Fall 12-2014

Utilizing Landsat TM and OLI in Predicting *Oncomelania Hupensis* Habitats Around Poyang Lake Before and After Three Gorges Dam Completion

Stephanie L. McCracken
University of Southern Mississippi

The University of Southern Mississippi

Utilizing Landsat TM and OLI in Predicting *Oncomelania Hupensis* Habitats Around Poyang Lake Before and After Three Gorges Dam Completion

by

Stephanie McCracken

A Thesis
Submitted to the Honors College of
The University of Southern Mississippi
in Partial Fulfillment
of the Requirements for the Degree of
Bachelor of Science
in the Department of Geography and Geology

December 2014

Approved by

David H. Holt, Ph.D., Thesis Adviser
Associate Professor of Geography

Andy Reese, Ph.D., Chair
Department of Geography and Geology

Ellen Weinauer, Ph.D., Dean
Honors College

Abstract

Schistosomiasis is a parasitic disease caused by the *Schistomsoma japonicum* flatworm that utilizes the *Oncomelania hupensis* snail as an intermediate agent. In the People's Republic of China, these amphibious snails contaminate freshwater systems infecting humans, bovines, and other mammals and have caused significant morbidity for over two thousand years (Wertheim, et al. 2012; Zhang, et al. 2012.) The gravity of this disease prompted the national government to initiate sizable public health programs, such as the World Bank Loan Project (WBLP.) In spite of WBLP's achievements, in 2004, after this program ended, a national survey acknowledged a resurgence of *schistosomiasis* in various regions including the Poyang Lake area in the Jiangxi Province (Zhang, et al. 2012.) Poyang Lake, since the completion of the Three Gorges Dam in 2009, has experienced significant changes in the lake's depth and water extent exposing increased land surface, which possesses the potential to alter snail habitats (McManus, et al. 2010.) In analyzing these changes, this project sought to apply remote sensing techniques using multispectral imagery from Landsat 7 and 8 in conjunction with spatial analysis tools offered by Geographic Information Systems (GIS.) Because snail habitats rely heavily on ecological factors, including location of water bodies, submersion periods, and vegetation coverage, this analysis observed these attributes using Modified Difference Water Index (MDWI) and Normalized Difference Vegetation Index (NDVI) calculations. These computations derived from images taken during 2000-2001 and 2013-2014 to observe Poyang Lake before and after Three Gorges Dam completion (Hui, et al. 2008.) Though the examination observed a drastic increase in potential *O. hupensis* habitats, further analysis that incorporates data unavailable during this project would establish a more complete suitability model (Chen & Lin 2004.)

Key Words: *Schistosomiasis*, *Oncomelania Hupensis*, Geographic Information Systems, Remote Sensing, Poyang Lake, Three Gorges Dam

Dedication

I dedicate my work to my loving husband, Sean, for his continuous support throughout my undergraduate education and keeping me uplifted during hard times. This thesis is also devoted to my friends, family, The Senior Honors College, the University of Southern Mississippi's Department of Geography and Geology, and peers within the department for continually inspiring me to learn and remain motivated in pursuing my goals.

I also wanted to give dedication in memory of my father, Kenneth John Reichardt Sr., for always encouraging me to dream big.

Acknowledgments

I would like to express my gratitude to all of the individuals who have offered me their guidance and support throughout the process of writing this thesis. My primary advisor, Dr. David Holt, taught me how to construct my thesis and applying my knowledge to this project thoughtfully. Throughout this research, Dr. Holt has also remained supportive, patient, and encouraging. Other individuals deserving special thanks include Dr. Gregory Carter, Bill Funderburk, Carlton Anderson, and Guy W. Jeter for offering me helpful advice and solutions. Finally, to my friends and family, thank you for your ceaseless support throughout the duration of this research.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Infectious diseases have delineated much of human history in that illnesses, such as smallpox and the bubonic plague decimated vast populations and was an important driver for social and economic transformations for centuries (Beltz 2011.) The fight against these diseases led to some of humankind's greatest accomplishments including vaccinations, innovative medical procedures, and a well-developed comprehension of causative agents spreading illnesses. Understanding that pathogenic organisms are responsible for spreading infectious diseases and possess a unique interaction with their environment was a paramount discovery in determining disease origin and high-risk areas (Nelson & Williams 2001.) One early epidemiologic case occurred in 1854 when Dr. John Snow established that London's cholera outbreak originated from a contaminated well (Figure 1.) Dr. Snow made his conclusions after mapping cholera deaths throughout the city and finding that more deaths occurred around this water source (Price 2012.) In addition, this study affiliated the location of water pumps in comparison with residencies, factories, and workplaces. This analysis was one of the first well-known studies to utilize mapping capabilities in determining an illness's spatial distribution and, over the years, these methods advanced with various innovations benefiting today's epidemiologists (Nelson & Williams 2001.)

One growing technical tool used to help predict infectious disease regions, behavior, and origin is geographical information systems (Price 2012.) Geographic Information Systems (GIS) refers to a collection of computer applications that permits analysts to tie data to specific locations. Though Dr. Snow's study represents an early example of how GIS could be utilized, enhancements with GIS software in recent years allow for far more sophisticated evaluations than basic mapping capabilities (Price 2012.) Another geospatial tool utilized in epidemiological inquiries is remote sensing or observing phenomena without being in direct, physical contact using particular instruments. In recent years, most remote sensing data collection has been accomplished through aerial or satellite imagery. Considering that many infectious microorganisms and their hosts thrive only under certain environmental conditions, incorporating remote sensing data into infectious disease analysis is favorable because of remote sensing's ability to observe various biophysical properties over a wide area. This information in conjunction with GIS's computational tools help create prediction and risk analysis models without extensive, laborious fieldwork, yet many question its legitimacy within epidemiology because of technological constraints, such as limited resolution from remote sensing imagery (Nelson & Williams 2001.) To address these applications' validity within disease models, this study intends to observe various geospatial model types in the context of *schistosomiasis* within China. In addition, this research will attempt to model the effects of the Three Gorges Dam on *schistosomiasis* distribution specifically within the Poyang Lake area. In essence, in spite of various uncertainties concerning geospatial tool's usefulness in analyzing disease distribution, remote sensing's unique capability in observing ecological features over wide geographical areas paired with statistical methodology could better predict *schistosomiasis* prevalence in China than current conventional field methods, especially in conjunction with a rapidly changing environment (Xianyi 2005.)

Schistosomiasis is a parasitic disease caused by the *Schistomsoma japonicum* flatworm that utilizes the *Oncomelania hupensis* snail as an intermediate agent. These amphibious snails contaminate freshwater systems, which potentially infect humans, bovines, and many other mammals within the local area. Those infected experience symptoms, such as liver damage, anemia, hypertension, and, potentially, a reduction in cognitive growth in children (Wertheim, et al. 2012.) This damaging disease also possesses a longstanding history in China as documentation from well over two thousand years ago vividly describe it devastation (Zhang, et al. 2012.) Additionally, during nationwide surveys conducted shortly after the inauguration of the People's Republic of China in 1949, epidemiological examinations recorded over 380 counties within 12 provinces experiencing a *schistosomiasis* endemic. Over twelve million citizens were infected while an extra one-hundred million lived within high-risk areas (Ross, et al. 2001.) The gravity of these findings prompted the national government to initiate sizable public health programs beginning in the mid-1950's to reduce *S. japonicum* prevalence throughout the landscape. From the 1950's to 1986, various initiatives successfully decreased endemic regions by eliminating disease transmission in 4 out of the 12 provinces. Further, more substantial contributions to this effort began in 1992 with China's collaboration with the World Bank to reduce *schistosomiasis* through the World Bank Loan Project (WBLP.) This project's mission was to decrease *schistosomiasis* morbidity rates by diminishing disease exposure in both humans and farm animals, such as cows, while focusing on decreasing infected snail populations (Xianyi 2005.) The WBLP addressed these obstacles primarily by prescribing praziquantel to infected human and livestock populations while spraying molluscicides in endemic areas to reduce *O.*

hupensis numbers. These efforts drastically diminished s*chistosomiasis* burden amongst the population largely due to massive declines in snail population as a few provinces experienced up to 90% reductions (Xianyi 2005.)

In spite of WBLP's achievements, in 2004, shortly after this program officially ceased, a national survey acknowledged a resurgence of *schistosomiasis* (Zhang, et al. 2012.) Districts previously pronounced with high *schistosomiasis* containment by China's Minister of Health possessed growing snail populations and infection rates whereas provinces south of the Yangzi River including, Hubei, Anhui, Hunan, Anhui, Jiangsu, continued to challenge public health workers as these districts remained endemic of this disease (Qian 2013) (Figure 2.) This comeback was directly related to the swelling of the *O. hupensis* snail populations throughout China perpetuated by a numerous set of endeavors. First, WBLP's discontinuance lead to substantial financial cutbacks resulting in drastic reductions in chemotherapy or molluscicide treatment making it difficult to maintain the same level of snail population management previously reached with this initiative. This process also requires a considerable work force, which is unfeasible, especially in hard-to-reach, countryside locations. In addition, compliance to *schistosomiasis* drug treatments in infected populations dwindled. Both factors played a central role in *schistosomiasis c*ontrol in the past, yet, with this evident lack of continuity, the WBLP's efforts appeared effective for short-term regulation while inadequate for long-term disease management (Zhang, et al. 2008.)

Additional factors affecting snail populations include a combination of variances in human activity and environmental components with ecological circumstances better delineating snail habitats. China's accelerated urbanization induced by its rapid economic growth within the past decade triggered the largest rural-to-urban migration in history consisting of about 120 million citizens. This human movement was perceived as a potential hindrance to *schistosomiasis* containment due to the knowledge that substantial migrations, in the past, played a considerable role in inducing infectious disease epidemics. As such, infected individuals from rural areas likely reintroduced this disease in urban regions previously believed to have reached disease containment. Unfortunately, due to the lack of spatial and temporal data concerning these population movements, their connection with this disease's re-emergence is ambiguous. Factors that are more advantageous and readily understood to predict snail population habitats are those dealing with environmental conditions, such as temperature, humidity, elevation, frequent floods along the Yangzi River along with many other elements (Zhou, Liang, and Jiang 2012.)

Because biophysical characteristics, at the moment, provide more coherent data than human migration patterns, a feasible research approach is to analyze how particular anthropogenic activities directly affect the environment. One of the most controversial projects due to its size and drastic effects on the Yangtze River patterns is the Three Gorges Dam (TGD) project. The premise for the TGD construction involved flood-prevention and various economic concerns. The Yangtze River basin has experienced devastating flood disasters since China's earliest settlements and with China's vast, most recent economic growth, possessed a lack of supply in electrical power to meet these demands. With these factors in mind, the TGD was meant to attenuate both issues. Currently, the TGD is the largest dam in the world at 185 meters tall and its reservoir, which is referred as the Three Gorges Reservoir Region (TGRR), expands across the region between Yichang, the location of the TGD, and Chongqing occupying an estimate of 600 kilometers of land (Figure 3.) The extent of this reservoir required the resettlement of approximately 1.2 million citizens because, after the TGRR was filled to impound the flow of the Yangtze River, 140 towns, 13 metropolitan areas, and 326 rural villages were submerged. Along with the obvious epidemiological concerns caused by a vast movement of populations, another concern that arose after the TGD was completed in 2009. Unquestionably, the dam succeeded in preventing flooding in the downstream regions of the Yangtze River, yet, in doing so, it drastically altered sand and water areas downstream inspiring great ecological change in various regions, especially those identified as endemic of *schistosomiasis* (McManus, et al. 2010.)

One region applicable to these issues is the Poyang Lake region in the Jiangxi Province. This lake is the largest freshwater lake in China and affects the lives of approximately 4.4 million residents (Guo, et al. 2005 & Hui, et al. 2008.) Once again, in spite of efforts at eliminating *schistosomiasis*, marshland areas surrounding lake areas, namely the Poyang and Dongting lakes, disease control efforts have proven significantly difficult (Ross et al. 2013.) One of the most significant contributors to the *O. hupensis* snail distribution that diffuses the disease in this area involves this lake's seasonal water fluctuations. Prior to the TGD, the Poyang Lake flooded during this region's rainy season beginning in April and ending in October. During November, around 90% of the Poyang Lake's water would recede due to the dry winter conditions, which, consequently, exposed vast expanses of grassy marshlands where snail populations thrive (Guo, et al. 2005.) However, since the TGD official began operation, this lake's natural inundation pattern has ceased altering its ecological system. The dam caused higher winter flows while notably reducing summer water levels, which increased exposed land area. In addition, researchers predict that the TGD also altered the two disease transmission periods into one long season as the preexisting floods, oftentimes, drowned the adult snail population (McManus, et al. 2010.) With these considerations in mind, a thorough environmental investigation would provide ample information in determining the most current *O. hupensis* snail distribution (Yang, et al. 2005.)

It is evident that forecasting high-risk regions based on environmental components facilitating snail populations is critical, yet common survey methods do not incorporate these factors and are, oftentimes, inaccurate in calculating current snail populations within a given area. These field analyses are unreliable because of inadequate survey methods and many examiners grow impatient, as these processes tend to be labor intensive and time-consuming. A study performed by Zhang, et al., discovered that these practices only account for approximately 20% of the total snail population within a given area. Another drawback is that these conventional methods lack the ability to produce a useable prediction model composed of applicable ecological factors that bolster snail densities (Zhang, et. al 2008.) GIS in conjunction with remote sensing and statistical analysis avert many of these inadequacies. These tools generate models covering wide spatial areas within a relatively short timeframe while aggregating various environmental attributes to account for ecological characteristics synonymous with suitable snail habitats (Qian 2013.)

Overall, this research project sought to explore the capabilities offered by remote sensing and GIS in modeling *schistosomiasis* in correlation with many recent changes occurring around Poyang Lake including those induced by the Three Gorges Dam. Throughout the data collection process, it was found in consideration with supply deficiency and time-constraints that the two pieces of ecological available for observation included Modified Difference Water Index (MNDWI) to delineate yearly water patterns and vegetation information derived from Normalized Difference Vegetation Index (NDVI) calculations using remotely sensed images of Poyang Lake. These computations derived from images taken during 2000-2001 and 2013-2014 to observe Poyang Lake before and after Three Gorges Dam completion (Hui, et al. 2008.) In analyzing these changes, this project sought to apply remote sensing techniques using multispectral imagery from Landsat 7 and 8 in conjunction with spatial analysis tools offered by Geographic Information Systems (GIS.) In developing these models, this study observed the potential in utilizing geospatial tools in *schistosomiasis* prediction while also interpreting limitations in using this technology in vector-borne disease prediction models. Though the examination observed a drastic increase between the two study periods in potential to *O. hupensis* habitats, further analysis that incorporates data unavailable during this project would establish a more complete suitability model (Chen & Lin 2004.)

Chapter 2: Literature Review

Remote sensing generally refers to the collection of data without the researcher being in direct contact with the object of interest. These systems possess the ability to observe the physical characteristics of the earth along with various natural or cultural practices. Investigating phenomena remotely, especially with aerial photography, possesses a considerable history dating even as far back as the Civil War as the Union Army utilized hot-air balloons to observe Confederate activity during the battle of Fair

Banks (Jensen 2007.) Historical events particularly crucial in advancing disease surveillance through aerial and satellite imagery began, most notably, in 1970. During that year, Barnett Cline wrote his "New Eyes for Epidemiologists: Aerial Photography and Other Remote Sensing Techniques" published in the *American Journal of Epidemiology* and his article promoted the potential of using remote sensing in conjunction with epidemiological studies (Cline 2006.) Less than a year later, the National Aeronautics and Space Administration (NASA) initiated the Health Application Office (HAO) that was responsible for organizing programs that applied remote sensing capabilities to disease research. Then, in 1972, the first Earth Resources Satellite (ERS), Landsat I, was launched without incident and represented a large step in utilizing remote sensing technology for the sake of human health (Zhang, et al. 2013.)

Cline’s prediction concerning the use of remote sensing techniques for epidemiological studies originated from the concept that infectious diseases represent associations between epidemic occurrence, human characteristics, and how each of these interplay with the natural environment. He pressed that remote sensing provides scientists a means of analyzing many of these variables over a large geographic space (Cline 2006.) The ways in which satellite imagery is useful in collecting ecological data such as elevation, vegetation, and soil composition, is by recording the electromagnetic energy diffused or reflected from an object of interest. Every substance or living thing with a temperature above absolute zero emits energy within the electromagnetic spectrum, which is the extent of all the possible frequencies an object can radiate. Reflectance percentages collected by remote sensors within specific portions of the spectrum, such as infrared and ultraviolent, serves as a surrogate for the physical properties of the variable under investigation (Jensen 2007.) For example, the infrared segment of the electromagnetic spectrum is commonly used in classifying various vegetation characteristics, including yield, stress, and many others, because the chlorophyll within plants absorb much of the infrared radiation (Jensen 2005.) Because data from remote sensing techniques offer updated, detailed environmental information, predictions concerning vector-borne illness behavior are achieved when used in conjunction with either disease prevalence data or knowledge pertaining to preferred parasite habitats (Nelson and Williams 2001.)

Though the phenomena, processes, or objects of interest all, in some way, affect sensor readings for any remote sensor system, it should be remembered that the internal workings and calibrations of any of these systems also delineate the final electromagnetic measurements. Each remote sensing mechanism maintains a specific set of resolutions that determine the data collected including spectral, spatial, temporal, and radiometric resolutions. Spectral resolution refers to the wavelength range and dimensions within the electromagnetic spectrum that could be readily examined by a remote sensing instrument. Then, spatial resolution is the smallest angular unit detected by the remote sensing system (Jensen 2007.) For instance, the Moderate Resolution Imaging Spectrometer's (MODIS) highest resolution is 250 meters whereas the Landsat 7 Enhanced Thematic Mapper can portray objects within a 30-meter grid (Jensen 2005.) Out of these two sensors, the Landsat 7 possesses a higher spatial resolution because it can resolve smaller angles leading to a more detailed image. Temporal resolution describes the time and frequency in which data collection occurs. Finally, radiometric resolution, also known as quantization, describes the extent in which the instrument detects differences in spectral strength as it records the radiation emitted, reflected, or absorbed by the terrain. Prior to beginning any investigation with any remote sensing system, a sensor's resolution in each of these categories must be considered to achieve a thorough analysis (Jensen 2007.)

The application of remote sensing and geographic information system technology for *schistosomiasis* tracking in China began roughly in the early 1980's, yet initial progress was relatively slow (Guo-Jing, et al. 2005.) In 1980, Qian et al. established the first *schistosomiasis* atlas containing historical disease data for every county by utilizing GIS technology, yet it was not until 1985 that a team of researchers began to implement remote sensing applications in tracking disease processes. Unfortunately, though the first examination of *Oncomelancia hupensis* identified snail habitats, it failed to inspire other scientists to use remote sensing imagery because their experiment was non-replicable as they failed to record their methodology for image processing and analysis. It was not until 1990, when Li et al. conducted a well-documented analysis on suitable *O. hupensis* habitats that attracted many other Chinese researchers to explore remote sensing potential in disease surveillance (Zhang, et al. 2013.) Since then, continuous advancements in Chinese *schistosomiasis* exploration with geospatial tools occurred in conjunction with developments in remote sensing technology, geographic information systems (GIS), and methodology (Zhou, et al. 2001.) The combined capabilities of both remote sensing and GIS make it possible for researchers to study interrelationships between disease processes, environmental factors, anthropological activity, and many other variable types in greater detail (Hu, et al. 2013.)

Collecting ecological data aggregated through remote sensing imagery analysis is commonplace in determining vector-borne illness activity (Zhi-Ying, et al. 2005.) Some universal environmental factors investigated through remote sensing include air and soil temperature, types of vegetation, land use, land cover, and elevation, especially with regard to *schistosomiasis*. This species is closely related to specific environmental conditions and are particularly sensitive to even minor changes in their ecosystem (Yang, et al. 2008.) For instance, both air and soil temperature are crucial considerations because the *O. hupensis* generally live in regions with average temperatures of 16-20 °C and around soils near swamps or rivers (Jia-Gang et al. 2005; Kameda & Kato 2011.) The higher the land surface temperature within the snail's preferred threshold, the faster the flatworm species completes significant lifecycle processes, such as mating and development, and higher disease incidence (Yang, et al. 2005.) Temperature data collection is generated using a number of methods but some of the most common methods include the use of meteorological satellites, previous weather reports, or measuring surface temperature by interpreting a specific portion of the electromagnetic spectrum, such as infrared (Jensen 2005; Zhi-Ying et al. 2005.) Similar to temperature, all other ecological factors could be obtained through various means. Considering there are no "best-practice" methods in remote sensing, determining data collection techniques relies on numerous considerations, such as the aim of the investigation, satellite capabilities mentioned previously, budget concerns and many other factors (Jensen 2007.)

As temperature is a fundamental concern in estimating and locating snail populations, it is crucial to distinguish the difference between weather and climate. Though these terms are similar, weather refers to the status of the atmosphere at a particular time and place whereas climate pertains to long-term, average values of weather phenomena, including, but not limited to, rainfall and temperature. In *schistosomiasis* research, both of these approaches have been explored, yet, as with any other geospatial models, various limitations lead to research gaps and uncertainty, particularly with this disease's relationship with global climate change. Global climate change is a recent scientific revelation that the Earth is currently going through a warming phase. Evidence suggests that these rising temperatures are in conjunction with increasing carbon dioxide levels, yet, with climate being a chaotic system, many other factors contribute at varying levels making future predictions for climate change effects exceedingly difficult. Furthermore, various regions may experience cooling effects as a result of the development of additional water vapor or cloud cover in response to rising temperatures in other regions making the changes uneven across the planet (Arbogast 2007.) Analysis on the effects of global climate change on disease incidence is essential in future surveillance as the World Health Organization (WHO) forecasts that a 2 °C increase would lead to a 50-100% increase in *O. hupensis* geographic distributions (Qian et al. 2013.) However, only a small quantity of investigations attempted to predict how global climate change would effect *schistosomiasis* transmission due to both complexity and sensor limitations (McCreesh & Booth 2013.)

Another environmental factor in relation to *O. hupensis* spatial distributions worth special mention is elevation. Oftentimes, study area selection is relatively straightforward because marshlands are regions throughout China generally known as endemic areas for *schistosomiasis* as these snails are amphibious (Zhi-Ying, et al. 2005.) Unfortunately, in 2004, *schistosomiasis* outbreaks resurfaced throughout mountainous areas within the Sichuan Province. Densities of live, infected snails increased by approximately 524%, in spite of a 17% infection rate reduction between the years 2000 and 2004 (Yang, et al. 2008.) Few examinations employed geospatial tools in mountainous regions within the Yunnan and Sichuan provinces making Yi Dong's and his group's work an innovative addition to *schistosomiasis* research. Their work demonstrated that elevation played a significant role in snail densities throughout hilly, mountainous landscapes (Zhou, et al. 2009.) In addition, a few studies found that *O. hupensis* populations in these regions are often clustered but these, somewhat, random distributions are still not understood. Most predictions utilizing GIS modeling in conjunction with remote sensing within mountainous areas use autocorrelation methods, which calculate an unknown area's snail density using a combination of statistics and neighboring values (Yang, et al. 2008; Zhou, et al. 2009.) Overall, in light of discovering increased *schistosomiasis* in hilly areas, research shows that the *O. hupensis* dwells in one of three specific ecological zones: waterway network in plain regions, mountainous regions, and lakeside areas (Zhang, et al. 2008.)

Finally, similar to the uncertainties introduced by global climate change, only a handful of research focuses on human lifestyle and activity and their correlation with *schistosomiasis* infection risk. As with global climate change, anthropogenic processes represent a crucial variable in *schistosomiasis* risk making it advantageous to incorporate these systems with previous environmental models (Hu, et al. 2010.) Some of the most significant human activity in recent years involved China's vast economic development and the completion of the Three Gorges Dam. Both activities generated the migration of millions of citizens while substantially altering the natural environment, especially near the Yangtze River (Qian, et al. 2013.) Unfortunately, the total implications for these activities in correlation with snail population are not well comprehended. Obstructions in this exploration include a lack of data pertaining to definite spatial patterns of recent human migration in China and most current modeling methods are not sophisticated enough to explore these simulations (Zhou, et al. 2010; Hu, et al, 2010.) However, some of the areas explored by geospatial techniques include the effects of farming activity and how the Three Gorges Dam altered river velocities. Geospatial studies done by Xiao-Hua Wu and colleagues found that domesticated bovines are significant players in disease transmission, which clearly suggests that future research and control efforts should focus on domestic farm animals, namely bovines. Then, other studies found that the Three Gorges Dam significantly slowed the Yangtze River water velocity producing livable areas around the areas that were previously uninhabitable to the *O. hupensis*. Overall, though both examples are considerable findings, more exploration is necessary in *schistosomiasis* model production with geospatial tools (Zhou, et al. 2009.)

Considering that there is an apparent gap in research done in anthropological activities like the Three Gorges Dam completion in conjunction with *schistosomiasis* *(*Zhou, et al. 2009; McCreesh & Booth 2013*)*, this study sought to explore whether or not remote sensing and GIS could predict disease distribution with these variables in mind. With this notion in mind, the environmental attributes collected for this study using remotely sensed images included water detection using MNDWI calculations and interpreting vegetation abundance through NDVI. Both methods have been utilized previous by various epidemiological studies interpreting *schistosomiasis* specifically.

The Normalized Difference Water Index (NDWI) was devised in order to highlight water bodies in remotely sensed images while, at the same time, removing reflectance values deriving from vegetation or soil features. Because of spectral noise caused by the original NDWI, the MNDWI was established to reveal features more efficiently by removing shadows on water and delineating more subtle characteristics throughout the body of water (Hui, et al. 2005.) The calculation on remotely sensed imagery is a ratio of the green band reflectance values subtracted from the near infrared divided by the total from both bands (Tseng et al. 2014.) This computation utilizes these two bands by enhancing the reflectance of the water body through the green wavelengths, using vegetation's inherent high reflectance within the NIR band, and minimizing any low-reflectance readings within the NIR band to augment water detection (Xu, 2006.) This index has been applied to other projects determining snail habitats, including Tseng et al. through the use of the Moderate Resolution Imaging Spectroradiometer (MODIS) and Hui, et al. using Landsat 7's Thematic Mapper (2014 & 2005.)

Then, the Normalized Difference Vegetation Index developed by Rouse et al. in 1974 to monitor vegetation coverage along with interpreting seasonal changes in growth. This ratio between the red and near infrared bands is effective at detecting vegetation indices due to the inherent spectral reflectance of the photosynthetic processes of plants as very active vegetation absorbs the majority of the red band while reflecting back most of the light around the NIR part of the electromagnetic spectrum (Jensen 2005.) This index is one of the most common remotely sensed calculations historically utilized for *schistosomiasis* prediction through data collected through remotely sensed imagery considering that snails mostly thrive in highly vegetative areas (Yang, et al. 2005.)

Vegetation and water-detection methods are collected during most schistosomiasis studies because both are firmly related to the *O. hupensis* life cycle as it correlates with the *S. japonicum* flatworm. Unlike other flatworm species causing schistosomiasis, such as the *S. mansoni*, which are primarily aquatic, the *S. japonicum* are amphibious making their means to contaminate areas more probable (Utzinger, et al. 2011.) Additionally, the survival of these snails equated with the annual flooding season, especially with snails living around Poyang Lake. Though snails live within close proximity of water, most are not found in areas where flooding submerges a region for more than eight months as adult snail populations drown. At the same time, it is rare to find the *O. hupensis* where the submergence duration is less than three months (Chen & Lin 2004.) This peculiarity about these snails account for the two *schistosomiasis* transmission periods experienced around Poyang Lake that were indicative to these lands before the Three Gorges Dam. The first period of transmission began around April and lasted until June as it coincided with China's flood season expanding the lake's volume. During this time, as temperatures became warmer and water started to raise, the female snails lay copious amounts of eggs, so, by June, the snail population increased considerably and each with the capability to contaminate. Then, a second transmission period occurred around September as lake waters receded and the young snails become fully grown (McManus, et al. 2010.) Finally, snails are also found in areas with high vegetation because plants may provide a micro environment and food essential for their survival. Most surveys found that areas with the highest snail populations occur in areas were vegetation coverage was over 60% whereas snails were unlikely to be found in areas with less than 20% coverage (Chen & Lin 2004.)

Overall, methodologies similar to those used by researchers such as Tseng et al. or Hui et al. were be incorporated to, essentially, develop a *schistosomiasis* prediction model. Both research teams utilized a mixture of climatic and ecological variables detected by remote sensing techniques to predict and delineate *O. hupensis* populations across the landscape (Tseng, et al. 2014 & Hui, et al. 2005.) In creating and assessing this model, this project will, not only explore remote sensing and GIS's capability in forecasting *schistosomiasis* in China, but also how this technology could potentially be applied to other tropical infectious diseases (Nelson and Williams 2001.)

Chapter 3: Methodology

3.1 Study Area

Poyang Lake is based in the northern segment of the Jiangxi province and is circumscribed by coordinates 28°N and 29°N latitude and 115 °E and 117 °E longitude. Its annual mean temperature is 17.6 °C with a humid subtropical climate (Jia-Gang, et al. 2005) (Figure 4.)The study area was analyzed with a combination of remote sensing imagery analysis and GIS spatial analyst tools. Considering that this analysis sought to explore the effects of the Three Gorges Dam on the *O. hupensis* densities around this lake, satellite imagery predating its completion of the dam was collected along with the most recent images available to enable comparison. The imagery preceding the Three Gorges Dam were taken between April 2000 and May 2001 while the most recent study period incorporated images taken between May 2013 and May 2014 (Jis, et al. 2010.)

3.2 Sensor Information

The sensor utilized to obtain the 2000-2001 data derived from the Landsat 7 imagery archive whereas the most current imagery originated from the Landsat 8 multispectral satellite. Landsat 7, which was launched April 1999, collects data through its Enhanced Thematic Mapper Plus with a spatial resolution of 30 x 30 meters for the seven spectral bands. This sensor also possessed a thermal and panchromatic band, which were not utilized for this study. With a 16-day temporal resolution, Landsat 7 provided ample ecological data from its launch date to 2003 after its Scan Line Corrector failed. This mechanical failure since then has caused significant loss of data per image that ultimately limited the study periods that could be utilized for this study (Jensen 2007.)

The second sensor applied to this study, Landsat 8, acquires spectral data using nine spectral bands, one panchromatic band through its Operational Land Imager (OLI), and a Thermal Infrared Sensor. In addition, the data collected by this sensor is stored using a 12-bit quantization meaning it possesses an enhanced capability in detecting variations in spectral signatures than its predecessor, Landsat 7. In spite of the obvious abundance of potential spectral information, this project solely utilized the green, red, and near infrared (NIR) bands to extract ecological data. Each of these bands maintains a 30x30 meter spatial resolution. Final format is as a GeoTiff file. These files were collected through the satellite image database on *http://earthexplorerusgs.gov* ("Landsat 8" August 2013.)

3.3 Image Rectification

Adequately covering seasonal changes throughout the lake area required the collection of monthly Landsat imagery covering the entire study area. Due to extensive cloud-cover in many of the GeoTiff files available for download, only fifteen images were applied to this study with seven images representing the 2000-2001 timeframe and eight images covering the 2013-2014 study period. The selected imagery, though primarily cloudless, possessed less than ten-percent cloud cover. Considering that the images are stored as a GeoTiff file, georectification to a specific coordinate system was not necessary because the World Geodetic System 1984 (WGS 84) is the assigned datum for every image ("Landsat 8" August 2013.)

Instead, each image was geometrically calibrated in respect to one another through image-to-image registration. This process rotates and converts images within the same geographic area and similar geometry to establish proper alignment of corresponding elements between images (Jensen 2005.) This process was necessary in analyzing modifications in lake coverage throughout the year because, without image registration, two or more images could not be examined properly without introducing preventable, irrelevant calculations. The software used for this image-to-image registration was the Environment for Visualizing Images (ENVI) through the "Image Registration" tool. After selecting a base image and a second image to geometrically adjust, ground control points were selected, which enabled the software to register the image to the base image. The maximum acceptable root mean squared error (RMSE) of an image registration is .5 pixels, yet, for this examination, the RMSE values were at or below 0.006 pixel units (Hui, et al. 2008) (Tables 1 & 2.)

Radiometric calibration was done so through two stages, transforming digital numbers (DN) stored within the Landsat GeoTiff files into surface reflectance values and histogram matching. Once again, any investigation aiming to analyze imagery from various dates should translate any value in DN format to surface reflectance to diminish radiometric contrasts and eliminating solar angle variations between images. To fulfill this procedure, the "Radiometric Calibration" tool within ENVI's "Radiometric

Correction" toolbox option was employed to each image individually. Selecting "Reflectance" under the "Calibration Type" drop-down box permitted the software to apply an algorithm using the metadata stored in the GeoTiff file to convert DNs into reflectance percentages (Figure 3.) After completing this process, a base image was selected once again to accomplish histogram matching throughout each of these images. Because of this examination's need to spectrally classify changes between these images, it was crucial to minimize tonal contrasts through histogram matching. This color-matching process was completed using ENVI's "Seamless Mosaic" tool (Exelis:Visual Information Solutions. "Seamless Mosaic." 2014.)

$$\rho_\lambda = \frac{\pi L_\lambda d^2}{ESUN_\lambda \sin\theta}$$

Where:

L_λ = Radiance in units of W/(m^2 * sr * μm)

d = Earth-sun distance, in astronomical units.

$ESUN_\lambda$ = Solar irradiance in units of W/(m^2 * μm)

θ = Sun elevation in degrees

Figure 3: ENVI Algorithm (USGS 2014.)

3.4 Water Detection

Delineating total lake coverage throughout the year was carried out by applying a simple band ratio method, the Modified Normalized Difference Water Index (MNDWI), to every image. The band algorithm for MNDWI is as follows:

$$\text{MNDWI} = \frac{Green - NIR}{Green + NIR}$$

With the "Band Math" tool in ENVI's toolbox, this equation was applied to each individual image. For Landsat 7 images representing the study period covering 2000-2001, bands 2 and 4 were inputted into the Band Math equation while images collected by Landsat 8 for the later study period utilized bands 3 and 5.

The reflectance values used to illustrate water bodies and later utilized in creating feature classes in ArcMap were found using an unsupervised method of classifying spectral values throughout a remotely sensed image referred as Iterative Self-Organizing Data Analysis Technique (ISODATA.) Unsupervised classification enables the analyst to extract land-cover data in an image without predetermining spectral value ranges that fit within a specific land-classification category. ISODATA allows the user to cluster spectral values into defined classifications with minimal input from the analyst, which includes inputting the image, determining number of iterations, and number of desirable classes. The spectral threshold for water throughout each image was recorded to be utilized later in ArcMap in creating water-body feature classes (Jenson 2005.)

3.5 Vegetation Index

The last ecological dataset analyzed in this analysis was the Normalized Difference Vegetation Index (NDVI), which was determined using the formula below:

$$\mathrm{NDVI} = \frac{NIR - Red}{NIR + Red}$$

The necessary bands in creating an NDVI image include Landsat 8's bands 4 and 5 and bands 3 and 4 from Landsat 7 (Jensen 2005.) Similar to retrieving water detection calculations, this process is accomplished using ENVI's "Band Math" tool and this equation was applied to each of the images by inserting the appropriate bands into the equation ("Frequently Asked Questions about the Landsat Missions" March 20, 2014.)

3.6 Converting Classified Lake Area into Feature Classes

With all of the images converted into a data file (.dat), ArcMap was able to process and display each of the photographs without conflict. Uploading the imagery onto ArcMap, not only permitted further analysis of the landscape by creating feature classes representing water-detection and vegetation values, but this software also allowed model-building to help characterize suitable snail habitats. The first tool utilized in creating water-body feature classes was the "Spatial Analyst" tool, "Raster Calculator." Utilizing the thresholds previously extracted from ENVI through ISODATA, these values were inserted into the calculator to create a raster calculation image highlighting regions meeting the water-detection spectral value criteria. After the new raster calculation file was created, the attribute table was opened and only the water regions were selected. The "Raster to Polygon" tool was then used to create a feature class representing the water bodies for that particular month. This procedure was repeated for all fifteen raster files.

With the lake-coverage feature classes established to represent water area for each image, the feature classes were separated according to the appropriate study period each represented. Beginning with the 2000-2001 study period, the exact dates in which Landsat acquired each image was used to determine the number of days between observations (Hui, et al. 2008.) By using ArcMap's "Intersect" option, the polygon features created after the raster calculations were inputted into this tool with only two images during a single iteration. The intersect function equates features based on their geographical relationship to one another. Portions of the features that overlap geometrically will be included in the newly created feature class while all other segments are omitted (Price 2006.)These intersect features represented the duration of water coverage throughout the lake area. The newly established intersect features classes were labeled according to the number of days each symbolized. This process was repeated for the 2013-2014 study period (Hui, et al. 2008.)

3.7 Classifying Areas Bordering Longstanding Water

Considering that villages at highest risk for infection usually lie less than 500 meters from bodies of water, this distance was also applied when finding the most suitable areas for snail habitats. However, a specific criterion that must be met for an inundated area to be adequate for snail reproduction is that the submerging period cannot last longer than eight months or less than three months. Accordingly, when applying a buffer feature class around the intersected water-detection feature classes, only those features representing submergence periods between 90 and 240 days obtained this proximity measure. These buffer features were clipped further only to remove edges that represented any water body outside of the acceptable criteria (Chen & Lin 2004.)

3.8 Vegetation Classification and Weighted Overlay

The lake inundation buffer zones created for both study periods were applied to the NDVI images belonging to their respective time-period. To accelerate processing, the NDVI imagery was clipped according to the boundaries of the buffer feature classes with the "Extract by Mask" tool. Following this extraction, the NDVI raster files were inputted into the "Weighted Overlay" analysis tool. Because *O. hupensis* snails often dwell in regions with at least sixty-percent coverage in vegetation and rarely found in areas with less than twenty-percent vegetation coverage, the created overlay respected these factors. Any NDVI value 0.6 or higher were rated with a 10 to represent "most likely" areas for snail habitat while NDVI values more than 0.2 and less than .6 were given a rating of 1 (Chen & Lin 2004.) Upon extracting the final calculations, the highest values representing "most suitable" habitats remained on the final map along with a visual representation of the seasonal lake area alterations.

Chapter 4: Results

Seasonal water patterns and extent differed substantially between the two study periods. During the 2000-2001 study period, there was a clear escalation of the lake's extent around June as it correlates with the Poyang Lake region's flood season. Water levels then decreased by approximately 18% during September 2000 and continued to decrease gradually throughout the winter with the lowest water levels occurring during the beginning of March 2001 with the total water area of 3216.38 km^2 (Figure 5; Table 3.) As far as the 2013-2014 time-frame, a similar enlargement of the lake's total area occurred during the Spring/Summer season with the highest water levels in July 2013. During October 2013, a dramatic decrease in the lake's total area took place equating to a 41% decrease in the Poyang Lake's coverage. The lowest levels recorded during the most recent study period ensued during November 2013 with a total water area of 2322.94 km^2. From that point until May 2014, the total water area possessed minor fluctuations (Figure 6; Table 4.) In addition, the average water coverage for the earlier study period was 3406.23 km^2 whereas the average for 2013-2014 was 2540.08 km^2 (Tables 3 and 4.)

In addition to the total area covered by the lake's extent, this study was also interested in flooding periods that were synonymous with an advantageous environment for snails. To that end, it was derived that the total area advantageous to snail populations based on specific inundation limits amounted to 941 km^2 during the 2000-2001 study period whereas 2013-2014 resulted in 1100km^2. In addition, year-around water coverage, which does not permit the survival of adult snails, summed to be 2,229.4 km^2 in 2000 while it was only 1081.6 km^2 during the second time period. Because the lake extent presented an apparent decrease, it came as no surprise that vegetation coverage increased as a result of exposed land area in regions previously submerged by lake water.

The extracted NDVI values were applied to the 500-meter buffer created around water features meeting the criteria for snail survival and it was found that, based on this methodology that the total area for suitable snail habitats increased significantly. While the suitable habitats, according to this model amounted to cover approximately 154.15 km^2 with the majority of these areas located on the eastern shoreline of the Poyang Lake, the suitable habitats calculated for the 2013-2014 period was 1,144.26 km^2. Furthermore, the areas sufficient for this species' survival, based on visual interpretation of the maps created using ArcMap 10.2, were scattered throughout the landscape making the distribution much different than those derived for the earlier time period (Figure 7.)

Chapter 5: Problems

5.1 Lack of Ground Data

The inability to extract field data for this study significantly hindered the potential of obtaining the necessary spectral values from the region of interest. Analysts possessing the capability of collecting ground spectral data through the use of a spectroradiometer could perform an absolute atmospheric radiometric correction known as the empirical line calibration. Assuming that in situ reflectance values could be collected around the same date and time as the data aggregated from the remotely sensed image, the image could be radiometrically corrected for atmospheric noise. Similar correction results could occur by using a spectral library, such as the library offered through USGS. Unfortunately, due to both time-constraints and the study area being outside of the United States, a spectral library including relevant landforms could not be found to successfully run an empirical line calibration. Instead, this project utilized a combination of turning digital numbers into reflectance values and histogram matching between images taken within the same study period. It should be noted that this methodology is not as accurate as utilizing an empirical line (Jensen 2005.)

Another piece of information that was not readily available was water-level data around the Poyang Lake area. According to previous studies exploring appropriate conditions for the *O. hupensis*, ninety-five percent of snails are found around water with water levels ranging from fourteen meters to seventeen meters. These areas with these water levels are referred as the “high-density zone,” which is typically inundated between April and October. Areas in which no snails are found are water bodies with levels above eighteen meters. It is apparent, then, that adequate classification of snail distribution highly relies on these measurements (Chen & Lin 2004.) Unfortunately, in spite of strenuous effort during the data collection process, the procedure was unsuccessful in extracting any water-level data that could be utilized in conjunction with the study area or study periods. Initial attempts included extensive queries in scientific journal search engines and Chinese government-run websites. Though water-level data was found on websites, such as *Jiangxi Hydrological Information Network*, the graphs and tables were in Chinese and the metrics used for measurements, such as meters, was not readily apparent on the document (2014.) Additional time and resources would have provided a means to incorporate this data into the final snail prediction model making the final product better resemble areas in which the snail populations could thrive (Chen & Lin 2004.) It should be noted, however, that the data collected by the Jiangxi provincial government did not cover the entire study area, so accurate assessment would be limited to those select areas (Jiangxi Hydrology Information Network 2014.)

Snail density data extracted from in-place surveys was another item that was unavailable for this analysis. A typical *schistosomiasis* prediction model, assuming that the research was performed throughout the Poyang Lake area, collects field data pertaining to the number of snails found throughout various sampling areas. Research projects seeking to explore *schistosomiasis* using this field data collection must adhere to standard procedures set forth by the Chinese National *Schistosomiasis* Control Programme (Guo, et al. 2005.) Furthermore, the conventional method of extracting snail densities is through what is referred as a "stratified sampling method," which, as it pertains to snails, requires the use of a small square frame that measures 0.11 m^2 and is placed systematically across the study area. Snails within the square frame are counted and most researchers mark this data collection coordinate by using a Global Positioning System (GPS) (Yang, et al. 2008.) Statistical analyses, which were, oftentimes, nonparametric correlations, including but not limited to the Spearman correlation or Bayesian modeling, were used in conjunction with snail populations and ecological measurements. These statistics were used to access the correlation between the models researchers created with data extracted from the field (Yang, et al 2005.) As it pertains to this study, there existed early hopes that in situ snail data could be derived from previous studies enabling statistical examination. However, the data found, such as surveys aggregated by Jia-Gang, et al, occurred outside of this study's study period and, as a result, could not be used for analysis (2005.)

5.2 Limited Temporal Resolution

The temporal resolution for the Landsat multispectral remote sensing systems is sixteen days for any given location on the Earth regularly covered by the sensor (Jensen 2007.) Though the resolution would be sufficient in interpreting fluctuations in water coverage around throughout Poyang Lake, images with excessive cloud-cover rendered many available images inadequate for any conceivable analysis. The duration between collected images ranged from 32 days to 112 days. This analysis was simply unable to investigate month-to-month changes in the lake's composition that may drastically change from significant weather phenomena, such as heavy rainfall. However, even sensors with higher temporal resolution, including Moderate-Resolution Imaging Spectroradiometer (MODIS) or Advanced Very High Resolution Radiometer (AVHRR), possess drawbacks as both maintain coarse spatial resolution when compared to Landsat (Hui, et al. 2008.)

Chapter 6: Discussion

Unquestionably, several past analyses interpreting large-scale patterns of *schistosomaisis* utilized remote sensing imagery analysis techniques coinciding with GIS spatial tools (Zhou, et al. 2001.) The advantages of employing these tools to epidemiological studies are appealing, because they include the ability to analyze ecological conditions across large study areas without needing to be in direct contact with phenomena and remote sensors collect data systematically preventing human error in data extraction or bias. However, it should be noted that remote sensing possesses several inherent disadvantages. The first drawback with using satellite imagery is limited resolution, whether it would be spatial, temporal, or spectral. Another disadvantage is that image analysis takes a certain level of skill making these endeavors difficult for those not readily exposed to image interpretation. In addition, though remote sensing could, ultimately, save time when compared to solely extracting data through field methods, the remote sensing process is still relatively time-consuming. Each of these considerations must be accounted for prior to beginning an investigation of this magnitude (Jensen 2007.)

The results ascertained during this study found vast changes around the Poyang Lake area since the completion of the Three Gorges Dam affecting the distribution of snails across the landscape. Based on the criteria appointed for this study, seasonal lake-boundary changes, 500-meter buffer around suitable submergence days, and dense vegetation, this suitability model found a compelling increase in potential snail habitats during the 2013-2014 study period. In fact, there appeared to be over a 600% increase in the total area that would be advantageous to this species. These regions are more opportune for snail-population increase because the annual flooding the occurred previously, killing the adult snail populace, dramatically dwindled in duration since the activation of the TGD (Chen & Lin 2004.) Though it is quite possible that this model overestimated the potential enlargement of snail habitat on the basis that more environmental data would need to be extracted to verify these counts, such as water-level information, these findings resemble previous predictions regarding ecological changes around the study area from the TGD. Inferences about the possible changes posited that the TGD would cause alterations to downstream flows that would increase total marshland areas conducive to the *O. hupensis.* This suitability model likely illustrated this particular development through an increase in overall vegetation density and, as the lake receded from the TGD impoundment, left small, isolated patches of water. Verifying this hypothesis would require visual analysis of the areas in question. In addition, the hundreds of secluded water areas are worth further mention as they introduce an air of unpredictability in snail distribution. Whereas the data collected for the earlier study period resulted in snail habitats primarily located along the western shoreline of Poyang Lake, the model representing 2013-2014 presented a much wider and haphazard snail habitat dispersal. Even supposing that there is an overestimation in the prediction model, the results do warrant further investigation to ensure the health and safety of the surrounding communities (McManus, et al. 2010)

Clearly, since the official activation of the Three Gorges Dam, significant changes have taken place throughout Poyang Lake with regards to water area and vegetation potentially increasing snail distribution. On that end, it should also be recognized that this dam has significantly sand and water distribution of many other areas along the Yangtze

River, as well. In fact, originally, prior to investigating Poyang Lake, this project sought to determine whether *schistosomiasis* could be introduced around the Three Gorges Reservoir Region (TGRR), which, as mentioned previously, extends 600 km (Ross, et al. 2001.) Because the dam considerably decelerated the Yangtze River's original, rapid velocity, various researchers grew concerned that this reduction in speed would establish an area conducive to snails that was historically absent of this species' presence. Additionally, raised concerns arose due to the fact that the TGRR is located between two key endemic areas with one located 500 km upstream whereas the other is based only 40 km downstream; it was feared that these two transmission areas would eventually merge into one large endemic region (McManus, et al. 2010) (Figure 3.) However, this study area was not explored because it was unlikely that snail populations could survive in the midst of the reservoir because it was 185 m in depth, which exceeded water levels in which this species could thrive. The only area with potential to harbor the *O. hupensis* would be along the shoreline but a suitability model adequately predicting their distribution must include water-level data, as the two variables used in this analysis would, potentially, overestimate snail habitats (Chen & Lin 2004.)

Even supposing that the TGRR could be analyzed, the models created using the methodology used with this project only predict the existence of snails based on environmental suitability and do not perfectly equate with *schistosomisis* risk to nearby communities (Friis & Sellers 2014.) First, it should be emphasized that not every snail is infected with the *S. japonica* flatworm. Past studies that inspected infection potential delineated which snails were infected by crushing each specimen while analyzing it under the microscope (Yang, et al. 1009.) As such, this study delineated suitable habitats for both infected and non-infected snails. Calculating *schistosomiasis* risk, in a true epidemiological study, would require some sort of observational-based or experimental study. In the majority of cases, observational studies are carried out due to budget considerations. These study designs interpret human populations in question, such as villages, subsets within a population, or entire regions. Although explanation for these studies is beyond the scope of this study, it is crucial to understand that an epidemiologic measure, such as relative risk, which is utilized in cohort studies, interprets risk by comparing incidence rates to groups exposed to a particular environment or circumstance compared to those that are unexposed. Exposures in epidemiology could range from living conditions, age, income, environmental hazards, along with many other potential variables. If this study strove to analyze risk, an enhanced understanding of the communities living throughout the Poyang region must be extracted prior to making meaningful conclusions about *schistosomiasis* risk. However, through the use of remote sensing and building a predictive model through a GIS, this project, at least, derived an area worth future epidemiological analysis (Friis & Sellers 2014.)

This study solely analyzed Poyang Lake's suitability for snail reproduction using two specific criteria, water boundaries, areas within 500 meters of these water sources, and vegetation coverage. These two attributes have proven quite crucial in analyzing snail reproduction, yet many analyses also incorporated other variables in conjunction with these two calculations creating a more comprehensive representation of suitability. Some other ecological features include precipitation, elevation, and land surface temperature (LST.) In fact, one of the original intents of this project was to incorporate LST into the final suitability model. Previous studies using LST concluded that landscapes with higher temperatures possessed a higher likelihood that the snail populations were infected with the *S. japonica* flatworm, which could not be defined this this project based on the environmental elements collected. Several analyses concluded that higher temperatures were synonymous with the rapid development of these flatworms helping them reach their reproduction stage in life sooner (Yang, et al 2005.) LST formulas were found that derive surface temperatures based on reflectance values throughout an image with the hopes to adding this attribute to the predictive model (Zhang, et al. 2005.) However, considering the Earth's diurnal cycle or temperature changes that occur within a 24-hour time period, which changes drastically throughout the day, more information regarding this cycle in China is required. Furthermore, the time of day the satellite collects imagery would have to be better understood to gain a comprehensive and useable understanding of the measurements derived from the LST formula. Ultimately, time-constraints forced this analysis to exclude this variable from the study (Jensen 2007.)

Another aim this project sought to explore relating to temperature was the possible effect of Global Climate Change (GCC) on this snail species' reproduction and residence. The interest in exploring this content was driven by the sheer lack of literature describing snail habitat models through the use of geospatial tools (McManus, et al. 2010.) It was quickly realized during data collection that few models attempted to incorporate GCC because this process does not represent equal warming of the earth but, instead, is an unequal cooling and warming of the Earth driven by changes likely caused by increased anthropogenic activities. Many General Circulations Models (GCMs) have been created over the years and, in spite of strenuous research, possess various uncertainties caused by a combination of variances of greenhouse gasses around the world and complex weather patterns, which is beyond this scope of this thesis (Osborne et al. 2013.) Because of the *O. hupensis* snail is sensitive to changes in the environment, there is no doubt that any change in climate would alter its activities and survival (Yang, et al. 2010.) According to discussions held among the World Health Organization (WHO), a 2ºC increase in temperature throughout the People's Republic of China would increase snail distribution by 50-100% (Qian, et al.) Incorporating this theory with a model would entail extracting long-term temperature data to make meaningful conclusions as to areas that would be most affected by GCC in terms of snail population. Though this aspect was worth exploring, much more research and an overall understanding of GCC must be derived (Slenning 2010.)

Chapter 7: Conclusion

This examination analyzed the feasibility in utilizing remote sensing and GIS tools in, not only observing the transformations around the Poyang lake region proceeding the completion of the Three Gorges Dam, but how these alterations influenced potential snail habitats. The two *O. hupensis* suitability models created indicated a significant transformation of the Poyang lake landscape likely caused by one of the original intents of the TGD construction, to prevent annual downstream flooding (McManus, et al. 2010.) According to the models built during this analysis, the potentiality for snail population expansion appeared probable in accordance with annual water boundaries, areas within close proximity of these water areas, and vegetation coverage (Tseng, et al. 2014.) Additionally, with the increases in marshlands and the distribution of many isolated pockets of water conducive to snail survival, further investigation interpreting this phenomenon is warranted to ensure that appropriate public health initiatives are taking place (McManus et al. 2010.) As such, utilizing satellite imagery to predict the location of species contributing to infectious diseases using known environmental preferences for these vectors is profitable for surveillance (Friis & Sellers 2014.)

Although the use of satellite imagery possesses a great potential forecasting vector habitats, this methodology possesses various limitations imposed by those inherent to remote sensing. The most notable restrictions in using the remote sensing process include mechanical inhibitions within the sensors themselves, such as limited temporal, spatial, or spectral resolutions. Furthermore, though remote sensing possesses the capability of reducing long, strenuous fieldwork, in situ data is still essential in calibration and statistical analysis of the findings. Another item to consider is that remote sensing can merely predict vector habitats using environmental data that could be extracted using reflectance values throughout a remotely sensed image (Jenson 2007.) This factor in itself is limiting because human interaction throughout these areas play a significant role in vector survival, as well. As such, a complete, comprehensive prediction model for vector habitats would be exceptionally difficult and would require extensive knowledge about activities carried out throughout the communities of interest. So long as the analysts comprehend that attributes will, most certainly, be missing from the final analysis, vector-borne illness analysis with geospatial tools could be done so thoughtfully (Nelson & Williams 2001.)

Overall, this application in utilizing Landsat 8's Operational Land Imager and Landsat 7's Thematic Mapper Plus in interpreting *schistosomiasis* in the People's Republic in China supported the potential that spatial modeling and satellite imagery could predict vector-borne illnesses. This application is crucial in consideration with the widespread reemergence of, not only *schistosomiasis* in various provinces in China, but also other infectious diseases affecting primarily developing countries around the world, including malaria or African *schistosomiasis* (Bertz 2011.) Using remote sensing, analysts could interpret large geographic areas in predicting suitability that could help allocate where resources or medical research should be warranted. In spite of the limitations discovered throughout this study, it is without question that as sensors and methodologies enhance, geospatial tools will offer a unique capability in infectious disease modeling that is efficient in time, cost allocation, and, perhaps, predict how climatic changes will affect vector distribution (Cline 2006.)

References

Arbogast, Alan. *Discovering Physical Geography*. Hoboken, NJ: John Wiley & Sons, Inc. , 2007.

Beltz, Lisa, A. *Emerging Infectious Diseases: A Guide to Diseases, Causative Agents, and Surveillance.* Wiley Publishers, 2011: 4.

Chen, Honggen, and Dandan Lin. "The Prevalence and Control of Schistosomiasis in Poyang Lake Region, China." *Parasitology International* 53 (2004): 115 & 117-18.

Cline, Barnett. "Invited Editorial For the Inaugural Issue of Geospatial Health." *Geospatial Health*. (2006): 3-4.

Friis, Robert H., and Thomas A. Sellers. *Epidemiology for Public Health Practice*. 4th ed. Sudbury, Mass.: Jones and Bartlett Publishers, 2009. 8-9, 12, & 280-281.

Guo-Jing, Yang, Penelope Vounatsou, Zhou Xiao-Nong, Jurg Utzinger, and Marcel Tanner. "A Review Of Geographic Information System and Remote Sensing with Applications of the Epidemiology and Control of Schistsomiasis in China." *Acta Tropica*. no. 96 (2005): 118.

Hu, Haitang, Peng Gong, and Bing Xu. "Spatially Explicit Agent-Based Modelling for Schistosomiasis Transmission: Human-Environment Interaction Simulation and Control Strategy Assessment." *Epidemics*. (2010): 49.

Hu, Yi, Zhije Zhang, Yue Chen, Zengliang Wang, Jie Gao, Bo Tao, Qiulin Jiang, and Qingwu Jiang. "Spatial Pattern of Schistosomiasis in Xingzi Jiangxi Province, China: The Effects of Environmental Factors." *Parasites & Vectors*. no. 214 (2013): 2.

Hui, Fengming, Bing Xu, Huabing Huang, Qian Yu, and Peng Gong. "Modelling Spatial-temporal Change of Poyang Lake Using Multitemporal Landsat Imagery." *International Journal of Remote Sensing*, 2008, 5767-772 & 5777.

Jensen, John. *Introductory Digital Image Processing: A Remote Sensing Perspective*. Upper Saddle River, NJ: Pearson Prentice Hall, 2005: 11, 13, 305.

Jensen, John. *Remote Sensing of the Environment: An Earth Resource Perspective*. Upper Saddle River, NJ: Pearson Prentice Hall, 2007: 7-10 & 14-18.

Jia-Gang, Guo, Penelope Vounatsou, Cao Chun-Li, Jurg Utzinger, Zhu Hong-Qing, DanielAnderegg, Zhu Rong, and He Zhan-Ying. "A Geographic Information and Remote Sensing Based Model for Prediction of Oncomelania Hupensis Habitats in the Poyang Lake Area, China." *Acta Tropica*. (2005): 214.

Jiangxi Hydrological Information Network. January 1, 2014. Accessed August 11, 2014. http://www.jxsw.cn.

Jis, Yun-Fei, Katy Siegel, June Mei, and . "Three Gorges Dam Migration." *College Art Association*. no. 3 (2010): 73.

Kameda, Yuichi, and Makoto Kato. "Terrestrial Invasion of Pomatipsid Gastropods in the Heavy-Snow Region of the Japanese Archipelago." *BMC Evolutionary Biology*. no. 118 (2011): 4.

McCreesh, Nicky, and Mark Booth. "Challenges in Predicting the Effects of Climate Change on Schistosoma Mansoni and Schistosoma Haematobium Transmission Potential." *Trends in Parasitology*. no. 11 (2013): 549.

Nelson, Kenrad, E and Williams, Carolyn. *Infectious Disease Epidemiology: Theory and Practice.* Gaithersburg, MA: Aspen Publishers, 2001: 3 and 231.

Osborne, Tom, Gillian Rose, and Tim Wheeler. "Variation in the Global-scale Impacts of Climate Change on Crop Productivity Due to Climate Model Uncertainty and Adaptation." *Agricultural and Forest Meteorology* 170 (2013): 183-84.

Price, Maribeth. *Mastering ArcGIS: Fifth Edition.* New York, NY: McGraw-Hill Companies, 2012: 1-3.

Qian, Ying-Jun, Li, Shi-Zhu, Xu, Jing, Yang, Kun, Huang, Yi-Xin, Cao, Zhi-Gho, Miu, Feng, Dang, Hui, Zhang, Li-Juan, Zhang, Li, Wang, Qiang, Bergquist, Robert, Zhou, Xiao-Nong. "Potential *Schistosomiasis* Foci in China: A Prospective Study For *Schistosomiasis* Surveillance and Response." *Acta Tropica.* 2013: 1-2.

Ross, Allen, G, Sleigh, Adrian, C., Lim Yuesheng, Davis, George, M., Williams, Gail, M., Jiangm Zheng, Feng, Zheng, and McManus, Donald. "Schistsomiasis in the People's Republic of China: Prospects and Challenges for the 21st Century." *Clinical Microbiology Reviews.*14(2) American Society for Microbiology, 2001: 271.

"Seamless Mosaic (Using ENVI) | Exelis VIS Docs Center." Seamless Mosaic (Using ENVI) | Exelis VIS Docs Center. January 1, 2014. Accessed September 1, 2014. http://www.exelisvis.co m /docs/MosaicSeamless.html.

Shiode, Shino. "Revisiting John Snow’s map: network-based Spatial Demarcation of Cholera Area." *International Journal of Geographical Information Science*. no. 1 (2012): 136.

Slenning, B. D. "Global Climate Change And Implications For Disease Emergence." *Veterinary Pathology* 47, no. 1 (2010): 28-31.

Tseng, Kuo-Hsin, Song Liang, Motomu Ibaraki, Hyongki Lee, and C.K. Shum. "Study of the Variation of Schistosomiasis Risk in Lake Poyang in the People's Republic of China Using Multiple Space-Borne Sensors for Monitoring and Modeling." *Geospatial Health* 8, no. 2 (2014): 353-59.

U.S. Geological Survey (USGS), "Frequently Asked Questions about the Landsat Missions." Last modified March 20, 2014. http://landsat.usgs.gov/band_designations_landsat_satellites.php.

U.S. Geological Survey (USGS), "Province Hubei Map," 2014. EarthExplorer. http://earthexplorer.usgs.gov/ (Accessed April 22, 2014).

U.S. Geological Survey (USGS), "Landsat 8." Last modified August 2013. Accessed http://pubs.usgs.gov/fs/2013/3060/pdf/fs2013-3060.pdf.

Utzinger J, Raso G, Steinmann P, Zhou XN, Vounatsou P, Brooker S, N'Goran EK (2011) Schistosomiasis. In: *Encyclopedia of Environmental Health*, 10-19. Vol. 5. Amsterdam: Elsevier Science, 2011.

Wertheim, Heiman, Peter Horby, and John Woodall. *Atlas of Human Infectious Diseases*. 2012.

Xianyi, Chen, Liying, Wang, Jiming, Cai, Xiaonong, Zhou, Jiang, Zheng, Jiagang, Guo, Xiaohua, Wu, Engels, D., and Minggang, Chen. "*Schistosomiasis* Control in China: The Impact of a 10-Year World Bank Loan Project (1992-2001). *Bulletin of the World Health Organization.* 83(1), 2005: 43-45.

Yang, Guo-Jing, Penelope Vounatsou, Xiao=Nong Zhou, Marcel Tanner, and Jurg Utzinger. "A Bayesian-Based Approach for Spatio-Temporal Modeling of County Level Prevalence of Schistosoma Japonicum Infection in the Jiangsu Province, China." *International Journal of Parasitology*. (2005): 6.

Yang, Kun, Xian-Hong Wang, Guo-Jing Yang, Xiao-Hua Wu, Yun-Liang Qi, Hong-Jun Li, and Xiao-Nong Zhou. "An Integrated Approach to Identify Distribution of Oncomelania Hupensis, The Intermediate Host of Schistosoma Japonicum in a Mountainous Region In China." *International Journal of Parasitology*. no. 38 (2008): 1007-1008.

Zhang, Zhiji, Carpenter, Tom, E., Chen, Yue, Clark, Allen, B., Lynn, Henry, S., Peng, Wenxiang, Zhou, Yibiao, Zhao, Genming, and Jiang, Qingwu, Jiang. "Identifying High-Risk Regions for *Schistosomiasis* in Guichi, China: A Spatial Analysis." *Acta Tropica.* 2008: 1-2.

Zhang, Zhijie, Ong, SengHuat, Peng, Wenxiang, Zhou, Yibiao, Zhuang, Jianlin, Zhoa, Genming, and Jiang, Qingwu. "A Model for the Prediction of *Oncomelania Hupensis* in the Lake and Marshland Regions, China." *Parasitology International*. Elseivier Publishers: 2008: 1-2.

Zhang, Zhijie, Ward, Michael, Jie Gao, Zengliang Wang, Baodong Yao, Tiejun Zhang, and Qingqu Jiang. "Remote Sensing and Disease Control in China: Past, Present, and Future." *Parasites & Vectors*. no. 11 (2013): 1-2.

Zhang, Zhijie, Zhu, Rong, Ward, Michael, P., Xu, Wanghong, Zhang, Lijuan, Guo, Jiagang, Zhao, Fei, and Jiang, Qingwu. "Long-Term Impact of the World Bank Loan Project for *Schistosomiasis* Control: A Comparison of the Spatial Distribution of *Schistosomiasis* Risk in China." *PLOS: Neglected Tropical Diseases.* 6(4), 2012: 1-2.

Zhi-Ying, Zhang, Xu De-Zhong, Zhou Xiao-Nong, Zhou Yun, and Liu Shi-Jun. "Remote Sensing and Spatial Statistical Analysis to Predict the Distribution of Oncomlania Hupensis In the Marshlands of China." *Acta Tropica*. no. 96 (2005): 206.

Zhou, X.N., J.B. Malone, T.K Kristensen, and N.R. Bergquist. "Application of Geographic Information Systems and Remote Sensing to Schistosomiasis Control in China." *Acta Tropica*. no. 79 (2001): 98.

Zhou, Yi-Biao, Liang, Song, and Jiang, Qing-Wu. "Factors Impacting on Progress Towards Elimination of Transmission of *Schistosomiasis* Japonica in China." *Parasites and Vectors*. 5(275), 2012: 1-4.

Figures

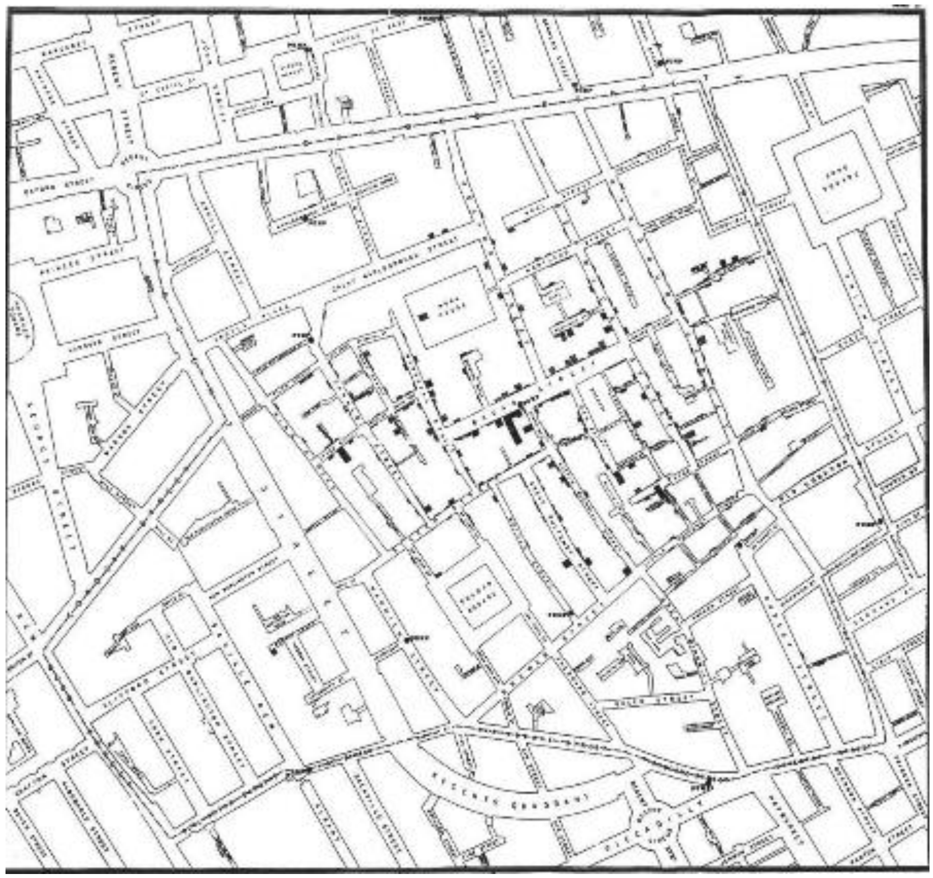

Figure 1: Dr. John Snow's Cholera Map showing cholera out breaks that occurred in London, England in 1854 (Shiode 2012.)

Figure 2: Map presenting the provinces in China.

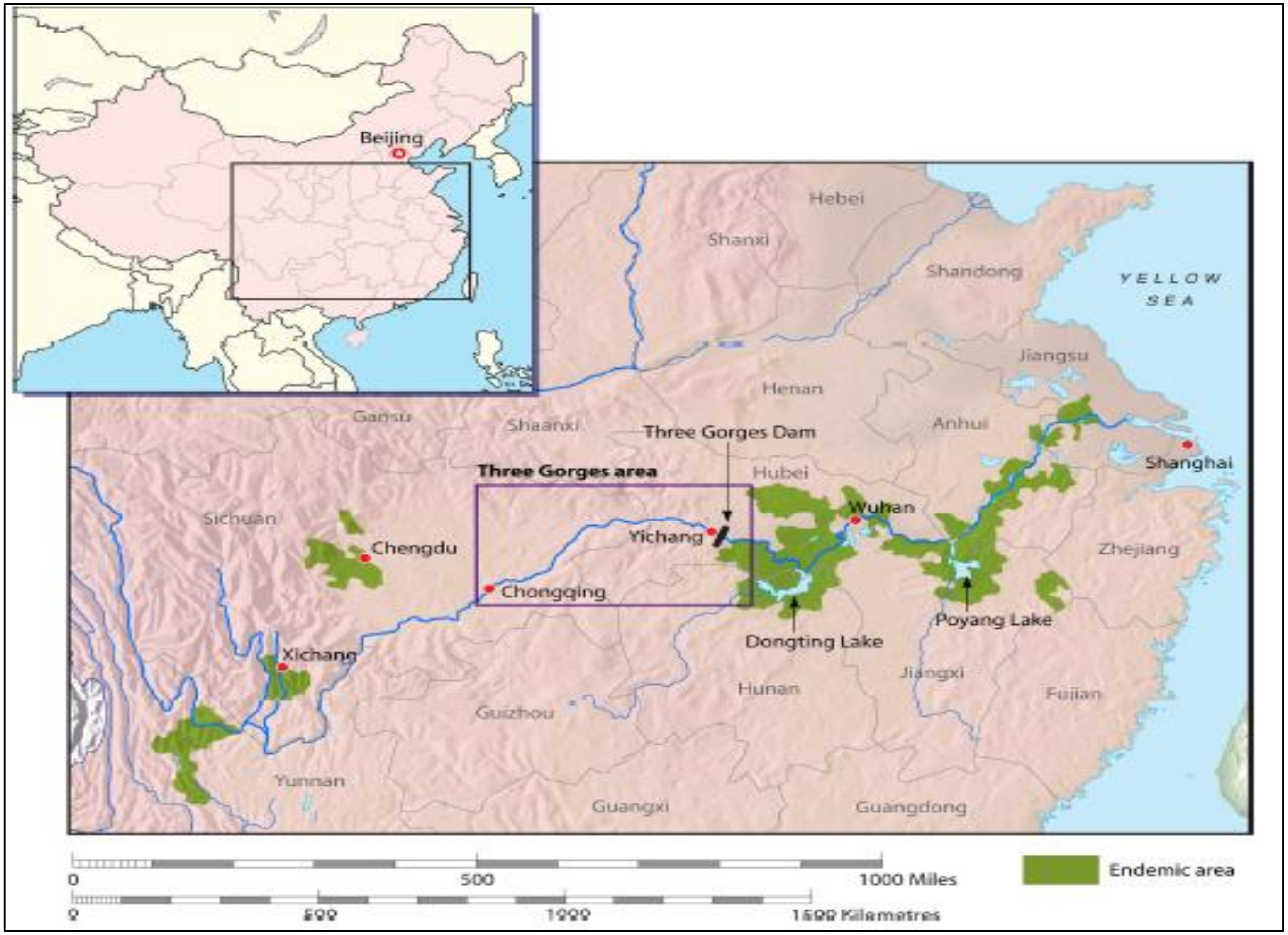

Figure 3: Map highlighting areas classified as endemic to *schistosomiasis*. This image also disignates the location of the Three Gorges Dam in association with these areas (McManus, et al. 2010).

Figure 4: Image Taken by Landsat 7of Poyang Lake, April 2000 (image credit: USGS 2014.)

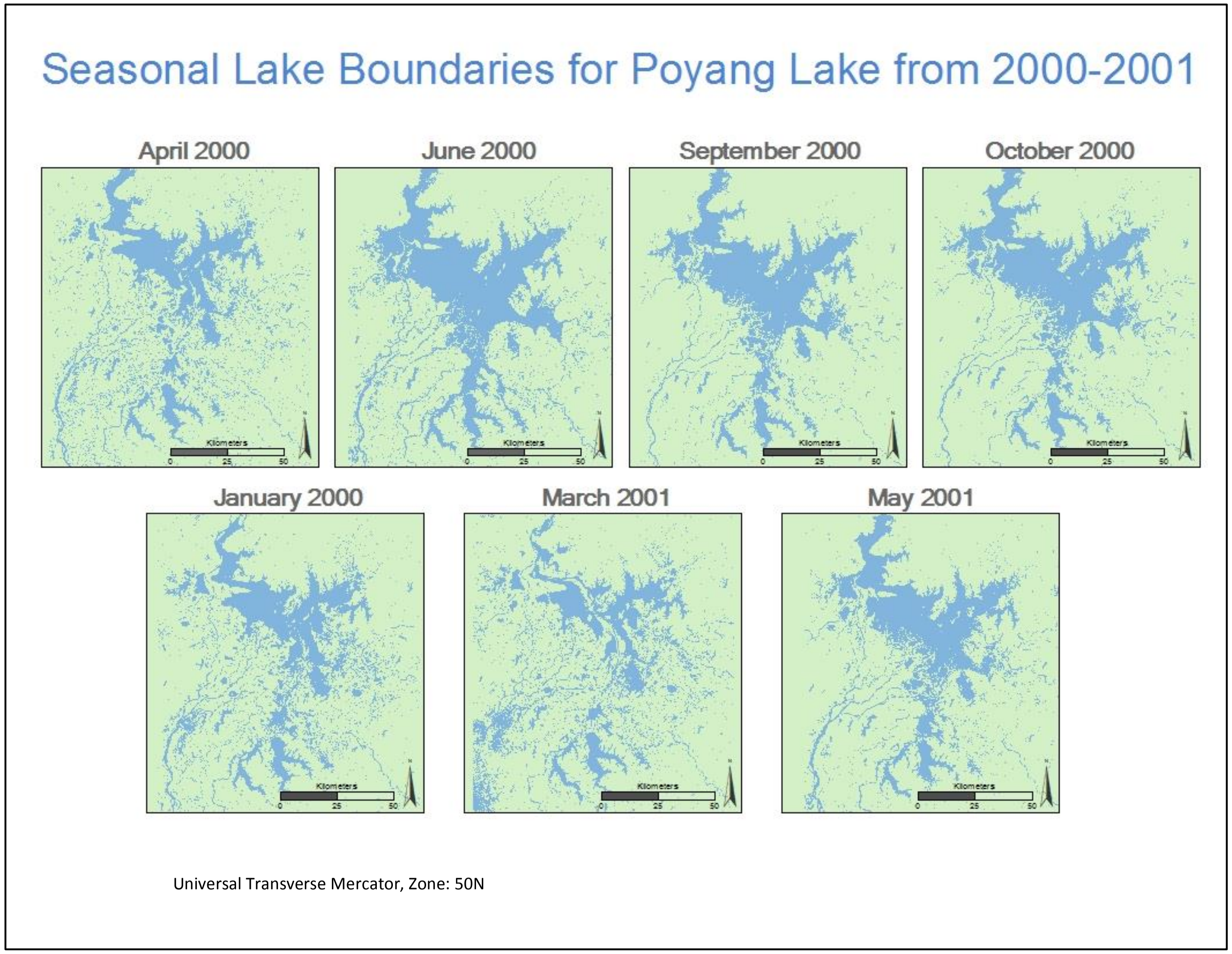

Figure 5: Monthly lake-boundary of Poyang Lake based on water-detection calculations from April 2000 to May 2001.

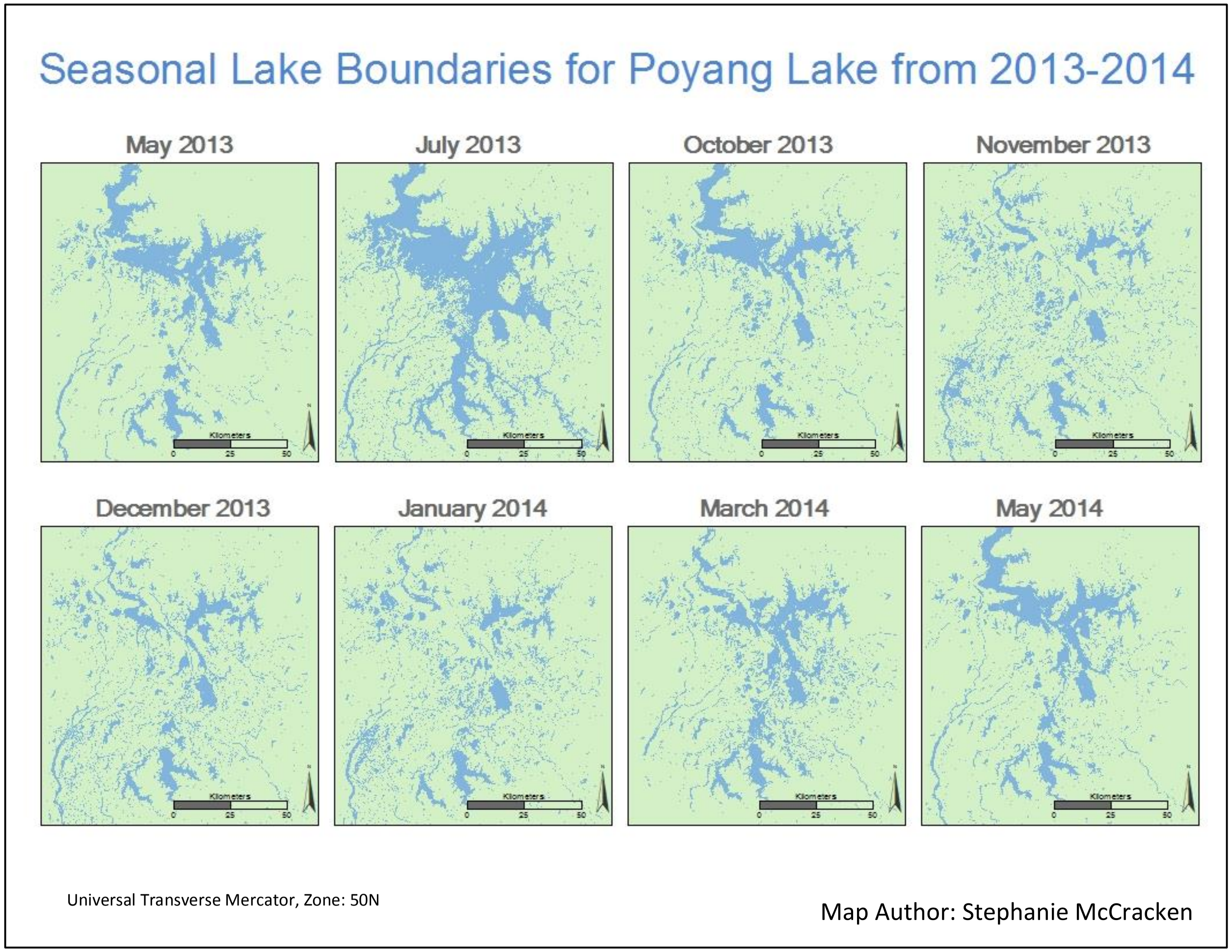

Figure 6: Monthly lake-coverage extent at Poyang Lake that was derived through using water-detection computations covering the time-period from May 2013 - May 2014.

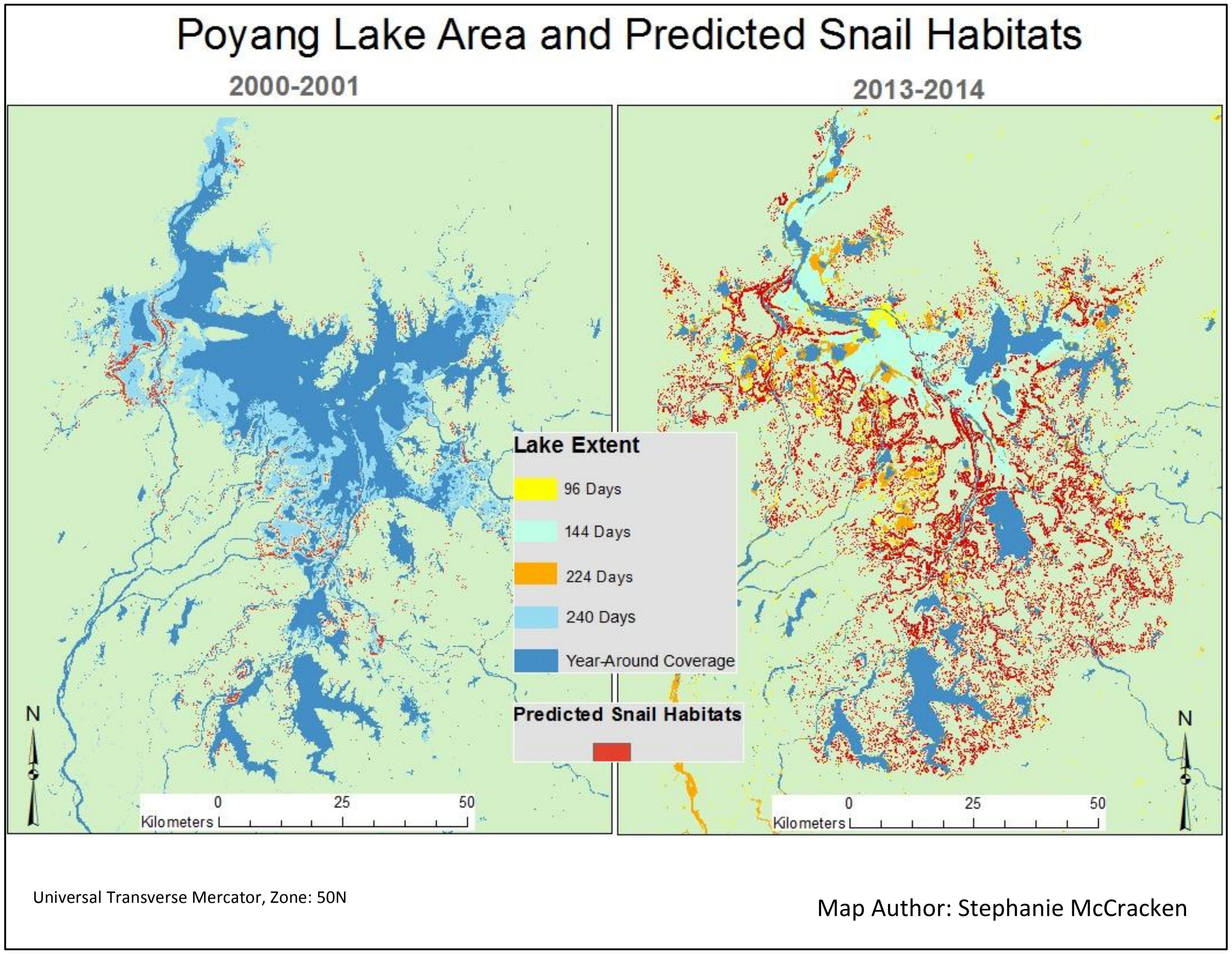

Figure 7: Map representing snail habitats during the 2000-2001 and 2013-2014 study periods based on weighted overlay spatial analysis using NDVI and MNDWI values.

Tables

Images Utilized to Examine Poyang Lake during 2000-2001 Study Period

Date Image Taken	Sensor (Landsat)	Spectral Bands	Percent Cloud Cover	Number of Ground Control Points	Root Mean Square Error (RMSE) unit = pixel
16 April 2000	ETM	1-5, 7	1%	76	0
3 June 2000	ETM	1-5, 7	5%	16	.006
23 September 2000	ETM	1-5, 7	1.7 %	96	0
9 October 2000	ETM	1-5, 7	0	58	.008
29 January 2001	ETM	1-5, 7	0	n/a	n/a
2 March 2001	ETM	1-5, 7	3.4%	12	.006
13 May 2001	TM	1-5,7	0	14	.004

Table 1: This table represents essential information describing the usefulness of the images selected for analysis during the 2000-2001 study period. Items including cloud-cover percentages and root mean square error were calculated manually using the specified region-of-interest only.

Images Utilized to Examine Poyang Lake during 2013-2014 Study Period

Date Image Taken	*Sensor (Landsat)*	*Spectral Bands*	*Percent Cloud Cover*	*Number of Ground Control Points*	*Root Mean Square Error (RMSE) unit = pixel*
14 May 2013	*OLI*	*1-5*	*1.5%*	*75*	*0*
1 July 2013	*OLI*	*1-5*	*9 %*	*30*	*.002*
5 October 2013	*OLI*	*1-5*	*0*	*91*	*0*
22 November 2013	*OLI*	*1-5*	*0*	*61*	*.003*
24 December 2013	*OLI*	*1-5*	*0*	*111*	*0*
25 January 2014	*OLI*	*1-5*	*0*	*n/a*	*n/a*
14 March 2014	*OLI*	*1-5*	*7%*	*91*	*0*
1 May 2014	*OLI*	*1-5*	*0*	*92*	*0*

Table 2: Covering the 2013-2014, this table portrays information regarding information including, date, sensor, and important data that is relevant for adequate remote sensing image analysis.

Monthly Total Water Area, Poyang Lake 2000-2001

Month and Year	Total Water Area (km^2)
April 2000	3279.83
June 2000	4072.12
September 2000	3306.14
October 2000	3346.38
January 2000	3286.69
March 2001	3216.38
May 2001	3336.07

Table 3: Data extracted from water-detection calculations for the months available during the study period derived from the "Calculate Geometry" feature in ArcMap 10.2. This table displays the total area the Poyang Lake covers within a given month within the 2000-2001 timeframe.

Monthly Total Water Area, Poyang Lake 2013-2014

Month and Year	Total Water Area (km^2)
May 2013	2350.1
July 2013	3971.96
October 2013	2338.89
November 2013	2322.94
December 2013	2324.84
January 2014	2323.57
March 2014	2343.39
May 2014	2344.93

Table 4: This table includes all of the water-coverage data extracted from imagery calculations for all of the available months within the 2013-2014 study period.