an augmented reality (AR) model would require (Cirulis & Brigmanis, 2013). Conclusions drawn from the urban stakeholder workshop, hosted as a part of this study, demonstrated a need amongst stakeholders for a better way to communicate and visualize spatial data in the complex urban infrastructure environments, paving the way to the creation of this VR tool.

Traditionally, urban planning stakeholders have been trained with tools such as computer-aided design (CAD) and geographical information system (GIS) drawings (Wu, He, & Gong, 2010). These tools are helpful in visualizing city streets and networks but do not show the dynamic and inter-connectivity of the different features in the UIS. VR tools, however, has the ability to facilitate a more comprehensive approach to urban planning and infrastructure management by showcasing all features in the infrastructure ecosystem and demonstrating how they interact with one another (Santos, Zarraonandia, Díaz, & Aedo, 2018). Creating a VR rendering of a city street, with both above and belowground infrastructure components, provides urban planners with the answers to

questions such as, what is the spatial relationship between sewage and drinking water supply? How are belowground utilities organized under the street? How vulnerable is the infrastructure network to collapse? Industry experts have a heuristic understanding of the placement of various utilities in relation to one another, but a VR realization could make this more concrete (Nelson, 2016). Ideally, having a tool that enables stakeholders to visualize co-located urban infrastructure features would allow for a more coordinated infrastructure management approach that could increase the resiliency and efficiency of the entire infrastructure system.

VR, while pioneered in the gaming industry, has transformed over time through innovations in application (Edler, Kühne, Keil, & Dickmann, 2019). Recent advancements in cartographic methods and GIS technology have allowed for data representation in the third- and fourth- dimension (height and time respectively), giving cartographers access to new realms of mapping (Wolfartsberger, 2019). Unfortunately, there are technical challenges to bridging a GIS database into an AR or VR database but when done successfully, adding the third- and fourth-dimensions, 3D GIS decision support systems can create three-dimensional scenarios from overlapping spatial datasets, e.g., street measurements of different infrastructure features. This integrative model is helpful in enabling urban planners to see the interconnectedness of the urban infrastructure system and has added depth representation to spatial data and enhanced visualization (Hruby, 2019). AR is useful for many of the same applications in urban planning as VR, however, with AR technology the user must travel to the real-world location in order to envision the virtual model (Carozza & Tingdahl, 2014). AR has

proven useful in urban applications in countless studies (Allen, Regenbrecht, & Abbott, 2011; Imottesjo & Kain, 2018; Ishii et al., 2002) but for the immersive and portable experience associated with VR, our team decided VR would be more applicable for communicating spatial interactions in the UIS.

VR models create an environment of spatial data that enables the user to visualize, interact, and immerse themselves into the unique map from anywhere in the world (Kersten et al., 2018). VR models have huge potential to revolutionize urban planning and the mapping of 'smart cities' because they allow planners to simulate future scenarios (Tao, 2013). This extensive immersion mapping technology was attractive to our team because it would allow stakeholders to envision the interactions between infrastructure features and see the dynamics of the UIS. Using VR for urban planning is not a novel idea. In a study by Fairbairn and Parsley (1997), the authors examined the use of VR and virtual reality modeling language for cartographic presentation, and provided several examples that demonstrate successful virtual campus construction. Prior studies by Batty, Dodge, Doyle, and Hudson-Smith (1998) and Doyle, Dodge, and Smith (1998) have described the 'Virtual London' project that marries a range of VR and Internet GIS technologies. Urban stakeholders benefit from VR technology because it is useful in exploring ways to plan, model, and simulate urban planning and aid in impact assessment (Kamel Boulos, Lu, Guerrero, Jennett, and Steed, 2017). The creation of these virtual models has enabled planners to interface with the complex physical and social data incorporated in planning and managing cities in a realistic and meaningful interactive way.

VR has additional applications in risk assessment and urban resiliency in a variety of contexts, including wind damage (Repetto et al., 2017), forest fires (Gaudreau, Perez, & Drapeau, 2016), and other natural disasters (Breunig et al., 2015). ESRI (the company making GIS software) has created a mobile VR solution for urban planners, architects, and GIS professionals called CityEngine that can create a VR tool to compare urban planning scenarios on a mobile device. Standard 3D GIS packages include 3D city modeling applications, such as City Engine (Neukom, 2018) and CityGML (issued by the Open Geospatial Consortium) to render and store digital 3D models of cities and landscapes (Pouliot, Larrivée, Ellul, & Boudhaim, 2018). The standard ArcGIS API enables users to build full-featured 3D applications powered by web scenes consisting of terrain, integrated mesh layers, and 3D objects. Additionally, the open-source JavaScript library Cesium can create web-based globes and maps, also useful for visualizing dynamic data. iTowns, written in JavaScript/WebGL, is frequently used for precise 3D visualization of street view images and terrestrial LiDAR point cloud. Unfortunately, due to the massive size of spatial data, web-based GIS applications can create network latency as well as bottlenecks when handling multiple users. Despite these difficulties, VR-GIS packages are becoming increasingly popular for addressing and solving urban problems because of their ability to incorporate the dynamics of aboveground and underground features (Boulos et al., 2017).

VRGIS has become an increasingly popular for tool for urban planners looking for an interactive way to model urban decision-making processes (Sameeh El halabi et al., 2019). VRGIS establishes a three-dimensional model in a virtual environment, and

operates via personal computers, mobile devices and smart glasses. Examples of VR technologies include Google Daydream View VR, and its' cheaper predecessor Cardboard (2014), which utilizes a smartphone's gyroscope for head tracking. VRGIS is almost seven decades old, however, recent innovations and developments in technology, such as big data, augmented reality, graphic processing units (GPUs), and the Internet of Things (IoT), has enabled VRGIS to have better performance and more intuitive human–computer interactive modes. These advancements in VRGIS have encouraged its applicability in visualizing, experiencing, and solving more complex, real-world problems (Boulos et al., 2017; Li et al., 2015).

5.2.2 Workshop

In order to provide a tool to aid in implementing a more coordinated infrastructure management approach, a baseline understanding of the current management practices amongst infrastructure stakeholders had to be established. To obtain this baseline data, our team hosted an urban infrastructure workshop in June of 2017 that brought urban stakeholders together to discuss current infrastructure management protocol. Elected officials, city planners, engineers, utility workers, students, concerned citizens, activist groups, academics, and several other parties invested in making cities work efficiently attended. The goal of the workshop was to encourage groups of people who did not typically interact to discuss the systematic difficulties, educational obstacles, and/or communication barriers in managing urban infrastructure in Massachusetts.

Baseline data was collected in the form of visual maps and talk-back sessions in response to two exercises. In an effort to establish a baseline understanding of the attending stakeholders in regard to their perceptions of above and belowground infrastructure, they were asked to draw a cross-section of a typical city street. The stakeholders were randomly assigned to different groups and presented a prompt (Figure. 5.1) that asked them to draw a cross-section of a city street in fictional Anytown, USA and to highlight the interactions between the features of infrastructure they included and the stakeholders responsible for managing those features. For example, were they to include gas pipes and water pipes in their drawing, our hope was that they would list the water company and the gas company as the administrators of those infrastructure features and also include that those features co-existed underneath the street? Additionally, we asked the stakeholders to include what barriers existed that prevented a more coordinated management approach. For example, if the water company and the gas company ever worked together to repair pipes to avoid traffic disruption along the same segment of the street.

In the second exercise, attendees were presented with a problem posed by the fictional community requesting more green space and buried utility lines. The attendees had to work together to identify the stakeholders that would need to be involved in such a project and to list any existing partnerships or communication tools that would be useful for such a project. Lastly, they were asked to identify any institutional, systematic, or functional barriers that existed in implementing such an infrastructure project.

The five groups completed the two exercises and presented their visual maps back to the larger group. The visual maps displayed cumulative team insight into current infrastructure design and management processes in and around Boston, MA.

"Market Street in Anytown, USA, is a busy commercial corridor recognized to have excellent economic potential but held back by a streetscape in disrepair and daily traffic jams. Market Street has relatively high vacancy and business turnover and little foot traffic on the sidewalks. The two-way street has two lanes in either direction; sidewalks; storefronts and a few frontage surface parking lots. Market street is a 1950's design built primarily for the automobile, although it has a bus line and potential for nearby transit connections. Above the street is a tangle of electrical wires and utility poles. Underneath the street is a haphazard array of gas, water and sewer pipes of varying age and condition."

EXERCISE 1: Network/Relationship Map: Map interplay among various owners of above and below ground infrastructure.

1. Who are all of the stakeholders associated with the above and below ground infrastructure represented on your schematic? Public, private, community, regulatory, etc.
2. What are their roles in the infrastructure management process?
3. Where are the connections amongst the stakeholders?
4. What tools or communications support those connections?
5. What are barriers to connections or communication?

Infrastructure Ecology: Residents call for undergrounding the unsightly electrical wires, a request that had not been previously made, and for tree plantings and green space. Moreover, above-ground electrical utility poles and fixtures seriously constrain space available for proper sidewalks and bike lanes, and green space would compete with space for pedestrians and bicyclists. How can cities better physically allocate above- and below-ground space for street infrastructure? Are infrastructure synergies possible, wherein above- and below-ground infrastructure can be spatially co-organized?

EXERCISE 2: Determine how the city can resolve the above challenge identified by the community.

Revisit the Network Map: Is the "network" equipped to address the problem? What enhancements/improvements are needed?

1. Who "owns" the problem?
2. Are there any stakeholders not represented who are critical to addressing the problem?
3. Are the existing connections, communication channels and tools sufficient to address the issue? What more is needed?

Figure. 5.1. Prompt with two exercises for five randomly chosen stakeholder groups to work through and discuss. Results were collected in visual maps that were presented back to the larger audience.

5.2.3 VR design

For the VR model, a neighborhood in South Boston was chosen as the study area, namely the Dorchester Ave corridor between the MBTA Red Line stations Broadway and

Andrew because of the planned redevelopment in this neighborhood. The study area (4.6 km2) can be covered by two USGS Lidar point cloud scenes, which were obtained from USGS 3DEP (U.S. Geological Survey, 2015a, U.S. Geological Survey, 2015b). Building height can be derived from LiDAR point cloud. Combining with the building footprint, we can populate the VR scene with buildings with the appropriate height. A schematic of the VR tool created for this research project is shown in Figure. 5.2.

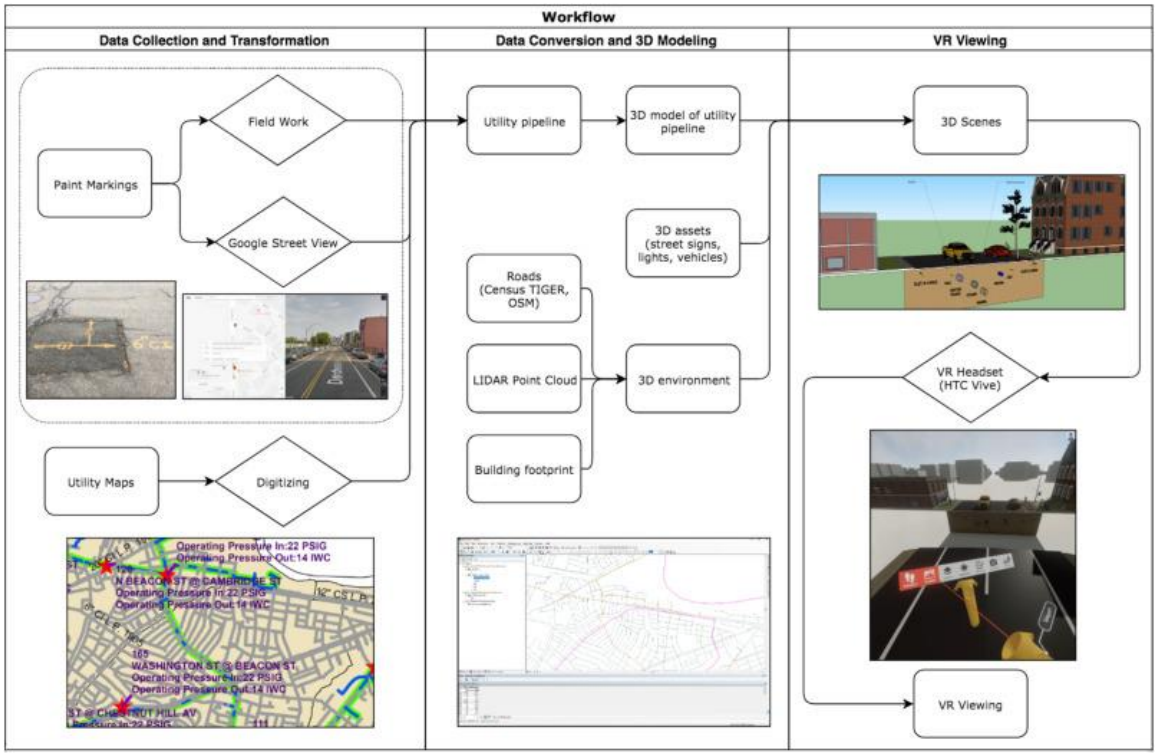

Figure. 5.2. Schematic representation and flow chart of the VR tool creation.

For this analysis, Google Street View was used. Street View is a service provided by Google that allows a user to view panoramic street-level images across the world. Google collects panoramic images using a vehicle-mounted 360-degree camera that are made publicly available on Google Maps. Google recollects Street View images every 3

to 4 years for populated areas. Information on the location and type of utilities was collected by locating pipeline locations marked on the street with spray paint. Cities around the world mark with spray paint color-coded pipeline locations on the streets indicating the utility type, the location, orientation, diameter, and material of the pipe (Figure. 3) (APWA, 2019). Obtaining spatial information about belowground utilities is difficult because utility companies limit the distribution of underground infrastructure data to the public. This is a national security issue, as utility companies do not want to put service areas in a vulnerable position, were the exact location of all infrastructure to be public knowledge.

Fig. 5.3. Example of a spray painted marking on a road by utility company.

Designing a VR environment to highlight above and belowground infrastructure features proved difficult due to these stakeholder regulations. However, by utilizing the street markings left behind by utility companies, we were able to collect enough street

markings to build a comprehensive model of the pipeline network. Using Google Street View for this analysis was beneficial because of the historical record of images Google has. The spray-painted utility markings can fade due to traffic and dust so looking at a selection of photos from a single vantage point enabled our team to collect as much data as possible. There are no existing tools that allow the user to collect the street marker information directly, so the tool "Underground Utility" was created for this purpose. Although time intensive, this method of collecting, converting, and visualizing natural gas infrastructure created for this study can be applied to other underground utilities.

"Underground Utility" was written in JavaScript, using the Google Maps JavaScript API. The main interface was split into two sections: the left side has the Google Maps, and the right side is Google Street View (Figure. 5.4). The tool allowed a user to place custom markers on the Google Street View panel that sync to the Street View panel. The markers could be customized with information such as utility type, pipe material, and pipe diameter. Addresses were reverse geocoded from the coordinates of the markers and after enough markers are placed for post-processing, the location and attributes of markers were exported to an Excel file and added to a GIS software to generate a spatial display of the underground utility data.

In some cases, utility maps were public domain, allowing us to convert these maps into shapefiles and import them directly into a GIS software. For this study, we obtained natural gas pipeline distribution maps from National Grid territory in Massachusetts (National Grid, 2010). These pipelines were represented in vector (as opposed to represented in pixel in raster maps) so it was possible to convert the polylines

directly to shapefiles. To convert a PDF to a shapefile, AutoCAD was used as the medium to extract the polylines and export them to ArcMap. Then the spatial information was added using the Georeference and Spatial Adjustment tools.

Combining the utility pipeline data along with spatial information allowed us to create a tool modeling a comprehensive pipeline network. On traditional GIS platforms, we could visualize pipeline data as latitude and longitude. However, because most of the utility pipeline was buried belowground, it was difficult to differentiate above versus belowground. Therefore, a third dimension was introduced. 3D models are an excellent way of visualizing data in three dimensions. Airborne LiDAR data can be used to create DSM (Digital Surface Model). The main difference between DSM and DEM is that DSM captures the surface height, that includes the building height, canopy height. Combined with building footprints, we could create 3D models of buildings. The LiDAR data we used were USGS Lidar Point Cloud MA Sndy (U.S. Geological Survey, 2015a, U.S. Geological Survey, 2015b), and the first return was used to estimate building height above mean sea level.

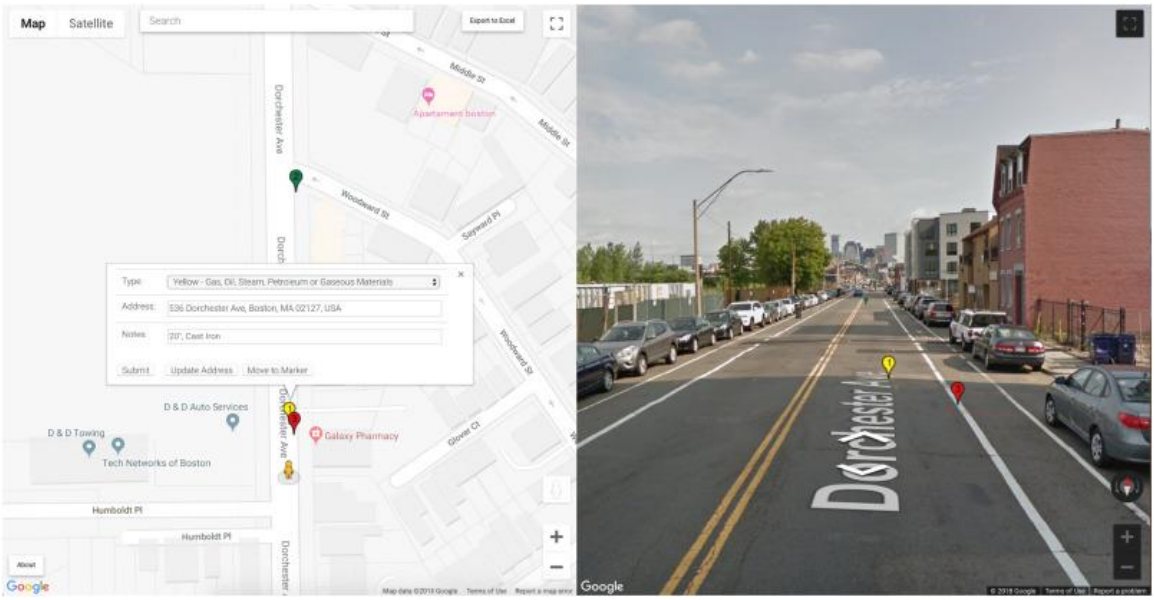

Figure. 5.4. Screenshot of Underground Utility interface. Google Maps display of Dorchester Ave, Boston, MA (left panel) and Google Street view of the same street (right side).

In order to display the 3D model and to add more details, the shapefile was imported to Trimble SketchUp using a modified plugin in which we could select the field containing building height information. This plugin can also be used to import roads as polylines to create road models that follow the polylines. Sketchup s useful for creating 3D models because of its access to the world's largest open-source assets library. 3D models limited to shapefiles are commonly plain-looking because shapefiles contain only buildings, utility pipelines, and roads. Adding auxiliary assets, such as ground cover, cars, humans, and streetlights to the 3D model created a more realistic user experience. A 3D model is useful to highlight the third dimension that distinguishes above and belowground infrastructure features, unfortunately, the models are still constrained to a flat computer monitor limiting user perception. VR became the logical next step to

creating an environment in which the user could observe and interact with objects as they are in the real world.

There are two requirements for a VR experience, hardware and software. For hardware, we used the HTC Vive platform. This platform contains a head mounted display (HMD) and a pair of hand-held controllers. Two wall-mounted lighthouses track the X, Y and Z position of the user in real time. It allows the user to move within the VR environment by moving their heads, body, and hands. A minimum of 2 m by 2 m of unobstructed space is recommended for a room scale setup because it provides the greatest user immersion. Unfortunately, rendering a VR environment is a heavy load on the computer because the graphics card needs to drive two full HD screens in the HMD at 90 frames per second (FPS). According to HTC, a GTX 970 or equivalent graphics card is the minimum requirement for VR.

There are multiple software packages possible for the user to view 3D models in a VR environment. SYMMETRY is a software tool that converts CADs, in this case SketchUp models, to VR and is currently available on Steam, a video game digital distribution service by Valve. The import feature converts .skp files along with SketchUp layers and textures into VR. Additionally, there are two viewer modes: the “Studio Mode” that provides the user with an overview, as if viewing a model inside a studio; and the “Immerse Mode” that brings the user inside the model where they can use the markup tool, camera, and memo to communicate with other users and exchange ideas.

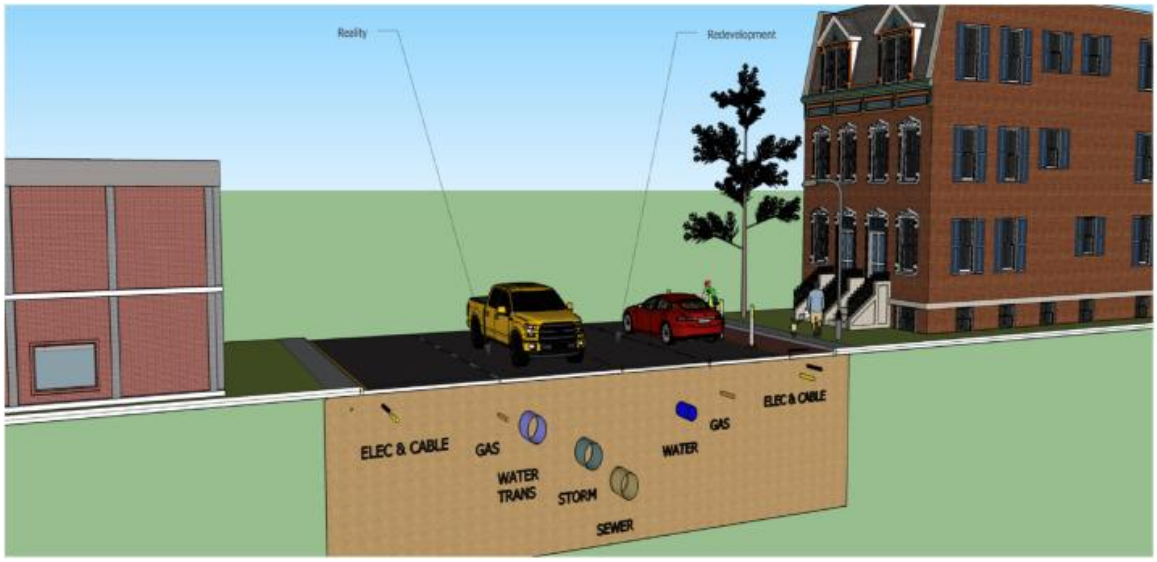

Figure. 5.5. Cross section of the VR model "Virtual Reality & Urban Ecology".

We created our 3D urban infrastructure experience using the HTC Vive platform. Our VR experience allowed users to dive beneath city streets and look at the variety of utilities that exist and interact with one another. The VR entitled "Virtual Reality & Urban Ecology" allowed the user not only to explore belowground utilities but also see a city block of aboveground infrastructure, including cars, bike lanes, public transportation lanes, buildings, sidewalks, and pedestrians (Figure. 5.5). With this VR environment, we aimed to teleport urban stakeholders to an environment which highlights the interconnections of above and belowground infrastructure features.

5.3 Results and discussion

5.3.1 Visual maps

From the five visual maps (Figure. 5.6) collected in response to the two Anytown, USA prompts, several common themes emerged, which suggested a variety of reasons as to why a coordinated infrastructure management plan is difficult to implement: 1) Elected officials expressed concern with aboveground infrastructure only; 2) Engineers and utility workers did not communicate efficiently outside of their particular utility to coordinate infrastructure repair and replacement projects; 3) There was a general lack of understanding and/or appreciation for how urban infrastructure functions as a system; 4) Budgetary and practical concerns exist, preventing future urban infrastructure innovations, like the utilidor, from being implemented.

Elected officials, city managers, and park officials concentrated on infrastructure elements people can see. Elected officials especially, focused their campaigns and time in office bettering what people can see in the aboveground environment. Many admitted to an underappreciation for how the aboveground built environment was influenced by belowground infrastructure. For example, when traffic is disrupted because of pipe repair and uneven streets result from trenching and cement patchwork. Green infrastructure, like parks and street trees are also affected by belowground systems. For example, leaky gas lines pollute street tree pits with methane and kill vegetation along sidewalks (Hendrick et al., 2016). Transitioning from an aboveground infrastructure mindset to one focusing on the system dynamics of all urban infrastructure became crucial for implementing a coordinated management plan.

Inefficient communication between belowground utilities has created problems in project efficiency and coordination, resulting in more street disruptions and more expensive projects. Even though belowground utilities work in the same space, there is commonly no notification across utility companies alerting to a street dig up for repair. If all utilities needing to do repairs on that street could do the repairs simultaneously, the street and traffic could be disrupted just once. This coordination could decrease excess noise, traffic disruption, and patchy/uneven streets.

Generally, workshop attendees admitted to considering above and belowground as separate entities rather than thinking about urban infrastructures as a system. However, when presented with the UIS approach many understood how more coordinated repair projects and management could benefit city functioning. The system dynamics of urban infrastructure highlights the interactions and impacts belowground infrastructure has on aboveground and vice versa. These interactions are worthwhile to educate urban stakeholders on in order to reconstruct the management of urban infrastructure.

The visual maps drawn by each group modeled what stakeholders imagined as the most ideal infrastructure system. Unfortunately, many of the features were idealistic because of concerns over budgetary constraints. For revolutionizing belowground utilities, most groups preferred a utilidor solution. A utilidor is a tunnel that consolidates and co-locates multiple utilities, with street access at an easy-to-access point (preferably on the sidewalk to discourage traffic interruption) for maintenance or repair (Hunt, Nash, & Rogers, 2014). Placing all utilities in a single corridor would enable companies to complete repairs through the sidewalk without disrupting traffic or disturbing another

utility. Unfortunately, discussion of utilidors amongst stakeholders uncovered safety, liability, and budgetary concerns making utilidors an unrealistic solution in the near future (Canto-Perello, Curiel-Esparza, & Calvo, 2016).

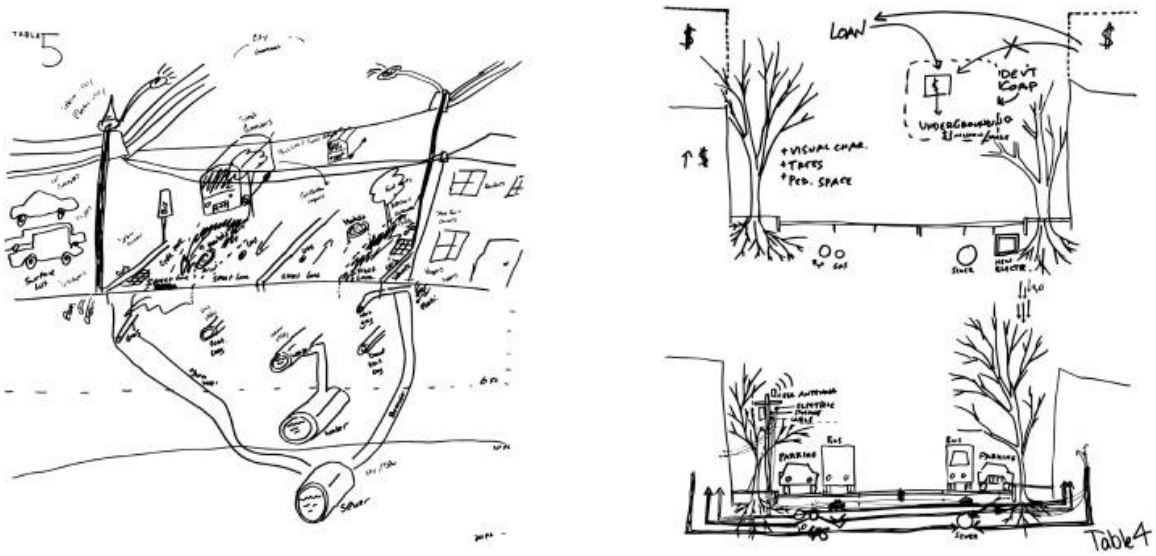

Fig. 5.6. Two examples of visual maps drawn by break out groups at stakeholder workshop.

After analyzing the elements stakeholders decided to include in their visual maps, some commonalities appeared. All of the groups included private buildings, sidewalks, and lanes for traffic and public transportation. Only one group included a parking lot and only two included a utilidor or on-street parking. Only three of the five groups included belowground utilities such as water, sewer, storm water, gas, electric, and/or cable. What groups included and what they omitted provided insight into what the stakeholders considered important elements of urban infrastructure.

From the talkback session, a clear need for a spatial communication tool emerged. Stakeholders understood the necessity of viewing urban infrastructure as a system and admitted there was a challenging lack of coordination in current management tactics.

Unfortunately, demonstrating the system dynamics of the UIS is a challenge because half of the system environment is out of sight below the streets. These results inspired the development of the VR tool to aid stakeholders in seeing the interactions within the UIS and to help them visualize belowground infrastructure and its influence on the street. A visualization tool would benefit all utility stakeholders because it can demonstrate how utilities interact and behave in the UIS. The VR program developed provided a good starting place because it emerged the user into an urban environment where they could interact with all features of infrastructure.

5.3.2 User responses: "Virtual Reality & Urban Ecology"

In response to the results gathered from the stakeholder workshop, the VR tool, "Virtual Reality & Urban Ecology" was created and user experience was collected. Two demonstrations were held in October and November of 2017 and participants included researchers, academics, students, non-profit people, businessmen, and lawyers. Each participant was fitted with the VR headset and hand-controllers and immersed in our model for a fifteen-minute session (Figure. 5.7). Aboveground, the user could interact with cars and cyclists, in addition to exploring the layout of sidewalks, roads, public transportation lanes, and bike lanes. By simply looking downwards, the participant could dive beneath the street and see the relative location of multiple belowground utilities. A user could explore gas, water, sewage, and other pipelines, as well as a rendering of a utilidor.

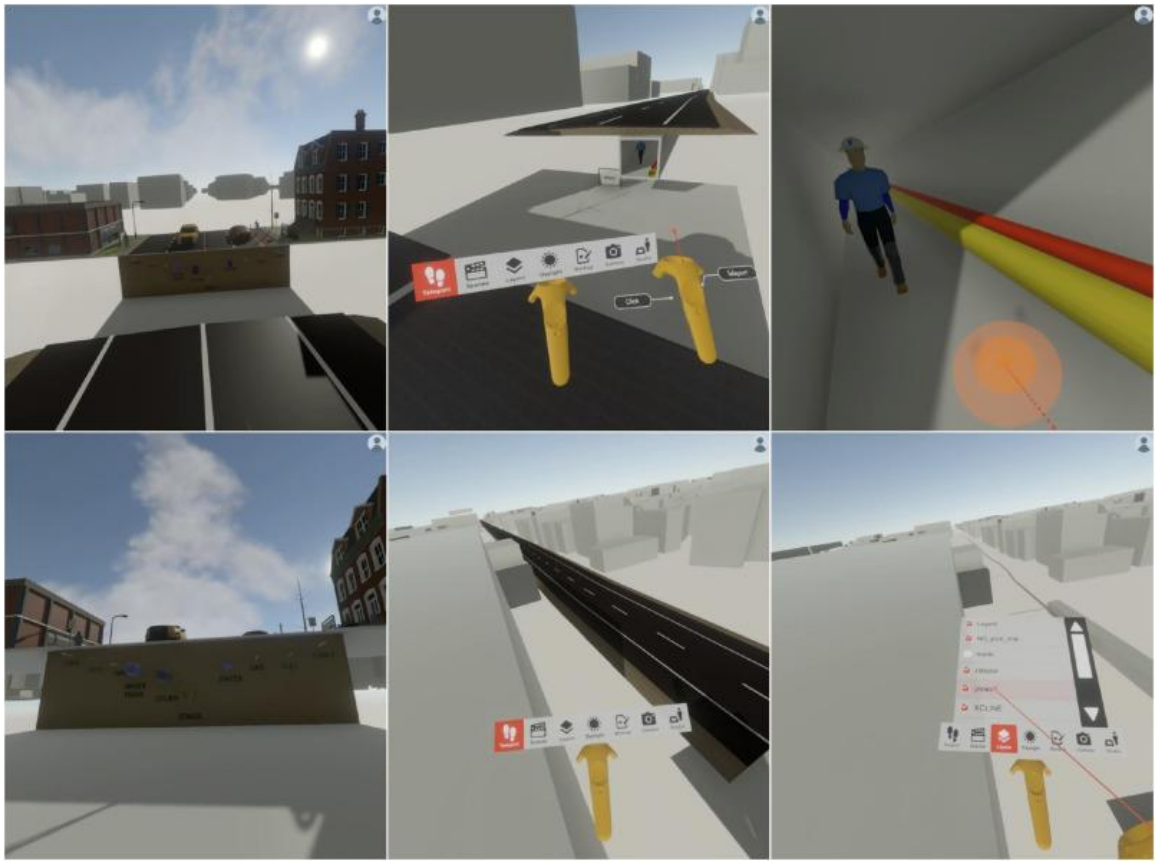

Fig. 5.7. VR user perspective, while using the "Virtual Reality & Urban Ecology" VR model during October 2017 test.

After using the VR, each user was asked to complete an exit survey where they answered questions about their likes/dislikes of the VR, what their overall satisfaction was, and whether or not they thought VR would be a helpful tool in urban planning. Nearly 60 participants were surveyed. Most users (90%) were excited about the VR and enjoyed the experience. The most common complaints included motion sickness, difficulty wearing the headset over glasses, and dizziness. Nearly all (95%) people surveyed were first time VR users and most were satisfied with the experience. Additionally, about 95% of our users encouraged the use of VR for urban planning and

thought the VR model helped them further understand the system dynamics of urban infrastructure.

VR has been used in urban planning contexts over the years. VR and VRGIS has been used in urban planning because of its “powerful immersive visualization approach… [that] can be used to better engage with, and collect the opinion of, stakeholders and citizens/communities about any proposed future city plans affecting the places they live and work in” (Kamel Boulos et al., 2017). Typically, more than 90% of information required for a city's administration has a spatial component, such as location of facilities, routing delivery and provision of facilities, meaning GIS has been seen as an essential technology for urban management. Our proposed VR application for urban infrastructure management along with GIS gives planners the potential and ability to make advised choices in the spatial decision-making framework by incorporating a combination of computer and information technology, urban growth models, and computer-based visualization techniques to support community-based planning. Planners, surveyors, utilities and engineers primarily rely on GIS technology to design and map facilities in the cities to assist in the urban planning process.

Recent developments in 3D-GIS and urban data modeling are leading to innovations in the representation, storage and analysis based on 3D city and landscape models (Breunig & Zlatanova, 2011; Tang & Zhang, 2008; Wang, 2005). Incorporating the belowground urban infrastructure environment will be critical moving forward with these technologies, especially when construction activity inadvertently but commonly disrupts and damages underground infrastructure. Cities are building better models to

address this problem and incorporating VR technology will advance these efforts even further. For example, the City of Las Vegas has developed a 3D CAD model of their above- and below-ground infrastructure in the core downtown area to improve safety and awareness of below ground utilities (Haala and Kada, 2010a, Haala and Kada, 2010b). The city wanted an accurate and up-to-date source of information for urban planning, designing, and maintenance of its infrastructure that included both above and belowground infrastructure components.

VR can be an exceptional tool for communicating spatial information and sparking excitement in areas of the urban environment previously unexplored. However, there are many shortcomings to this technology that can prevent its implementation on a wide scale. In the VR community, it is "widely acknowledged that creating [VR models] is challenging, and requires carefully-crafted research and technological progress" (Çöltekin, Oprean, Wallgrün, & Klippel, 2019). Additionally, because of the complexities of the interaction modalities, implementing an intuitive large-scale VR model is cost-intensive and time-consuming for urban stakeholders already preoccupied with other concerns. An ongoing area of research includes improving the accessibility of large-scale VR environments so that these communication tools can be more widely dispersed (Çöltekin et al., 2019). VR models are impressively time demanding, especially in an urban environment where an accurate model requires a large amount of geometric, satellite, LiDAR, and aerial or street-level data (Kamel Boulos et al., 2017). VR is also made to be an interactive technology and require a relatively advanced rendering technology to gain the full effect, that is not always available to people interested in

utilizing the technology (Kamel Boulos et al., 2017). However, there are limitations as to what can be successfully communicated using VR. With a lack of data provided from the utilities, characteristics such as material, diameter, age, and pressure are important data that cannot be included in a VR model. Additionally, infrastructure is fragile and there are environmental factors beyond a municipality's control that can cause damage and can alter the integrity of the infrastructure. The unpredictability of the UIS is difficult to display in a VR model but can be crucial information for urban stakeholders to understand.

The VR model, “Virtual Reality & Urban Ecology”, is a tool meant to be shared. It's success in our team's initial pilot launch suggested that this tool would be useful to share with stakeholders involved in all branches of urban planning. Applying this VR as an educational tool would ideally pique the curiosity of people involved in aboveground infrastructure to learn how their infrastructure features play a role in the overall UIS and vice-versa.

5.4 Conclusion

Providing key stakeholders with spatial information about the system dynamics of urban infrastructure will be key for managing aging urban infrastructure in the most efficient and coordinated way (Pandit et al., 2015). Unfortunately, communicating spatial data to stakeholders that highlights the interactions of infrastructure elements above and below the ground is challenging. The complex system that exists below city streets is difficult to visualize but its impact on aboveground infrastructure is critical and its role cannot be

overlooked when making urban planning decisions. VR has been used in urban planning for many years as an effective method for conveying spatial data in the built environment. Our team utilized VR to visualize and communicate the specific spatial and dynamic relationships between infrastructure features above and below city streets in response to a clear lack of such a tool in current management practices. Our VR tool benefitted stakeholders by helping to address the issues uncovered in our stakeholder workshop relating to a lack of foundational understanding of the interactions of above and belowground infrastructure features. VR enabled us to provide an interactive experience to promote a better understanding of the built urban environment and the system dynamics of the infrastructure ecosystem (Billger, Thuvander, & Stahre Wästberg, 2016). Using VR as a spatial communication tool will be beneficial in informing urban stakeholders about how infrastructure features work together and encourage the implementation of a more coordinated urban infrastructure management plan (Howard & Gaborit, 2007). To build upon this research into the future, a more complete neighborhood could be modeled in the next VR, perhaps highlighting a proposed infrastructure reconstruction plan. Climate change models and growing population metrics could be incorporated into a VR to help plan a more efficient infrastructure project that would be resilient into the future.

Acknowledgements: Thanks to Conor Leblanc and Katharine Lusk of the Boston University Initiative on Cities for organizing the stakeholder workshop described in this paper.

Funding: This material is based upon work supported by the National Science Foundation under Grant No. 1617053 and the Boston University Initiative on Cities Early-Stage Urban Research Award.

CHAPTER 6 – CONCLUSION

Continued population growth and development have increased the intensities of natural resource use and environmental degradation, threatening the stability and security of our planet. Coupled Human and Natural System (CHANS) offers a valuable framework to focus on understanding the process and pattern that characterizes the two-way dynamical interactions between human and natural systems. These interactions in CHANS are complex, where the degree of heterogeneity is highly dependent on the scale of the study. CHANS research often integrates data and methods across multiple spatial and temporal scales. The goal of CHANS is to find a pathway highlighting the development process with less impact on the biosphere and ecosystems to ensure a sustainable future. CHANS research results in analysis and insights to inform policy and decision-making in the form of spatially explicit models.

My dissertation research aims to build a set of methodologies and spatially explicit models and tools essential for operationalizing and monitoring CHANS, in studies ranging from local to regional scales across time span and multiple levels of interactions. In the second chapter, we explored the potential impact of energy investments on biodiversity and deforestation. A weighted site-wise biodiversity metric was employed to address the spatial scale mismatch due to the coarseness of IUCN geographic ranges. In addition to traditional deforestation area/rate measurement, we conducted a morphological spatial pattern analysis on Landsat derived annual tree cover. We found that energy investments with Chinese development finance disproportionately impact marine vertebrate species, which is caused by a higher number of coal power

plants in China's energy investment portfolio in Southeast Asia. Coal power plants have less geographic restriction in site location and are more likely to be within close proximity to the coastline near major population centers. Also, we discovered that deforestation area/rate doesn't exhibit all information on changes in land cover. Morphological Spatial Pattern Analysis (MSPA) shows that the construction of hydropower plants has a significantly higher impact on core forests. The morphology of forest loss is of great biological importance and could lead to better policy implementation in species conservation. Such as protecting connecting pathways and conservation of forest core dependent species.

The third chapter demonstrated the effect of spatial non-stationary models. Tobler's first law of geography says that "everything is related to everything else, but near things are more related than distant things". Residuals from Ordinary Least Square (OLS) regression exhibit spatial autocorrelation, indicating that the stationary model over or underestimates the pattern of clusters. Non-stationary model (geographically weighted regression) show lower RSS (Residual Sum of Squares) and AIC (Akaike information criterion) values compared with OLS. However, not all variables exhibit spatial non-stationary at a certain scale. Treating these variables as spatial non-stationary would cause over-fitting of the model. A mixed geographic weighted regression model is introduced to solve this issue. A Monte Carlo approach was used to determine whether a variable should be treated as a stationary variable.

In the fourth chapter, we examined the dynamics of aquatic vegetation in a shallow freshwater lake in East Africa. Aquatic vegetation is a critical refugium for

native endangered fish species. Previous studies show that there are difficulties in accurately classifying certain classes of aquatic vegetation given that they have similar spectral signatures and are often mixed with water. We proposed a workflow that utilized the entire Landsat archive to automatically extract water bodies from each Landsat image based on various water indices, adaptive binarization, and majority voting rules. Then we proposed two data products: aggregated water occurrence and aggregated maximum extent from extracted water bodies. Finally, we added spatial and temporal metrics derived from two data products that can capture the characteristics of different types of aquatic vegetation to Landsat spectral bands in order to improve classification accuracy. Our results show increased overall accuracy compared with traditional methods, from 60.2% to 86.8%.

We explored the emerging geospatial visualization technologies in the last chapter. We showed that there's complex CHANS coupling of infrastructures in urban scenarios. From our workshop, elected officials expressed concern with aboveground infrastructure, but there was a general lack of understanding and/or appreciation for how urban infrastructure functions as a system. We learned that we need an education and outreach tool to enable people to visualize the invisible. A virtual reality environment was built from remote sensing and GIS data; our survey shows that most users were excited about the VR and enjoyed the experience. About 95% of our users encouraged the use of VR for urban planning and thought the VR model helped them further understand the system dynamics of urban infrastructure. We showed that VR provides key stakeholders with spatial information about the system dynamics of urban

infrastructure. Using VR as a spatial communication tool will be beneficial in informing urban stakeholders about how infrastructure features work together and encourage the implementation of a more coordinated urban infrastructure management plan.

My research demonstrates how geospatial analysis and spatial statistics can be integrated to spatialize the CHANS framework to find sustainable solutions to produce relevant analysis and provide decision and policy insights to the stakeholders, including the public. The geospatial methods presented in this dissertation facilitate translating data-driven geospatial analysis into policy and community action through spatial, temporal, and organizational coupling.

APPENDIX A

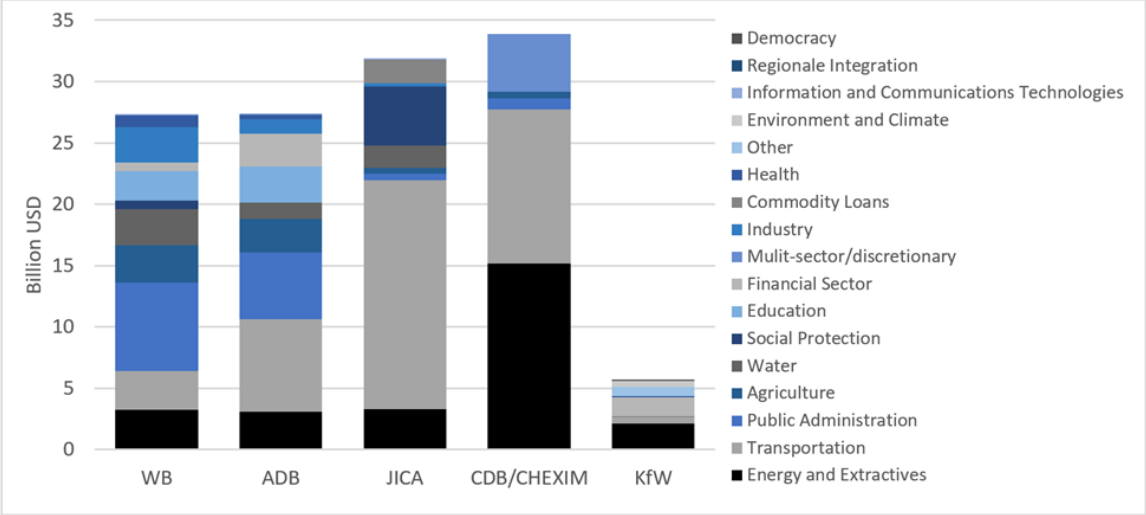

Figure A1. Development finance commitment from main DFIs in Southeast Asia (2013 - 2019). Compiled from development finance institution project databases & China's Overseas Development Finance Database, Global Development Policy Center, Boston University.

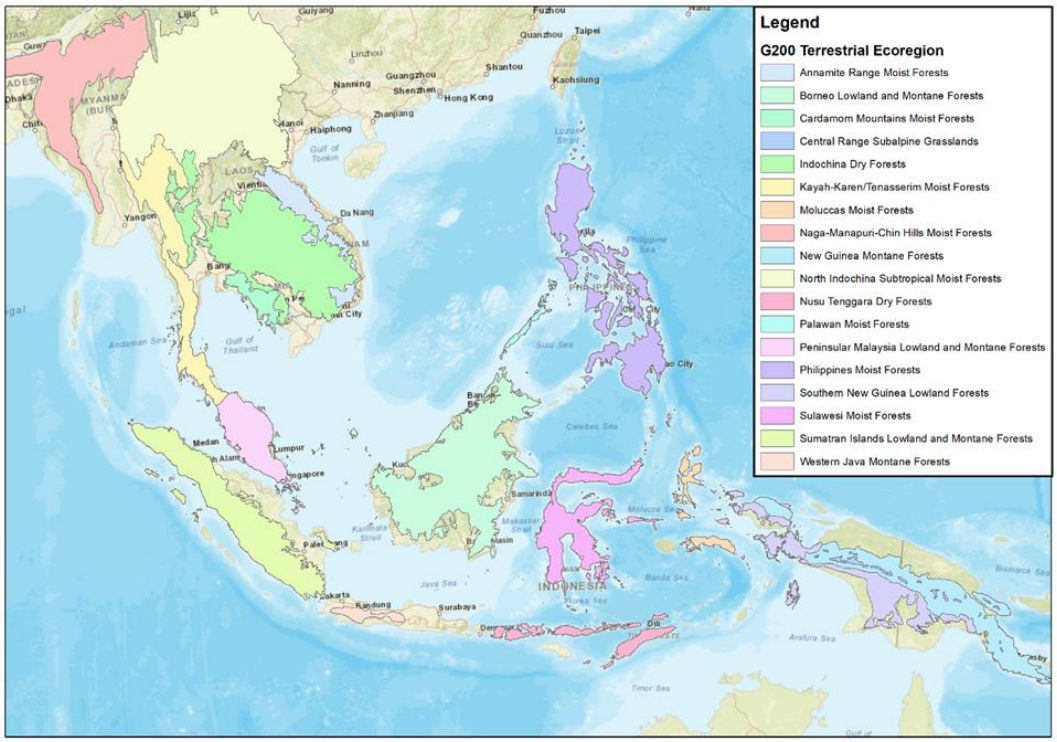

Figure A2. WWF Global 200 terrestrial ecoregions in Southeast Asia.

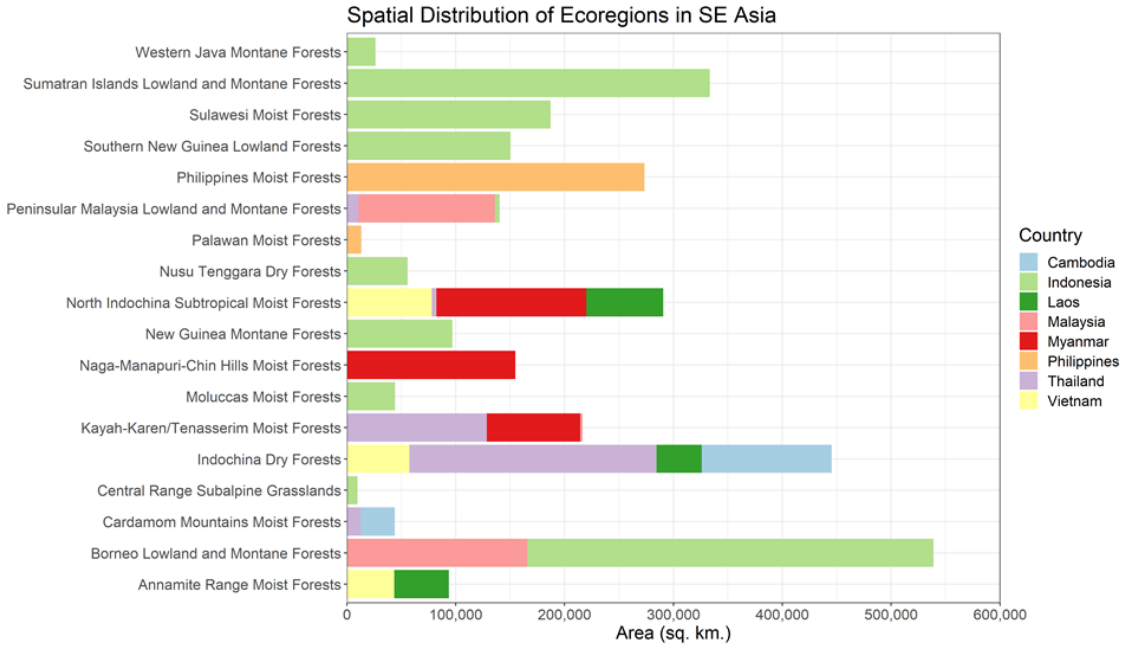

Figure A3. Spatial distribution of ecoregions in Southeast Asia. area (Sq Km) by country.

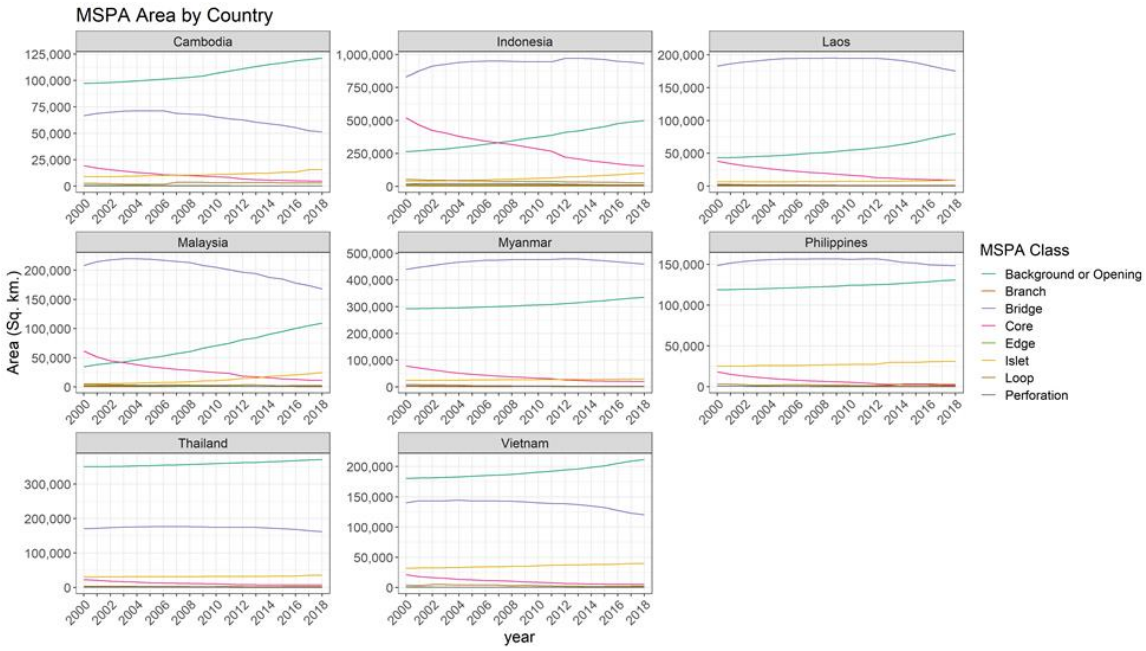

Figure A4. MSPA results from 2000-2018 showing changes in forest morphology in the 8 countries of Southeast Asia.

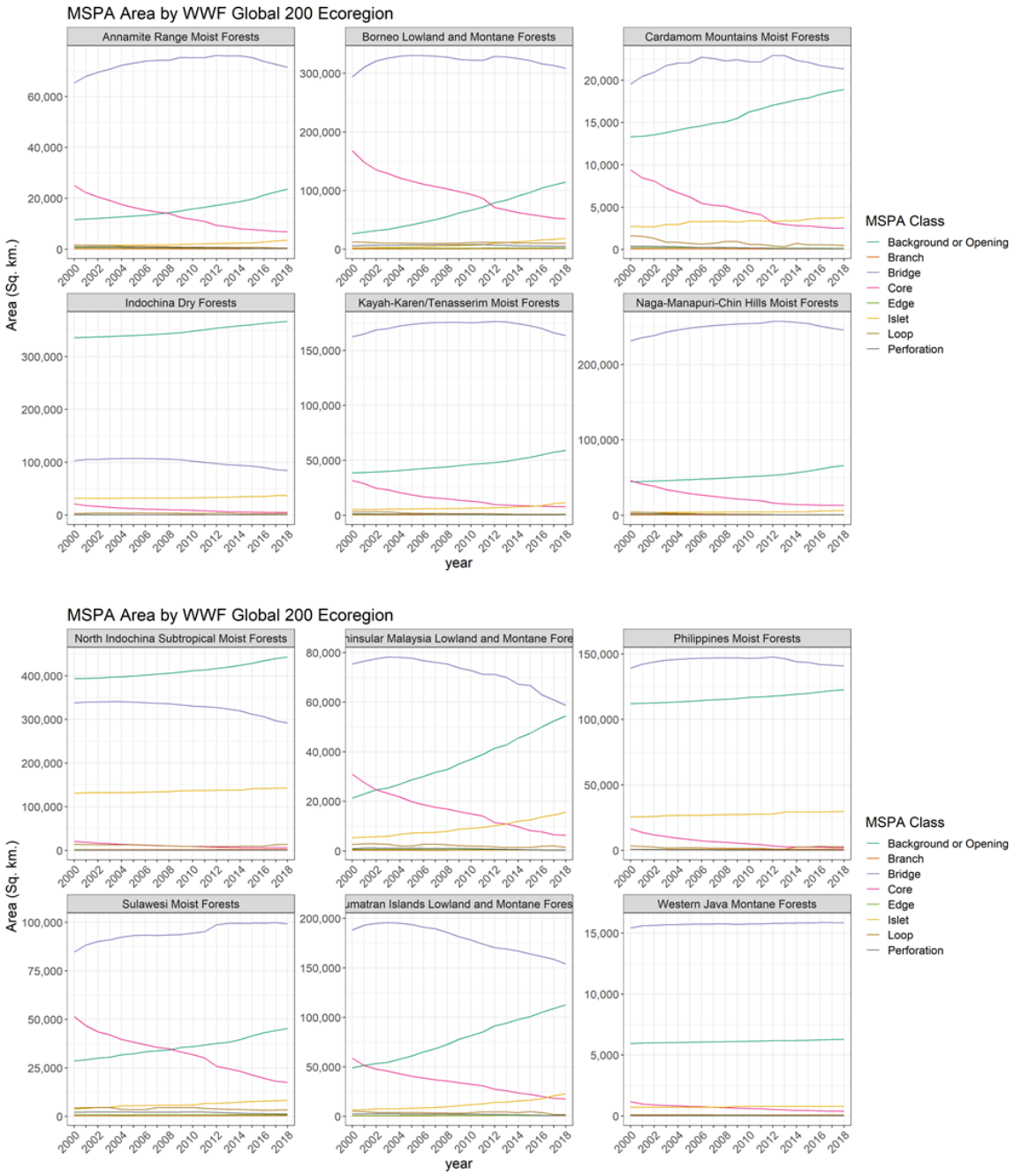

Figure A5. MSPA results from 2000-2018 showing changes in forest morphology in the 12 ecoregions of Southeast Asia.

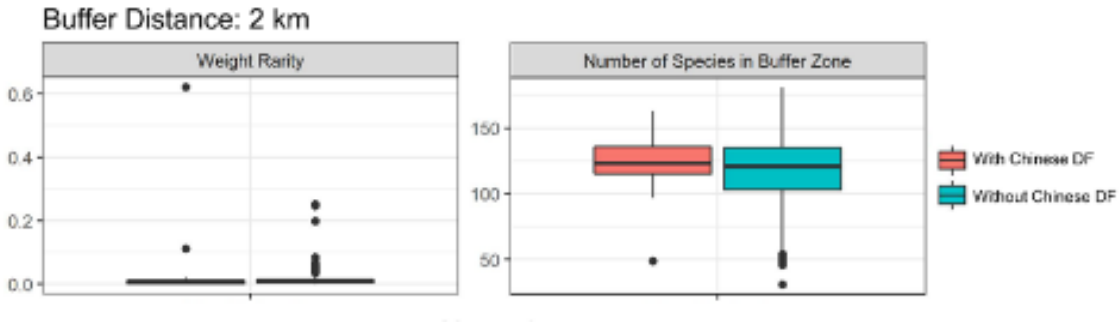

Figure A6. Differences in weighted range size rarity and number of species of mammals in 2-kilometer buffer zones of power plants with and without Chinese Development Finance.

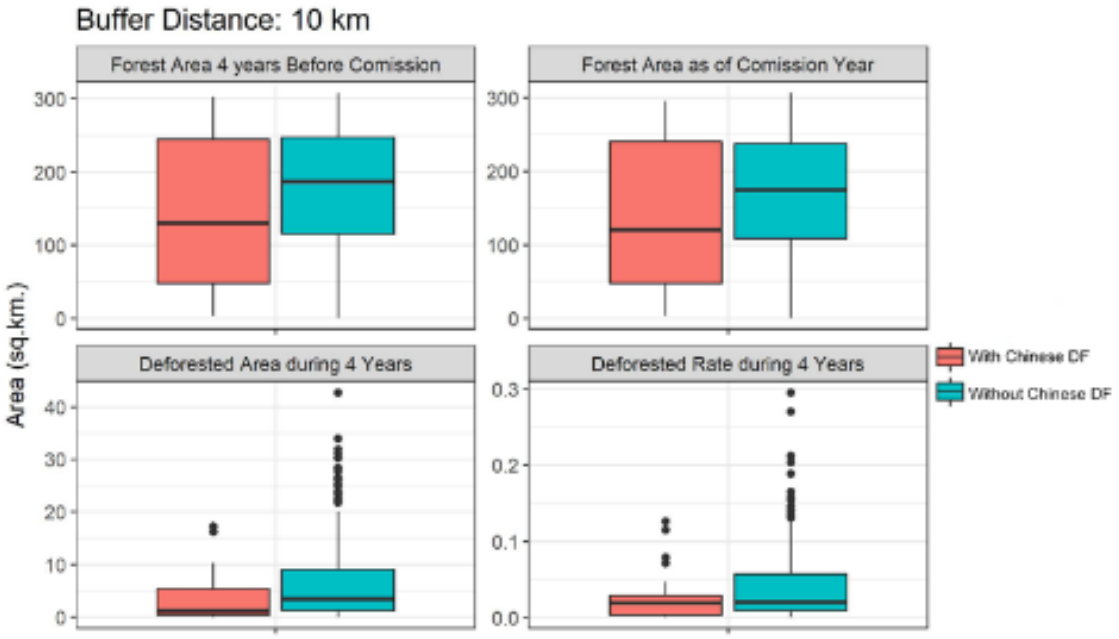

Figure A7. Forest area four years prior to investment and deforestation area during and four years after the commission year of the power plants.

Table A1. Energy investments in Southeast Asia with and without Chinese development finance

Countries	Primary Fuel	with Chinese DF	without Chinese DF
Cambodia	Coal	0	2
	Hydro	3	2
Indonesia	Coal	19	59
	Hydro	1	41
Laos	Coal	0	1
	Hydro	13	15
Malaysia	Coal	0	7
	Hydro	0	9
Myanmar	Coal	0	2
	Hydro	2	17
Philippines	Coal	1	22
	Hydro	0	17
Thailand	Coal	0	5
	Hydro	0	10
Vietnam	Coal	12	15
	Hydro	0	174
Sum	Coal	32	113
	Hydro	19	285

Table A2. The total weighted rarity of species at PLTU Tanjung Kasam in Indonesia

IUCN Categories	Species Count	Weight
Least Concern	24	1
Near Threatened	4	2
Vulnerable	7	4
Endangered	3	6
Critically Endangered	1	8
Data Deficient	10	2

APPENDIX B

Tables from B1 to B6: The first 3 columns refer to range of coefficients and median values of the intercepts and 10 independent variables. The % positive and % negative refer to the sign of coefficients. p-Values of the F3 and the significance refer to the coefficients (Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.').

Median Home Price							
Basic GWR Model (2000)							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (F3)**	**Sig.**
Intercept	−212.70	243.24	1190.76	90.77	9.23	1.47×10^{-11}	***
Population Density	−11.92	0.68	28.78	60.71	39.29	0.5142965	-
Unprotected Forest	−217.68	68.88	388.90	84.52	15.48	0.3060173	-
Unemployment Rate	−189.08	−17.66	16.92	11.31	88.69	$<2.2 \times 10^{-16}$	***
Residential Area	−756.43	−73.39	300.69	35.71	64.29	0.0004993	***
Vehicle ownership	−1323.53	−162.52	229.86	9.82	90.18	$<2.2 \times 10^{-16}$	***
Higher Education	297.39	1110.62	1985.25	100	0	6.85×10^{-13}	***
Senior Population	−522.05	226.41	1470.80	84.82	15.18	1.04×10^{-11}	***
Dist. to Stations	−8.48	−0.18	16.72	39.88	60.12	3.59×10^{-6}	***
Property Tax	−20.63	−3.48	1.38	10.12	89.88	6.51×10^{-12}	***
CPI	−1.37	1.38	6.78	80.65	19.35	0.040276	*

Table B1. GWR results for home prices in MA towns 2000.

Median Home Price							
Basic GWR Model (2010)							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (F3)**	**Sig.**
Intercept	−130.54	347.02	1501.33	97.92	2.08	3.07×10^{-2}	*
Population Density	−11.22	2.30	25.15	78.27	21.73	0.531882	-
Unprotected Forest	−61.80	100.84	745.17	77.08	22.92	1.90×10^{-13}	***
Unemployment Rate	−114.97	−15.33	3.10	2.68	97.32	$<2.2 \times 10^{-16}$	***
Residential Area	−781.96	−107.39	246.64	33.63	66.37	4.49×10^{-9}	***
Vehicle ownership	−872.92	−274.53	64.77	2.08	97.92	1.57×10^{-7}	***
Higher Education	−2067.55	626.96	2327.99	88.99	11.01	$<2.2 \times 10^{-16}$	***
Senior Population	−409.73	660.41	2574.71	83.63	16.37	$<2.2 \times 10^{-16}$	***
Dist. to Stations	−18.26	−2.00	3.39	25	75	$<2.2 \times 10^{-16}$	***
Property Tax	−29.32	−7.03	2.51	4.17	95.83	3.94×10^{-3}	**
CPI	−3.72	3.59	17.98	87.5	12.5	0.043231	*

Table B2. GWR results for home prices in MA towns 2010.

Median Home Price							
Basic GWR Model (2009)							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (F3)**	**Sig.**
Intercept	−1507.61	−34.44	433.69	44.94	55.06	3.91×10^{-5}	***
Population Density	−10.47	5.07	30.88	89.88	10.12	0.2161	-
Unprotected Forest	−59.97	141.74	812.44	93.45	6.55	$<2.2 \times 10^{-16}$	***
Unemployment Rate	−29.38	−3.14	17.12	21.43	78.57	7.21×10^{-8}	***
Residential Area	−842.60	−139.32	242.05	15.18	84.82	4.40×10^{-5}	***
Vehicle ownership	−1018.52	−300.08	126.04	2.98	97.02	5.52×10^{-7}	***
Higher Education	−545.12	1083.32	2578.53	97.92	2.08	$<2.2 \times 10^{-16}$	***
Senior Population	−625.56	537.45	2741.07	82.14	17.86	$<2.2 \times 10^{-16}$	***
Dist. to Stations	−12.16	−2.29	3.16	26.19	73.81	1.90×10^{-5}	***
Property Tax	−37.52	−7.63	8.01	7.44	92.56	7.09×10^{-8}	***
CPI	0.09	5.47	28.54	100	0	$<2.2 \times 10^{-16}$	***

Table B3. GWR results for home prices in MA towns 2009.

Median Home Price							
Basic GWR Model (2011)							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (F3)**	**Sig.**
Intercept	−1213.56	1.45	339.49	50.6	49.4	5.81×10^{-1}	-
Population Density	−7.34	6.93	31.59	92.86	7.14	0.2586052	-
Unprotected Forest	−154.56	97.60	857.46	79.76	20.24	$<2.2 \times 10^{-16}$	***
Unemployment Rate	−35.99	−7.29	2.88	1.79	98.21	1.72×10^{-8}	***
Residential Area	−950.93	−169.73	147.89	9.23	90.77	1.03×10^{-3}	**
Vehicle ownership	−860.20	−253.15	99.18	2.38	97.62	1.93×10^{-3}	**
Higher Education	−1494.37	800.34	2474.53	94.94	5.06	$<2.2 \times 10^{-16}$	***
Senior Population	−567.12	539.88	2330.52	87.5	12.5	$<2.2 \times 10^{-16}$	***
Dist. to Stations	−9.94	−1.14	3.57	24.4	75.6	1.58×10^{-1}	-
Property Tax	−25.86	−5.60	3.56	3.87	96.13	9.45×10^{-4}	***
CPI	−0.22	4.12	25.70	99.4	0.6	1.20×10^{-13}	***

Table B4. GWR results for home prices in MA towns 2011.

Median Home Price							
Basic GWR Model (2012)							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (F3)**	**Sig.**
Intercept	−918.93	−27.37	341.96	46.13	53.87	4.86×10^{-1}	-
Population Density	−8.03	7.73	40.81	92.56	7.44	0.005703	**
Unprotected Forest	−191.22	94.72	770.02	75	25	$< 2.2 \times 10^{-16}$	***
Unemployment Rate	−40.78	−7.78	4.25	4.46	95.54	2.99×10^{-16}	***
Residential Area	−1178.58	−203.18	144.67	9.23	90.77	3.64×10^{-6}	***
Vehicle ownership	−742.86	−222.69	93.19	13.39	86.61	6.33×10^{-5}	***
Higher Education	−1489.48	837.34	2527.95	94.35	5.65	$< 2.2 \times 10^{-16}$	***
Senior Population	−477.13	481.55	2309.47	87.5	12.5	5.66×10^{-15}	***
Dist. to Stations	−11.23	−0.94	10.82	29.17	70.83	9.25×10^{-5}	***
Property Tax	−26.11	−5.16	5.01	3.57	96.43	8.22×10^{-9}	***
CPI	0.76	4.62	23.77	100	0	$< 2.2 \times 10^{-16}$	***

Table B5. GWR results for home prices in MA towns 2012.

Median Home Price							
Basic GWR Model (2013)							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (F3)**	**Sig.**
Intercept	−748.02	−159.96	279.56	30.95	69.05	9.71×10^{-1}	-
Population Density	−7.81	7.60	41.48	93.15	6.85	1.43×10^{-6}	***
Unprotected Forest	−63.64	120.01	768.53	65.77	34.23	$<2.2 \times 10^{-16}$	***
Unemployment Rate	−32.81	−6.94	2.92	6.55	93.45	3.50×10^{-12}	***
Residential Area	−937.70	−202.99	148.44	13.1	86.9	9.43×10^{-6}	***
Vehicle ownership	−796.82	−259.97	519.41	5.65	94.35	1.62×10^{-8}	***
Higher Education	−537.89	843.47	1975.59	97.62	2.38	$<2.2 \times 10^{-16}$	***
Senior Population	−408.29	656.82	2340.80	91.96	8.04	$<2.2 \times 10^{-16}$	***
Dist. to Stations	−10.98	−0.95	10.05	25	75	8.52×10^{-6}	***
Property Tax	−21.75	−5.72	1.91	8.33	91.67	7.10×10^{-9}	***
CPI	1.05	6.42	20.66	100	0	2.01×10^{-5}	***

Table B6. GWR results for home prices in MA towns 2013.

From table B7 to B12: The first 3 columns refer to range of coefficients and median values of the intercepts and non-stationary independent variables. The % positive and % negative refer to the sign of coefficients. p-Values of the Monte Carlo test and the significance refer to the coefficients (critical value 0.05). Global variables are stationary are shown in the lower panel of the table along with their coefficient estimates. (Signif. Codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.').

Median Home Price							
Mixed GWR Model (2000)							
Local Variables							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (MC)**	**Sig.**
Intercept	−47.18	212.41	1440.56	94.64	5.36	0.04	.
Population Density	−8.63	1.26	25.85	63.39	36.61	0.00	***
Unemployment Rate	−244.13	−24.55	13.49	23.21	76.79	0.00	***
Residential Area	−730.09	−98.82	279.04	34.23	65.77	0.00	***
Vehicle ownership	−1352.88	−208.48	304.02	11.61	88.39	0.00	***
Senior Population	−528.56	179.11	1395.85	77.38	22.62	0.00	***
Dist. to Stations	−8.04	−0.27	14.43	34.23	65.77	0.00	***
Property Tax	−17.89	−3.25	0.84	10.42	89.58	0.04	.
Global Variables	-	-	-	-	-	-	-
Unprotected Forest	-	-	-	72.49	-	0.48	-
Higher Education	-	-	-	1063.50	-	0.11	-
CPI	-	-	-	1.08	-	0.32	-

Table B7. MGWR results for home prices in MA towns 2000.

Median Home Price							
Mixed GWR Model (2009)							
Local Variables							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (MC)**	**Sig.**
Population Density	1.68	6.92	14.98	100	0	0.01	.
Unprotected Forest	−18.47	136.56	877.81	97.02	2.98	0.01	.
Unemployment Rate	−32.56	−4.52	5.24	9.52	90.48	0.04	.
Higher Education	−46.09	1027.44	2162.85	99.7	0.3	0.01	.
Senior Population	−717.27	543.89	2525.48	84.82	15.18	0.00	***
Dist. to Stations	−17.40	−2.75	4.61	26.79	73.21	0.00	***
Property Tax	1.29	5.19	9.59	100	0	0.01	.
Global Variables	-	-	-	-	-	-	-
Intercept	-	-	-	−22.98	-	0.23	-
Residential Area	-	-	-	−192.74	-	0.08	-
Vehicle ownership	-	-	-	−263.5063	-	0.21	-
Property Tax	-	-	-	−6.2983	-	0.19	-

Table B8. MGWR results for home prices in MA towns 2009.

Median Home Price							
Mixed GWR Model (2010)							
Local Variables							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (MC)**	**Sig.**
Population Density	−10.61	3.01	23.37	75.6	24.4	0.00	***
Unprotected Forest	−117.24	81.15	1128.29	82.14	17.86	0.02	.
Unemployment Rate	−78.05	−17.19	−3.43	0	100	0.00	***
Residential Area	−693.29	−89.47	238.38	27.38	72.62	0.01	.
Higher Education	−50.24	901.48	1603.10	99.7	0.3	0.00	***
Senior Population	−676.64	641.03	3500.10	77.68	22.32	0.00	***
Dist. to Stations	−20.46	−1.46	4.61	20.54	79.46	0.00	***
Global Variables	-	-	-	-	-	-	-
Intercept	-	-	-	273.3707	-	0.45	-
Vehicle ownership	-	-	-	−243.0027	-	0.13	-
Property Tax	-	-	-	−4.75	-	0.51	-
CPI	-	-	-	273.37	-	0.53	-

Table B9. MGWR results for home prices in MA towns 2010.

Median Home Price							
Mixed GWR Model (2011)							
Local Variables							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (MC)**	**Sig.**
Population Density	−3.65	7.91	31.49	95.54	4.46	0.01	.
Unprotected Forest	−283.68	121.03	783.88	91.37	8.63	0.00	***
Residential Area	−970.74	−192.49	132.06	6.55	93.45	0.00	***
Higher Education	−150.67	936.71	2298.48	99.11	0.89	0.00	***
Senior Population	−456.07	496.82	2273.67	88.1	11.9	0.00	***
Dist. to Stations	−13.77	−1.91	3.66	18.15	81.85	0.00	***
CPI	−1.25	4.93	9.95	97.62	2.38	0.00	***
Global Variables	-	-	-	-	-	-	-
Intercept	-	-	-	−18.07	-	-	-
Unemployment Rate	-	-	-	−7.38	-	-	-
Vehicle ownership	-	-	-	−183.00	-	-	-
Property Tax	-	-	-	−6.56	-	-	-

Table B10. MGWR results for home prices in MA towns 2011.

Median Home Price							
Mixed GWR Model (2012)							
Local Variables							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (MC)**	**Sig.**
Population Density	−7.00	9.36	42.00	97.62	2.38	0.00	***
Unprotected Forest	0.77	85.71	712.42	100	100	0.00	***
Residential Area	−1101.51	−262.61	198.37	5.65	94.35	0.00	***
Higher Education	−183.99	927.83	1689.29	99.7	0.3	0.00	***
Senior Population	−474.79	444.50	2876.72	84.23	15.77	0.00	***
Dist. to Stations	−17.50	−1.10	4.17	20.83	79.17	0.00	***
Global Variables	-	-	-	-	-	-	-
Intercept	-	-	-	7.78	-	0.77	-
Unemployment Rate	-	-	-	−7.8404	-	0.06	-
Vehicle ownership	-	-	-	−196.24	-	0.64	-
Property Tax	-	-	-	−6.95	-	0.06	-
CPI	-	-	-	4.7264	-	0.05	-

Table B11. MGWR results for home prices in MA towns 2012.

Median Home Price							
Mixed GWR Model (2013)							
Local Variables							
Variables	**Minimum**	**Median**	**Max**	**% of Positive**	**% of Negative**	***p*-Value (MC)**	**Sig.**
Population Density	−5.25	8.75	32.65	94.94	5.06	0.00	***
Unprotected Forest	−71.06	100.28	960.42	64.58	35.42	0.00	***
Unemployment Rate	−32.11	−7.66	0.85	2.68	97.32	0.02	.
Residential Area	−867.56	−232.16	119.82	6.25	93.75	0.02	.
Higher Education	73.76	800.33	1532.83	100	0	0.01	.
Senior Population	−606.89	584.40	2809.65	86.9	13.1	0.00	***
Dist. to Stations	−11.99	−1.68	8.03	17.26	82.74	0.00	***
Property Tax	−19.62	−7.01	1.37	5.06	94.94	0.03	.
Global Variables	-	-	-	-	-	-	-
Intercept	-	-	-	−105.84	-	0.90	-
Vehicle ownership	-	-	-	−206.13	-	0.24	-
CPI	-	-	-	6.03	-	0.37	-

Table B12. MGWR results for home prices in MA towns 2013.

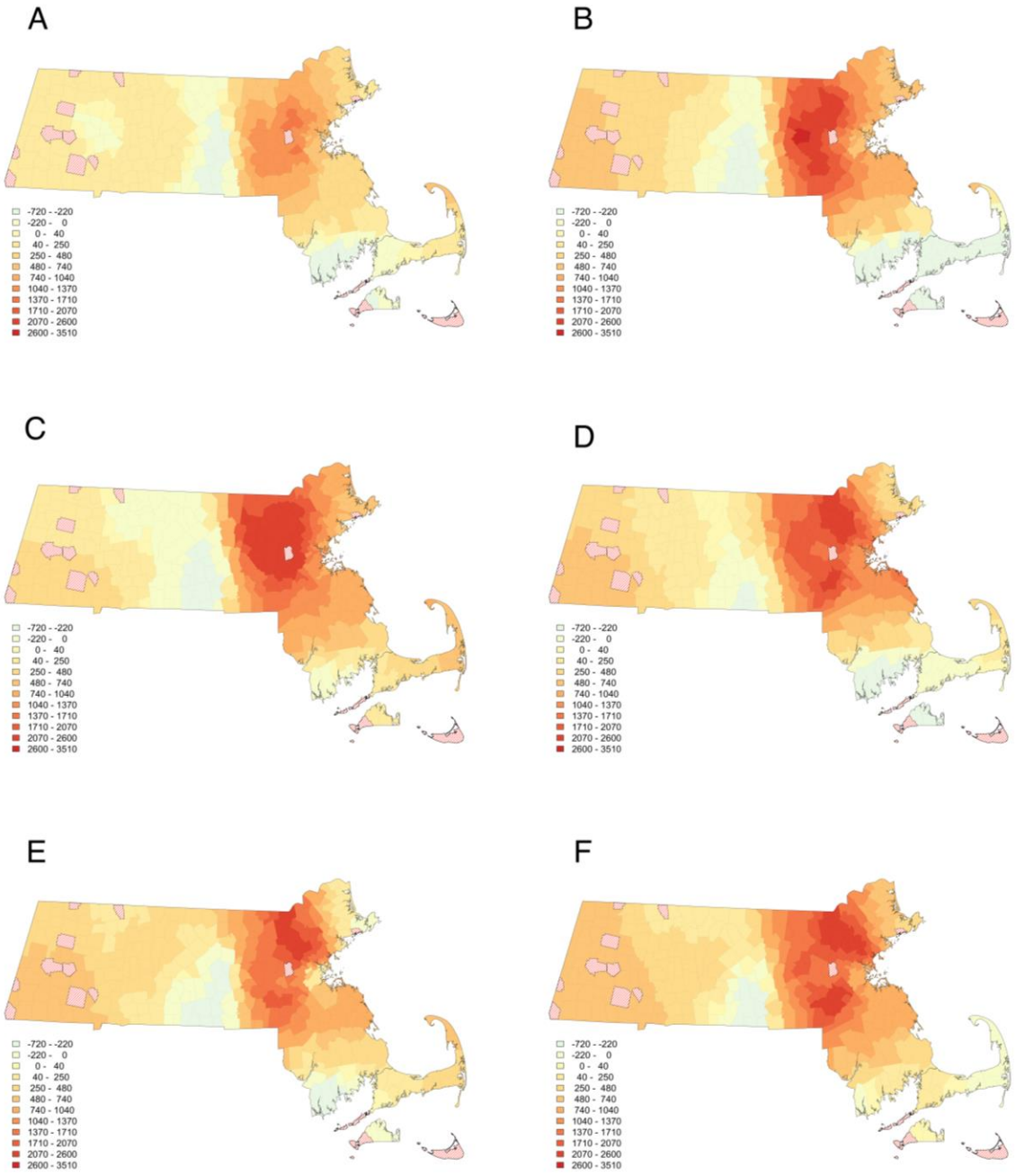

Figure B1. Coefficient estimates of basic GWR for senior population: (A) 2000; (B) 2009; (C) 2010; (D) 2011; (E) 2012; (F) 2013.

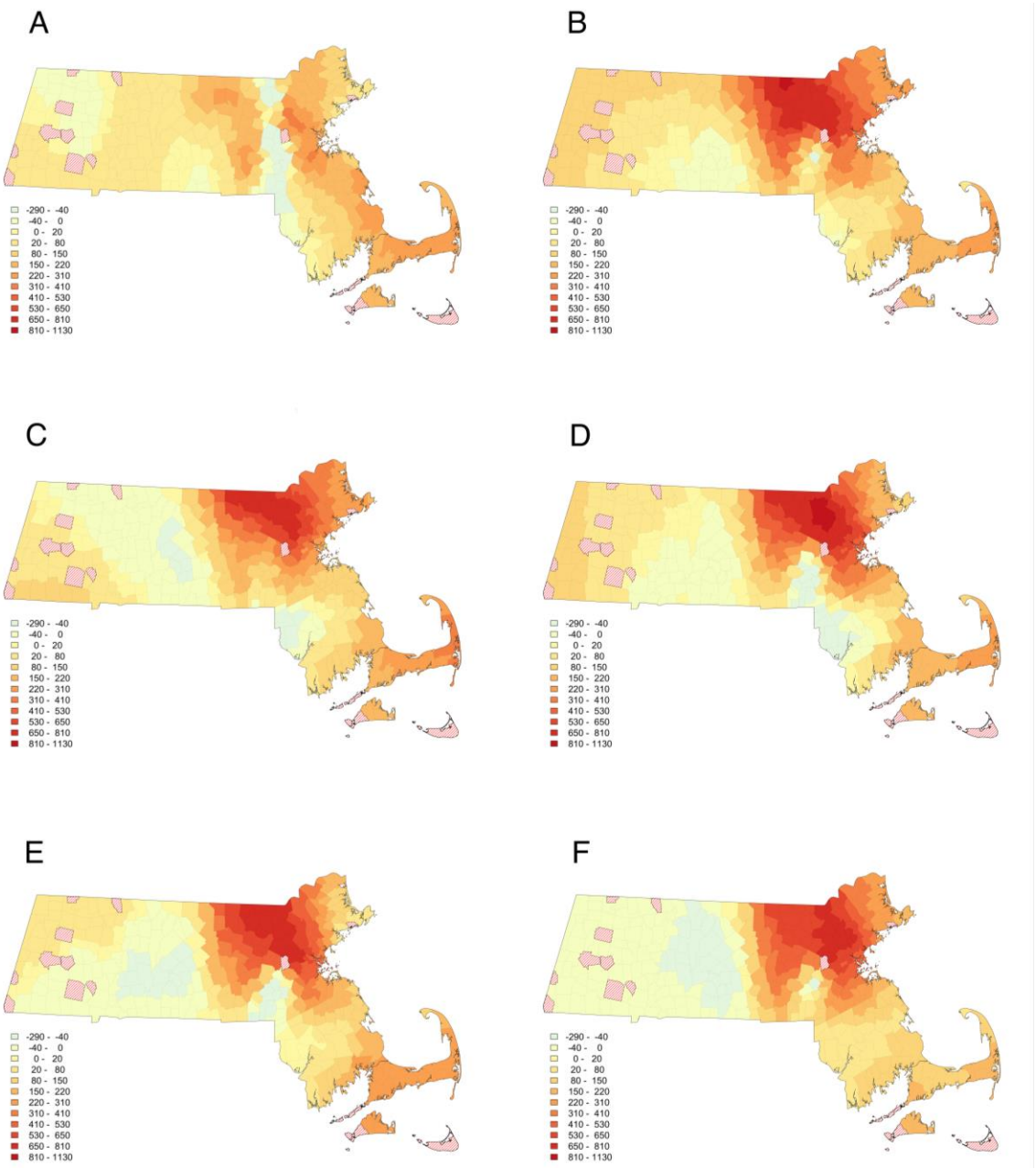

Figure B2. Coefficient estimates of basic GWR for unprotected forest: (A) 2000; (B) 2009; (C) 2010; (D) 2011; (E) 2012; (F) 2013.

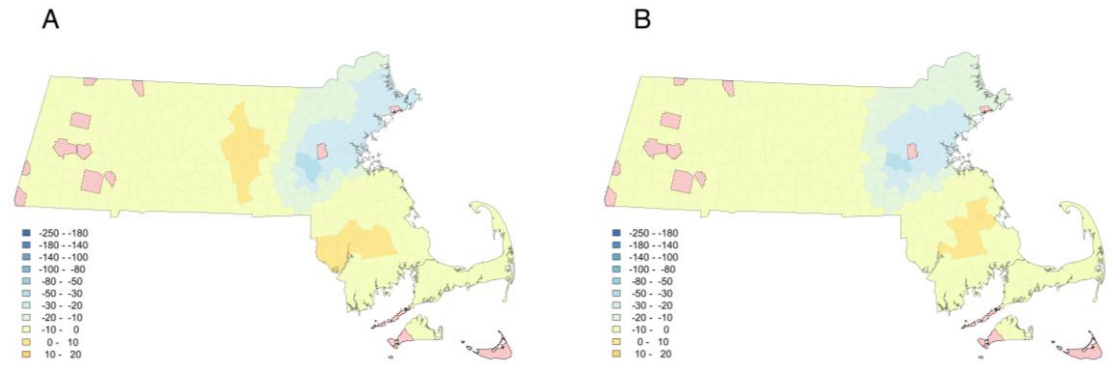

Figure B3. Unemployment coefficient for 2013: (A) Basic GWR; (B) MGWR.

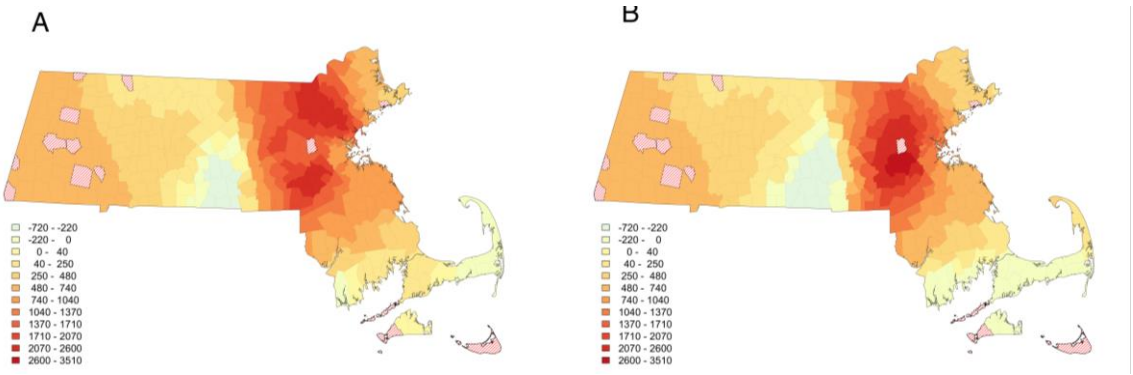

Figure B4. Senior population coefficient for 2013: (A) Basic GWR; (B) MGWR.

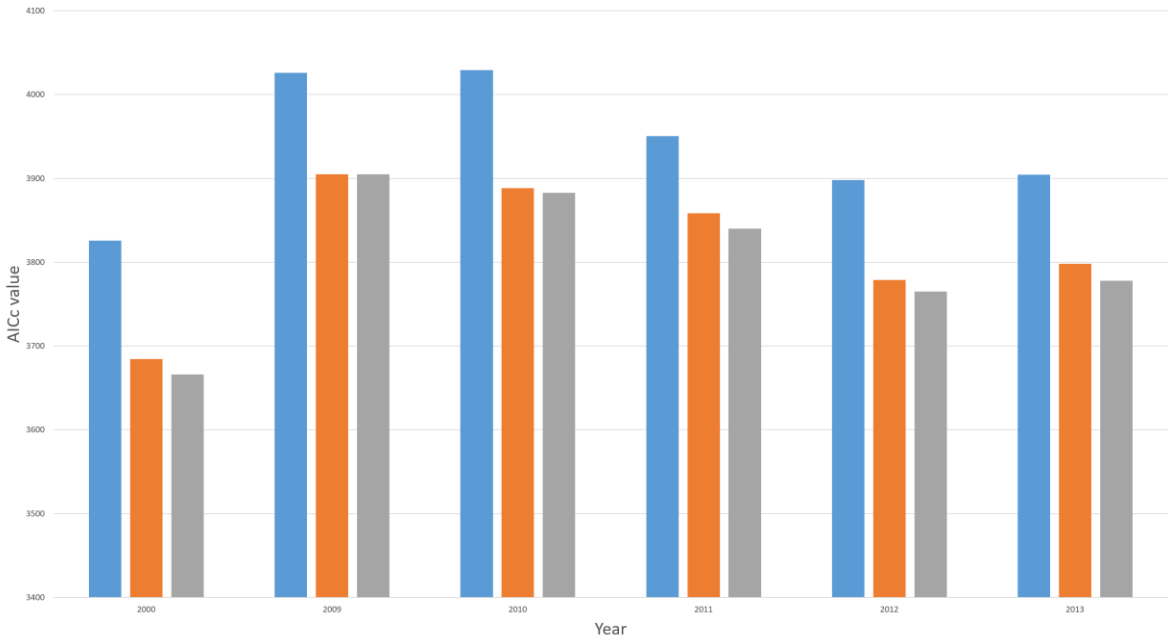

Figure B5. Senior population coefficient for 2013: OLS (Blue); Basic GWR (Brown); MGWR (Gray).

REFERENCES

2010 Glossary of AYP Reporting Terms. (2010). https://profiles.doe.mass.edu/ayp/ayp_report/glossary2010.html#cpi

Achieng, A. O., Masese, F. O., & Kaunda-Arara, B. (2020). Fish assemblages and size-spectra variation among rivers of Lake Victoria Basin, Kenya. *Ecological Indicators*, *118*, 106745. https://doi.org/10/gk8h6r

Adams, W.M., Aveling, R., Brockington, D., Dickson, B., Elliott, J., Hutton, J., Roe, D., Vira, B., and Wolmer, W., 2004. Biodiversity conservation and the eradication of poverty. *Science (New York, N.Y.)*, 306 (5699), 1146–1149.

ADB, n.d. Projects and Tenders [online]. *Projects and Tenders*. Available from: https://www.adb.org/projects [Accessed 15 Feb 2021].

AIIB, 2016. Environmental and Social Framework - Framework Agreements - AIIB [online]. Available from: https://www.aiib.org/en/policies-strategies/framework-agreements/environmental-social-framework.html [Accessed 28 Dec 2020].

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10/d98qkw

Alamgir, M., Campbell, M.J., Sloan, S., Suhardiman, A., Supriatna, J., and Laurance, W.F., 2019. High-risk infrastructure projects pose imminent threats to forests in Indonesian Borneo. *Scientific Reports*, 9 (1), 140.

Alberti, M. (1999). Modeling the urban ecosystem: A conceptual framework. *Environment and Planning B: Planning and Design*, *26*(4), 605–630. https://doi.org/10/chdhxq

Albright, T. P., & Ode, D. J. (2011). Monitoring the dynamics of an invasive emergent macrophyte community using operational remote sensing data. *Hydrobiologia*, *661*(1), 469–474. https://doi.org/10/bvsn6c

Ali, R., Barra, A.F., Berg, C., Damania, R., Nash, J., and Russ, J., 2015. *Highways to Success or Byways to Waste: Estimating the Economic Benefits of Roads in Africa*. Washington, DC : International Bank for Reconstruction and Development/The World Bank: World Bank Publications.

Alkon, A. H., & Agyeman, J. (2011). *Cultivating Food Justice: Race, Class, and Sustainability*. The MIT Press. https://muse.jhu.edu/book/23088

Allen, M., Regenbrecht, H., & Abbott, M. (2011). Smart-phone augmented reality for public participation in urban planning. *Proceedings of the 23rd Australian computer-human interaction conference* (pp. 11–20). https://doi.org/10.1145/2071536.2071538

Alroy, J., 2017. Effects of habitat disturbance on tropical forest biodiversity. *Proceedings of the National Academy of Sciences*, 114 (23), 6056–6061.

Andersson, E., Barthel, S., Borgström, S., Colding, J., Elmqvist, T., Folke, C., & Gren, A. (2014). Reconnecting cities to the biosphere: Stewardship of green infrastructure and urban ecosystem services. *AMBIO*, *43*(4), 445–453. https://doi.org/10.1007/s13280-014-0506-y

Anselin, L. (1998). GIS Research Infrastructure for Spatial Analysis of Real Estate Markets. *Journal of Housing Research*, *9*(1), 113–133. https://doi.org/10/gk33mn

Anselin, L. (2013). *Spatial Econometrics: Methods and Models*. Springer Science & Business Media.

APWA (2019). American public workers association - uniform color code. Web page. Date accessed: September 24, 2019. URL https://www3.apwa.net/content/library/colorcc.pdf

Ascensão, F., Fahrig, L., Clevenger, A.P., Corlett, R.T., Jaeger, J.A.G., Laurance, W.F., and Pereira, H.M., 2018. Environmental challenges for the Belt and Road Initiative. *Nature Sustainability*, 1 (5), 206–209.

Aughinbaugh, A. (2013). *Patterns of Homeownership, Delinquency, and Foreclosure Among Youngest Baby Boomers*. https://ecommons.cornell.edu/handle/1813/77459

Baird, I.G., 2016. Non-government Organizations, Villagers, Political Culture and the Lower Sesan 2 Dam in Northeastern Cambodia. *Critical Asian Studies*, 48 (2), 257–277.

Barber, C.P., Cochrane, M.A., Souza, C.M., and Laurance, W.F., 2014. Roads, deforestation, and the mitigating effect of protected areas in the Amazon. *Biological Conservation*, 177, 203–209.

Batty, M., Dodge, M., Doyle, S., & Hudson-Smith, A. (1998). *Modelling virtual urban en- vironments.* Centre for Advanced Spatial Analysis – University College London: Working paper series, paper 1. http://discovery.ucl.ac.uk/219/1/modelvue.pdf

Bebbington, A.J., Humphreys Bebbington, D., Sauls, L.A., Rogan, J., Agrawal, S., Gamboa, C., Imhof, A., Johnson, K., Rosa, H., Royo, A., Toumbourou, T., and Verdum, R., 2018. Resource extraction and infrastructure threaten forest cover and community rights. *Proceedings of the National Academy of Sciences of the United States of America*, 115 (52), 13164–13173.

Benítez-López, A., Alkemade, R., and Verweij, P.A., 2010. The impacts of roads and other infrastructure on mammal and bird populations: a meta-analysis. *Biological conservation*, 143 (6), 1307–1316.

Berger, M., & Bill, R. (2019). Combining VR visualization and sonification for immersive exploration of urban noise standards. *Multimodal Technologies and Interaction, 3*(2), 34. https://doi.org/10.3390/mti3020034

Billger, M., Thuvander, L., & Stahre Wästberg, B. (2016). In search of visualization challenges: The development and implementation of visualization tools for sup- porting dialogue in urban planning processes. *Environment and Planning B. Urban Analytics and City Science, 44*(6), https://doi.org/10.1177/0265813516657341

Birk, S., & Ecke, F. (2014). The potential of remote sensing in ecological status assessment of coloured lakes using aquatic plants. *Ecological Indicators, 46,* 398–406. https://doi.org/10.1016/j.ecolind.2014.06.035

Birot, H., Campera, M., Imron, M.A., and Nekaris, K. a. I., 2020. Artificial canopy bridges improve connectivity in fragmented landscapes: The case of Javan slow lorises in an agroforest environment. *American Journal of Primatology*, 82 (4), e23076.

Bitter, C., Mulligan, G. F., & Dall'erba, S. (2007). Incorporating spatial variation in housing attribute prices: A comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*, *9*(1), 7–27. https://doi.org/10/fh7vrd

Borda-de-Água, L., Barrientos, R., Beja, P., and Pereira, H.M., 2017. *Railway Ecology*. Springer.

Boulos, M. N. K., Lu, Z., Guerrero, P., Jennett, C., & Steed, A. (2017). From urban plan- ning and emergency training to Pokémon Go: Applications of virtual reality GIS (VRGIS) and augmented reality GIS (ARGIS) in personal, public and environmental health. *International Journal of Health Geographics, 16*, 7. https://doi.org/10.1186/s12942-017-0081-0

Bowman, A. W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, *71*(2), 353–360. https://doi.org/10.1093/biomet/71.2.353

Bradley, D., & Roth, G. (2007). Adaptive Thresholding using the Integral Image. *Journal of Graphics Tools*, *12*(2), 13–21. https://doi.org/10/cpgshv

Brasington, D., & Haurin, D. R. (2006). Educational Outcomes and House Values: A Test of the value added Approach*. *Journal of Regional Science, 46*(2), 245–268. https://doi.org/10/bsbfrk

Breunig, M., & Zlatanova, S. (2011). 3D geo-database research: Retrospective and future directions. *Computers & Geosciences, 37*(7), 791–803. https://doi.org/10.1016/j.cageo.2010.04.016

Breunig, M., Al-Doori, M., Butwilowski, E., Kuper, P. V., Benner, J., & Haefele, K. H. (2015). *3D geoinformation science: The selected papers of the 3D GeoInfo 2014. Vol. 6*, Springer International Publishing258. https://doi.org/10.1007/978-3-319-12181-9

Brunsdon, C., Fotheringham, S., & Charlton, M. (2000). *Geographically Weighted Regression as a Statistical Model* [Monograph]. University of Newcastle-upon-Tyne. http://mural.maynoothuniversity.ie/5975/

Buchanan, G.M., Parks, B.C., Donald, P.F., O'Donnell, B.F., Runfola, D., Swaddle, J.P., Tracewski, Ł., and Butchart, S.H.M., 2016. The Impacts of World Bank Development Projects on Sites of High Biodiversity Importance. *docs.aiddata.org*.

Buntaine, M.T., 2011. Does the Asian Development Bank Respond to Past Environmental Performance when Allocating Environmentally Risky Financing? *World Development*, 39 (3), 336–350.

Bureau, U. C. (n.d.-a). *American Community Survey (ACS)*. The United States Census Bureau. Retrieved July 3, 2021, from https://www.census.gov/programs-surveys/acs

Bureau, U. C. (n.d.-b). *Decennial Census Data*. The United States Census Bureau. Retrieved July 3, 2021, from https://www.census.gov/programs-surveys/decennial-census/data.html

Bush, E.R., Baker, S.E., and Macdonald, D.W., 2014. Global trade in exotic pets 2006-2012. *Conservation Biology: The Journal of the Society for Conservation Biology*, 28 (3), 663–676.

Butler, B. J. (2016). *Forests of Massachusetts, 2015*. https://doi.org/10/gk33mh

Byers, L., Friedrich, J., Hennig, R., Kressig, A., Li, X., McCormick, C., and Valeri, L.M., 2018. A global database of power plants. *World Resour. Inst*, 18.

Byun, K. (2010). The US housing bubble and bust: Impacts on employment. *Bureau of Labor and Statistics Monthly Review (December), 3&17*.

Cameron, R.D., Reed, D.J., Dau, J.R., and Smith, W.T., 1992. Redistribution of Calving Caribou in Response to Oil Field Development on the Arctic Slope of Alaska. *Arctic*, 45 (4), 338–342.

Canto-Perello, J., Curiel-Esparza, J., & Calvo, V. (2016). Strategic decision support system for utility tunnel's planning applying A'WOT method. *Tunneling and Underground Space Technology, 55*, 146–152. https://doi.org/10.1016/j.tust.2015.12.009

Carozza, L., Tingdahl, D., & Bosche´, F. (2014). Markerless vision-based augmented reality for urban planning. *Computer-Aided Civil and Infrastructure Engineering, 29*, 2–17. https://doi.org/10.1111/j.1467-8667.2012.00798.x

Case, K. E., & Mayer, C. J. (1996). Housing price dynamics within a metropolitan area. *Regional Science and Urban Economics*, *26*(3–4), 387–407. https://doi.org/10/b2bjjk

Case, K. E., & Shiller, R. J. (2003). Is There a Bubble in the Housing Market? *Brookings Papers on Economic Activity*, *2003*(2), 299–362. https://doi.org/10/b65cb5

Cavalli, R. M., Laneve, G., Fusilli, L., Pignatti, S., & Santini, F. (2009). Remote sensing water observation for supporting Lake Victoria weed management. *Journal of Environmental Management*, *90*(7), 2199–2211. https://doi.org/10/dnzsmw

Chapman, L. J., Chapman, C. A., Ogutu-Ohwayo, R., Chandler, M., Kaufman, L., & Keiter, A. E. (1996). Refugia for Endangered Fishes from an Introduced Predator in Lake Nabugabo, Uganda. *Conservation Biology*, *10*(2), 554–561. https://doi.org/10/b24s5w

Cheyne, S.M., Sastramidjaja, W.J., Muhalir, Rayadin, Y., and Macdonald, D.W., 2016. Mammalian communities as indicators of disturbance across Indonesian Borneo. *Global Ecology and Conservation*, 7, 157–173.

Cirulis, A., & Brigmanis, K. B. (2013). 3D outdoor augmented reality for architecture and urban planning. *Procedia Computer Science, 25*, 71–79

Clapp, J. M., Nanda, A., & Ross, S. L. (2008). Which school attributes matter? The influence of school district performance and demographic composition on property values. *Journal of Urban Economics*, *63*(2), 451–466. https://doi.org/10/d2vgqm

Coffin, A.W., 2007. From roadkill to road ecology: A review of the ecological effects of roads. *Journal of Transport Geography*, 15 (5), 396–406.

Cohen, A. S., Kaufman, L., & Ogutu-Ohwayo, R. (1996). Anthropogenic Threats, Impacts and Conservation Strategies in the African Great Lakes: A Review. In *The Limnology, Climatology and Paleoclimatology of the East African Lakes*. Routledge.

Colding, J., & Barthel, S. (2017). An urban ecology critique on the "smart city" model. *Journal of Cleaner Production, 164*, 95–101. https://doi.org/10.1016/j.jclepro.2017.06.191

Collins, J. B., & Woodcock, C. E. (1996). An assessment of several linear change detection techniques for mapping forest mortality using multitemporal landsat TM data. *Remote Sensing of Environment*, *56*(1), 66–77. https://doi.org/10/d44gr6

Çöltekin, A., Oprean, D., Wallgrün, J. O., & Klippel, A. (2019). Where are we now? Re-visiting the digital earth through human-centered virtual and augmented reality geovisualization environments. *International Journal of Digital Earth*, *12*(2), 119–122. https://doi.org/10.1080/17538947.2018.1560986

Crespo, R., Fotheringham, S., & Charlton, M. (2007). Application of Geographically Weighted Regression to a 19-year set of house price data in London to calibrate local hedonic price models. *Proceedings of the 9th International Conference on GeoComputation*. http://www.geocomputation.org/2007/

Cumulative Cost Mapping | Google Earth Engine. (n.d.). Google Developers. Retrieved July 20, 2021, from https://developers.google.com/earth-engine/guides/image_cumulative_cost

Cunningham, S., Rogan, J., Martin, D., DeLauer, V., McCauley, S., & Shatz, A. (2015). Mapping land development through periods of economic bubble and bust in Massachusetts using Landsat time series data. *GIScience & Remote Sensing*, *52*(4), 397–415. https://doi.org/10/gk33mc

Cushman, S.A., Macdonald, E.A., Landguth, E.L., Malhi, Y., and Macdonald, D.W., 2017. Multiple-scale prediction of forest loss risk across Borneo. *Landscape Ecology*, 32 (8), 1581–1598.

Cutler, F. original by L. B. and A., & Wiener, R. port by A. L. and M. (2018). *randomForest: Breiman and Cutler's Random Forests for Classification and Regression* (4.6-14) [Computer software]. https://CRAN.R-project.org/package=randomForest

Cutts, B. B., Darby, K. J., Boone, C. G., & Brewis, A. (2009). City structure, obesity, and environmental justice: An integrated analysis of physical and social barriers to walkable streets and park access. *Social Science & Medicine*, *69*(9), 1314–1322. https://doi.org/10/bwpmj8

D. (2017). Using virtual reality for assessing the role of noise in the audio-visual design of an urban public space. *Landscape and Urban Planning*, *167*, 98–107. https://doi.org/10.1016/j.landurbplan.2017.05.018

Damtew, Y. T., Verbeiren, B., Assayie, A., & Triest, L. (2021). Satellite Imageries and Field Data of Macrophytes Reveal a Regime Shift of a Tropical Lake (Lake Ziway, Ethiopia). *Water*, *13*. https://doi.org/10/gj6kpc

Demšar, U., Fotheringham, A. S., & Charlton, M. (2008). Exploring the spatio-temporal dynamics of geographical processes with geographically weighted regression and geovisual analytics. *Information Visualization*, *7*(3–4), 181–197. https://doi.org/10/c9nmt2

Derrible, S. (2017). Urban infrastructure is not a tree: Integrating and decentralizing urban infrastructure systems. *Environment and Planning B. Urban Analytics and City Science*, *44*(3), https://doi.org/10.1177/0265813516647063

Dohong, A., Aziz, A.A., and Dargusch, P., 2017. A review of the drivers of tropical peatland degradation in South-East Asia. *Land Use Policy*, 69, 349–360.

Donchyts, G., Schellekens, J., Winsemius, H., Eisemann, E., & Van de Giesen, N. (2016). A 30 m Resolution Surface Water Mask Including Estimation of Positional and Thematic Differences Using Landsat 8, SRTM and OpenStreetMap: A Case Study in the Murray-Darling Basin, Australia. *Remote Sensing*, *8*(5), 386. https://doi.org/10/ggcct4

Doyle, S., Dodge, M., & Smith, A. (1998). The potential of web-based mapping and virtual reality technologies for modelling urban environments. *Computers, Environment and Urban Systems*, *22*(2), 137–155. https://doi.org/10.1016/S0198-9715(98)00014-3.

Drollette, D., 2013. *Gold rush in the jungle: the race to discover and defend the rarest animals of Vietnam's 'lost world'*. First edition. New York: Crown Publishers.

Duckworth, J.W., Batters, G., Belant, J.L., Bennett, E.L., Brunner, J., Burton, J., Challender, D.W.S., Cowling, V., Duplaix, N., Harris, J.D., Hedges, S., Long, B., Mahood, S.P., McGowan, P.J.K., McShea, W.J., Oliver, W.L.R., Perkin, S., Rawson, B.M., Shepherd, C.R., Stuart, S.N., Talukdar, B.K., Dijk, P.P. van, Vié, J.-C., Walston, J.L., Whitten, T., and Wirth, R., 2012. Why South-east Asia should be the world's priority for averting imminent species extinctions, and a call to join a developing cross-institutional programme to tackle this urgent issue. *S.A.P.I.EN.S. Surveys and Perspectives Integrating Environment and Society*, (5.2).

Echevarria Sanchez, G. M., Van Renterghem, T., Sun, K., De Coensel, B., & Botteldooren,

Edler, D., Kühne, O., Keil, J., & Dickmann, F. (2019). Audiovisual cartography: Established and new multimedia approaches to represent soundscapes. *KN - Journal of Cartography and Geographic Information*, *69*(1), 5–17. https://doi.org/10.1007/s42489-019-00004-4

wekaKMeans | Google Earth Engine. (n.d.). Google Developers. Retrieved July 20, 2021, from https://developers.google.com/earth-engine/apidocs/ee-clusterer-wekakmeans

Elsby, M. W., Hobijn, B., & Sahin, A. (2010). *The Labor Market in the Great Recession* (Working Paper No. 15979; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w15979

Estoque, R.C., Ooba, M., Avitabile, V., Hijioka, Y., DasGupta, R., Togawa, T., and Murayama, Y., 2019. The future of Southeast Asia's forests. *Nature Communications*, 10 (1), 1829.

Fairbairn, D., & Parsley, S. (1997). The use of VRML for cartographic presentation. *Computers & Geosciences*, *23*, 475–481

Ferrer, A. L. C., Thomé, A. M. T., & Scavarda, A. J. (2018). Sustainable urban infra-structure: A review. *Resources, Conservation and Recycling*, *128*, 360–372. https://doi.org/10.1016/j.resconrec.2016.07.017

Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, *140*, 23–35. https://doi.org/10/f5p7wd

Forman, R.T.T. and Alexander, L.E., 1998. Roads and Their Major Ecological Effects. *Annual Review of Ecology and Systematics*, 29 (1), 207–231.

Forman, R.T.T., Sperling, D., Bissonette, J.A., Clevenger, A.P., Cutshall, C.D., and Dale, V.H., 2002. *Road Ecology: Science and Solutions*. Island Press.

Fotheringham, A. S., Brunsdon, C. F., & Charlton, M. E. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. *Https://Eprints.Ncl.Ac.Uk*. https://eprints.ncl.ac.uk

Fotheringham, A. S., Crespo, R., & Yao, J. (2015). Geographical and Temporal Weighted Regression (GTWR): Geographical and Temporal Weighted Regression. *Geographical Analysis*, *47*(4), 431–452. https://doi.org/10/f7wwsn

Fugitive methane emissions from leak-prone natural gas distribution infrastructure in urban environments. *Environmental Pollution*, *213*, 710–716. https://doi.org/10.1016/j.envpol.2016.01.094

Fusilli, L., Collins, M. O., Laneve, G., Palombo, A., Pignatti, S., & Santini, F. (2013). Assessment of the abnormal growth of floating macrophytes in Winam Gulf (Kenya) by using MODIS imagery time series. *International Journal of Applied Earth Observation and Geoinformation*, *20*, 33–41. https://doi.org/10/bgq7fs

Gabr, B., Ahmed, M., & Marmoush, Y. (2020). PlanetScope and Landsat 8 Imageries for Bathymetry Mapping. *Journal of Marine Science and Engineering*, *8*(2), 143. https://doi.org/10/gjghrt

Gallagher, K.P., 2019. China's Global Energy Finance [online]. Available from: http://www.bu.edu/gdp/ [Accessed 15 Feb 2021].

Gaudreau, J., Perez, L., & Drapeau, P. (2016). BorealFireSim: A GIS-based cellular automata model of wildfires for the boreal forest of Quebec in a climate change para- digm. *Ecological Informatics, 32*, 12–27. https://doi.org/10.1016/j.ecoinf.2015.12.006

Gibbons, S., & Machin, S. (2008). Valuing school quality, better transport, and lower crime: Evidence from house prices. *Oxford Review of Economic Policy*, *24*(1), 99–119. https://doi.org/10/dth4c2

Gillespie, T.W., Lipkin, B., Sullivan, L., Benowitz, D.R., Pau, S., and Keppel, G., 2012. The rarest and least protected forests in biodiversity hotspots. *Biodiversity and Conservation*, 21 (14), 3597–3611.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software*, *63*(17). https://doi.org/10.18637/jss.v063.i17

Goodman, A. C., & Thibodeau, T. G. (2003). Housing market segmentation and hedonic prediction accuracy. *Journal of Housing Economics*, *12*(3), 181–201. https://doi.org/10/bv499t

Goodman, C. J., & Mance, S. M. (2011). Employment loss and the 2007–09 recession: An overview. *Monthly Labor Review*, 3–12.

Grantham, H.S., Duncan, A., Evans, T.D., Jones, K.R., Beyer, H.L., Schuster, R., Walston, J., Ray, J.C., Robinson, J.G., Callow, M., Clements, T., Costa, H.M., DeGemmis, A., Elsen, P.R., Ervin, J., Franco, P., Goldman, E., Goetz, S., Hansen, A., Hofsvang, E., Jantz, P., Jupiter, S., Kang, A., Langhammer, P., Laurance, W.F., Lieberman, S., Linkie, M., Malhi, Y., Maxwell, S., Mendez, M., Mittermeier, R., Murray, N.J., Possingham, H., Radachowsky, J., Saatchi, S., Samper, C., Silverman, J., Shapiro, A., Strassburg, B., Stevens, T., Stokes, E., Taylor, R., Tear, T., Tizard, R., Venter, O., Visconti, P., Wang, S., and Watson, J.E.M., 2020. Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity. *Nature Communications*, 11 (1), 5978.

G-REALM - Kyoga. (n.d.). Retrieved July 20, 2021, from https://ipad.fas.usda.gov/cropexplorer/global_reservoir/gr_regional_chart.aspx?regionid=eafrica&reservoir_name=Kyoga

Grimm, N. B., Morgan Grove, J., Pickett, S. T. A., & Redman, C. L. (2000). Integrated Approaches to Long-TermStudies of Urban Ecological Systems. *BioScience*, *50*(7), 571. https://doi.org/10/fk5tjx

Gunnarsson, C. C., & Petersen, C. M. (2007). Water hyacinths as a resource in agriculture and energy production: A literature review. *Waste Management, 27*(1), 117–129. https://doi.org/10/fd62m8

Haala, N., & Kada, M. (2010). An update on automatic 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*, 570–580. https://doi.org/10.1016/j.procs.2013.11.009

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., and Townshend, J.R.G., 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342 (6160), 850–853.

Harris, D. R. (1999). "Property Values Drop When Blacks Move in, Because...": Racial and Socioeconomic Determinants of Neighborhood Desirability. *American Sociological Review, 64*(3), 461. https://doi.org/10.2307/2657496

Hasse, J., & Lathrop, R. G. (2003). A Housing-Unit-Level Approach to Characterizing Residential Sprawl. *Photogrammetric Engineering & Remote Sensing, 69*(9), 1021–1030. https://doi.org/10/gg3dpb

Hecht, J.S., 2017. Making multi-stakeholder water resources decisions with limited streamflow information. PhD Thesis. Tufts University.

Helbich, M., Brunauer, W., Vaz, E., & Nijkamp, P. (2014). Spatial Heterogeneity in Hedonic House Price Models: The Case of Austria. *Urban Studies, 51*(2), 390–411. https://doi.org/10/f5k9p3

Hendrick, M. F., Ackley, R., Sanaie-Movahed, B., Tang, X., & Phillips, N. G. (2016).

Higgins, C. D., & Kanaroglou, P. S. (2016). Forty years of modelling rapid transit's land value uplift in North America: Moving beyond the tip of the iceberg. *Transport Reviews, 36*(5), 610–634. https://doi.org/10/ggw575

Hoffmann, M., Hilton-Taylor, C., Angulo, A., Böhm, M., Brooks, T.M., Butchart, S.H.M., Carpenter, K.E., Chanson, J., Collen, B., Cox, N.A., Darwall, W.R.T., Dulvy, N.K., Harrison, L.R., Katariya, V., Pollock, C.M., Quader, S., Richman, N.I., Rodrigues, A.S.L., Tognelli, M.F., Vié, J.-C., Aguiar, J.M., Allen, D.J., Allen, G.R., Amori, G., Ananjeva, N.B., Andreone, F., Andrew, P., Ortiz, A.L.A., Baillie, J.E.M., Baldi, R., Bell, B.D., Biju, S.D., Bird, J.P., Black-Decima, P., Blanc, J.J., Bolaños, F., Bolivar-G, W., Burfield, I.J., Burton, J.A., Capper, D.R., Castro, F., Catullo, G., Cavanagh, R.D., Channing, A., Chao, N.L., Chenery, A.M., Chiozza, F., Clausnitzer, V., Collar, N.J., Collett, L.C., Collette, B.B., Fernandez, C.F.C., Craig, M.T., Crosby, M.J., Cumberlidge, N., Cuttelod, A., Derocher, A.E., Diesmos, A.C., Donaldson, J.S., Duckworth, J.W., Dutson, G., Dutta, S.K., Emslie, R.H., Farjon, A., Fowler, S., Freyhof, J., Garshelis, D.L.,

Gerlach, J., Gower, D.J., Grant, T.D., Hammerson, G.A., Harris, R.B., Heaney, L.R., Hedges, S.B., Hero, J.-M., Hughes, B., Hussain, S.A., M, J.I., Inger, R.F., Ishii, N., Iskandar, D.T., Jenkins, R.K.B., Kaneko, Y., Kottelat, M., Kovacs, K.M., Kuzmin, S.L., Marca, E.L., Lamoreux, J.F., Lau, M.W.N., Lavilla, E.O., Leus, K., Lewison, R.L., Lichtenstein, G., Livingstone, S.R., Lukoschek, V., Mallon, D.P., McGowan, P.J.K., McIvor, A., Moehlman, P.D., Molur, S., Alonso, A.M., Musick, J.A., Nowell, K., Nussbaum, R.A., Olech, W., Orlov, N.L., Papenfuss, T.J., Parra-Olea, G., Perrin, W.F., Polidoro, B.A., Pourkazemi, M., Racey, P.A., Ragle, J.S., Ram, M., Rathbun, G., Reynolds, R.P., Rhodin, A.G.J., Richards, S.J., Rodríguez, L.O., Ron, S.R., Rondinini, C., Rylands, A.B., Mitcheson, Y.S. de, Sanciangco, J.C., Sanders, K.L., Santos-Barrera, G., Schipper, J., Self-Sullivan, C., Shi, Y., Shoemaker, A., Short, F.T., Sillero-Zubiri, C., Silvano, D.L., Smith, K.G., Smith, A.T., Snoeks, J., Stattersfield, A.J., Symes, A.J., Taber, A.B., Talukdar, B.K., Temple, H.J., Timmins, R., Tobias, J.A., Tsytsulina, K., Tweddle, D., Ubeda, C., Valenti, S.V., Dijk, P.P. van, Veiga, L.M., Veloso, A., Wege, D.C., Wilkinson, M., Williamson, E.A., Xie, F., Young, B.E., Akçakaya, H.R., Bennun, L., Blackburn, T.M., Boitani, L., Dublin, H.T., Fonseca, G.A.B. da, Gascon, C., Lacher, T.E., Mace, G.M., Mainka, S.A., McNeely, J.A., Mittermeier, R.A., Reid, G.M., Rodriguez, J.P., Rosenberg, A.A., Samways, M.J., Smart, J., Stein, B.A., and Stuart, S.N., 2010. The Impact of Conservation on the Status of the World's Vertebrates. *Science*, 330 (6010), 1503–1509.

Holly, S., Pesaran, M. H., & Yamagata, T. (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics*, *158*(1), 160–173. https://doi.org/10.1016/j.jeconom.2010.03.040

Holt, T.V., Binford, M.W., Portier, K.M., and Vergara, R., 2016. A stand of trees does not a forest make: Tree plantations and forest transitions. *Land Use Policy*, 56, 147–157.

Howard, T., & Gaborit, N. (2007). Using virtual environment technology to improve public participation in the urban planning process. *Journal of Urban Planning and Development*, *133*(4), https://doi.org/10.1061/(ASCE)0733-9488(2007)133:4(233)

Hruby, F. (2019). The sound of being there: Audiovisual cartography with immersive virtual environments. *KN - Journal of Cartography and Geographic Information*, *69*(1), 19–28. https://doi.org/10.1007/s42489-019-00003-5

Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International Journal of Geographical Information Science*, *24*(3), 383–401. https://doi.org/10/dcb8vq

Hughes, A.C., 2018. Have Indo-Malaysian forests reached the end of the road? *Biological Conservation*, 223, 129–137.

Hughes, A.C., 2019. Understanding and minimizing environmental impacts of the Belt and Road Initiative. *Conservation Biology*, 33 (4), 883–894.

Hui, F., Xu, B., Huang, H., Yu, Q., & Gong, P. (2008). Modelling spatial-temporal change of Poyang Lake using multitemporal Landsat imagery. *International Journal of Remote Sensing*, *29*(20), 5767–5784. https://doi.org/10/bnc85q

Hunt, D. V. L., Nash, D., & Rogers, C. D. F. (2014). Sustainable utility placement via multi- utility tunnels. *Tunneling and Underground Space Technology*, *39*, 15–26. https://doi.org/10.1016/j.tust.2012.02.001

Hurvich, C. M., Simonoff, J. S., & Tsai, C.-L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *60*(2), 271–293. https://doi.org/10/cxwsnw

Husson, E., Hagner, O., & Ecke, F. (2014). Unmanned aircraft systems help to map aquatic vegetation. *Applied Vegetation Science*, *17*(3), 567–577. https://doi.org/10/f23rxj

Imottesjo, H., & Kain, J. (2018). The urban CoBuilder – A mobile augmented reality tool for crowd-sourced simulation of emergent urban development patterns: Requirements, prototyping and assessment. *Computers, Environment and Urban Systems, 71*, 120–130. https://doi.org/10.1016/j.compenvurbsys.2018.05.003

Ingalls, M., Diepart, J.-C., Truong, N., Hayward, D., Neil, T., Phomphakdy, C., Bernhard, R., Fogarizzu, S., Epprecht, M., Nanhthavong, V., Vo, D.H., Nguyen, D., Nguyen, P.A., Saphangthong, T., Inthavong, C., Hett, C., and Tagliarino, N., 2018. *State of Land in the Mekong Region*. Bern Open Publishing.

Ishii, H., Ben-Joseph, E., Underkoffler, J., Yeung, L., Chak, D., Kanji, Z., & Piper, B. (2002). Augmented urban planning workbench: Overlaying drawings, physical models, and digital simulation. *Proceedings of the 1st international symposium on mixed and augmented reality* (pp. 203).

Japan International Cooperation Agency, n.d. ODA Loan Project DATA | Our Work | JICA | Japan International Cooperation Agency [online]. *ODA Loan Project DATA*. Available from: https://www2.jica.go.jp/en/yen_loan/index.php [Accessed 15 Feb 2021].

Jeon, S. B., Olofsson, P., & Woodcock, C. E. (2014). Land use change in New England: A reversal of the forest transition. *Journal of Land Use Science*, *9*(1), 105–130. https://doi.org/10/ggfq2r

Jiang, H., Zhao, D., Cai, Y., & An, S. (2012). A Method for Application of Classification Tree Models to Map Aquatic Vegetation Using Remotely Sensed Images from Different Sensors and Dates. *Sensors*, *12*(9), 12437–12454. https://doi.org/10/f94hd2

Kamel Boulos, M. N., Lu, Z., Guerrero, P., Jennett, C., & Steed, A. (2017). From urban planning and emergency training to Pokémon Go: Applications of virtual reality GIS

(VRGIS) and augmented reality GIS (ARGIS) in personal, public and environmental health. *International Journal of Health Geographics, 16*(7), https://doi.org/10.1186/s12942-017-0081-0

Kaufman, L. S., Chapman, L. J., & Chapman, C. A. (1997). Evolution in fast forward: Haplochromine fishes of the Lake Victoria region. *Endeavour*, *21*(1), 23–30. https://doi.org/10/d5qspd

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sensing of Environment*, *114*(12), 2897–2910. https://doi.org/10/c2v3fc

Kersten, T., Deggim, S., Tschirschwitz, F., Lindstaedt, M. U., & Hinrichsen, N. (2018). Segeberg 1600 – Eine stadtrekonstruktion in virtual reality. *Kartographische Nachrichten, 68/4*, 183–191

KfW, n.d. KfW promotes development programmes in Africa, Asia, Latin America and South East Europe on behalf of the Federal Government. [online]. *Global Engagement.* Available from: https://www.kfw-entwicklungsbank.de/International-financing/KfW-Development-Bank/Projekte/ [Accessed 15 Feb 2021].

Koebel, C. T., McCoy, A. P., Sanderford, A. R., Franck, C. T., & Keefe, M. J. (2015). Diffusion of green building technologies in new housing construction. *Energy and Buildings, 97*, 175–185. https://doi.org/10/f7h3qc

Lamoreux, J., Resit Akçakaya, H., Bennun, L., Collar, N.J., Boitani, L., Brackett, D., Bräutigam, A., Brooks, T.M., da Fonseca, G.A.B., Mittermeier, R.A., Rylands, A.B., Gärdenfors, U., Hilton-Taylor, C., Mace, G., Stein, B.A., and Stuart, S., 2003. Value of the IUCN Red List. *Trends in Ecology & Evolution*, 18 (5), 214–215.

Landsat Collection 1 Level-1 Quality Assessment Band. (n.d.). Retrieved July 19, 2021, from https://www.usgs.gov/core-science-systems/nli/landsat/landsat-collection-1-level-1-quality-assessment-band?qt-science_support_page_related_con=0#qt-science_support_page_related_con

Laurance, W.F., Croes, B.M., Guissouegou, N., Buij, R., Dethier, M., and Alonso, A., 2008. Impacts of Roads, Hunting, and Habitat Alteration on Nocturnal Mammals in African Rainforests. *Conservation Biology*, 22 (3), 721–732.

Laurance, W.F., Goosem, M., and Laurance, S.G.W., 2009. Impacts of roads and linear clearings on tropical forests. *Trends in Ecology & Evolution*, 24 (12), 659–669.

Lautzenheiser, T., Collins, J., Ricci, E., & Clarke, J. (2014). Losing ground: Planning for resilience. *Massachusetts Audubon Society, Lincoln, Massachusetts, USA.*

Lechner, A.M., Chan, F.K.S., and Campos-Arceiz, A., 2018. Biodiversity conservation should be a core value of China's Belt and Road Initiative. *Nature Ecology & Evolution*, 2 (3), 408–409.

LeSage, J. P. (2008). An Introduction to Spatial Econometrics. *Revue d'économie Industrielle*, *123*, 19–44. https://doi.org/10/gf2ztv

Leung, Y., Mei, C.-L., & Zhang, W.-X. (2000). Statistical Tests for Spatial Nonstationarity Based on the Geographically Weighted Regression Model. *Environment and Planning A: Economy and Space*, *32*(1), 9–32. https://doi.org/10/d7m6zc

Li, X., Lv, Z., Hu, J., Zhang, B., Shi, L., & Feng, S. (2015). XEarth: A 3D GIS platform for managing massive city information. *Computational intelligence and virtual environments for measurement systems and applications, 2015*IEEE International. https://doi.org/10. 1109/CIVEMSA.2015.7158625

Li, Y., Gong, X., Guo, Z., Xu, K., Hu, D., & Zhou, H. (2016). An index and approach for water extraction using Landsat–OLI data. *International Journal of Remote Sensing*, *37*(16), 3611–3635. https://doi.org/10/gd54tc

Liu, X., Blackburn, T.M., Song, T., Li, X., Huang, C., and Li, Y., 2019. Risks of Biological Invasion on the Belt and Road. *Current Biology*, 29 (3), 499-505.e4.

Lovett, A., Appleton, K., Warren-Kretzschmar, B., & Von Haaren, C. (2015). Using 3D visualization methods in landscape planning: An evaluation of options and practical issues. *Landscape and Urban Planning*, *142*, 85–94. https://doi.org/10.1016/j.landurbplan.2015.02.021

Lu, B., Charlton, M., Harris, P., & Fotheringham, A. S. (2014). Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *International Journal of Geographical Information Science*, *28*(4), 660–681. https://doi.org/10/gkphct

Lu, B., Harris, P., Charlton, M., & Brunsdon, C. (2014). The GWmodel R package: Further topics for exploring spatial heterogeneity using geographically weighted models. *Geo-Spatial Information Science*, *17*(2), 85–101. https://doi.org/10/gdsksj

Mace, G.M., Collar, N.J., Gaston, K.J., Hilton-Taylor, C., Akçakaya, H.R., Leader-Williams, N., Milner-Gulland, E.J., and Stuart, S.N., 2008. Quantification of Extinction Risk: IUCN's System for Classifying Threatened Species. *Conservation Biology*, 22 (6), 1424–1442.

Maliene, V., & Malys, N. (2009). High-quality housing—A key issue in delivering sustainable communities. *Building and Environment*, *44*(2), 426–430. https://doi.org/10/bbjv4d

Marshall, B. E. (2018). Guilty as charged: Nile perch was the cause of the haplochromine decline in Lake Victoria. *Canadian Journal of Fisheries and Aquatic Sciences*, *75*(9), 1542–1559. https://doi.org/10/gk8h6x

MassGIS (Bureau of Geographic Information) | Mass.gov. (n.d.). Retrieved July 3, 2021, from https://www.mass.gov/orgs/massgis-bureau-of-geographic-information

Mbabazi, D., Ogutu-Ohwayo, R., Wandera, S. B., & Kiziito, Y. (2004). Fish species and trophic diversity of haplochromine cichlids in the Kyoga satellite lakes (Uganda). *African Journal of Ecology*, *42*(1), 59–68. https://doi.org/10/bw789c

McFEETERS, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10/fkkz88

McLellan, B.N. and Shackleton, D.M., 1989. Grizzly Bears and Resource-Extraction Industries: Habitat Displacement in Response to Seismic Exploration, Timber Harvesting and Road Maintenance. *Journal of Applied Ecology*, 26 (2), 371–380.

Mei, C.-L., Wang, N., & Zhang, W.-X. (2006). Testing the Importance of the Explanatory Variables in a Mixed Geographically Weighted Regression Model. *Environment and Planning A: Economy and Space*, *38*(3), 587–598. https://doi.org/10.1068/a3768

Middleton, C., Elmhirst, R., and Sūphang Chanthawānit, 2019. *Living with floods in a mobile Southeast Asia a political ecology of vulnerability, migration and environmental change*.

Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *The American Economic Review*, *57*(2), 197–210.

Mittermeier, R.A., Myers, N., Mittermeier, C.G., and Robles Gil, P., 1999. Hotspots: Earth's biologically richest and most endangered terrestrial ecoregions. *Hotspots: Earth's biologically richest and most endangered terrestrial ecoregions.*

Moilanen, A., Pouzols, F., Meller, L., Veach, V., Arponen, A., Leppänen, J., and Kujala, H., 2014. Zonation—Spatial conservation planning methods and software. *Version 4. User Manual*, 290.

Montesino Pouzols, F., Toivonen, T., Di Minin, E., Kukkala, A.S., Kullberg, P., Kuusterä, J., Lehtomäki, J., Tenkanen, H., Verburg, P.H., and Moilanen, A., 2014. Global protected area expansion is compromised by projected land-use and parochialism. *Nature*, 516 (7531), 383–386.

Morgado, N.C. and Taşkın, Ö., 2019. Managing environmental risks in development banks and development finance institutions–what role for donor shareholders?

Morris, R.J., 2010. Anthropogenic impacts on tropical forest biodiversity: a network structure and ecosystem functioning perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365 (1558), 3709–3718.

Mullaney, J., Lucke, T., & Trueman, S. J. (2015). A review of benefits and challenges in growing street trees in paved urban environments. *Landscape and Urban Planning, 134*, 157–166. https://doi.org/10.1016/j.landurbplan.2014.10.013

Mulley, C. (2014). Accessibility and Residential Land Value Uplift: Identifying Spatial Variations in the Accessibility Impacts of a Bus Transitway. *Urban Studies*, *51*(8), 1707–1724. https://doi.org/10/f54p32

Muth, R. F. (1969). *CITIES AND HOUSING; THE SPATIAL PATTERN OF URBAN RESIDENTIAL LAND USE.* https://trid.trb.org/view/545388

Myers, D., & Ryu, S. (2008). Aging Baby Boomers and the Generational Housing Bubble: Foresight and Mitigation of an Epic Transition. *Journal of the American Planning Association*, *74*(1), 17–33. https://doi.org/10/fcwmrb

National Grid (2010). *D.P.U. 10-55 attachment AG-46-36*.

Nellemann, C., Vistnes, I., Jordhøy, P., Strand, O., and Newton, A., 2003. Progressive impact of piecemeal infrastructure development on wild reindeer. *Biological Conservation*, 113 (2), 307–317.

Nelson, P. P. (2016). A framework for the future of urban underground engineering. *Tunneling and Underground Space Technology*, *55*, 32–39. https://doi.org/10.1016/j.tust.2015.10.023

Neukom, B. (2018). *CityEngine VR experience for unreal studio.* GeoNet The Esri Community. Web page. Last edited: June 19, 2019. Date accessed: September 24, 2019. URL https://community.esri.com/docs/DOC-11563-cityengine-vr-experience- for-unreal-studio

Newman, P., & Kenworthy, J. (1998). *SUSTAINABILITY AND CITIES: OVERCOMING AUTOMOBILE DEPENDENCE*. https://trid.trb.org/view/503432

NICFI DATA Program. (2021). Planet.com. https://assets.planet.com/docs/NICFI_UserGuidesFAQ.pdf

Njiru, J., Knaap, M. van der, Kundu, R., & Nyamweya, C. (2018). Lake Victoria fisheries: Outlook and management. *Lakes & Reservoirs: Science, Policy and Management for Sustainable Use*, *23*(2), 152–162. https://doi.org/10/gk8h6q

O'Shea, T. (2021, March 2). *Universal Access To Satellite Monitoring Paves The Way To Protect The World's Tropical Forests*. https://www.planet.com/pulse/universal-access-to-satellite-monitoring-paves-the-way-to-protect-the-worlds-tropical-forests/

Ogutu-Ohwayo, R. (1990). The decline of the native fishes of lakes Victoria and Kyoga (East Africa) and the impact of introduced species, especially the Nile perch, Lates niloticus, and the Nile tilapia, Oreochromis niloticus. *Environmental Biology of Fishes*, *27*(2), 81–96. https://doi.org/10/dx3ksn

Olson, D.M. and Dinerstein, E., 2002. The Global 200: Priority Ecoregions for Global Conservation. *Annals of the Missouri Botanical Garden*, 89 (2), 199.

Ortega, Y.K. and Capen, D.E., 1999. Effects of Forest Roads on Habitat Quality for Ovenbirds in a Forested Landscape. *The Auk*, 116 (4), 937–946.

Ostrom, E. (2009). A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science*, *325*(5939), 419–422. https://doi.org/10/br5gnw

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10/bfq5pv

Outa, N. O., Yongo, E. O., Keyombe, J. L. A., Ogello, E. O., & Wanjala, D. N. (2020). A review on the status of some major fish species in Lake Victoria and possible conservation strategies. *Lakes & Reservoirs: Science, Policy and Management for Sustainable Use*, *25*(1), 105–111. https://doi.org/10/gk8h6b

Pace, R. K., LeSage, J., & Zhu, S. (2009). Impact of Cliff and Ord on the Housing and Real Estate Literature: Impact of Cliff and Ord on the Housing and Real Estate Literature. *Geographical Analysis*, *41*(4), 418–424. https://doi.org/10/b7937j

Pandit, A., Lu, Z., & Crittenden, J. C. (2015). Managing the complexity of urban systems. *Journal of Industrial Ecology*, *19*(2), 201–204. https://doi.org/10.1111/jiec.12263

Paulsen, K. (2013). The Effects of Growth Management on the Spatial Extent of Urban Development, Revisited. *Land Economics*, *89*(2), 193–210. https://doi.org/10/f5gvdt

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10/f9gkxn

Pickens, A. H., Hansen, M. C., Hancher, M., Stehman, S. V., Tyukavina, A., Potapov, P., Marroquin, B., & Sherani, Z. (2020). Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment*, *243*, 111792. https://doi.org/10/ggszjz

Pickett, S. T. A., Burch, W. R., Dalton, S. E., Foresman, T. W., Grove, J. M., & Rowntree, R. (1997). A conceptual framework for the study of human ecosystems in urban areas. *Urban Ecosystems*, *1*(4), 185–199. https://doi.org/10/c35sf5

Platts, S.G., 2018. *World Electric Power Plants Database, 2018*. McGraw-Hill New York.

Pouliot, J., Larrivée, S., Ellul, C., & Boudhaim, A. (2018). Exploring schema matching to compare geospatial standards: Application to underground utility networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-4/W10*, 157–164. https://doi.org/10.5194/isprs-archives-XLII-4-W10-157-2018

Projections and Implications for Housing a Growing Population: Older Households 2015-2035 | Joint Center for Housing Studies. (2016, December 13). https://www.jchs.harvard.edu/research-areas/reports/projections-and-implications-housing-growing-population-older-households

Qiang, Y. (2019). Disparities of population exposed to flood hazards in the United States. *Journal of Environmental Management*, *232*, 295–304. https://doi.org/10.1016/j.jenvman.2018.11.039

Quintero, J.P., 2007. Mainstreaming conservation in infrastructure projects : case studies from Latin America.

Rauterkus, S., Thrall, G., & Hangen, E. (2010). Location Efficiency and Mortgage Default. *Journal of Sustainable Real Estate*, *2*(1). https://doi.org/10/gk33mk

Ray, R., N. Kring, W., Pitts, J., and Simmons, B.A., 2021. Geolocated Dataset of Chinese Overseas Development Finance. *Nature Scientific Data*.

Redford, K. and Fearn, E., 2007. *PROTECTED AREAS AND HUMAN LIVELIHOODS*.

Ree, R. van der, Smith, D.J., and Grilo, C., 2015. *Handbook of Road Ecology*. John Wiley & Sons.

Reed, J., Vianen, J.V., Deakin, E.L., Barlow, J., and Sunderland, T., 2016. Integrated landscape approaches to managing social and environmental issues in the tropics: learning from the past to guide the future. *Global Change Biology*, 22 (7), 2540–2554.

Repetto, M. P., Burlando, M., Solari, G., De Gaetano, P., Pizzo, M., & Tizzi, M. (2017). A web-based GIS platform for the safe management and risk assessment of complex structural and infrastructural systems exposed to wind. *Advances in Engineering Software, 117*, 29–45. https://doi.org/10.1016/j.advengsoft.2017.03.002

Rinaldi, S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems, 21*(6), https://doi.org/10.1109/37.969131

Rivera, S., Landom, K., & Crowl, T. (2013). Rivera, S., K. Landom and T. Crowl, 2013. Monitoring macrophytes cover and taxa in Utah Lake by using 2009-2011 Landsat digital imagery. Revista de Teledeteccion (Spain) 39, 106-115. ISSN: 1988-8740. *Revista de Teledeteccion*, *39*, 106–115.

Rodrigue, J.-P., Comtois, C., & Slack, B. (2016). *The Geography of Transport Systems* (4th ed.). Routledge. https://doi.org/10.4324/9781315618159

Rodrigues, A.S.L., Pilgrim, J.D., Lamoreux, J.F., Hoffmann, M., and Brooks, T.M., 2006. The value of the IUCN Red List for conservation. *Trends in Ecology & Evolution*, 21 (2), 71–76.

Rogers, A. S., & Kearney, M. S. (2004). Reducing signature variability in unmixing coastal marsh Thematic Mapper scenes using spectral indices. *International Journal of Remote Sensing*, *25*(12), 2317–2335. https://doi.org/10/fwn39w

Rogers, W., & Winkler, A. (2013). The relationship between the housing and labor market crises and doubling up: An MSA-level analysis, 2005–2011. *Monthly Labor Review*. https://doi.org/10/gg36kj

Sabo, J.L., Ruhi, A., Holtgrieve, G.W., Elliott, V., Arias, M.E., Ngor, P.B., Räsänen, T.A., and Nam, S., 2017. Designing river flows to improve food security futures in the Lower Mekong Basin. *Science*, 358 (6368).

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *Quarterly Journal of Economics*, *125*(3), 1253–1296. https://doi.org/10/dckwvv

Salas, J., & Yepes, V. (2018). Urban vulnerability assessment: Advances from the strategic planning outlook. *Journal of Cleaner Production*, *179*, 544–558. https://doi.org/10.1016/j.jclepro.2018.01.088

Sameeh El halabi, A., El Sayad, Z. T., & Ayad, H. M. (2019). VRGIS as assistance tool for urban decision making: Rafah – Gaza – Palestine. *Alexandria Engineering Journal, 58*, 367–375.

Santos, A., Zarraonandia, T., Díaz, P., & Aedo, I. (2018). A virtual reality map interface for geographical information systems. *Proceedings of the 2018 international conference on advanced visual interfaces. No. 83.* https://doi.org/10.1145/3206505.3206580

Saphores, J.-D., & Li, W. (2012). Estimating the value of urban green areas: A hedonic pricing analysis of the single family housing market in Los Angeles, CA. *Landscape and Urban Planning*, *104*(3–4), 373–387. https://doi.org/10/ck7xhw

Schipper, J., Chanson, J.S., Chiozza, F., Cox, N.A., Hoffmann, M., Katariya, V., Lamoreux, J., Rodrigues, A.S.L., Stuart, S.N., Temple, H.J., Baillie, J., Boitani, L., Lacher, T.E., Mittermeier, R.A., Smith, A.T., Absolon, D., Aguiar, J.M., Amori, G., Bakkour, N., Baldi, R., Berridge, R.J., Bielby, J., Black, P.A., Blanc, J.J., Brooks, T.M., Burton, J.A., Butynski, T.M., Catullo, G., Chapman, R., Cokeliss, Z., Collen, B., Conroy, J., Cooke, J.G., Fonseca, G.A.B. da, Derocher, A.E., Dublin, H.T., Duckworth, J.W., Emmons, L., Emslie, R.H., Festa-Bianchet, M., Foster, M., Foster, S., Garshelis, D.L., Gates, C., Gimenez-Dixon, M., Gonzalez, S., Gonzalez-Maya, J.F., Good, T.C., Hammerson, G., Hammond, P.S., Happold, D., Happold, M., Hare, J., Harris, R.B., Hawkins, C.E., Haywood, M., Heaney, L.R., Hedges, S., Helgen, K.M., Hilton-Taylor, C., Hussain, S.A., Ishii, N., Jefferson, T.A., Jenkins, R.K.B., Johnston, C.H., Keith, M., Kingdon, J., Knox, D.H., Kovacs, K.M., Langhammer, P., Leus, K., Lewison, R., Lichtenstein, G., Lowry, L.F., Macavoy, Z., Mace, G.M., Mallon, D.P., Masi, M., McKnight, M.W., Medellín, R.A., Medici, P., Mills, G., Moehlman, P.D., Molur, S., Mora, A., Nowell, K., Oates, J.F., Olech, W., Oliver, W.R.L., Oprea, M., Patterson, B.D., Perrin, W.F., Polidoro, B.A., Pollock, C., Powel, A., Protas, Y., Racey, P., Ragle, J., Ramani, P., Rathbun, G., Reeves, R.R., Reilly, S.B., Reynolds, J.E., Rondinini, C., Rosell-Ambal, R.G., Rulli, M., Rylands, A.B., Savini, S., Schank, C.J., Sechrest, W., Self-Sullivan, C., Shoemaker, A., Sillero-Zubiri, C., Silva, N.D., Smith, D.E., Srinivasulu, C., Stephenson, P.J., Strien, N. van, Talukdar, B.K., Taylor, B.L., Timmins, R., Tirira, D.G., Tognelli, M.F., Tsytsulina, K., Veiga, L.M., Vié, J.-C., Williamson, E.A., Wyatt, S.A., Xie, Y., and Young, B.E., 2008. The Status of the World's Land and Marine Mammals: Diversity, Threat, and Knowledge. *Science*, 322 (5899), 225–230.

Schleuning, M., Farwig, N., Peters, M.K., Bergsdorf, T., Bleher, B., Brandl, R., Dalitz, H., Fischer, G., Freund, W., Gikungu, M.W., Hagen, M., Garcia, F.H., Kagezi, G.H., Kaib, M., Kraemer, M., Lung, T., Naumann, C.M., Schaab, G., Templin, M., Uster, D., Wägele, J.W., and Böhning-Gaese, K., 2011. Forest Fragmentation and Selective Logging Have Inconsistent Effects on Multiple Animal-Mediated Ecosystem Processes in a Tropical Forest. *PLOS ONE*, 6 (11), e27785.

Schweizer, D., Cisneros, R., Traina, S., Ghezzehei, T. A., & Shaw, G. (2017). Using National Ambient Air Quality Standards for fine particulate matter to assess regional wildland fire smoke and air quality management. *Journal of Environmental Management, 201*, 345–356. https://doi.org/10.1016/j.jenvman.2017.07.004Get

Shen, L., & Li, C. (2010). Water body extraction from Landsat ETM+ imagery using adaboost algorithm. *2010 18th International Conference on Geoinformatics*, 1–4. https://doi.org/10/fvfgqr

Shi, X. and Yao, L., 2019. Prospect of China's Energy Investment in Southeast Asia under the Belt and Road Initiative: A Sense of Ownership Perspective. *Energy Strategy Reviews*, 25, 56–64.

Shiller, R. J. (2006). Long-Term Perspectives on the Current Boom in Home Prices. *The Economists' Voice*, *3*(4). https://doi.org/10/ft4qtq

Shiller, R. J. (2007). *Understanding Recent Trends in House Prices and Home Ownership* (Working Paper No. 13553; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w13553

Silva, T. S. F., Costa, M. P. F., Melack, J. M., & Novo, E. M. L. M. (2008). Remote sensing of aquatic vegetation: Theory and applications. *Environmental Monitoring and Assessment*, *140*(1), 131–145. https://doi.org/10/bjsc9m

Sodhi, N.S., Koh, L.P., Brook, B.W., and Ng, P.K.L., 2004. Southeast Asian biodiversity: an impending disaster. *Trends in Ecology & Evolution*, 19 (12), 654–660.

Sodhi, N.S., Koh, L.P., Clements, R., Wanger, T.C., Hill, J.K., Hamer, K.C., Clough, Y., Tscharntke, T., Posa, M.R.C., and Lee, T.M., 2010. Conserving Southeast Asian forest biodiversity in human-modified landscapes. *Biological Conservation*, 143 (10), 2375–2384.

Stevens, L., Anderson, B., Cowan, C., Colton, K., and Johnson, D., 2017. The footprint of energy: Land use of US Electricity production. *STRATA: Logan, UT, USA*.

Sun, Y., & Cui, Y. (2018). Evaluating the coordinated development of economic, social and environmental benefits of urban public transportation infrastructure: Case study of four Chinese autonomous municipalities. *Transport Policy, 66*, 116–126. https://doi.org/10.1016/j.tranpol.2018.02.006

Tang, F., & Zhang, X. (2008). A GIS-based 3D simulation for occupant evacuation in a building. *Tsinghua Science and Technology, 13*, 58–64. https://doi.org/10.1016/S1007-0214(08)70127-5

Tao, W. (2013). Interdisciplinary urban GIS for smart cities: Advancements and opportunities. *Geo-spatial Information Science*, *16*(1), 25–34. https://doi.org/10.1080/10095020.2013.774108

Thamaga, K. H., & Dube, T. (2018). Remote sensing of invasive water hyacinth (Eichhornia crassipes): A review on applications and challenges. *Remote Sensing Applications: Society and Environment, 10*, 36–46. https://doi.org/10/gjhx74

Thomaz, S., Esteves, F., Murphy, K., Santos, A., Caliman, A., & Guariento, R. (2008). *AQUATIC MACROPHYTES IN THE TROPICS: ECOLOGY OF POPULATIONS AND COMMUNITIES, IMPACTS OF INVASIONS AND USE BY MAN* (pp. 1252–1280).

Tsatsaronis, K., & Zhu, H. (2004). *What Drives Housing Price Dynamics: Cross-Country Evidence* (SSRN Scholarly Paper ID 1968425). Social Science Research Network. https://papers.ssrn.com/abstract=1968425

Tu, C. C., & Eppli, M. J. (2001). An Empirical Examination of Traditional Neighborhood Development. *Real Estate Economics, 29*(3), 485–501. https://doi.org/10/bfkvdd

Tungaraza, C., Eliapenda, E., Osewe, K. O., & Palapala, P. M. (2012). *Long-term climate impact on the Lake Victoria region influences water level fluctuation and resource availability*.

Turner, I.M., 1996. Species Loss in Fragments of Tropical Rain Forest: A Review of the Evidence. *Journal of Applied Ecology*, 33 (2), 200–209.

Twongo, T., Bugenyi, F. W. B., & Wanda, F. (1995). The potential for further proliferation of water hyacinth in Lakes Victoria, Kyoga and Kwania and some urgent aspects for research. *African Journal of Tropical Hydrobiology and Fisheries*, *6*(1 & 2), 1–10.

U.S. Geological Survey (2015a). *20150529, USGS lidar point cloud MA Sndy-CMPG 201319TCG300875 LAS.* U.S. Geological Survey

U.S. Geological Survey (2015b). *20150529, USGS lidar point cloud MA Sndy-CMPG 201319TCG300890 LAS.* U.S. Geological Survey

Uhm, D.P. van, 2016. *The Illegal Wildlife Trade: Inside the World of Poachers, Smugglers and Traders*. Springer International Publishing.

Upadhyaya, J. K., Biswas, N., & Tam, E. (2014). A review of infrastructure challenges: Assessing stormwater system sustainability. *Canadian Journal of Civil Engineering, 41*, 483–492. https://doi.org/10.1139/cjce-2013-0430

Vardon, M., Burnett, P., and Dovers, S., 2016. The accounting push and the policy pull: balancing environment and economic decisions. *Ecological Economics*, 124, 145–152.

Vetter, D., Hansbauer, M.M., Végvári, Z., and Storch, I., 2011. Predictors of forest fragmentation sensitivity in Neotropical vertebrates: a quantitative review. *Ecography*, 34 (1), 1–8.

Villa, P., Bresciani, M., Bolpagni, R., Pinardi, M., & Giardino, C. (2015). A rule-based approach for mapping macrophyte communities using multi-temporal aquatic vegetation indices. *Remote Sensing of Environment*, *171*, 218–233. https://doi.org/10/gk8gq9

Vogt, P. and Riitters, K., 2017. GuidosToolbox: universal digital image object analysis. *European Journal of Remote Sensing*, 50 (1), 352–361.

Wachsmuth, D., Cohen, D. A., & Angelo, H. (2016). Expand the frontiers of urban sustainability. *Nature*, *536*(7617), 391–393. https://doi.org/10/gfkqwz

Wang, S., Fang, C., Wang, Y., Huang, Y., & Ma, H. (2015). Quantifying the relationship between urban development intensity and carbon dioxide emissions using a panel data analysis. *Ecological Indicators*, *49*, 121–131. https://doi.org/10/f6tt4z

Wang, X. (2005). Integrating GIS, simulation models, and visualization in traffic impact analysis. *Computers, Environment and Urban Systems*, *29*(4), 471–496. https://doi.org/10.1016/j.compenvurbsys.2004.01.002

Wei, C.-H., & Qi, F. (2012). On the estimation and testing of mixed geographically weighted regression models. *Economic Modelling*, *29*(6), 2615–2620. https://doi.org/10/f4gk9n

Wheeler, D. C. (2007). Diagnostic Tools and a Remedial Method for Collinearity in Geographically Weighted Regression. *Environment and Planning A: Economy and Space*, *39*(10), 2464–2481. https://doi.org/10/b4b5t9

Wheeler, D. C. (2010). Visualizing and Diagnosing Coefficients from Geographically Weighted Regression Models. In B. Jiang & X. Yao (Eds.), *Geospatial Analysis and Modelling of Urban Structure and Dynamics* (pp. 415–436). Springer Netherlands. https://doi.org/10.1007/978-90-481-8572-6_21

Wild, T.B., Loucks, D.P., Annandale, G.W., and Kaini, P., 2016. Maintaining Sediment Flows through Hydropower Dams in the Mekong River Basin. *Journal of Water Resources Planning and Management*, 142 (1), 05015004.

Wilkie, D., Shaw, E., Rotberg, F., Morelli, G., and Auzel, P., 2000. Roads, Development, and Conservation in the Congo Basin. *Conservation Biology*, 14 (6), 1614–1622.

Wolch, J. R., Byrne, J., & Newell, J. P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough.' *Landscape and Urban Planning*, *125*, 234–244. https://doi.org/10/f28drr

Wolfartsberger, J. (2019). Analyzing the potential of virtual reality for engineering design review. *Automation in Construction, 104*, 27–37. https://doi.org/10.1016/j.autcon.2019.03.018

World Bank, n.d. Projects [online]. *Projects*. Available from: https://projects.worldbank.org/en/projects-operations/projects-list [Accessed 15 Feb 2021].

Wu, B., Li, R., & Huang, B. (2014). A geographically and temporally weighted autoregressive model with application to housing prices. *International Journal of Geographical Information Science*, *28*(5), 1186–1204. https://doi.org/10/ghzxnv

Wu, H., He, Z., & Gong, J. (2010). A virtual globe-based 3D visualization and interactive framework for public participation in urban planning processes. *Computers, Environment and Urban Systems*, *34*, 291–298. https://doi.org/10.1016/j.compenvurbsys.2009.12.001

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, *27*(14), 3025–3033. https://doi.org/10/bc3sqh

Yang, W. (2014). *An extension of geographically weighted regression with flexible bandwidths* [Thesis, University of St Andrews]. https://research-repository.st-andrews.ac.uk/handle/10023/7052

Yongo, E., Agembe, S., Outa, N., & Owili, M. (2018). Growth, mortality and recruitment of Nile perch (Lates niloticus) in Lake Victoria, Kenya. *Lakes & Reservoirs: Science, Policy and Management for Sustainable Use*, *23*(1), 17–23. https://doi.org/10/gk8h6v

Yu, D., Wei, Y. D., & Wu, C. (2007). Modeling Spatial Dimensions of Housing Prices in Milwaukee, WI. *Environment and Planning B: Planning and Design*, *34*(6), 1085–1102. https://doi.org/10/b7f67f

Yunana, D. A., Shittu, A. A., Ayuba, S., Bassah, E. J., & Joshua, W. K. (2017). Climate change and lake water resources in Sub-Saharan Africa: Case study of lake Chad and lake Victoria. *Nigerian Journal of Technology*, *36*(2), 648–654. https://doi.org/10/gk8h6k

Zadek, S., 2019. Financing a Just Transition. *Organization & Environment*, 32 (1), 18–25.

Zhang, Y., Liu, X., Qin, B., Shi, K., Deng, J., & Zhou, Y. (2016). Aquatic vegetation in response to increased eutrophication and degraded light climate in Eastern Lake Taihu: Implications for lake ecological restoration. *Scientific Reports*, *6*(1), 23867. https://doi.org/10/f8gqrg

Zhao, D., Jiang, H., Yang, T., Cai, Y., Xu, D., & An, S. (2012). Remote sensing of aquatic vegetation distribution in Taihu Lake using an improved classification tree with modified thresholds. *Journal of Environmental Management*, *95*(1), 98–107. https://doi.org/10/bmp78v