GEOSPATIAL ANALYSIS OF THE ANCIENT MAYA WETLAND AGRICULTURAL FIELDS OF NORTHERN BELIZE

by

Tanya Catignani

A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Earth Systems & Geoinformation Science

Committee:

_____	Dr. Paul Houser, Dissertation Director
_____	Dr. Sheryl Beach, Committee Member
_____	Dr. Ronald Resmini, Committee Member
_____	Dr. Terrance Slonecker, Committee Member
_____	Dr. Giuseppina Kysar-Mattietti, Committee Member
_____	Dr. Donna M. Fox, Associate Dean, Office of Student Affairs & Special Programs, College of Science
_____	Dr. Peggy Agouris, Dean, College of Science
Date: _____	Spring Semester 2017 George Mason University Fairfax, VA

Geospatial Analysis of the Ancient Maya Wetland Agricultural Fields of Northern Belize

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

by

Tanya Catignani
Master of Arts
State University of New York at Buffalo, 2004

Director: Paul Houser, Professor
Department of Geography & Geoinformation Science

Spring Semester 2017
George Mason University
Fairfax, VA

DEDICATION

This dissertation is dedicated to my parents, Grazyna and Nicola, who built the foundation of my success.

ACKNOWLEDGEMENTS

I would like to thank my family and friends who supported me in this endeavor over the last several years. I've appreciated your encouragement, advice, patience, and love more than I can express. I would particularly like to thank Dr. Alexey Castrodad who critically reviewed several drafts of my dissertation, and generously spent many hours helping me reach my goal. My dear friend, Dr. Marci Zacherl, has been a sounding board since the beginning of this process, and always made sure that I took time to relax and enjoy life. My favorite part of this research was getting to spend time with my mother, who helped me collect soil samples in Ohio; I am fortunate to have been raised by a strong woman who enjoys hard work, the outdoors, and was never afraid of getting dirty. I'm very grateful to my committee for their guidance and enthusiasm in my research. Each one played an integral part in my dissertation journey. Finally, my research was generously funded and/or supported by a number of organizations, each of which provided a critical element to the final product. In no particular order, they are the National Geospatial-Intelligence Agency, the United States Geospatial Intelligence Foundation, the US Geological Survey, PANalytical, the Cosmos Club Foundation, the Maya Research Program, the National Science Foundation, the University of Texas at Austin, and George Mason University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

ABSTRACT

GEOSPATIAL ANALYSIS OF THE ANCIENT MAYA WETLAND AGRICULTURAL FIELDS OF NORTHERN BELIZE

Tanya Catignani, Ph.D.

George Mason University, 2017

Dissertation Director: Dr. Paul Houser

This research analyzes the spatial distribution of ancient Maya wetland agricultural fields using remotely sensed and geospatial data. The Maya constructed wetland agricultural fields during the cultural periods known as the Preclassic and the Classic. The extent of these fields is still unclear, as is the approximate size of the population they once supported. The study of these wetland agricultural fields will help to understand how the Maya civilization was able to thrive in an inhospitable environment for thousands of years, and why their population sharply declined around 900 AD. Since 1981, remotely sensed imagery such as synthetic aperture radar (SAR), multispectral, and lidar has been used to study Maya wetland agriculture. This research will employ several types of satellite imagery and a number of geospatial data sources to locate and predictively model the ancient fields, and analyze the spectral reflectance of their soils.

CHAPTER ONE - INTRODUCTION

Belize is a small country in Central America, situated between Honduras, Guatemala, and the Mexican state of Quintana Roo, and is well known for its beaches, blue waters, and fascinating cultural history. The landscape in the north changes drastically as one travels inland and away from the resort-lined barrier reef along the coast, to the flat coastal plains to rolling hills blanketed in tropical forest and dotted with sleepy towns where the people represent a broad range of ethnic backgrounds. Many of the faces beg the question of whether there was truly ever a "Maya collapse" as one-third of the population is of mixed Maya and European descent, or Mestizo. An additional 10% are full Maya descendants. The population also consists of one-third Creole, and smaller percentages of Afro-Amerindian, East Indian, Chinese, and Europeans and more. In many ways Belize has a similar essence that once existed during the height of ancient Maya civilization, as the cultural mix influences ideas, design, and technology to create a new regional identity.

Belize is also home to a large number of researchers and scientists who are drawn to the exceptional opportunities to explore subjects that are rare and pristine. The ancient karst landscape and tropical environment has attracted legions of academics who spend seasons, and often years of their lives, exploring the country's many spectacular mysteries. Some of the more common research interests are the beautiful exotic flora and

fauna, the towering stone temples, or remnants of the Permian extinction that have been forever captured in the local stratigraphy. Less well known in comparison are the ancient Maya wetland agricultural fields that sprawl along the landscape of northern Belize, peculiar, occasionally unassuming, but always radiating the story of the ancient Maya and their expert knowledge of the environment. Drs. Sheryl Luzzadder-Beach and Timothy Beach have spent two decades unraveling the mystery of ancient Maya wetland agriculture in northern Belize and the surrounding areas. They, and many others, have worked to shift the opinion among academics away from the idea that the fields were nothing more than natural formations, or that the ancient Maya were simple slash-and-burn agriculturalists with little understanding of the world around them. These ideas, downplaying the resourcefulness of the Maya, kept the complex identity of these ancient people trapped in a seemingly unmovable mindset. The Maya's message to future generations had previously been misinterpreted at best and at worst it was dangerously silenced. Although many questions still swirl around the drastic population decline around 900 A.D., the story of the ancient Maya holds many valuable lessons for modern day society, regardless that many of the details have yet to be laid out. The work that Drs. Sheryl Luzzadder-Beach and Timothy Beach continue to contribute adds a crucial perspective to the Maya's existence, one that can only be seen through the lens of wetland fields which acute proficiency in environmental adaptation. It is this perspective that returns the rightful esteem back to the ancient Maya by forcing the modern world to acknowledge their true identity.

So much of the Maya identity exists in the ancient wetland agricultural fields, and those located in northern Belize are some of the most expansive throughout Mesoamerica. Researchers have identified, mapped, and studied in a number of field systems along the Rio Hondo and New River. During their use, canals were dug to transport water to the farthest reaches of the fields. In some places the canals can still be seen as a pattern of shallow ditches under dense tropical canopy. In open fields they appear as series of irregularly colored lines of vegetation.

Figure 1: Wetland field canals in an open field near Blue Creek, Belize are identified by dark green vegetation

The canals divide up the land into smaller units, the smallest of which are typically at least 20 by 20 meters in size. Due to their large proportions, and the natural effect they have on the vegetation and topography, they've become excellent candidates for geospatial analysis and remote sensing. This dissertation examines three different approaches using geospatial and remote sensing techniques for analyzing wetland fields, all of which hold promise for future research.

The first approach is an in-depth interpretation of the visual characteristics displayed by wetland fields in high-resolution multispectral and panchromatic satellite imagery. For this, imagery was acquired over most of northern Belize, down to the foothills of the Maya Mountains, and Google Earth was supplemented for areas where cloud-free imagery was not available. Each image was reviewed for signatures of wetland fields. These signatures were determined by first examining areas with previously confirmed fields that had been visited on foot or observed by airplane. The similarities in size, shape, and color were noted and used to discover additional fields throughout the study area. Many of the fields visually identified in imagery were in locations where wetland fields had been noted by researchers in the past, such as Albion Island, located at the midsection of the Rio Hondo, and Pulltrouser Swamp, located at the midsection of the New River. This helped to confirm the methods being used to identify and map the location of additional fields. Eventually, the entire study area was inspected and a database of field locations, containing specific attributes for each location, was developed. This work will help support future analysis of field patterns in space since it

will be a comprehensive database of all visible fields in northern Belize along with important attributes such as acreage, shape, and pattern.

The second geospatial approach is a predictive model, based on the database of identified fields. A number of environmental variables that could have influenced the placement of fields were researched and spatial datasets were acquired from online sources, such as soil or bedrock type, developed specifically for this study area, such as surface temperature and a couple different vegetation-based remote sensing indices. A new sample point data set was created randomly within the boundaries of identified fields and also outside these boundaries. Values for each variable were attributed to each sample point and the final dataset was rigorously evaluated using logistic regression analysis to determine the variables that best correlated with field locations. The regression coefficients were then used to develop a spatial data layer to produce a series of probability maps that can be built upon in future research or utilized to ground-truth highly probable field locations.

Finally, the third approach is an analysis of wetland soils using reflectance spectroscopy, which is the study of electromagnetic radiation interacting with matter. Soil samples were collected in northwest Belize, while a parallel sample was collected in northeast Ohio, and analyzed to determine whether reflectance spectroscopy can be used to detect soil phosphorus and if a difference exists between different locations. Analyzing soil spectra in Belize presented a number of challenges such as using fresh samples that were not previously dried and sieved, and operating the field spectrometer in a semi-outdoor environment with fluctuating light and wind, introduced greater opportunities for

error. Therefore, the samples collected in Ohio served as a control since they were processed in a more stable environment. Understanding phosphorus in archaeological soils has been a focus among researchers for several decades and it is widely accepted that levels of soil phosphorus are often correlated with human activity. The use of reflectance spectroscopy to measure soil phosphorus is less well understood and so while this part of the dissertation was more risky, it is important to keep working to improve these methods. Although the tropical forests of Belize abound with research opportunities for biologists and ecologists, they present a significant challenge in the field of remote sensing. Therefore, reflectance spectroscopy provides a way to directly analyze the spectral properties of archaeological soils to a degree that may never be possible using satellite imagery.

Improvements in imagery resolution and technology in recent decades are creating new opportunities for geospatial research in Maya archaeology. Combined with advancements in our understanding of the Maya world, and the ability to build regional datasets that are digital and shareable, we are able to make discoveries that were not possible just 10 years ago. This is an exciting time for Maya archaeology, and especially for geographers who have made it their home.

CHAPTER TWO – MAPPING THE ANCIENT FIELDS

The Maya landscape is known for its dense tropical forests which impede archaeological investigation as well as remote sensing research. Researchers have struggled over the last century to work around these natural impediments and locate archaeological remains, including wetland agricultural fields. In the last fifty years, advancements in remote sensing technology have led to new methods of studying the Yucatan terrain which has often been difficult to access. In the current decade, Google Earth freely provides high-resolution multispectral imagery throughout most of northern Belize, marking a new milestone in Maya research. This new opportunity will allow archaeologist to improve our knowledge of ancient Maya wetland agriculture at a regional scale.

Introduction

Ancient Maya wetland fields have been located in several Central American countries. This study focuses on the high concentration of wetland fields in northern Belize and immediately surrounding areas (Merlín-Uribe, et al. 2013, Dunning, Beach and Rue 1997). Northern Belize is a flat landscape with three major rivers: Rio Hondo, New River, and the Belize River. In contrast, southern Belize is dominated by the Maya Mountains and lacks the large, slower moving rivers of the north. Consequently, the areas surrounding northern Belize have significantly different environments and opportunities for human occupation. This study focused on a small part of the Maya landscape and will help to identify local environmental conditions that were preferred for wetland farming.

Modern Farming Practices

Today, about 265,000 acres in Belize are dedicated to farming, of that 146,000 is cropland. Northern Belize agriculture consists mainly of small farms that produce crops such as sugar cane, and citrus. Western Belize produces corn, beans, beef, poultry, pork, and dairy. Milpa farming is an older practice that is also found in Northern Belize but is more common in southern Belize. The word "milpa" comes from a Nahuatl phrase which can be translated as "to the field." This type of farming consists of a cycle beginning with two years of cultivation followed by eight years of allowing the land to lie fallow and recuperate. This system is designed to produce high yields without the use of fertilizer. Although it can be a sustainable practice, milpa farming has the potential to become unstable at higher levels of intensity. Over time it has become less common due to the high poverty level of the mainly indigenous people who still follow this practice.

The modern day inhabitants of Belize have modified the landscape, sometimes destroying ancient field patterns. Once ancient fields are plowed, it is often difficult to distinguish ancient field patterns from naturally forming drainage patterns. Figures 3 and 4 depict this modern alteration of the landscape. In the first image taken in 2002, the area is forested and has no major roads. The second image is from 2010 and modern agriculture has taken over. In one section of the modern field, ancient field patterns can be seen which are very consistent compared to the naturally occurring drainage pattern.

Figure 2: Forested land in study area; image acquired in 2002.

Figure 3: Same area shown in Figure 3. Note that modern agriculture exposes and degrades ancient field patterns. Image acquired in 2010.

History of Research on the Ancient Maya Wetland Agricultural Fields

Ancient Maya wetland agricultural fields were first discovered by Alfred Siemens in 1968 along the Candelaria River in southern Mexico, but were not initially accepted by the archaeological community who suggested the canals were not for irrigation but served as fish conservancies or waterways for the movement of timber (M. Pohl 1990, Siemens and Puleston 1972). It was later proposed that the fields were more likely the result of natural gilgai soil formation (Puleston 1978). Gilgai are small depressions of a few meters in size that form in expanding clay soils (Schaetzl and Anderson 2005). They are often found in vertisols, a soil type that is prevalent in northern Belize (Garrity and Soller 2009). Figures 5, 6, and 7 are examples of a gilgai formation in Texas, an ancient Maya wetland field with distinct lines, and a wetland field with less distinct lines, respectively.

Figure 4: Gilgai formation in Texas (McDaniel 2013).

Figure 5: Distinct ancient Maya wetland fields located on the Mexican side of the Rio Hondo.

Figure 6: Ancient Maya wetland fields with amorphous shape located on the Mexican side of the Rio Hondo.

Alfred Siemens and Dennis Puleston led excavations and aerial surveys to map wetland fields and determine their extent in the Yucatan peninsula (Siemens and Puleston 1972). Their work would eventually be used to understand Maya settlement patterns in the region as well. In addition, a group of researchers joined Siemens and Puleston in a multidisciplinary approach to understanding Maya wetland agriculture. The collaborative group was known as the Rio Hondo Project and included geologists, paleo-environmentalists, botanists, and soil scientists. Since that time, research has continued on this subject by a number of workers (M. Pohl 1990).

Still, skepticism about the origins of the wetland fields persisted. In the late 1970's (Harrison and Turner 1978) again attempted to answer the question of whether the fields were natural or anthropogenic. They conducted aerial explorations of fields at a site called Pulltrouser Swamp in northern Belize. This work supported the original theory that the canals had been constructed by humans and their findings were reinforced through ground surveys. Additional archaeological excavations of several field locations eventually produced similar conclusions (M. Pohl 1990).

After the wide-spread acceptance of the fields' agricultural purpose, researchers began debating on the population size the fields could have supported and their overall significance in Maya subsistence. One argument stated they provided minor support to Maya populations since fields appeared to be confined to certain areas and that swidden cultivation of the uplands was of far greater importance in the region (Sanders 1979, Pope and Dahlin 1989). Swidden, also known as "slash-and-burn", refers to the method

of shifting between cultivation and fallow in order to allow the soil to regenerate (Harrison and Turner 1978). In later years, others have portrayed wetland agriculture as one of many adaptations in a patchwork of methods throughout the multifaceted Maya territory (N. Dunning, T. Beach and P. Farrell, et al. 1998). More recently, researchers have become confident that wetland agriculture probably provided adequate support to dense populations (Renard, et al. 2012)

Physical Characteristics of Mesoamerican Wetland Agricultural Fields

Ancient Maya wetland agriculture fields in Mesoamerica have been referred to as raised fields, ditched fields, chinampas, etc. However, the main function is always the same: manipulation of an otherwise limited landscape by regulating soil moisture and maintaining a steady supply of nutrients to crops. These systems are dependent on the water table and allowed the ancient farmers to extend the range of time during which food could be produced. Farmers could also transform inundated land into a productive field by adding soil to the growing surface and raising it above the surrounding area, allowing ancient Mesoamericans to intensify their food production (Sluyter 1994).

Wetland fields exhibit a variety of shapes, patterns, and a direct relationship between canal depth and the number of canals per field (Figure 13). Geography and cultural influence led to the distinction of five regional methods of wetland agriculture that spanned three cultural time periods. The regions are the Mesa Central, the Southern Highlands, the Gulf Coast, the Maya Lowlands, and the Maya Highlands. These regions practiced wetland agriculture during the Preclassic (2500 BC to AD 1), the Classic (AD 1 to 900) and the Postclassic (AD 900 to 1521), with each period of time coinciding with

environmental changes that likely influenced development of wetland agriculture. Over time, fields increased in overall length, width, frequency, and amplitude of ditching (Sluyter 1994).

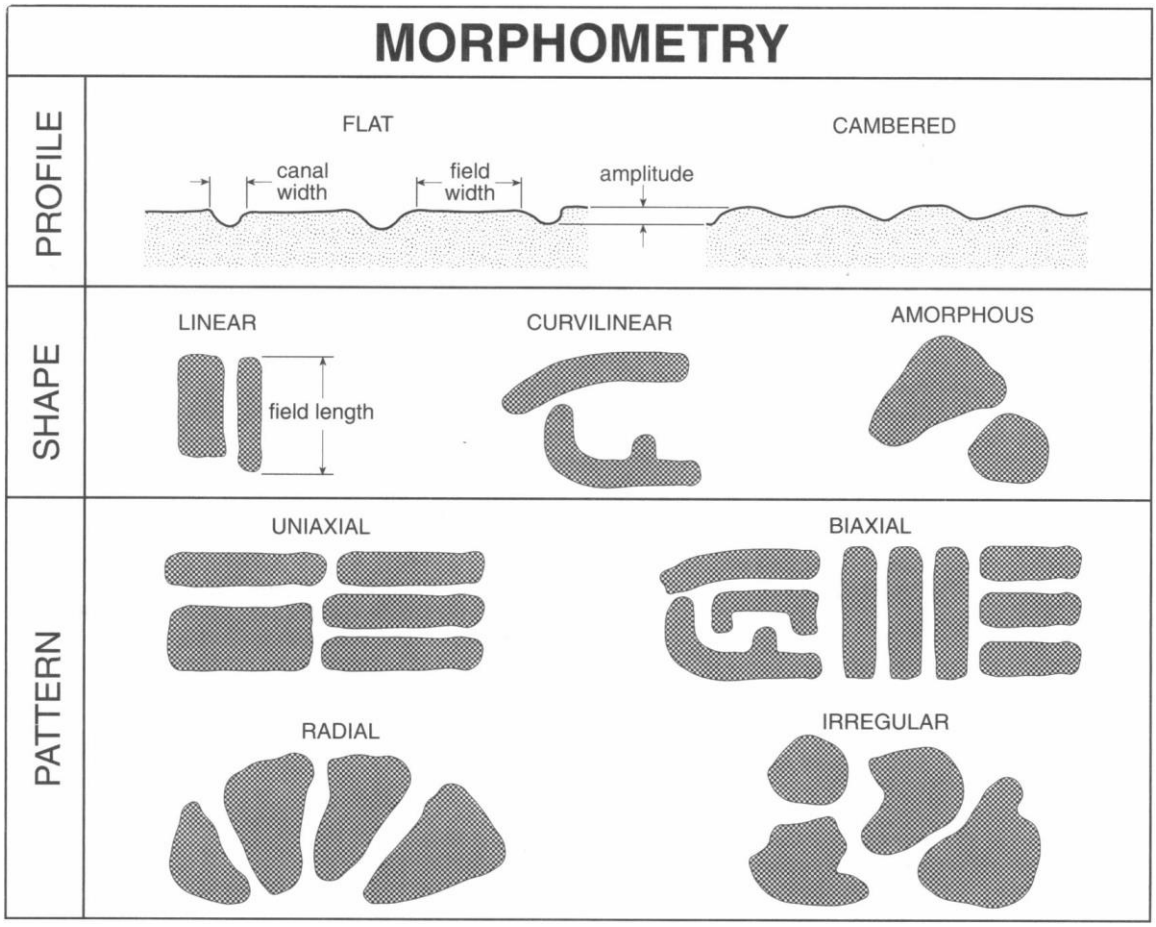


Figure 7: Known field shapes and patterns (Sluyter 1994).

The climate and topography in which wetland agricultural fields exist varies greatly in Mesoamerica. Wetland agriculture increased during the Preclassic period, which began with a cool and wet climate. By the late Preclassic, the climate became

drier. This may have put greater reliance on the wetlands which in turn would have provided increased security of food production. This shift in reliance may have grown during the Classic Period due to initial success of wetland fields during the Preclassic (Sluyter 1994).

The fields found in the Maya Lowlands exhibit several patterns and shapes as described in previous research (Sluyter 1994). Investigation of multispectral satellite imagery and imagery available in Google Earth show three common wetland field patterns in this region: linear, amorphous, and the larger canals that feed the smaller systems (Figures 14 and 15).

Figure 8: Linear-shaped fields

Figure 9: Amorphous-shaped fields.

In several places, large canals can be seen that extend from rivers and lead to smaller canals nearby. In figure 16, the small canals that make up wetland fields are roughly 700m from the river. This may indicate that in the past, the wetland fields could have been more extensive and filled the area between the river and existing fields.

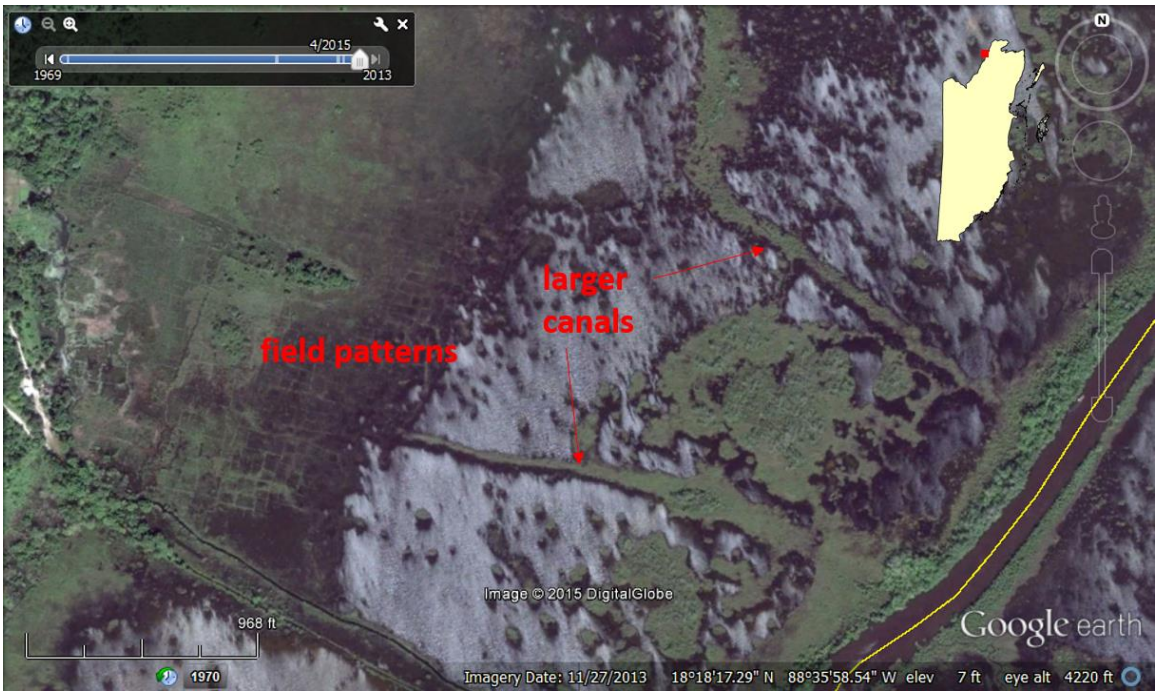


Figure 10: Large canals feed river water to smaller systems which may no longer be visible.

Northern Belize is one of the few locations in Mesoamerica where intensive wetland agriculture occurred during all three cultural time periods. Coincidentally, it is also one of three locations civilization began in Mesoamerica. In northern Belize, the terrain consists of low elevation, karst topography, and a hot and humid climate, where fields are often found in bajos and back swamps of streams. According to one estimate, the Maya Lowlands region may have 22,900 hectares of wetland fields, the most of any region in Mesoamerica, and could have supported up to 255,000 people a year. Even the best estimates of field acreage may be skewed since many fields were buried long ago by aggradation. Since population estimates are also unclear, it becomes increasingly difficult to estimate the importance of wetland agriculture in ancient Maya culture (Sluyter 1994).

Remote Sensing in Maya Research

Since the 1960s, remote sensing has advanced our knowledge of ancient Maya wetland agriculture by allowing archaeologists to locate fields and to conduct research in spite of the Yucatan Peninsula's dense tropical vegetation. A timeline of Maya remote sensing research, in relation to advances in commercial earth imaging capabilities, can be seen in the table below (Bauman 2009, GISGeography n.d., Lavers 2015, Satellite Imaging Corporation 2014a, National Aeronautics and Space Administration 2012,

Jeffrey 2013, Satellite Imaging Corporation 2014b, Satellite Imaging Corporation 2014c, Young 2009)

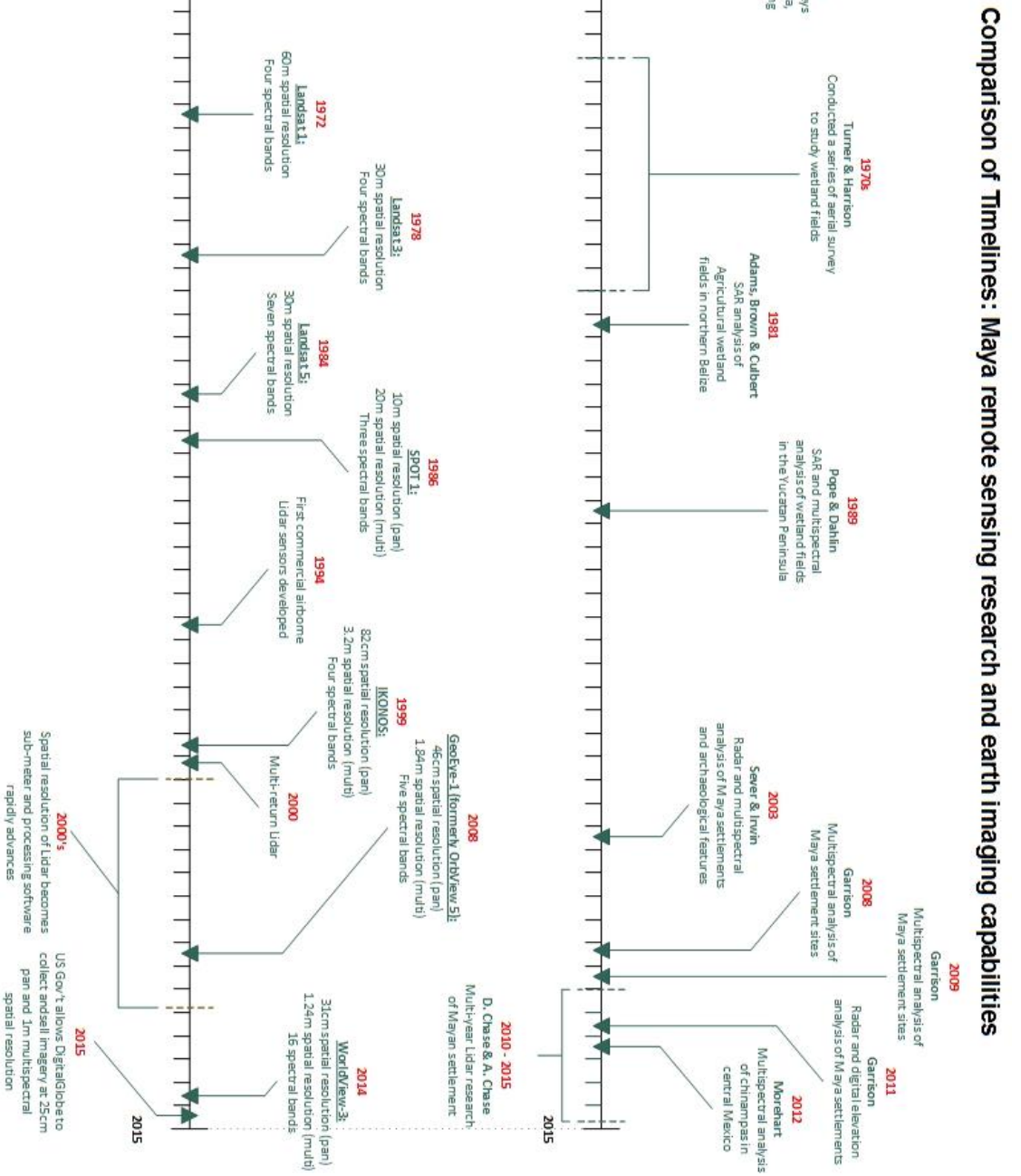


Figure 11: Timeline of research and imaging capabilities

The first attempt at what can be considered remote sensing research in Maya archaeology began in 1929. Oliver Ricketson and Alfred Kidder carried out the first aerial investigation of the Maya region, along with Charles A. Lindbergh, by flying over the countries of Mexico, Guatemala, and Belize (Figure 9). The expedition led to the discovery of Maya towns and temples, and their findings were documented in an articles published in *The Scientific Monthly* and the *Geographic Review* in 1930. (Kidder 1930, Ricketson and Kidder 1930).

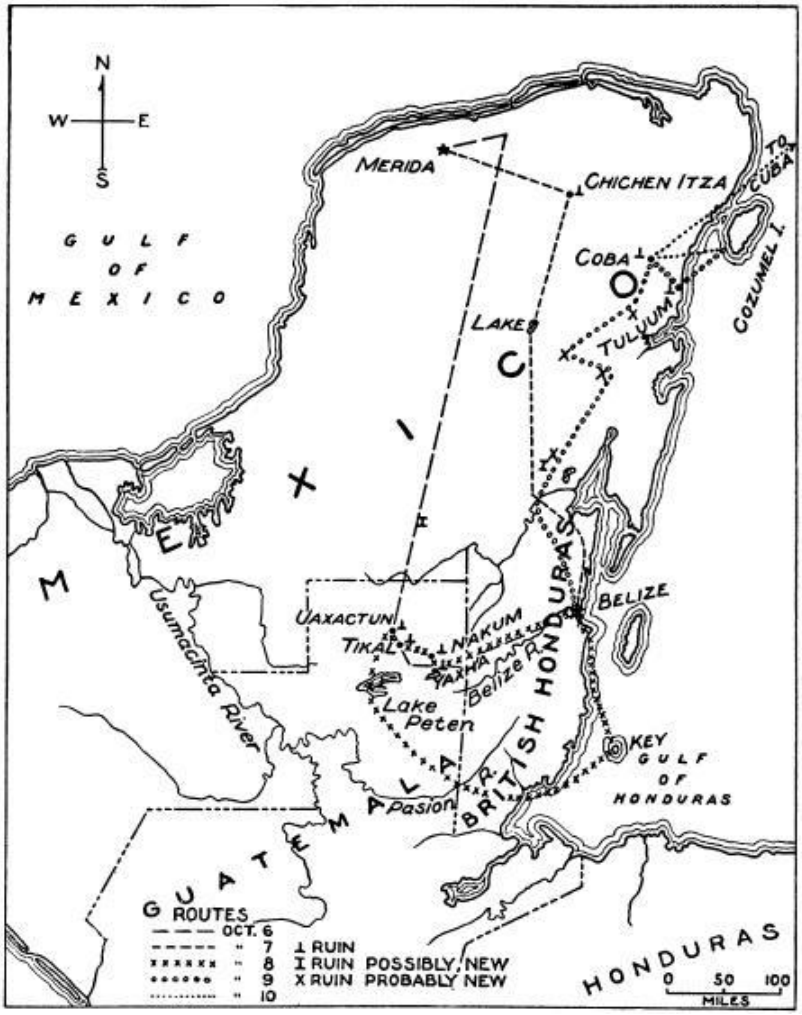


Figure 12: Lindbergh and Kidder's flight route and site locations (Kidder, 1930).

The pair traveled from Belize City to northern Guatemala, photographing sites and physical terrain features (Figure 10) and then flew north to southern Mexico, discovering new sites along the way. After returning to their base in Belize City, they discussed their initial findings and planned a second trip to focus on specific terrain features that led them to believe there were Maya sites below the forest canopy. The

results of their work highlighted the utility of aerial investigations in the exploration of unknown geographic locations and archaeological landscapes (Kidder 1930).

Figure 13: Aerial image taken by Lindbergh and Kidder of a temple at Tikal, Guatemala.

Several archaeologists have used radar imagery to study Maya sites and wetland agricultural fields with varying degrees of success. The first of these attempts was led by Richard E. W. Adams in 1981.The goal of this research was to understand the discrepancy between the seemingly large ancient Maya population and the questionably low potential food sources in the southern lowland region of the Yucatan Peninsula.

According to Adams et al., synthetic aperture radar (SAR) surveys performed by NASA revealed a more extensive wetland agricultural system than previously known. Ground-truth of several areas also verified the SAR findings. Previously, archaeologists had thought the Maya relied mainly on swidden agriculture. However, shortly after the discovery of the wetland fields, evidence began growing that suggested this method was unlikely to have been able to support large ancient populations. A more probable situation was that the Maya took part in extensive wetland agricultural farming (Adams, Brown Jr. and Culbert 1981).

Since aerial surveys alone were considered a somewhat dubious and subjective method for locating the agricultural fields, it was thought that radar imagery would provide a more objective and acceptable method. Unfortunately, this study created new questions about the validity of the identified fields. It is uncertain if all of the detected features are truly human-produced or are actually the result of non-Maya features such as roads, trails, or abandoned air strips. There were also questions about whether the imagery was able to detect some of the smaller wetland fields or if they were missed due to the large spatial resolution (30m) of the SAR imagery (Figure 11). The results of Adams (1981) study were critiqued Kevin Pope and Bruce Dahlin (1989). These researchers contested that the previous estimates of the extent of wetland agricultural fields in the Maya lowlands were exaggerated. They suggested that the fields are only concentrated in three areas: northern Belize, southern Quintana Roo in Mexico, and the upper Candelaria River in Campeche, Mexico. They based their arguments on their

analysis of the same imagery used by Adams et al. and found that only the largest canals could be positively identified with the radar imagery (Pope and Dahlin 1989).

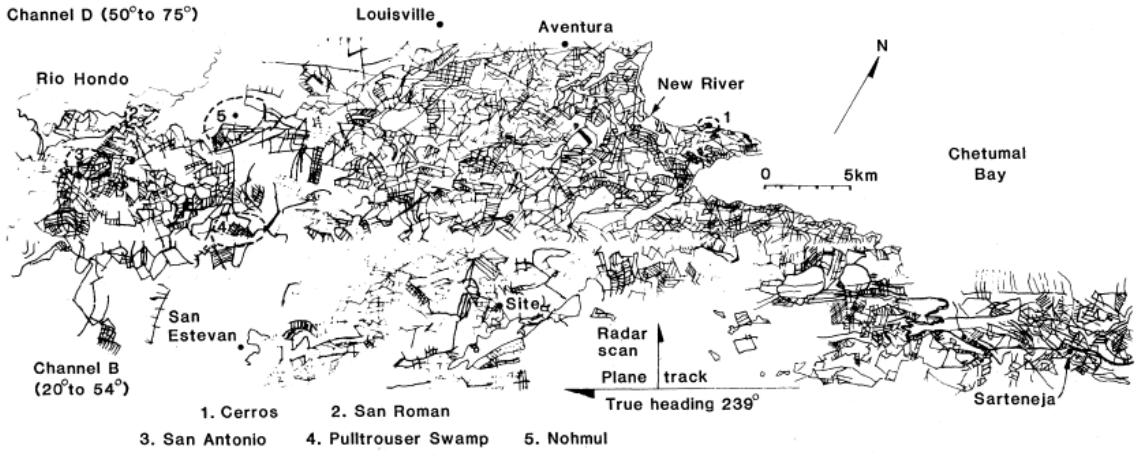


Figure 14: Fields detected by Adams, Brown Jr., and Culbert (1981).

In their own analysis, Pope and Dahlin used multispectral satellite imagery to detect the locations of agricultural fields. They were attempting to locate canal features that had been detected on the ground nearly two decades earlier (Siemens and Puleston 1972), and aimed to establish a quantifiable method for determining their existence on a regional scale. This work was further supported through the analysis of soils, hydrology, and vegetation studies. They used Landsat Thematic Mapper (TM) and radar imagery to locate wetlands and ancient canals across the Yucatan Peninsula. False color composites of the Landsat TM made it possible to detect some of the larger canal features. Small lattice patterns were not detectable by the available imagery and were only located while flying overhead in an airplane. Canals were detected along the karstic Rio Hondo river

but were absent in the non-karstic rivers to the south which are faster-moving and have greater fluctuation in their water levels. Also, the back-swamps of karstic rivers are perennially flooded whereas the backwaters of non-karstic rivers are only seasonally flooded, allowing them to desiccate and crack during the dry season. The findings of Pope and Dahlin definitively showed the large canals also crossed over different land types, connecting the river channel with karst depressions that were located farther away. They were also able to determine that the locations of fields did not coincide with areas of high population and so population density was not a likely influence on the selection of agricultural field locations. Additionally, they were able to compare regional differences in vegetation among the different types of wetlands in a more quantifiable way (Pope and Dahlin 1989).

In 2003, Tom Sever and Daniel Irwin of the NASA Marshall Space Flight Center in Huntsville, Alabama, used multispectral imagery to map Maya archaeological features and natural terrain features in neighboring northern Guatemala to determine whether bajos had been used by the ancient Maya for wetland farming. They began by mapping linear features, vegetation, and drainage patterns using Landsat TM and Landsat Enhanced Thematic Mapper Plus (ETM+) imagery, and correlated their findings with locations of known settlements. They were able to identify canals and also used the Normalized Difference Vegetation Index (NDVI) to highlight areas of stressed vegetation that may have been manipulated for agriculture or habitation by the ancient Maya. High-resolution Ikonos multispectral imagery was then used to locate small-scale cultural features, and to explore the surrounding terrain. Finally, STAR-3i radar imagery was also

used to locate bajo islands (see Figure 12) which are elevated areas within a bajo that are not seasonally inundated. Their integrative analytical approach also allowed them to map different types of vegetation of which ancient Maya had been aware of and had adapted for their use. Overall, the combination of multiple imagery datasets allowed for a more thorough understanding of the study area (Sever and Irwin, 2003).

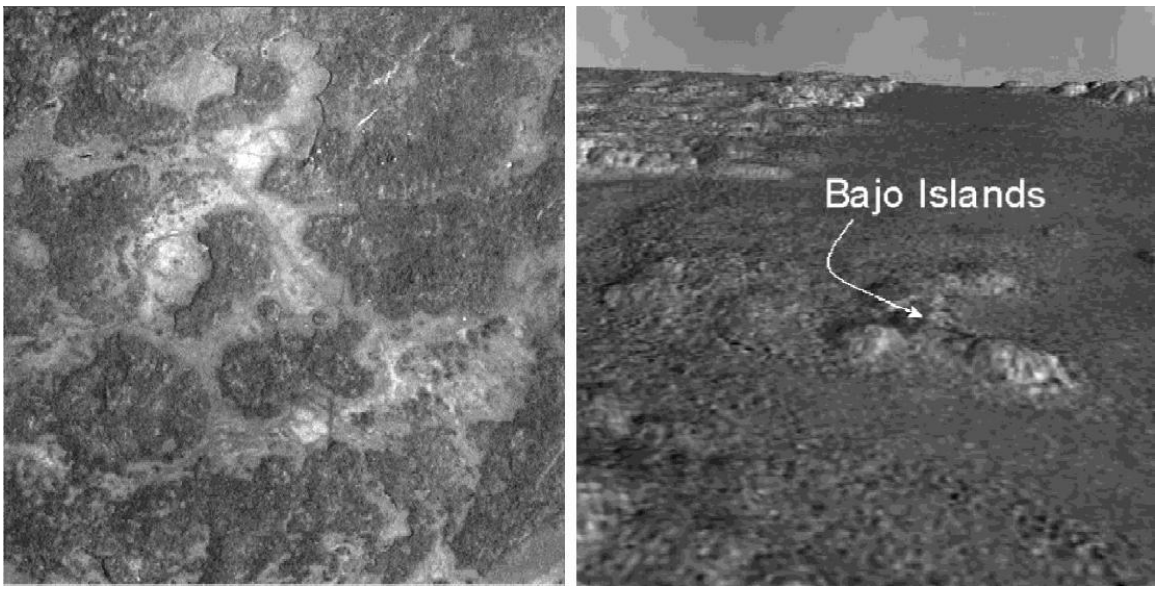


Figure 15: Canals detected with Landsat TM, bajo islands detected with STAR-3i radar (Sever and Irwin, 2003).

Garrison (2010) used ground survey data from Sever and Irwin's 2003 research to create a land classification map to aid future research in the region. The goal was to produce results similar to the ground survey in order to reduce the amount of future field work necessary for this type of research using high-resolution imagery. Previous ground surveys were done by Sever and Irwin as part of their 2003 study to map Maya settlement

sites and related features in the region. For part of their work they used false color composites of high resolution Ikonos imagery. They assessed that this type of imagery was only useful for discovery and so they relied mainly on 30m spatial resolution Landsat imagery for much of their analysis. However, Garrison attempted to find other uses for high-resolution imagery that might be more successful and also focus and reduce the amount of field work necessary in the future. An unsupervised classification was conducted on the Ikonos data but due to an unknown error in the imagery, it produced erroneous and unusable results and Quickbird multispectral high-resolution imagery was used instead. An unsupervised classification of this data, which incorporated an NDVI as a fifth band, produced a map of microenvironments in the region that are known to be associated with Maya sites. The result was an overall accuracy of 73% with one category reaching nearly 100% accuracy. Since most of the known archaeological features occur within the two most accurately classified land types, this method was deemed successful. (T. G. Garrison 2010).

More recently, (Garrison et al., 2011 took part in a radar survey of Maya sites throughout Guatemala. They used AIRSAR imagery, collected by a NASA airborne sensor which surveyed the Maya territory in 2004, to make a 3-dimensional model of the landscape in order to investigate the area in search of new sites. The intent of NASA's 2004 mission was to build a SAR database over the Maya expanse so that researchers could develop procedures for studying Maya archaeology and for future exploration of the region (Garrison, et al. 2011).

To build the elevation model, the AIRSAR imagery was first processed by performing an image coregistration from two separate radar images to produce and interferogram from which topography can be estimated. The AIRSAR-derived digital elevation model was compared to Shuttle Radar Topography Mission (SRTM) imagery and was found to have an overall difference of half a meter, verifying its accuracy (Garrison, et al. 2011). Using a GIS, the team applied a color ramp and performed a stretch of two standard deviations which allowed them to quickly locate major temples at El Zotz and El Palmar and two new sites: Hilltop Site and La Avispa. At El Zotz, they also found expansion sites consisting of basal platforms that held residential mounds, the likely locations of non-elite settlements (Garrison, et al. 2011).

Over the last five years, bare-earth LiDAR imagery has been used to analyze Maya sites in fine detail (Chase, Chase and Weishampel 2010, Chase, Chase and Weishampel, et al. 2011, Chase, Chase and Fisher, et al. 2012, Chase, Chase and Weishampel 2013, A. F. Chase, D. Z. Chase and J. J. Awe, et al. 2014, A. F. Chase, D. Z. Chase and J. Awe, et al. 2014). Since vegetation spectra do not always indicate what lies beneath it, places that are densely forested require the use high-resolution elevation data, such as LiDAR, to see features and structures that are below the surface of the canopy. LiDAR can be used to produce elevation data but provides a much higher level of information than a typical DEM. It allows researchers to find the faint outlines of features that are neither visible through multispectral imagery nor field surveys. Chase and Chase have used this technology extensively since 2010 to map features of Maya settlements, revealing structures, terracing, and other features obscured by vegetation. LiDAR has

proven to be extremely powerful in providing highly detailed maps of ancient Maya cities (Chase, Chase and Weishampel 2013, Chase, Chase and Weishampel 2010, Chase, Chase and Weishampel, et al. 2011, Chase, Chase and Fisher, et al. 2012).

In a 2012 study related to the topic of wetland agricultural fields, Morehart (2012) performed a remote analysis of chinampas, which are a version of wetland agricultural fields that were created by the neighboring Aztec civilization. The chinampas in this study are located near the ancient Aztec city of Xaltocan in southern Mexico. Chinampas are very similar to the fields in northern Belize, both consist of a system of large canals with smaller canals, or ditches, that branch off, and were discovered around the same time as the wetland fields of Belize. Although the people that built chinampas were of a separate cultural group, they likely shared knowledge of agricultural practices. Morehart (2012) used high resolution Quick Bird imagery, aerial photos from the 1950s and also Landsat imagery to detect the locations of canals. Their analysis consisted of an initial visual inspection of the imagery, creation of false color composites, and also the creation of an NDVI image. The latter two methods did not appear to enhance the ability to detect canals much more than the original true color images or aerial photos. Their final product was a map of a small section of chinampas. Although this research was at a more local level, it provides a comparison for similar remote sensing research that has been carried out for Maya wetland agricultural fields at a larger scale.

A Case for Visual Imagery Analysis and Field Detection

Maya researchers have ultimately employed nearly every major type of imagery available in order to detect sites and field systems: aerial photography, SAR,

multispectral, and LiDAR. Each attempt identified strengths and weaknesses of the imagery and methods involved and from those studies there have been several lessons learned. First, imagery with a low spatial resolution will be unlikely to reveal the location of canals except for the largest ones. Manipulating the data to create an NDVI does not significantly help to identify canal features since the fields span a variety of vegetation types and microenvironments resulting in varying success. Although comprehensive methods to identify wetland fields throughout this region remain elusive, visual imagery interpretation is now possible through high-resolution panchromatic and multispectral satellite imagery. Freely available through Google Earth, this resource has become a powerful tool in archaeological research and has been used in a number of recent studies around the world (Kennedy and Bishop 2011, Sadr and Rodier 2012, Thakuria, et al. 2013). Due to their large size, the fields are an ideal candidate for archaeological prospection using Google Earth.

In the last five years, Google has made high-resolution multispectral imagery available for much of northern Belize through their Google Earth product. Although some areas have yet to be updated, large field systems are easily discoverable while visually scanning the imagery at large scale, typically at least 1:5,000. While this may seem time consuming when compared to automated methods, it is extremely effective. By using verified ancient Maya wetland fields as a reference, it is possible to establish a set of parameters for identifying new fields and distinguishing them from naturally occurring phenomena.

Google Earth imagery has a spatial resolution that ranges from 15m to sub-meter, depending on the public interest of a given area (Google 2015, Mellen 2014). Imagery is continually updated as well as digital elevation data, provided by NASA's Shuttle Radar Topography Mission (SRTM) imagery (Rusli, Majid and Md Din 2014). The program allows users to add data and upload hand-held photos which further aid geographical analysis. Google Earth's historical imagery makes it possible to find features that may no longer exist or to understand change over time (Google 2015).

Methods

Imagery Analysis of Wetland Field Patterns

In 2013, the Maya Research Program (MRP) conducted an aerial survey along the Rio Hondo and identified 15 areas that contained visible field patterns. These locations were later visited on foot and verified as actual ancient fields. In order to identify additional wetland fields, the patterns and shapes of verified fields were reviewed in imagery.

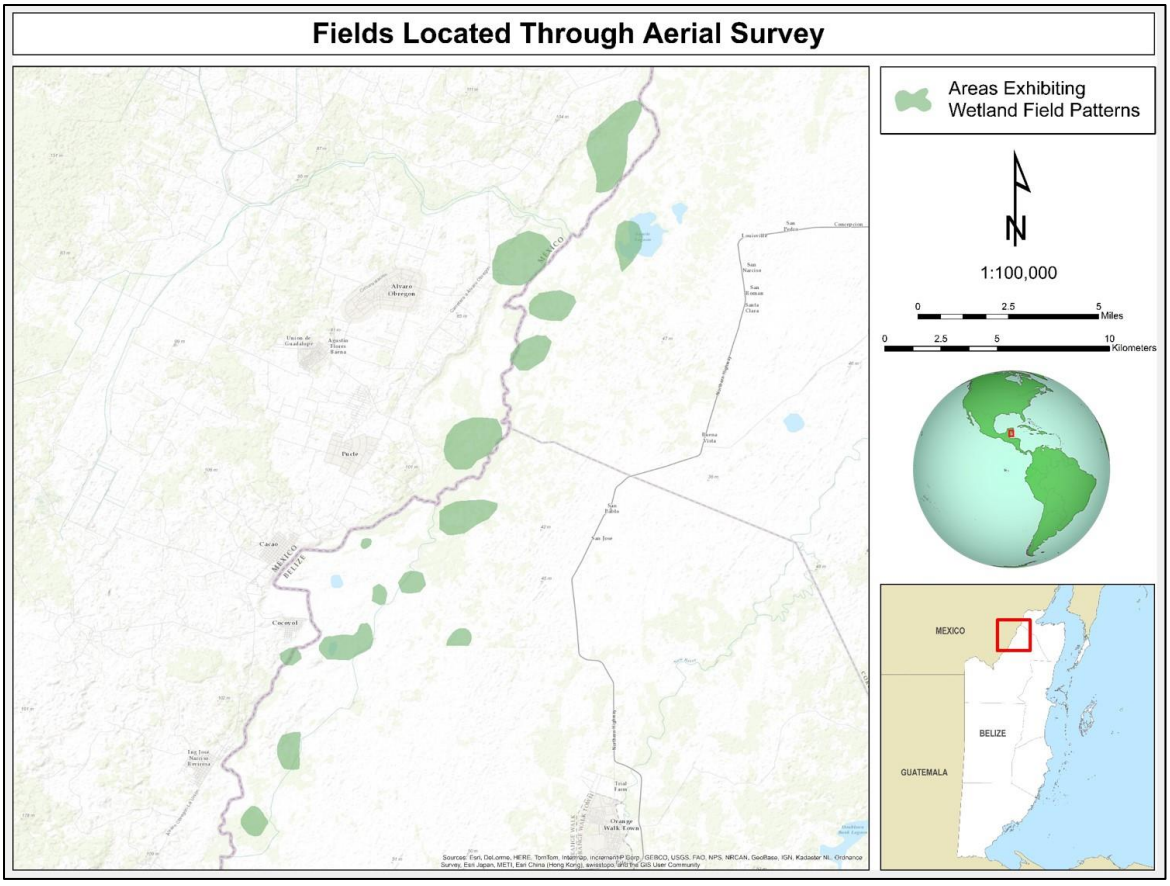


Figure 16: Fields located by air. Map created by Tanya Catignani.

New fields were first located using Google Earth and additional multispectral images were acquired from various sources (Ikonos multispectral provided by the Maya Research Program, and WorldView, OrbView, and QuickBird multispectral provided by The National Geospatial-Intelligence Agency) to further determine if they were natural or man-made. The footprints of these additional satellite images can be seen in Figure 18, below.

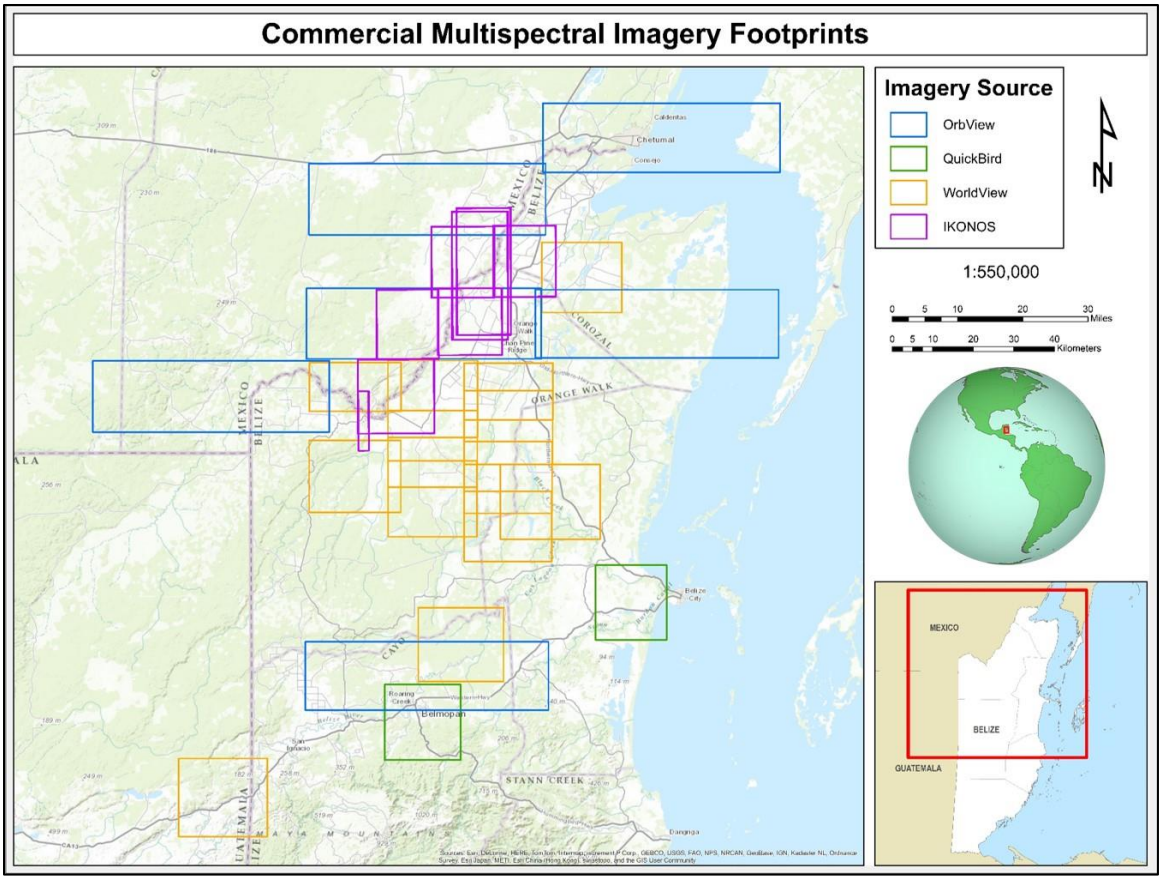


Figure 17: Commercial Multispectral Imagery Footprints. Map created by Tanya Catignani.

The fields mostly follow large rivers and are typically within one to two kilometers from their banks. Some still contain defined linear patterns that are consistent for hundreds of acres but often only small sections are visible. They are mostly found in low-lying flat areas which were irrigated by larger canals. Modern farms can often be found adjacent to ancient ones and in several instances, ancient field patterns can be seen surrounded by modern fields. In Figure 19, a road cuts through an ancient field but was avoided by modern farmers who ploughed around the ancient wetland field. This suggests at least some of the land which was valued by the ancient Maya continues to be

valued by modern farmers, probably for similar reasons including nutrient-rich soils and stable access to water.

Figure 18: Modern road cutting through ancient field.

Based on the commonalities found among known field locations, the following criteria were selected and used to locate additional ancient Maya wetland fields:

- defined trenches
- rectangular framework
- continuous (multiple trenches observed)
- size of smallest individual units ranging from ~250 m^2 to ~1250 m^2

Defined trenches are the obvious starting point for identifying field patterns but unless several of them are visible, it is hard to determine if a trench is truly part of an ancient field. Typically, fields were identified by observing a rectangular framework of trenches that continued for at least two or more defined field units. In ArcMap, a geographic information system (GIS) software created by Environmental Systems Research Institute (ESRI), the known fields with the most defined trench lines were defined so that the length and width of field units could be accurately measured. In total, 70 individual field units were identified, each of which was examined in imagery to determine their size range. Using the historic imagery in Google Earth, it was possible to look at the units with varying degrees of vegetation growth and determine if they consisted of multiple units. An example of this phenomenon can be seen in Figures 20 and 21.

Figure 19: This field unit appears to be singular during periods of increased vegetation.

Figure 20: Reviewing historical imagery shows the unit is actually three field units.

The length of individual units was measured to determine a range which was found to be ~250m^2 to ~1250m^2. In figure 22, both sides of this range can been seen at one location along the Rio Hondo.

Figure 21: Clear example of the smallest and largest field units typically found in northern Belize.

Results

Imagery analysis identified 104 locations and a total of 16,016 acres of visible field patterns throughout northern Belize and surrounding areas. Figure 23 shows the locations of wetland agriculture fields that have been located using visual interpretation of Google Earth and commercial multispectral imagery.

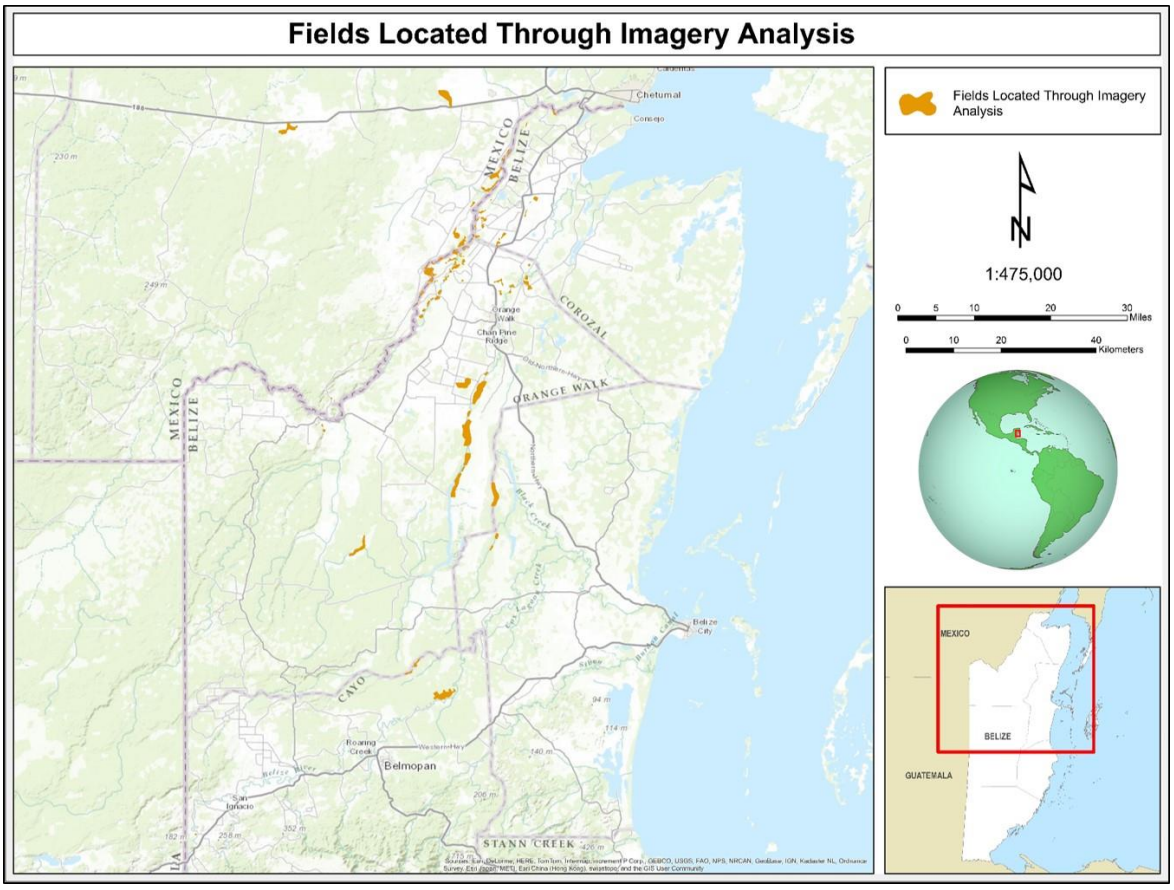


Figure 22: Fields located through imagery analysis. Map created by Tanya Catignani.

Figure 24 contains the locations of previous wetland field research areas which closely align with the imagery-identified field locations. This correlation supports the methods used here and may even have identified additional fields in the southern part of the study area that may be new discoveries.

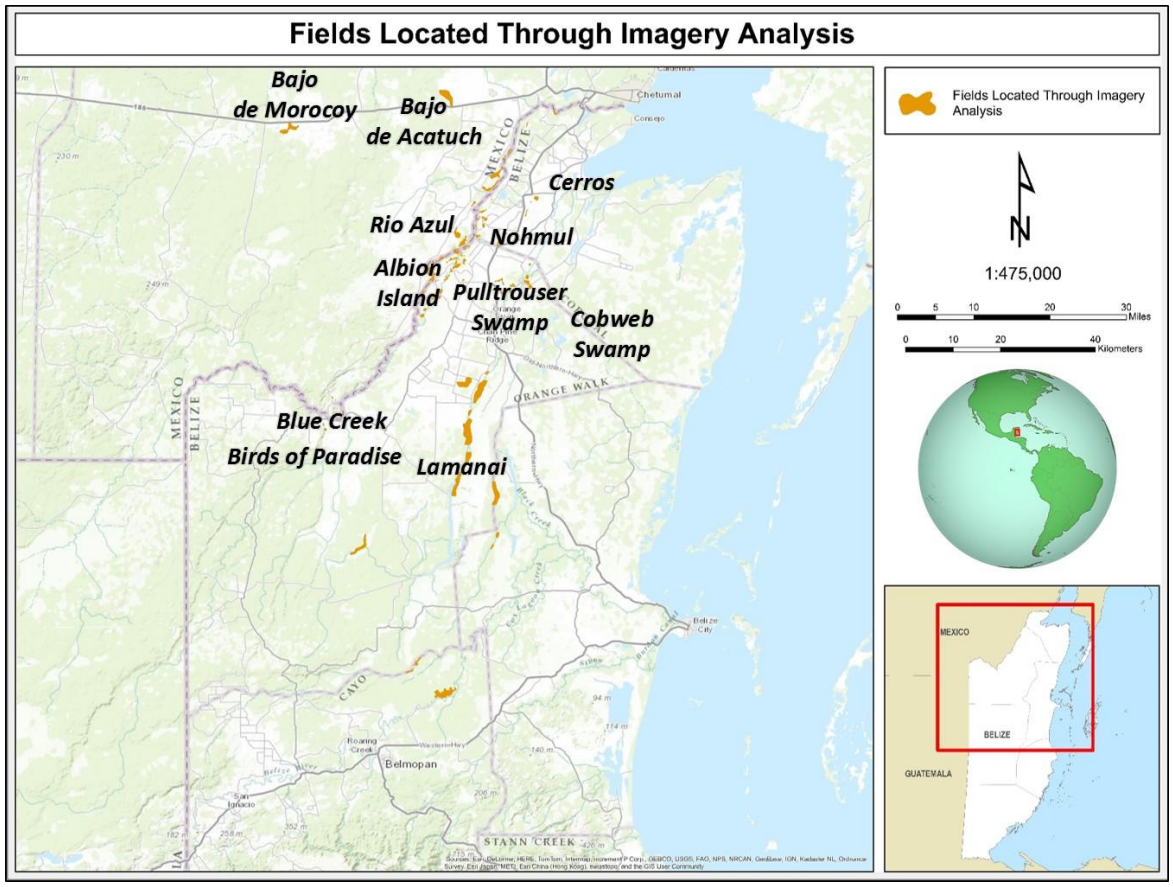
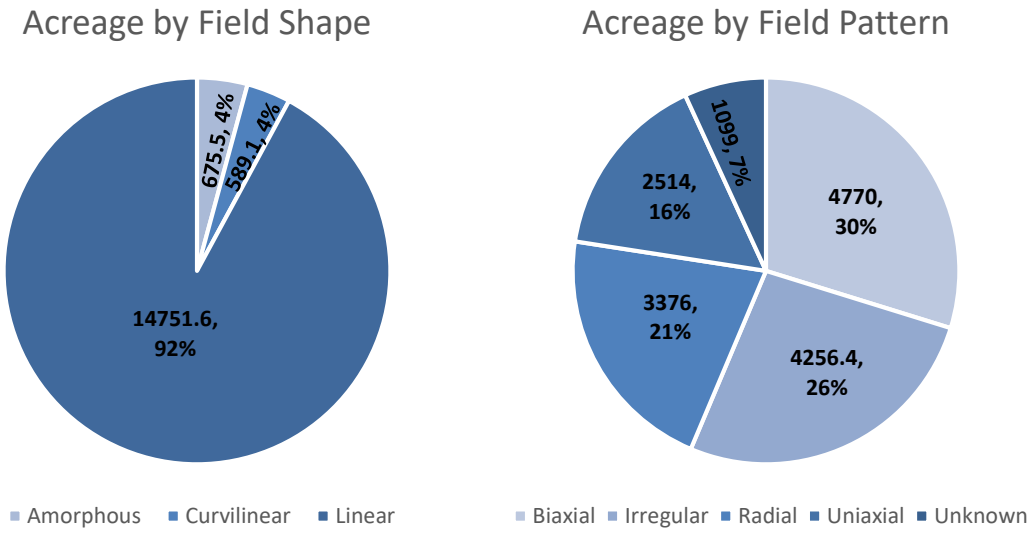


Figure 23: Identified fields and previous research sites. Map created by Tanya Catignani.

Based on Sluyter's classifications, the large majority of fields discovered through imagery analysis appear to be linear in shape but exhibit a variety of patterns: biaxial, irregular, radial, and uniaxial (Sluyter 1994).

Table 1: Breakdown of visually identified fields by shape and pattern.



The compilation of ancient Maya wetland field locations and imagery-based attribute data has provided a broader understanding of this agricultural practice and will act as foundational data for further regional analysis. Regional-scale analysis will also help to answer overarching questions about how the Maya learned to masterfully manipulate their landscape in support of large populations

Conclusions

Google Earth, high-resolution multispectral and panchromatic imagery, SAR and Lidar have provided new sources of information and avenues for ancient Maya wetland agriculture research. The results of analysis show that fields can be visually identified in a reasonable amount of time and new wetland field locations can continue to be discovered as spatial resolution improves. Belize's landscape is steadily changing due to modern farming practices which may hinder future research. By studying high-resolution

multispectral imagery it is possible to witness these changes as they occur and to preserve our understanding of ancient Maya agricultural practices while the physical evidence still exists.

CHAPTER THREE – PREDICTIVE ANALYSIS OF FIELD LOCATIONS

Introduction

The Maya culture thrived for millennia despite a lack of water and food sources, and learned to exploit the few opportunities that existed in their region. For example, the karst topography of the Yucatan peninsula provided them with flat landscape and shallow soils that could be manipulated through irrigation canals. It is clear from the analysis in the previous chapter that proximity to a water source was a highly influential factor for the Maya when planning the location of these fields, but there were likely additional environmental factors that influenced their placement. This article will attempt to identify the main environmental variables that were used to select areas for wetland agriculture, and use them to estimate the spatial extent of these fields.

Predictive Modeling in Archaeology

In archaeology, predictive analysis can be used as a statistical tool to develop an understanding of the behavior and choices made by humans in the past. As can be expected, it has inherent limitations that should be acknowledged and adhered to but this does not mean it should be avoided entirely. Unfortunately, this practice has often come under scrutiny by those who feel predictive analysis applied to archaeology is misleading at best, and at worst may be fundamentally wrong. This discrepancy is sometimes due to the perception of the results which should be used to identify areas of potential archaeological significance, and not a literal map of undiscovered sites. (Carleton, 2012) Regardless, the statistical analysis of ancient sites and landscape modifications can

undoubtedly reveal patterns among the practices of past cultures and begin to explain their ways of life.

Ernestene L. Green (1973) created a statistical model of Maya sites using Stepwise Multiple regression and a number of environmental factors including soil type and distanced to water. This type of regression attempts to create a best-fit line through the data, rotating variables in and out of the equation, to find the best combination of variables which can predict additional datasets. Her sample size was unfortunately small (22 known sites) but she did find significant relationships between the location of Mayan sites and both of these variables. She also noted that their appeared to be a preference for lime-enriched soils which offer high agricultural potential. Her results were not very exceptional, and only 22% of the variation within the sample could be explained by the model. However, it did provide a foundation for future types of analysis in this region. (Green, 1973)

Fedick (1996) created a predictive spatial model of terrace agriculture in the upper Belize River area of west central Belize. The model incorporated variables such as soil type, parent material, topography, land capability, and seasonal capability. Each variable was categorized, ranked, and analyzed in a geographic information system (GIS). The final model spatially depicted the locations most likely to contain evidence of ancient Maya terrace agriculture and was later tested through ground surveys. The survey results were incorporated back into the model in order to refine and improve results. In general, the model showed that areas with low slopes and specific soil types that were conducive to hand-cultivation were the most likely to contain evidence of ancient terracing.

More recently, Sallie Vaughn and Tom Crawford created another statistical model, in northwest Belize, using similar variables in a GIS. They chose to use Binary Logistic Regression since their data included a range of types including nominal, interval and ratio. This type of regression aims to calculate a probability whether data points will fall into one of two predetermined groups. The study included 50 sites in their sample as well as 50 randomly selected points. Their analysis found that east-facing slopes, and greenness (a measure calculated using a Normalized Difference Vegetation Index) were correlated with the presence of known sites. Distance to water was not found to be significant in this study. The resulting model was able to correctly predict the location of 66% of known sites and 60% of random points (Vaughn & Crawford, 2009).

Carleton, Conolly, and Ianonne (2012) outlined the potential pitfalls of predictive analysis for archaeological investigations and proposed methods to avoid these past miscalculations. In their analysis, they created a raster dataset for each variable that contains weighted values, ranging from 0 to 1, indicating site potential. The four resulting rasters were then compiled into one model of archaeological site potential. The authors claim that this method is superior for several reasons, but mainly that it produces a model by assigning each pixel a unique value of site potential, thereby preventing environmental phenomena from being parameterized. However, the study did not include any statistical measures of accuracy that would allow it to be directly compared to the studies completed by Green or Vaughn and Crawford.

Several concerns are found with Carleton et al.'s work. First, they acknowledge that due to the nature of raster datasets, the edges of the model may contain falsely low

values as some sites exist on the boundary of pixels. Therefore they are incorrectly omitted from the pixel, resulting in inaccurate values of site potential. Second, only four environmental variables were taken into consideration, of which soil type, arguably the most important, was based on digitized maps that were between 20 and 54 years old. Thirdly, this study creates values of site potential based only on known sites, with no statistical measures of significance. The resulting model may be biased towards the variables of those known sites based on the authors' incomplete knowledge of site locations.

Collectively, these studies laid valuable groundwork for future predictive models of site potential. This research will attempt to consider the successes and pitfalls of these previous works, in order to produce predictive model for ancient Maya wetland agricultural fields. This study will use a sample size of 521 points randomly selected within the boundaries of known field locations in northern Belize and the bordering areas of southern Quintana Roo, Mexico. An additional 521 points were randomly selected outside the location of known field boundaries for a total point sample size of 1042.

Geologic History of the Yucatan Peninsula

In order to understand the environmental challenges the Maya faced while living in northern Belize, and the way in which they selected land for wetland agriculture, we will explore how this landscape was created. The main force throughout Belize's geological history has been its position near the southern edge of the North American tectonic plate. Due to this proximity, the south half of the country is more structurally

complex than the north. The Maya Mountains in the south are made of Triassic granites and Late Paleozoic siliclastic sediments which have been metamorphosed (Jordan 2002).

Three main geologic events have led to the creation of this unique region, long before it became habituated by humans. The first defining event was the Paleozoic era deposition of clastic material and subsequent intrusion of Triassic period granite. The second event was the deposition of carbonates during the Cretaceous period. The third most influential event was the uplift and faulting during the Tertiary period, due to plate tectonic activity (Miller 1996). Though some geologic activity has occurred during the Quaternary Period, the events that had the most influence in forming the Maya landscape occurred millions of years prior (Bundschuh and Induni 2007). The figure below shows these key periods in the context of the geologic time scale.

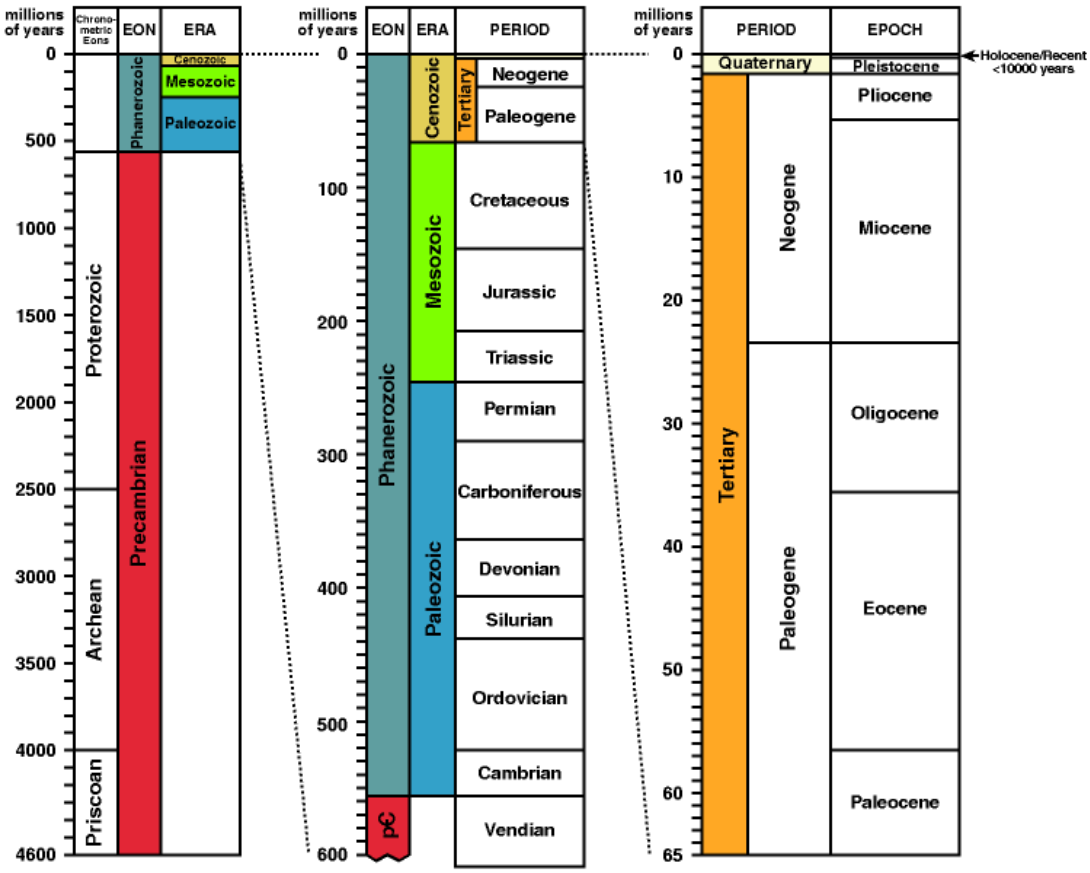


Figure 24: Geologic Time Scale (MacRae 1996)

Today, southern Belize is part of a physiographic region known as the Maya Highlands. This region consists of morphologically distinct mountain ranges, fault-controlled canyons and a few broad alluvial valleys. The Maya Highlands exist across a Cretaceous-Paleogene fold belt with a crystalline basement and sedimentary cover. The geomorphology of this region is controlled by lithologic variations that are exposed in the eroding high-altitude mountains (Bundschuh and Induni 2007). The south is physically distinct from the north which likely led to cultural differences in agriculture, architecture,

and art, based on archaeological evidence (Coe 2011, Perry, Velazquez-Oliman and Socki 2003).

The karst in northern Belize has been developing since the late Miocene, similar to much of the Yucatan Peninsula. Northern Belize also has significant surface and subsurface drainage that allows precipitation to directly infiltrate the ground water reservoirs or flows as surface runoff (T. Beach, S. Luzzadder-Beach, et al., Human and natural impacts on fluvial and karst depressions of the Maya Lowlands 2008). The region also has few permanently flowing rivers, in the west and southeast including Belize River, New River, and the Rio Hondo (Coe 2011). Karst topographies develop a system of internal drainage over time (Jordan 2002) and so the geomorphology of the Yucatan Platform has a strong relationship with its structure and the depth of the groundwater table (Bundschuh and Induni 2007).

Since the Tertiary, geologic activity in northern Belize has mostly been slow subsidence and carbonate deposition. In the south, uplift in the Maya mountain region raised the carbonates to 50m above sea level. Additional uplift since that time can be witnessed at river terraces, cave levels and marine faunal remains found in inland sediments. More recently, Quaternary alluvium from the Maya Mountains filled in waterways that had previously extended into the country (Miller 1996).

Northern Belize is part of physiographic region known as the Yucatan Platform consisting of expansive carbonate lowlands (Bundschuh and Induni 2007). More commonly known as the Yucatan Peninsula, it is basically a shelf of limestone uplifted over time (Coe 2011). Figure 2 is a map of the geologic units throughout the Yucatan

peninsula, along with the fault lines that have caused much of the uplift (Garrity and Soller 2009). The map also shows the older geologic units in southern Belize, (PZ (Paleozoic), MZg (Mesozoic) and K (Cretaceous), which form the Maya Mountains, as well as the series of small fault lines that traverse from east to west. Heading north, the geologic units shift from older volcanic and plutonic to geologic units that are younger and mainly sedimentary, formed after centuries of erosion. These younger units date to the Tertiary and Quaternary periods.

The unique geological setting has greatly influenced the region's ecosystem as well as human inhabitants. Weathering of carbonate landforms has produced minimal soil cover and insufficient nutrients. In addition, though rainfall is abundant, drinking water is scarce as there are few surface streams by which to transport the vital resource. In pre-colonial times, natural depressions and hand-dug wells were the main water sources (Perry, Velazquez-Oliman and Socki 2003).

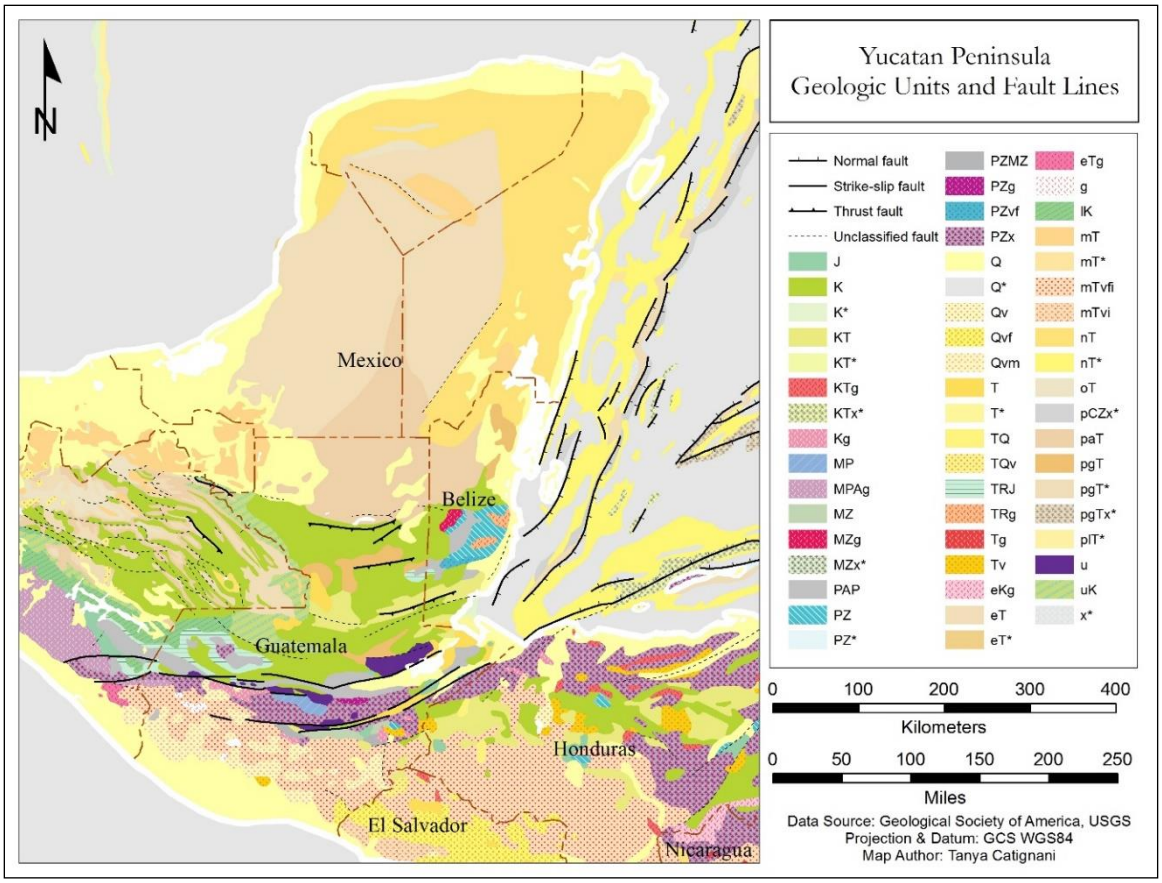


Figure 25: Yucatan Peninsula Geologic Units and Fault Lines.

The peninsula is also notably characterized by karst landforms such as sinkholes, cenotes, and cave networks. A cenote (Figure 26) is a circular sinkhole formed by the collapse of underground caves. They are currently the main source of drinking water since they are perennially filled with water percolating through limestone and are often served as focal points for native settlements. Another common feature is the bajo (Figure 27), a wetland depression that fills up in the summer but becomes dry by the winter. Additionally, this area also exhibits more distinctly Maya architecture and artwork (Coe 2011).

Figure 26: A cenote in Mexico (Kehnel 2008)

Figure 27: Bajo near the site of El Mirador in northern Guatemala (Authentic Maya 2005).

Environmental Processes of the Yucatan Peninsula during the Holocene

Since the last glacial maximum, the landscape of the Maya people was impacted by natural and human-induced processes. The Yucatan peninsula has undergone several climatic changes over the last 10,000 years, causing the human population to find new ways to adapt. Although erosion has been a pervasive force, human activity has actually been the main driver of surface processes in more recent times (Beach, Dunning, et al. 2006, T. Beach, S. Luzzadder-Beach, et al. 2008).

The first inhabitants of the Maya lowlands arrived 5,000 years ago, as determined by pollen analysis. As early as 4,000 years ago agriculture began to spread, which led to

severe erosion as a result of deforestation. The severe erosion also led to sedimentation of the lowlands which peaked in the late Preclassic. However, population levels did not peak until the late Classic period. After the well-known population decline of the late Classic period, there is evidence that reforestation began to occur, though some areas did not experience this process until after European contact (T. Beach, S. Luzzadder-Beach and N. Dunning, et al. 2009).

During the Late Preclassic period, the water table rose and was a likely motivation for farmers to begin using ditches to manage the excess water. By the Late Classic, the fields were completely flooded, pushing cultivation efforts further upland. At the same time, population increased and there was enhanced soil erosion from growing deforestation. Also, calcium carbonate and gypsum were being deposited in the Rio Hondo, creating layers of silt over the flood plains. When the water level receded during the Terminal Classic and Early Postclassic, farmers again created ditches along the fertile flood plains (T. Beach, S. Luzzadder-Beach and N. Dunning, et al. 2009).

Adapting to a Changing Environment

The ancient Maya and their predecessors were constantly adapting to changes in the climate and their environment during the Holocene. Starting with their most basic need, the Maya found ways to mitigate the challenge of having scarce water sources within the interior of the Yucatan peninsula. Areas away from the coast are limited in lakes and rivers, but the Maya developed additional methods of acquiring water by constructing reservoirs and water control features. These systems became larger and more complex in the late Classic Period when the population was at its peak. For example, the

reservoir at Tikal was able to provide 4.8 liters of water per day to an estimated 45,000 to 62,000 people. They also were able to keep their reservoirs clean by transforming them into wetland biospheres and used plants and algae to maintain its purity (Lucero, Gunn and Scarborough 2011).

One of their methods of developing a stable food supply was through wetland agriculture. In the Maya lowlands, two types of wetlands have been utilized and transformed during the Holocene: bajos, which often have Maya sites nearby, and the perennial wetlands. By studying sediment deposits in the wetlands, it is possible to identify two periods of slope instability and wetland aggradations which occurred at the beginning of the Holocene and at the beginning of the Maya Classic Period (AD 250 – 900), respectively. On the coast, wetlands may have been modified by the creation of raised fields, another form of wetland agriculture where soil is piled up in order to create a raised growing surface. Conversely, the existence of additional sediment on coastal fields may have been due to natural aggradations through erosion and gypsum precipitation that buried the raised fields (T. Beach, S. Luzzadder-Beach and N. Dunning, et al. 2009). Maya agriculture was likely a significant factor in the increased soil erosion of the region. Stone terraces and slope management practices allowed this process to be controlled in some locations. Although these features appear to be limited in number, it is possible that additional terraces could exist but are concealed by dense vegetation (Dunning and Beach 1994). Essentially, their manipulation of the landscape became so intense that the subsequent erosion and deforestation made significant changes to their landscape (Luzzadder-Beach and Beach 2009). The environmental changes that occurred

at the end of the Classic period may have arrived too quickly for them to make the necessary adjustments to their land use methods, and were left unable to deal with them effectively (T. Beach, S. Luzzadder-Beach and R. Terry, et al. 2011).

Paleoclimate evidence shows that during much of the Preclassic (~ 2,000 BCE – 250 AD), the climate was wet and the bajos were once shallow lakes, but by 250 AD these conditions shifted to become drier. This shift contributed to the demise of the large Preclassic cities in the northern region of Guatemala. At 585 AD there was a major dry event, marking the boundary between Early and Late Classic. By the end of the late Classic, population and intensive land use were at their peaks (Coe 2011). During this period there is also evidence of widespread drought and simultaneous abandonment of many regional Maya sites. In the centuries following this event, populations became smaller and dispersed along waterways (T. Beach, S. Luzzadder-Beach and N. Dunning, et al. 2009). One drought lasted from AD 800 to 1050 and was the driest period in 3,000 years. This period saw the collapse of the Classic period in the southern lowlands, after the Maya had degraded their environment to the point where it could no longer support high populations (Coe 2011).

Data and Methods

As shown in Chapter 2, improvements in the resolution of commercial multispectral imagery during recent years have led to the identification of ancient Maya wetland agricultural fields along the Rio Hondo and several other rivers in northern Belize, providing insight into how field locations were selected. In addition, field research and excavations have made it clear that certain environmental conditions would

have likely influenced the Mayans in their selection of areas in which to construct wetland fields (N. Dunning, T. Beach and P. Farrell, et al. 1998, Dunning, Luzzadder-Beach, et al. 2002). Like modern farmers, the Maya would have required a nearby water source. Additionally, they may have also preferred conditions that were conducive to farming such as soil type. Slope and elevation are also important factors to consider, especially since wetland agriculture was reliant on the ability to manipulate the water table. In addition, this type of research may lead to the identification of previously unknown environmental influences.

The data acquired for this study includes vector and raster datasets that describe the environmental characteristics of the northern Belize landscape. These vector datasets were acquired from multiple sources including The Electronic Atlas of Ancient Mayan Sites, The Biodiversity & Environmental Resource Data System of Belize, Data Basin, and Mexico's National Institute of Statistics and Geography (Witschey and Brown 2010, Biodiversity and Environmental Resource Data System of Belize (BERDS) 2012, Instituto Nacional de Estadistica y Geografia n.d., Selva Maya Consortium 2016). Many of these variables are common among similar archaeological predictive modeling studies (slope, elevation and distance to water) since these are some of the most limiting factors on where humans can live. Other variables, such as soil type, have been used by Green (1973) and Vaughn & Crawford (2009) in their earlier models of ancient Mayan sites and so for that reason were included in this analysis.

The spatial model will be developed from the initial database of identified field locations, along with values from the variable datasets as attributes. Attributes will be

both qualitative and quantitative. The full list of attributes will include distance to water, slope, elevation, soil type, bedrock type, and soil moisture.

Defining the Study Area

This research will focus on a defined area within the Maya Lowlands. The country boundary of Belize, acquired from the online Biodiversity and Environmental Resource Data System of Belize (BERDS), is used to create the western and eastern extents (minus the islands and the Belize Barrier Reef) of the study area. The southern extent is the foothills of the Maya Mountains. The Northern extent was created using the outline of the Rio Hondo watershed, obtained from the World Wildlife Fund's Conservation Science Program and created using Shuttle Radar Topography Mission elevation data. Since the Rio Hondo has many ancient Maya wetland agriculture fields along its sides, in both Belize and Quintana Roo, Mexico, the watershed was selected as a more natural boundary. Finally, the western extent of Belize was followed directly north to connect to the Rio Hondo Watershed extent. The full study area boundary is shown in the map below, highlighted in yellow.

Figure 28: Study area boundary. Map by Tanya Catignani.

Logistic Regression

Logistic Regression will be used to determine which of several environmental factors were the most influential in selecting land for wetland agriculture. In general, this is a regression method that is used when the dependent variable is binary or categorical, and can accommodate variables that are both continuous and or categorical. The independent variables used in this study include distance to water, slope, elevation, soil type, bedrock type, greenness, brightness temperature, and soil moisture. This type of regression analysis uses the logistic function to assign a probability score between 0 and 1 to each sample point. The probability can then be used to determine which category the point likely belongs to. The logistic regression model is shown below where β_0 is the y intercept and the following values represent the explanatory variables and coefficients.

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_{1x1} + \beta_{2x2} + \cdots \beta_{mxm}$$

Equation 1: Logistic regression

In order to solve for probability, the model is rewritten as the following.

$$p = \frac{e^{(\beta_0+\beta_{1x1}+\beta_{2x2}\ldots\beta_{mxm})}}{1 + e^{(\beta_0+\beta_{1x1}+\beta_{2x2}\ldots\beta_{mxm})}}$$

Equation 2: Logistic regression probability model

To begin the logistic regression analysis, a sample point dataset of 1042 points, half of which occur in areas of known wetland fields and the other were randomly selected throughout the study area, was created using the Create Random Points tool in ESRI's ArcMap. The dataset was then overlaid onto spatial layers representing each of

the selected variables, and the values at each point location were added to the dataset as attributes. The "glm" package in R was used to analyze the final dataset.

The null hypothesis of this study is that the environmental factors selected do not statistically differentiate between field and non-field locations. If the null hypothesis is correct, the Maya's selection of land for wetland agriculture was completely random. The alternative hypothesis is that the environmental factors do in fact statistically differentiate between field and non-field locations. If the alternative hypothesis is correct, one or more of the variables were influential in selecting land for wetland agriculture and may be used to find similar areas in northern Belize.

Ho: Environmental factors did not influence the ancient Maya's selection of land for wetland agriculture in the study area.

Ha: Environmental factors did influence the ancient Maya's selection of land for wetland agriculture in the study area.

Variables

Belize lies in zone 16N of the Universal Transverse Mercator coordinate (UTM 16N) system. All data layers used for this study were converted to this system, in accordance to the World Geodetic System 1984 (WGS84) datum.

Distance to Water

Many of the wetland fields identified in the previous study are near a currently existing body of water, and so distance to water was very likely an influential factor in the selection of land for agriculture. Using the "Near" tool in ArcMap, the distance from each sample point to the nearest body of water is calculated by locating the closest segment of the line or polygon which represents the water feature, and then measuring a

straight perpendicular line from the point to the segment (ESRI 2013). Data layers for rivers and waterbodies were acquired from BERDS (2012) and from The National Institute of Statistics and Geography (Instituto Nacional de Estadistica y Geografia n.d.). The distance to water variable was also used in Green (1973) for the predictive model of Maya settlement sites.

Distance to Sites

The distance from each sample point to the nearest known Maya site was also calculated using the "Near" tool in ArcMap. The site location data was acquired online from the "Electronic Atlas of Ancient Maya Sites" which was published by Walter R. T. Witschey and Clifford T. Brown (2010). The database contains over 6,000 sites, from small villages to large urban centers, spanning the ancient Maya world. Since it is likely that ancient Maya farmers would have lived near where they worked, this variable may help to identify areas that probably experienced wetland agriculture.

Slope

Since wetland fields have often been found in areas where the water table is near the ground surface, another likely influential factor is terrain slope or lack thereof. Slope was calculated for each sample point, using an ASTER 30-meter spatial resolution raster digital elevation model (DEM) and the "Extract Values to Points" tool in ESRI's ArcMap. In the tool's interface, the setting "Interpolate values at point locations" was selected so that surrounding pixel values would be considered. This variable was use by Vaughn and Crawford (2009) in their predictive model of Maya settlement sites. The ASTER L1B data product was retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC), USGS/Earth

Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access/data_pool.

Elevation

Many of the known wetland fields are also located at lower elevation levels, but this may or may not have precluded fields from having been created at a variety of elevations in northern Belize. A high water table may have been a more significant factor to wetland field development than elevation alone. This variable will be included the model so that its influence can be objectively compared with the rest. Elevation will also be determined based on the ASTER 30-meter DEM and the "Extract Values to Points" tool in ArcMap. In the tool's interface, the setting "Interpolate values at point locations" was selected so that surrounding pixel values would be considered (ESRI 2016).

Soil Type

The Belize soil dataset was acquired online at databasin.org and was developed by the Selva Maya consortium. It is one of several datasets created for the Selva Maya Ecoregional Planning effort and created at a scale of 1:250,000. Due to a lack of geological data for Belize in 2004, the mapping was supplemented by correlating the data with known geological features. Although the authors admit the data is not perfectly accurate, this data appears to be the most detailed geospatial soil layer publically available for Belize (Selva Maya Consortium 2016). Since many of the fields identified straddle the Rio Hondo, soil data for the land adjacent to Belize, in Mexico, was acquired from INEGI. It was created in 2002 at a scale of 1:1,000,000 (Instituto Nacional de Estadistica y Geografia n.d.).

Bedrock Geology

The Belize bedrock dataset was also acquired online at databasin.org, and developed by the Selva Maya consortium for the 2004 Selva Maya Ecoregional Planning effort. The consortium developed the dataset by studying the soil geomorphology of relief formations in Belize, which allowed for the translation of terrestrial processes at work as well as the origin of the soil formation. The data was created at a scale of 1:250,000 (Selva Maya Consortium 2016). A similar dataset for Mexico was created in 2002 at a scale of 1:1,000,000 (Instituto Nacional de Estadistica y Geografia n.d.).

Greenness / NDVI

Greenness, estimated using the Normalized Difference Vegetation Index (NDVI), is a measure of vegetation health. It has been used for a wide variety of remote sensing applications since the 1970s. This variable of greenness was found to be significant in predictive model by Vaughn and Crawford (2009). Here, it will be measured using two Landsat 8 images, acquired in November and December 2014. This index is a simple ratio of the difference in values between the near infrared (NIR) band and the red band per pixel, and the total of NIR and red values per pixel. The equation for NDVI is shown below.

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$$

Equation 3: Normalized Difference Vegetation Index (NDVI)

The Landsat 8 L1T data product was retrieved from the online Data Pool, courtesy of the NASA Land Processes Distributed Active Archive Center (LP DAAC),

USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, https://lpdaac.usgs.gov/data_access/data_pool.

Brightness Temperature / T_s

Brightness Temperature was derived from Landsat 8 thermal bands 10 and 11. This value is a proxy for land surface temperature and was calculated in the ENVI software program for imagery analysis and processing. First, the two thermal bands were converted from Digital Numbers (DN) to radiance values using an equation provided by USGS and ENVI's band math tool.

$$L_\lambda = M_L Q_{cal} + A_L$$

Equation 4: Digital numbers to radiance

In this equation, L_λ represents top of atmosphere (TOA) spectral radiance in Watts/(m2 * srad * μm), M_L represents the band-specific multiplicative rescaling factor found in the image metadata, and A_L represents the band-specific additive rescaling factor also from the metadata, and Q_{cal} represents the quantized and calibrated standard product pixel values, or DN (USGS 2015).

The new radiance bands were then made into a 3D cube, wavelengths were applied to each band in the header file, and then the cube was processed using ENVI's Thermal Atmospheric Correction module. The module uses an algorithm which plots in-scene maximum brightness temperatures against the radiance for each wavelength in order to fit a line along the highest values. The slope and offset of this line can then be used to adjust pixel values for a scene-specific thermal atmospheric correction (Harris Geospatial Solutions 2016).

Finally, the atmospherically corrected 3D cube was converted to temperature in Kelvin. ENVI has three methods for this, the one used in this study is the Emissivity Normalization technique. Since thermal radiation is a function of both surface temperature and emissivity, the Emissivity Normalization technique assumes a fixed value for emissivity and then uses this to calculate temperature for each pixel (Harris Geospatial Solutions 2016).

Soil Moisture / TVDI

The last variable in this model, soil moisture, was measured using the Temperature-Vegetation Dryness Index (TVDI). This index combines the NDVI with surface temperature (T_s), using only satellite derived data, and has been proven to be a straightforward and accurate model of soil moisture. It is ideal for use in regional studies, and given that Maya fields were often constructed near bajos, this index should help to identify prime agricultural land that may not be in close proximity to a river. The figure below explains theoretically where different land surface types exist in T_s/NDVI space, and how they relate to soil moisture (Sandholt, Rasmussen and Andersen 2002).

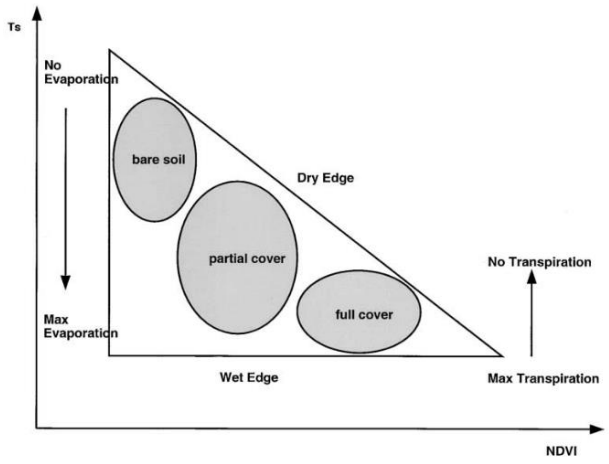


Figure 29: Simplified Ts/NDVI space (Sandholt, Rasmussen and Andersen 2002).

In the next image, each of the values used in the TVDI equation are defined. "$T_s$min" is the lower edge of the T_s/NDVI space and is also known as the Wet Edge. The Wet Edge is typically given the value of 0, but the slope of the line can also be used for improved accuracy. "$T_s$max" is the hypotenuse of the T_s/NDVI space and is also known as the Dry Edge. This value is always derived from the slope of the line (Sandholt, Rasmussen and Andersen 2002).

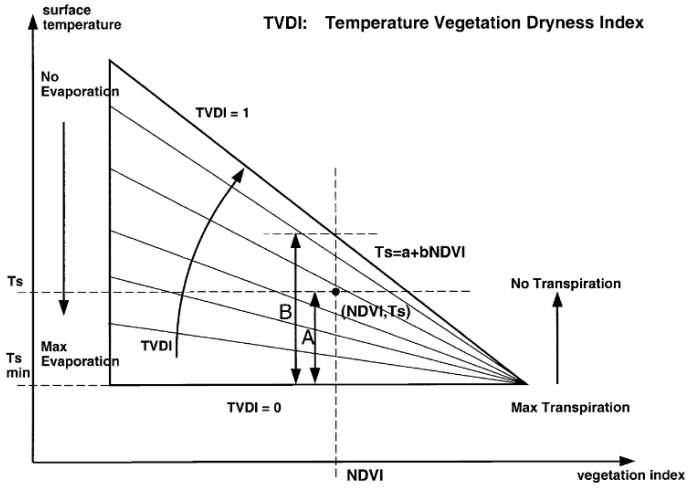


Figure 30: TVDI variables explained (Sandholt, Rasmussen and Andersen 2002).

The Landsat thermal bands 10 and 11, which have a spatial resolution of 100m, were used to create the TVDI. For this research, the TDVI algorithm was calculated using the slope of the Wet Edge for $T_s$min and the slope of the Dry Edge for $T_s$max. The equation is written as

$$TVDI = \frac{T_s - (a_2 + b_2 NDVI)}{(a_1 + b_1 NDVI) - (a_2 + b_2 NDVI)}$$

Equation 5: TVDI equation (Sandholt, Rasmussen and Andersen 2002)

Values from the previously calculated NDVI and T_s data are assigned to each sample point, resulting a scatterplot which is referred to as Ts/NDVI space. Since Ts tends to be more sensitive to water stress than NDVI, the combination of these measurements creates a more accurate depiction of soil moisture. The scatterplot is shown below.

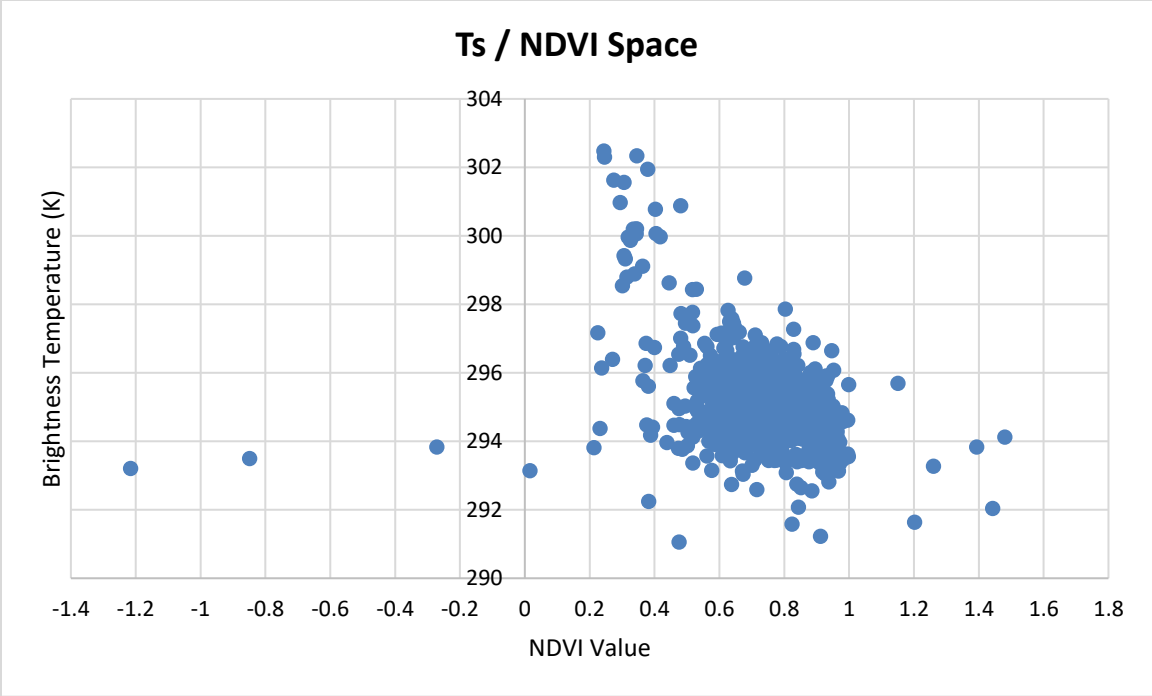


Figure 31: Ts / NDVI Space

In order to calculate the slope of the dry edge, the NDVI values were binned (.2-.39, .4-.59, .6-.79, .8-1) and the highest 10 T_s values for each bin was plotted. The slope of these points were used for the dry edge equation.

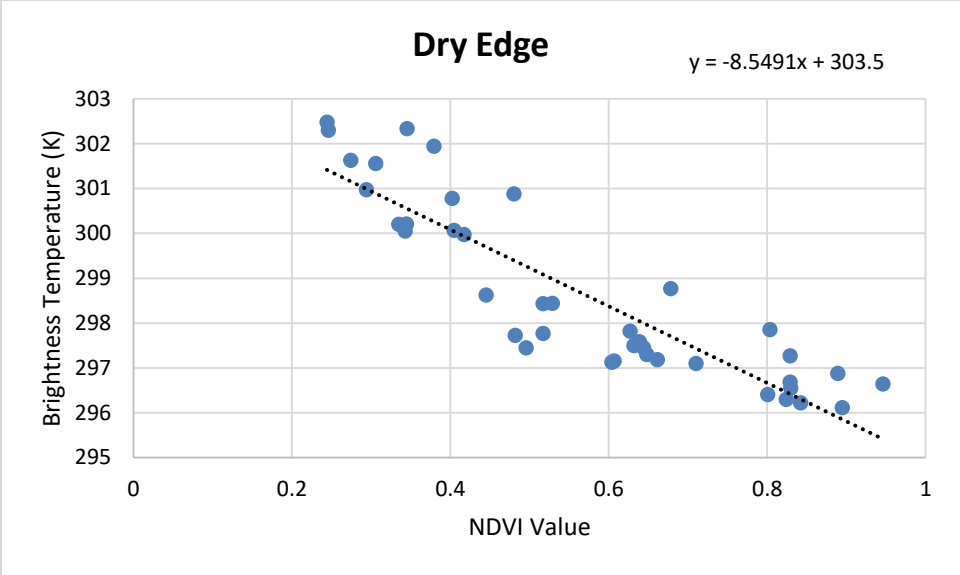


Figure 32: Ts / NDVI Space, Dry Edge

The wet edge was calculated in the same manner, but instead using the lowest 10 Ts values per bin. This plot is shown below.

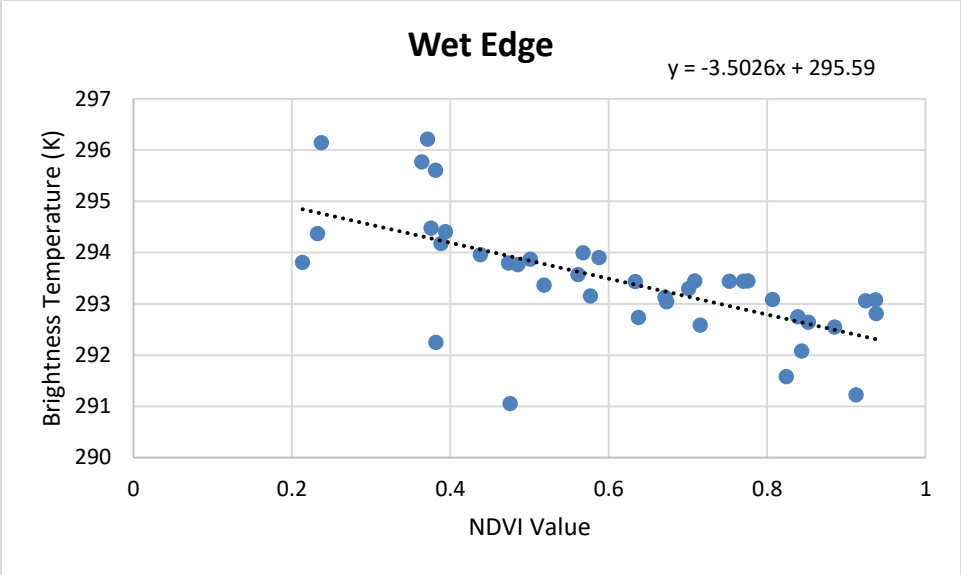


Figure 33: Ts / NDVI Space, Wet Edge

The final TVDI value for each sample point was calculated in Excel.

Results

The dataset was imported into R for the logistic regression analysis. After creating a separate binary variable for each soil and bedrock type, there were 17 independent variables in total. They are listed below with a description of their values.

1) NEAR_Water – The distance in meters from a sample point to the nearest body of water.
2) NEAR_Site – The distance in meters from a sample point to the nearest known Maya site.
3) Slope_deg – The land surface slope in degrees.
4) Elev_m – Elevation in meters.
5) NDVI – The greenness index which ranges from 0 to 1.

6) TVDI – The soil moisture index which ranges from 0 to 1.

7) Ts – Brightness temperature, or surface temperature, in degrees Kelvin.

8) Arenosol – The presence or absence of arenosol soil type (1 or 0).

9) Cambisol – The presence or absence of cambisol soil type (1 or 0).

10) Fluvisol – The presence or absence of fluvisol soil type (1 or 0).

11) Gleysol – The presence or absence of gleysol soil type (1 or 0).

12) Leptosol – The presence or absence of leptosol soil type (1 or 0).

13) Vertisol – The presence or absence of vertisol soil type (1 or 0).

14) Volcanic – The presence or absence of volcanic bedrock (1 or 0).

15) Sedimentary – The presence or absence of sedimentary bedrock (1 or 0).

16) Caliza – The presence or absence of caliza (limestone) bedrock (1 or 0).

17) Alluvial – The presence or absence of alluvial bedrock (1 or 0).

The "glm" package in R was used to create a logistic model and explore the significance of all variables. The 17 independent variables were used to predict the dependent variable, MayaField, which indicates the presence or absence of a known ancient Maya wetland field. The following command was run to produce the initial results.

```
> log_model <-glm(MayaField ~
NEAR_Water+Arenosol+Cambisol+Fluvisol+Gleysol+Leptosol+Vertisol+Alluvia
l+Caliza+Volcanic+Sedimentary+Elev_m+Slope_deg+NDVI+TVDI+Ts+NEAR_Site,
family=binomial(logit), data=model_data)
```

The results provided a table of coefficients and diagnostic values that provide insight to the quality of the initial model, as well as an error message that indicated there was separation in the variables. Separation occurs when one or multiple variables predicts

the dependent variable perfectly, for a part or all of the data, causing the model to fail (Zumel and Mount 2014).

```
> summary(log_model)
Call:
glm(formula = MayaField ~ NEAR_Water + Arenosol + Cambisol +
 Fluvisol + Gleysol + Leptosol + Vertisol + Alluvial + Caliza +
 Volcanic + Sedimentary + Elev_m + Slope_deg + NDVI + TVDI +
 Ts + NEAR_Site, family = binomial(logit), data = model_data)
Deviance Residuals:
 Min 1Q Median 3Q Max
-2.4937 -0.4614 0.1006 0.5418 3.7773
Coefficients: (2 not defined because of singularities)
 Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.934e+02 6.490e+02 -0.298 0.76565
NEAR_Water -2.544e-04 6.298e-05 -4.039 5.37e-05 ***
Arenosol -3.805e-01 1.219e+03 0.000 0.99975
Cambisol 2.049e+01 3.455e+03 0.006 0.99527
Fluvisol 2.673e-01 1.815e+03 0.000 0.99988
Gleysol 1.523e+01 6.467e+02 0.024 0.98121
Leptosol 1.509e+01 6.467e+02 0.023 0.98139
Vertisol NA NA NA NA
Alluvial -2.321e+00 4.174e-01 -5.561 2.68e-08 ***
Caliza -7.918e-01 3.419e-01 -2.316 0.02057 *
Volcanic -1.389e+01 2.400e+03 -0.006 0.99538
Sedimentary NA NA NA NA
Elev_m -5.501e-02 6.869e-03 -8.008 1.17e-15 ***
Slope_deg 3.159e-03 3.344e-02 0.094 0.92474
NDVI 2.993e+00 1.049e+00 2.853 0.00433 **
TVDI -5.373e+00 9.453e-01 -5.683 1.32e-08 ***
Ts 6.178e-01 1.861e-01 3.320 0.00090 ***
NEAR_Site -1.360e-04 2.888e-05 -4.709 2.49e-06 ***
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Dispersion parameter for binomial family taken to be 1)
 Null deviance: 1444.52 on 1041 degrees of freedom
```

```
Residual deviance: 741.76 on 1026 degrees of freedom
AIC: 773.76
Number of Fisher Scoring iterations: 15
```

The output provided a table of coefficients along with standard error, z value, and p-value or significance. P-values that are less than 0.5 are considered significant (Zumel and Mount 2014). Significant and non-significant variables are also indicated by the presence or absence of "*"in the far right column. The coefficients for two variables, vertisol and volcanic, were not generated since their presence within the dataset was too low. Also, the number of Fisher Scoring iterations was 15 which is too high, suggesting that the model is not valid. Ideally, this value should not be higher than 8 (Zumel and Mount 2014).

The Akaike Information Criterion (AIC) is a measure of model quality and is comparable to the r-squared used in linear regression. This value can be used to compare the quality of two models (Zumel and Mount 2014).

The model was then run a second time, after removing all non-significant variables. Although the AIC dropped to 768.87 and the Fisher Scoring iterations dropped to 7, the model still generated and error warning message indicating separation within the data. It was suspected that the cause of the error was the relationship between NDVI, Ts, and TVDI, since the first two were used to create the later. A scatterplot matrix was created to explore the relationships between variables.

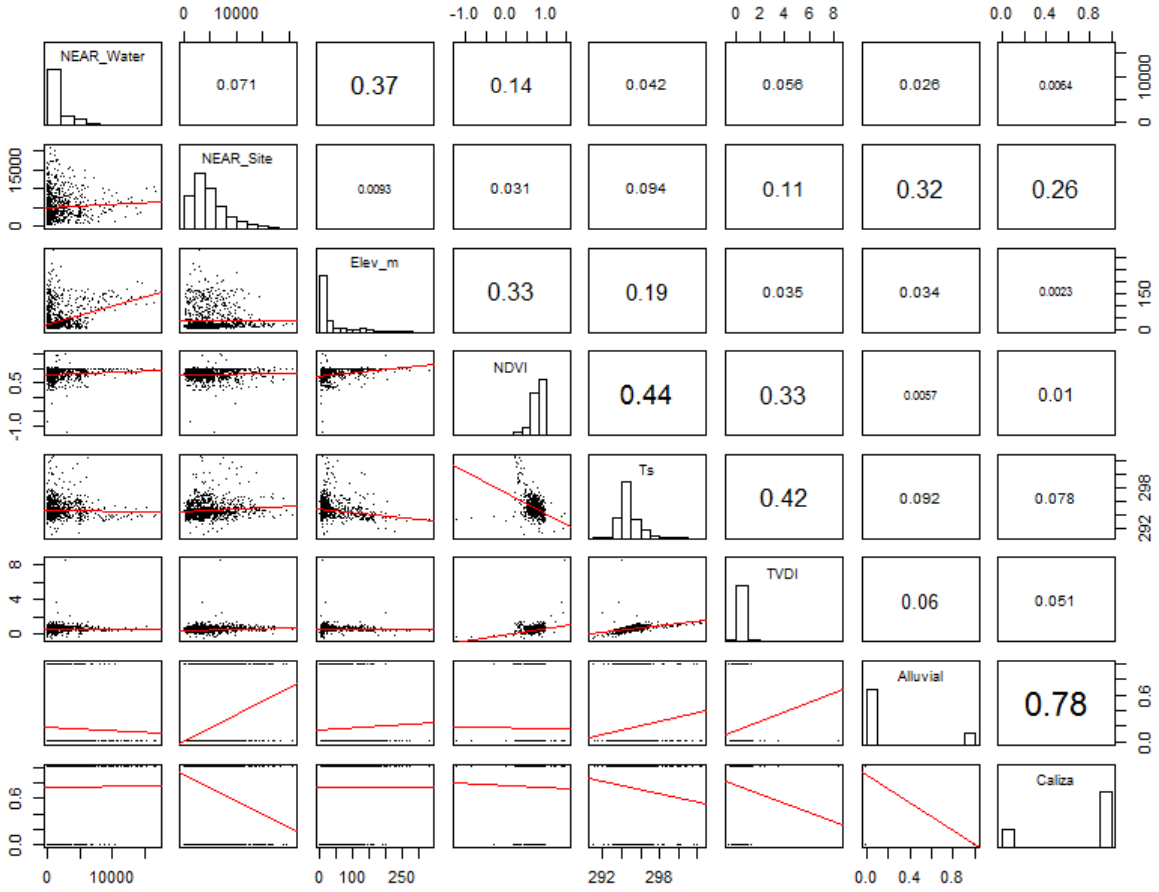


Figure 34: Scatterplot of all significant variables from logistic regression.

The scatter plot matrix shows that the relationships between NDVI, Ts, and TVDI is high and so the first two were removed from the list of variables. In addition, alluvial and caliza were very highly correlated and so alluvial was also removed. The final model paired down the original 17 variables to five: NEAR_Water, NEAR_Site, Elev_m, TVDI, and Caliza. The results are shown below.

```
> summary(log_model)
Call:
glm(formula = MayaField ~ NEAR_Water + NEAR_Site + Elev_m + TVDI +
 Caliza, family = binomial(logit), data = model_data)
```

```
Deviance Residuals:
 Min 1Q Median 3Q Max
-3.0333 -0.5457 0.1059 0.5968 3.7221
Coefficients:
 Estimate Std. Error z value Pr(>|z|)
(Intercept) 3.315e+00 3.248e-01 10.206 < 2e-16 ***
NEAR_Water -1.616e-04 5.876e-05 -2.750 0.00597 **
NEAR_Site -1.534e-04 2.650e-05 -5.791 7.02e-09 ***
Elev_m -5.995e-02 6.143e-03 -9.758 < 2e-16 ***
TVDI -2.777e+00 3.704e-01 -7.499 6.44e-14 ***
Caliza 7.967e-01 2.012e-01 3.959 7.53e-05 ***
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Dispersion parameter for binomial family taken to be 1)
 Null deviance: 1444.52 on 1041 degrees of freedom
Residual deviance: 802.07 on 1036 degrees of freedom
AIC: 814.07
Number of Fisher Scoring iterations: 7
```

The final model, consisting of five variables, shows a slightly higher AIC value than the original. However, all errors and diagnostic indicators have been addressed and the final result can be presumed to be valid. A scatter plot matrix of the significant variables is shown below.

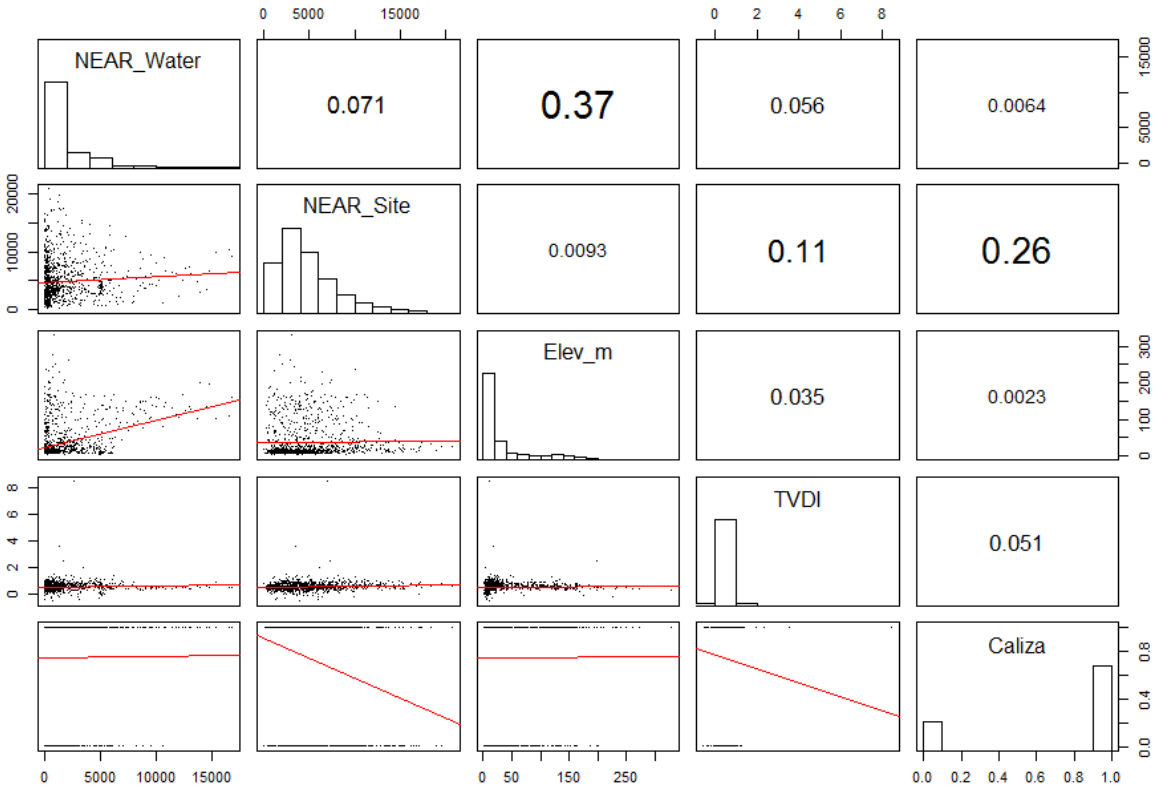


Figure 35: Scatterplot of the final five variables used in the logistic regression analysis.

The final list of variables and their coefficients were entered into the probability equation provided in the previous section, to create a probability map for the study area. To do this, each of the variables, except for elevation, needed to be calculated as a continuous data layer. First, both the "NEAR_Water" and "NEAR_Site" layers were generated using Euclidean Distance tool, resulting in two raster layers where each pixel contained the distance to the nearest water feature or site respectively. The TVDI layer was calculated by using the Ts and NDVI layers, along with the Wet Edge and Dry Edge line equations in TVDI equation. The Caliza data was already in shapefile format so the Belize and Mexico layers were merged and then converted to raster where the presence of

Caliza contained cell values of 1 and absene of Caliza had a value of 0. The final calculation of the probability layer was performed using ArcMap's raster calculator. The final output is shown below.

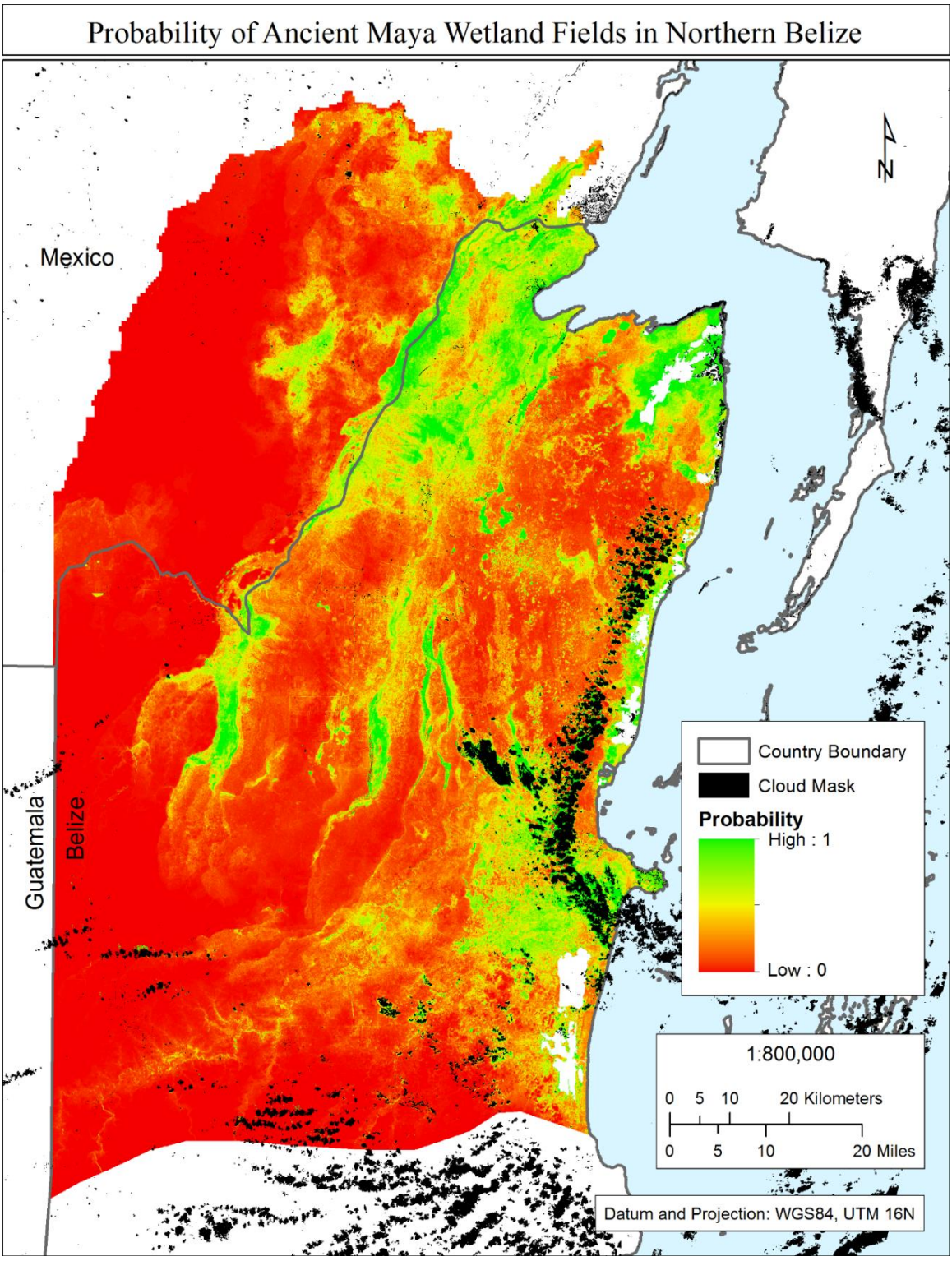


Figure 36: Probability map of ancient Maya wetland fields in northern Belize.

The map depicts the probability of an ancient Maya wetland fields existing in any part of the study area, and ranges from 0 to 100%. Appendix A contains a series of gridded maps that provide the results of the logistic regression analysis in greater detail. The visually identified wetland fields are also included in these maps so that their locations can be compared to the model's ability to predict additional field locations. The area within the top ranges of probability were calculated at 5% increments. A cloud mask was created from the original Landsat-8 images, by conducting an unsupervised classification in ArcMap, so that these areas could be removed from this analysis. The total study area consists of 3,691,132.05 acres, or 14,937.5 km^2, disregarding places with cloud cover.

Table 2: Acreage by probability range and running total.

Total Area Per Probability Range

Probability	Acres	Km^2
95 - 100%	9,270.91	37.52
90 - 95%	16,583.26	67.11
85 - 90%	19,429.43	78.63
80 - 85%	25,084.18	101.51
75 - 80%	32,462.11	131.37
70 - 75%	38,498.43	155.80
65 - 70%	44,662.05	180.74
60 - 75%	52,915.54	214.14
0 - 60%	3,452,226.12	13,970.68

Total Area Per Range - Running Total

Probability	Acres	Km^2
95 - 100%	9,270.91	37.52
90 - 100%	25,854.17	104.63
85 - 100%	45,283.61	183.26
80 - 100%	70,367.79	284.77
75 - 100%	102,829.90	416.14
70 - 100%	141,328.33	571.94
65 - 100%	185,990.39	752.68
60 - 100%	238,905.93	966.82
0 - 100%	3,691,132.05	14,937.50

The predictive model was tested using five-meter resolution elevation data, derived from SAR imagery, at two locations. The first location, La Union, contains a previously identified field; the second location, Rio Bravo Conservation Area, has a high likelihood of containing fields, as suggested by the model. A preliminary visual inspection of the elevation data was conducted for both locations. At La Union, fields could not be visually identified, however, at Rio Bravo Conservation Area the changes in elevation align with characteristic field patterns found through multispectral image analysis, and support the results of the predictive model. The fields at La Union are likely too shallow to be visually identified using this data. In Figure 37 below, the La Union and Rio Bravo Conservation Areas are shown in the top left and right Google Earth screen captures, respectively. The bottom two scenes show the corresponding elevation data as shown in ArcGIS.

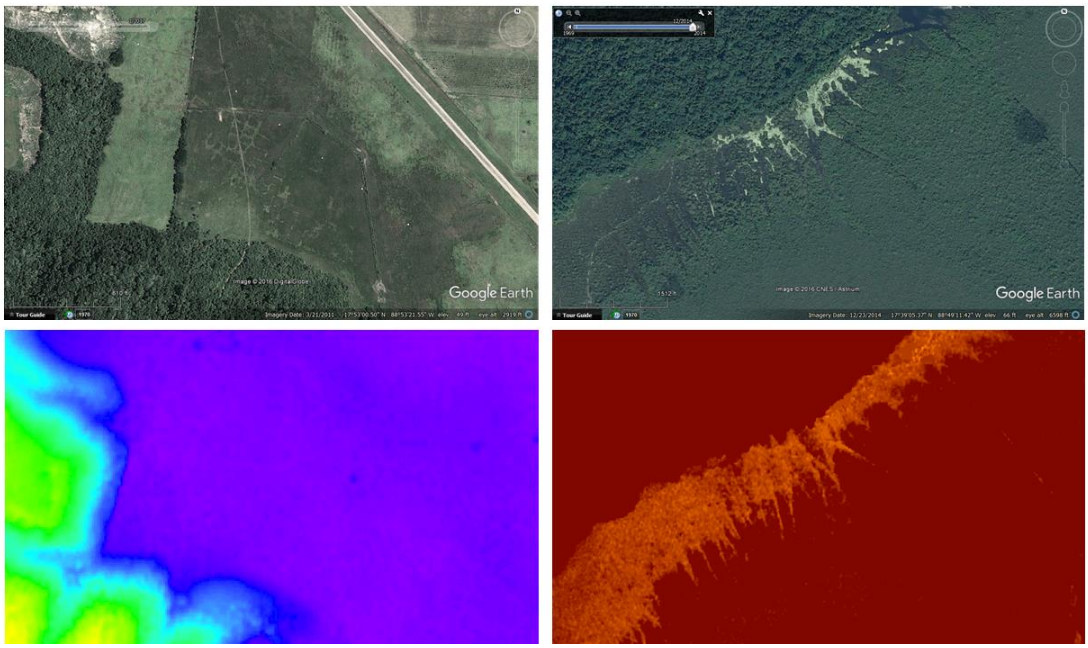


Figure 37: Google Earth image of La Union at top left; Google Earth image of Rio Bravo at top right; Original elevation data for La Untion at bottom left; Original elevation data for Rio Bravo at bottom right

Edge detection algorithms were also applied to the SAR-based elevation data, in order to determine if any linear field patterns could be extracted beyond those that were visually identifiable. These algorithms identify significant changes in pixel values to highlight edges within an image, providing information on the strength of an edge, their orientation, and their scale. Edges detection algorithms can be divided into two main groups: those that do not use prior knowledge about a scene and those that do. The former category is flexible and can be used on any image since they are based solely on neighboring pixels. They are appropriate for general purpose edge detection. The latter category can only be used in precise instances and are used when images repeatedly include the same objects (Jain, Kasturi and Schunck 1995).

Edge detection algorithms essentially aim to emphasize real shapes within an image and at the same time suppress or remove false edges. This is a challenging task since false edges are often the result of general noise within an image or can be created during the image acquisition process. Conversely, algorithms may not always recognize true edges, which are then dropped from the final output (Jain, Kasturi and Schunck 1995).

The ENVI image analysis software program contains several edge detection algorithms and more basic smoothing filters (Harris Geospatial Solutions 2017). Three were selected for this research including the Sobel, Roberts, and Laplacian algorithms. Image 38 below shows the results of these algorithms for the La Union elevation dataset.

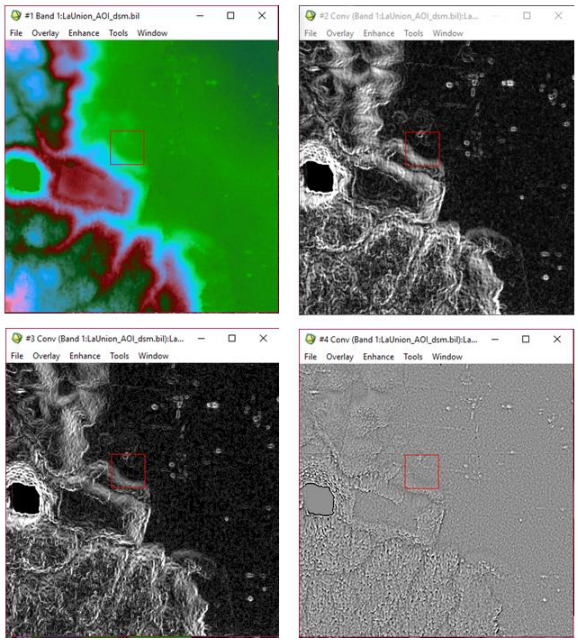


Figure 38: Original elevation data for La Union at top left; Sobel top right; Roberts bottom left; Laplacian bottom right

The edge detection algorithms did not extract any additional information about the field patterns beyond that provided by multispectral image analysis. This may be due to the fact that the ancient fields have been plowed over by modern ones, or that the range of elevation in the scene overpowers them and discarded as noise. These fields may be detectable using even higher resolution data such as lidar but for a large scale study it may not be cost effective or feasible.

Image 39 below shows the results of edge detection analysis at the Rio Bravo Conservation Area. The original data shows that elevation changes exist which align with the field patterns identified through multispectral image analysis. The edge detection algorithms aid in understanding the shape of these fields but do not detect any additional features.

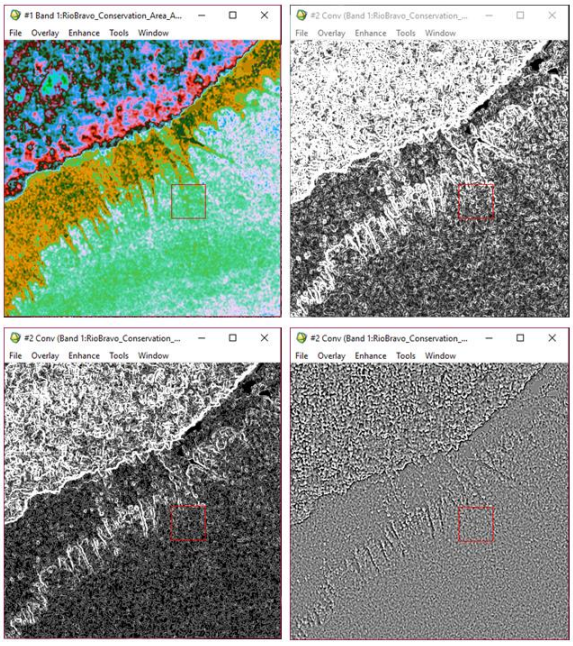


Figure 39: Original elevation data for Rio Bravo Conservation Area at top left; Sobel top right; Roberts bottom left; Laplacian bottom right

Discussion

Based on the analysis performed in Chapter 2, over 16,000 acres of ancient Maya wetland agricultural fields were identified. The probability analysis shows that there is a much higher amount of land within the study area that could have supported this type of agriculture. By itself, the top 10% range of the probability layer contains nearly 10,000 more acres than the discovery made through visual imagery analysis. At the 60% probability level there are over a quarter million acres, or nearly a thousand km^2, of suitable land. Many of the visually identified fields exist within the areas of high probability and several exist within the areas of moderate probability, suggesting that this model may even be conservative in its estimates. There is one area of wetland fields, above the border with Mexico, which exists in an area of low probability. This may be due to a difference in data quality between Belize and Mexico. Since river and waterbody layers for each country were made by different organizations (INEGI and BERDS) who likely have different data standards, it is possible that some features are missing. A map book of visually identified wetland fields and the probability layer was created so that the final output could be examined in greater detail.

Conclusions

This analysis provides a regional estimate of the extent of ancient Maya wetland field in northern Belize. The findings can be used as a baseline for future spatial statistical analysis of these features. The variables incorporated here were nearly all environmental, aside for the distance to known sites. Although these variables likely played an important factor in the selection of land for wetland agriculture, there are likely

many more socioeconomic variables that were influential but are outside the scope of this paper.

There are several ways the analysis could be improved for future research. First, the data used to calculate distance to sites contained information about the importance of each site, ranging from small villages to large cities. Instead of calculating distance to site, it may have been possible to create a population density, weighted by the size of each site. However, since this may have led to erroneous assumptions about population, it was safer to treat each point equally and calculate a simple distance instead. Once more detailed and accurate data regarding population estimates becomes available, this may improve predictive modelling attempts. In addition, although the soil and bedrock datasets used here were more current than those used by other predictive models in this region, they are still lacking in detail and are inconsistent across international boundaries. It would have been possible to use a worldwide data source such as the ISRIC World Soil Information Service (ISRIC World Soil Information Service 2016) which would have ensured that standards across boundaries are consistent. However, although this dataset is more consistent, it was created at a much lower spatial resolution. This paper acknowledges the differences in data quality between Belize and Mexico and presumes that the additional spatial resolution of the country-based datasets outweighs the standardization of the low-resolution worldwide datasets.

This study is the first attempt to use geoinformation science in a regional predictive model of ancient Maya wetland agricultural fields. It considers a range of environmental variables used by previous research and included new raster-based

variables. The final model is not the end goal for understanding the distribution of wetland fields but rather an opportunity to critically examine the relationships between variables that influenced the ancient Maya. Predictive models such as these should be continually compared, reconsidered, and improved upon to provide real value.

CHAPTER FOUR – REFLECTANCE SPECTROSCOPY OF WETLAND FIELD SOIL PHOSPHOROUS

Introduction

The analysis of sediments and soils helps to build a holistic understanding of past cultures and their environment. For example, human behavior such as agricultural activity will impact soil content by raising or lowering the amount of nutrients including phosphorus, nitrogen, potassium, and magnesium. Soils are produced when sediments are formed from the parent rock, and subsequently impacted by organisms, topography, climate and time. As they develop, horizontal layers, or stratigraphy, are formed which contain varying levels of biologic material and weathering. Eventually, a full soil profile develops. Different climates and ecosystems throughout the world tend to produce soil profiles that are unique to their environment. Both environmental and human induced events impact the development of soils, often resulting in distinct markings within the soil profile that may last for millennia (Rapp and Hill 2006). Reflectance spectroscopy is a common tool for the analysis of soils but has only been used to a limited extent to understand human impacts. This study is one of the first to apply reflectance spectroscopy specifically to the impact of agriculture on soil phosphorus (P). For this research, a FieldSpec 4 was used to collect spectral measurements of soil samples in northern Belize in July, 2014, from which P values would be predicted. This device measures spectra in the range of 350nm to 2500nm. A parallel analysis was conducted in 2015 on soil samples collected in northeast Ohio, using a FieldSpec 4 at the USGS headquarters office in Reston, Virginia.

Literature Review

Soil History of Northern Belize

In Mesoamerica, the transition from the Pleistocene to the Holocene is marked by a change in climate, hydrology, and vegetation, which impacted the soils. Human-induced modifications of the soil began with the hunter-gatherer populations that inhabited this region during the Early Hunter Phase (11,000 to 7,000 years ago), by participating in early domestication of crops. A more significant impact began to occur during the Archaic Period (7,000 to 4,000 years ago) when large scale clearing and more intense agriculture began (Piperno and Pearsall 1998). Eventually, farming, land clearance, and specifically wetland agriculture in northern Belize, led to the creation of the Maya Clays in the lowlands between roughly 3,000 and 400 years ago (Curtis, et al. 1998).

The Yucatan Peninsula consist of five main soil types: entisols and inceptisols, rendolls, alfisols, histosols, and vertizols. The first three of these types can be described as young, fertile, thin, clayey soils that can be up to 50 cm in depth. They are formed from organic matter and high amounts of carbonates and silicate minerals, supplied by the local limestone bedrock. Alfisols are older, red, iron-rich, less fertile, and generally thin, clayey soils. They also form from the limestone bedrock but organic matter decomposes more in these soils (Beach, Luzzadder-Beach and Dunning 2006).

In northern Belize, along the Rio Hondo where a large number of wetland fields have been discovered, soils are predominantly vertisols (Beach, Dunning, et al. 2003, Dunning, Jones, et al. 2003). Typically found between 45° north and 45° south latitude,

vertisols are dark clays which are known to shrink and swell in response to seasonal fluctuations in moisture. Over time, the change in moisture, which causes constant expanding and contracting, leads to large cracks forming in the soil profile. However, vertisols are excellent for agriculture due to their ability to retain large amounts of water. (Beach, Dunning, et al. 2003).

The vertisols of northern Belize, on which the Maya once farmed, now comprise a paleosol that is buried under a layer of sediment. The term paleosol is used to describe an ancient living surface and generally exist in two different states: on the surface of the landscape or buried within stratigraphic layers. Serving as indicator of human occupation, buried paleosols often appear as dark, well developed soil horizons between less-well developed stratigraphic layers. The dark color, resulting from an increase in organic material, may also indicate human activity such as trash pits and fire hearths. They tend to be wetter than soil layers above or below, since they are usually below the water table. They also tend to be more yellow in color, resulting from reduction in oxidation, a process known as gleying. They also often have poor horizonation due to their expanding and contracting nature (Beach, Dunning, et al. 2003). In Belize, insoluble minerals have accumulated in the soil which further classifies them as calcisols. Calcisols have a pronounced calcic horizon that contains calcretes and caliche and are commonly found in soils along the Rio Hondo (Beach, Dunning, et al. 2003, Rapp and Hill 2006).

The Rio Hondo, which starts in the highlands of Guatemala and Mexico then flows to the Caribbean Sea, forms the border between Belize and Mexico and contains a bifurcation at its middle known as Albion Island. The length of the river's banks are

marked by sinkhole lakes and wetlands. Past excavation of the Rio Hondo's stratigraphy revealed a complex history. It is constructed of a base made of gleyed clay, overlaid by peat and silt bands that alternate between dark and light colors. Above this is a layer of clay that is rich in calcium carbonate and Maya artifacts. The organic middle of the soil profile is a remnant of Maya agricultural activity. Nearby canal systems are a snapshot of the most recent Maya farming since the canals were regularly maintained and excess soil was removed. Therefore it is difficult to determine when they were originally constructed (Stein 1990).

Although it has been presumed that the Maya practice of swidden agriculture adequately allowed for soil nutrients to be restored, it has been a surprise to some that there has been no evidence of an organic fertilizer. One theory is that they utilized algae, also known as periphyton, as a fertilizer. Periphyton is abundant in the wetlands of the region and is an important nutrient source. Since phosphorus levels in calcaric soils are often low, this would provide the needed nutrients for wetland farming (Beach, Dunning, et al. 2003).

Phosphorus in the Environment

Phosphorus (P) is essential for plant growth and is one of the main ingredients in fertilizers. P is used by the plant to transfer energy and is constantly being removed from the soil when land is used for agriculture. It is not toxic by itself, but it can have a negative effect on water quality. Soils often retain most P but some can be released into water. Since it is found at higher levels in finer sediments, this suggests erosion will

move phosphorus out of the soil, potentially polluting nearby water bodies (Busman, et al. 2002).

P, as found in nature, is often part of a chemical compound. Soil can contain a number of phosphorus compounds, or phosphates, which together make up the total amount of inorganic phosphorus. P cycles through the environment much like other mineral nutrients. It is absorbed by plant roots, transferred to animals by consumption, and then returned back to the soil as waste. It exists in the soil as several different forms. The term "total phosphorus" refers to all forms collectively. Solution P is the small amount that is used by the plants. Active P refers to that which is returned or added to the soil. Active P is a much larger percentage of the total P and its role is to replenish the solution P as it is removed by plants. Fixed P is the result of microorganisms breaking down organic P and becomes mineralized. This form of P is often crystalline and undergoes minimal conversion back to organic forms. Therefore it adds little value to the soil's nutrient levels (Busman, et al. 2002).

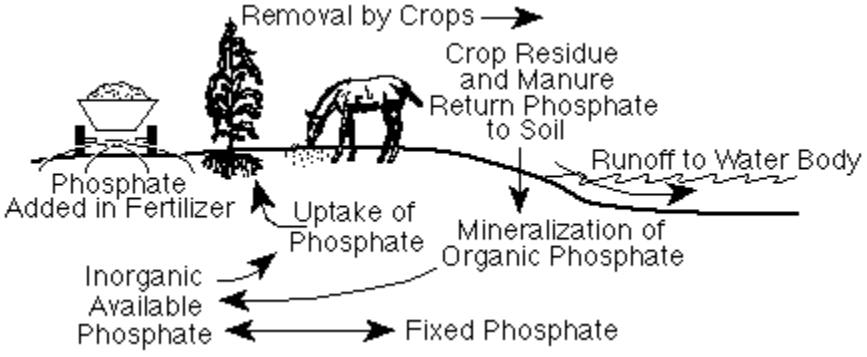


Figure 40: Phosphorus cycle (Busman, et al. 2002)

As P travels through this cycle, it exists as either organic (Porg) or inorganic (Pin), depending on where it is in the cycle. This is significant to archaeologists because the amounts of these two categories of P can provide clues to the type and intensity of human activity that may have occurred in a specific location. Humans add P to the soil through the disposal of garbage and waste, burials, and fertilizer. Subsequently, if fertilizer is not continuously added to soil then agricultural activity will result in a depletion of P. When it is first deposited in the soil, P exists as phosphate which is stable and mostly immobile once it has been absorbed in the soil. It is also less susceptible to leaching than other chemical elements used as indicators of human activity, so for this reason it is one of the most used by archaeologists (Holiday and Gartner 2007).

Analysis of Phosphorous in Archaeology

Anthropogenic biogeochemistry can be mapped out to create a three-dimensional picture that identifies human-induced soil features such as graves, trash pits, and even agricultural areas. One way this is done is through soil coring which allows for the measurement of soil nutrients at various depths. A number of soil nutrients are critical for plant life, such as nitrogen, phosphorus, potassium, calcium, magnesium, and sulfur. When humans remove vegetation, the nutrient level in soil decreases. Nutrients are added to the soil where humans discard food and waste. P is especially interesting to archaeologists since it doesn't fluctuate in soil as much as much as other nutrients, aside from soil erosion due to mass wasting and removal, and is a strong indicator of human activity (Rapp and Hill 2006).

Soil P is categorized by archaeologists into three types or physical states: easily extractable (associated with agriculture), tightly bound (associated with human activity), and natural geologic phosphate. Although not always a straightforward interpretation, the analysis of P in a soil is a valuable tool that continues to be improved by archaeologists (Rapp and Hill 2006). Soil P analysis in archaeology can be grouped into three main categories: prospection to locate or delineate sites, the delineation of specific activity areas, and to examine past agricultural practices.

There are several methods for measuring phosphorus in agricultural soils, one of the most acceptable is to extract the soluble phosphorus and then measure it quantitatively. The Mehlich-III test employs this methodology and has been used in recent years by archaeologists in Mesoamerica, as well as earlier versions (Mehlich and Mehlich-II) (Terry, et al. 2000). However, while the Mehlich-III method can be conducted fairly easily and inexpensively in the field, a faster and potentially less costly method is reflectance spectroscopy. Using a hand-held spectrometer, such as the FieldSpec 4, instant results may be provided and unlimited tests conducted. The results of this testing may also produce more accurate data as well, considering the reduction of steps and processing required in comparison to chemical tests (Bogrekci and Lee 2005).

Reflectance Spectroscopy and Phosphorous

Reflectance Spectroscopy is the study of electromagnetic radiation and its interaction with matter. It is a powerful scientific tool for determining material composition. It may be applied in the field with a portable device as well as through the use of remotely sensed imagery, specifically hyperspectral, which contains hundreds of

spectral bands throughout the electromagnetic spectrum. Geologic material (e.g., rocks and sediments) often possesses unique chemical compositions that produce equally unique spectral signatures. Hyperspectral imaging is therefore an ideal tool for geologic mapping, especially in areas with little or no vegetation. In regions such as northern Belize, where dense vegetation prohibits the use of hyperspectral imaging for the analysis of soil, reflectance spectroscopy can still be conducted by collecting spectral measurements in the field using a hand-held spectrometer (Eismann 2012).

The electro-optical (EO) region of the electromagnetic spectrum ranges from 400 nm to 14,000 nm (0.4 μm to 14μm) wavelength. The two main sources of energy for passive remote sensing are reflected sunlight and thermal emissions. The propagation of electromagnetic radiation, or radiative transfer, from material in an image scene to the sensor is affected by a number of factors such as other objects in the scene, and the atmosphere. The spectral radiance is defined in Equation 6 (Eismann 2012).

Equation 6: Spectral radiance as measured by an electro-optical satellite sensor

$$L_p(\lambda) = \frac{\tau_a(\lambda)\rho(\lambda)}{\pi} [E_s(\lambda) + E_d(\lambda)] + L_a(\lambda)$$

A benefit of the use of a hand-held spectrometer is that a number of these factors may be removed or reduced if there are minimal environmental influences, such as in a lab environment, allowing for a more accurate spectral measurement (Clark, et al. 2003).

In the last two decades, advances in reflectance spectroscopy have allowed for improved analysis of phosphates in soil. During this time, near- and mid-infrared

wavelengths specifically have proven to be successful in detecting soil compounds, including phosphates. Bogrekci and Lee (2005) report on a study to detect four different phosphates from soil samples taken from the Lake Okeechobee basin in Florida. The samples were leached using HCl and de-ionized water to remove any existing phosphorus. Different amounts of the phosphate compounds were then added: $CaPO_4$ (calcium phosphate), $AlPO_4$ (aluminum phosphate), $FePO_4 2H_2O$ (iron-III phosphate dihydrate, and $Mg(PO_4)2H_2O$ (magnesium phosphate dihydrate). After seven days of incubation, spectral measurements were taken of the samples. The samples were then oven-dried and sent to a lab for chemical analysis of total phosphorus concentrations. The results showed that absorbance spectra for the four compounds each had distinct absorbance bands in the near-infrared and ultraviolet regions of the electromagnetic spectrum. The spectral measurements were also found to be positively correlated with the results of chemical tests through the use of a partial least squares statistical analysis test (Bogrekci and Lee 2005).

Methodology

This study has two main research goals – and thus two hypotheses to test. The first is to determine if there is a significant difference in P levels between ancient wetland agricultural fields and areas where this type of agriculture likely did not exist. The second goal is to determine if this difference can be detected using reflectance spectroscopy. If both null hypotheses are rejected, this will indicate that reflectance spectroscopy has the potential to be a valuable tool for additional research on ancient wetland agriculture – in the field and in the laboratory.

Hypothesis #1:
H_o: There is no significant difference in P between ancient wetland fields and non-fields.
H_a: There is a significant difference in P between ancient wetland fields and non-fields.

Hypothesis #2:
H_o: Reflectance spectroscopy cannot detect P in soil samples.
H_a: Reflectance spectroscopy can detect P in soil samples.

Soil samples were collected in northern Belize in 2014, and also from three modern farms in northeast Ohio in 2014. Spectral measurements were then collected from each sample prior to undergoing the Mehlich-III chemical test. Both the Belize and Ohio samples underwent the same analytical process, the main difference being that the Ohio samples were collected and spectrally measured in a more controlled environment. Finally, the spectra and P levels were analyzed using partial least squares regression (PLSR) to answer the research questions.

Soil Sampling Methods in Ohio

Forty-five soil samples were collected from three modern agricultural fields in Trumbull County, Ohio, and were used as a control sample for the Belize soil analysis. Each field has been managed differently, growing different crops and using different fertilizers, and was expected to have phosphorus levels that were significantly different. The three locations are annotated on the map below.

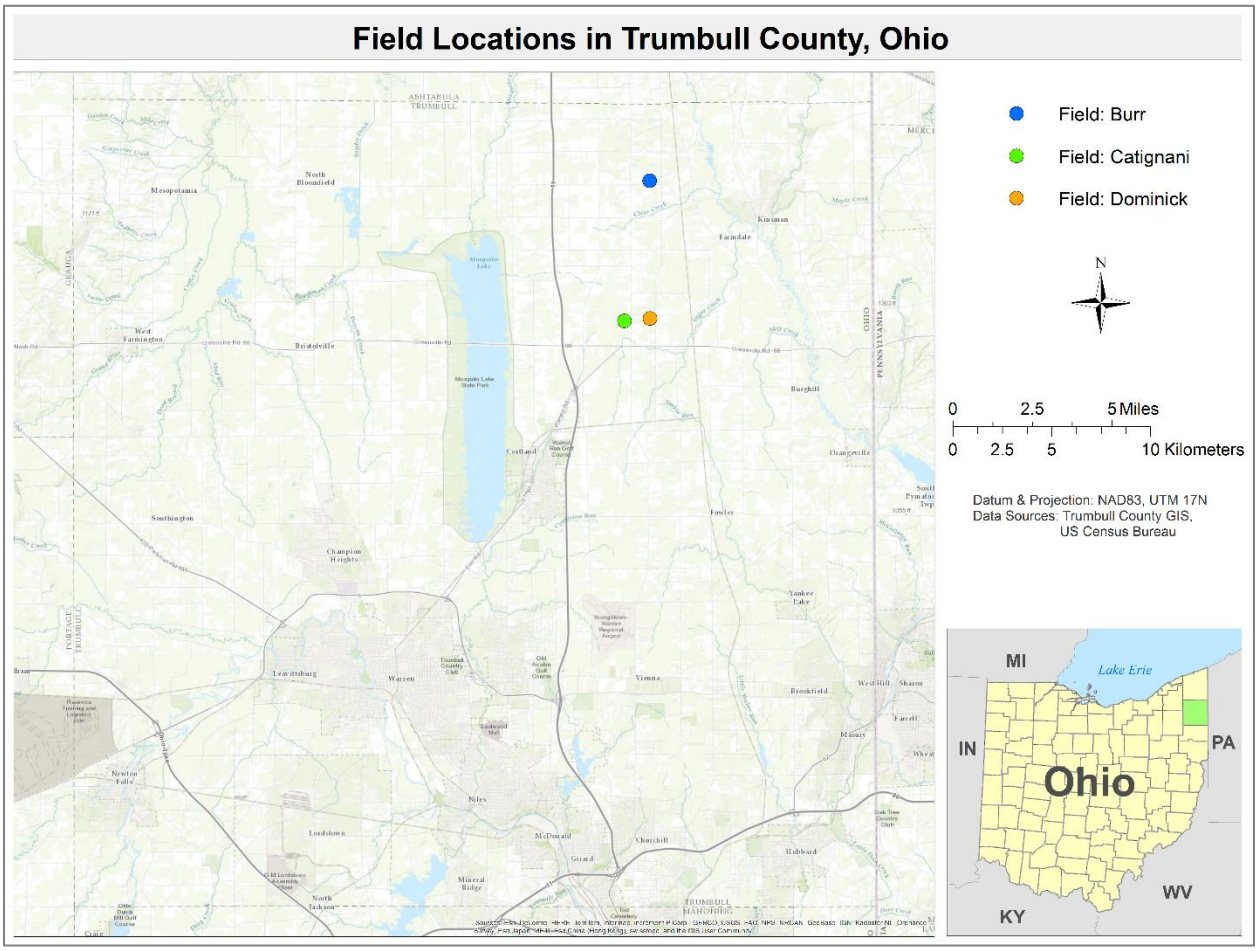


Figure 41: Soil sample field locations in Trumbull County, Ohio.

The Catignani farm was forested until the early 1970s when it was cleared and the top soil removed. It was left fallow until the mid-1990s when it began being used as an agricultural field to grow soy beans, corn, and hay. For the last seven to ten years it has again laid fallow. During its use, the agricultural fields were fertilized using manure only. The property contains a number of land types (agriculture, forest, residential) and so soil samples were collected from the fields and surrounding areas using a stratified random sampling method, to ensure representation from each type. Stratified random sampling

divides samples among subsets, or in this case land types, and sample locations are selected randomly within each subset. For the Catignani farm samples, eight were taken from the agricultural land, another eight were taken from the residential area, and four were taken from just outside the agricultural land along the pond and in the forested area. Another version of this method is known as stratified systematic sampling which also divides samples among subsets, but then samples are selected along a grid or transect. Stratified sampling can generate the most accurate results if accurate subset sizes are known but if the sizes are incorrect then this can lead to unintentionally biased results (Zhang 2007). The two types of stratified sampling are shown below in Figure 42.

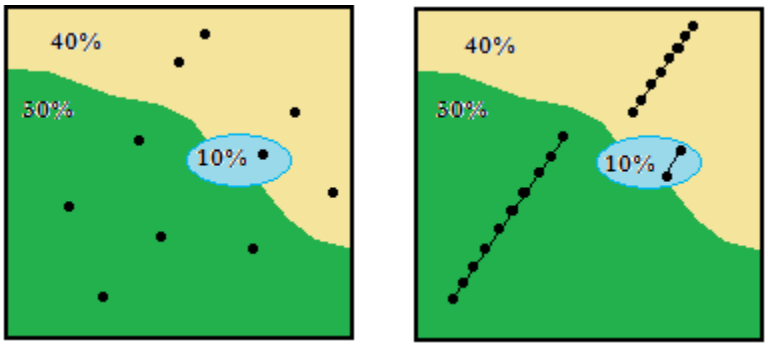


Figure 42: Stratified random and stratified systematic sampling methods.

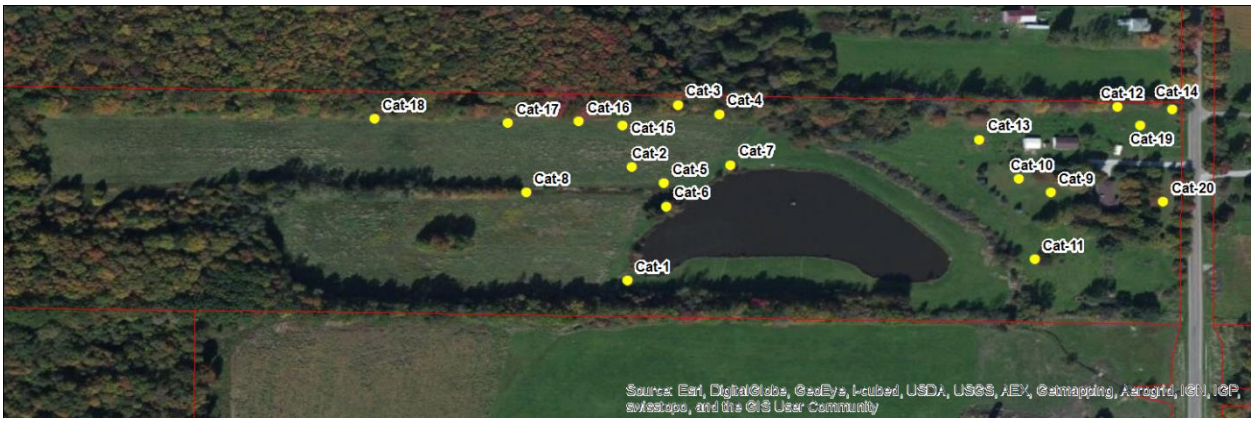


Figure 43: Soil sample locations at Catignani field and residential property.

The Burr farm is currently used to each year to grow hay for livestock. The land is fertilized using manure and chemical fertilizer. The property is mostly made up of agricultural land, and soil samples were collected from the fields and surrounding areas using a random point sampling method. To create a random line, one might produce pairs of random numbers and sample the line that connects the pair of locations. To create a random area, for example, the random selection point might delineate the bottom-left corner of a square of pre-determined size. The benefits of random sampling are that it is unbiased and can be used to sample populations of any size. The disadvantages are that if the population is too large it may produce a poor representation of the total population (Zhang 2007). An illustration of the three types of random sampling is show below in Figure 44.

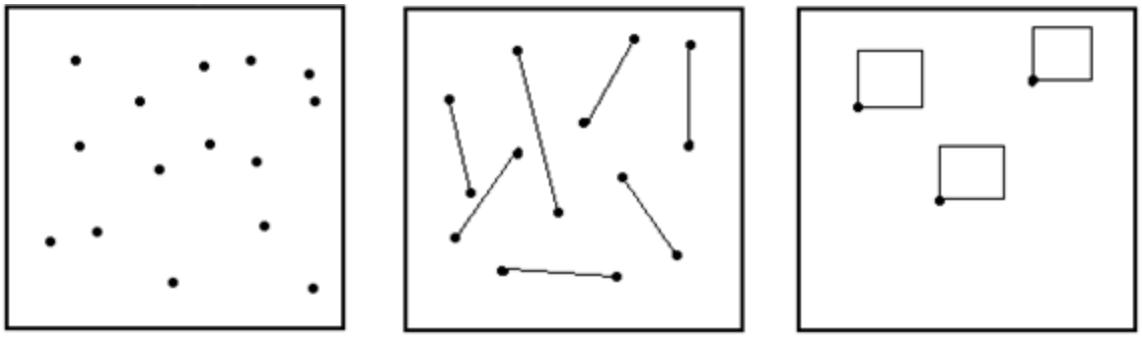


Figure 44: Random point, random line, and random area sampling methods

Figure 45: Soil sample locations at the Burr farm.

The Dominic farm is currently used to each year to grow corn. This farm has been completely organic for over 40 years and has only been fertilized using manure. Since this property was solely agricultural land and not mixed-use, soil samples were collected here using a systematic line sampling method. Systematic sampling is done by selecting locations that are evenly distributed such as on a grid or transect. This method can also be broken down into point, line, and area sampling. Systematic point sampling is done by producing a grid over the study area, then taking samples at each intersection of grid lines. Systematic line sampling uses the same principle but is done using a single straight line. Systematic area sampling is simply using either one of the previous techniques but sampling an entire area at regular intervals rather than points. The benefits of systematic sampling are that it provides good coverage of the entire population or study area. The disadvantage is that not all locations in a study area have an equal chance of being sampled since the regular interval will inherently skip areas and so this method is subject to the resolution of the grid or transect (Zhang 2007). An illustration of systematic sampling is show below in Figure 46.

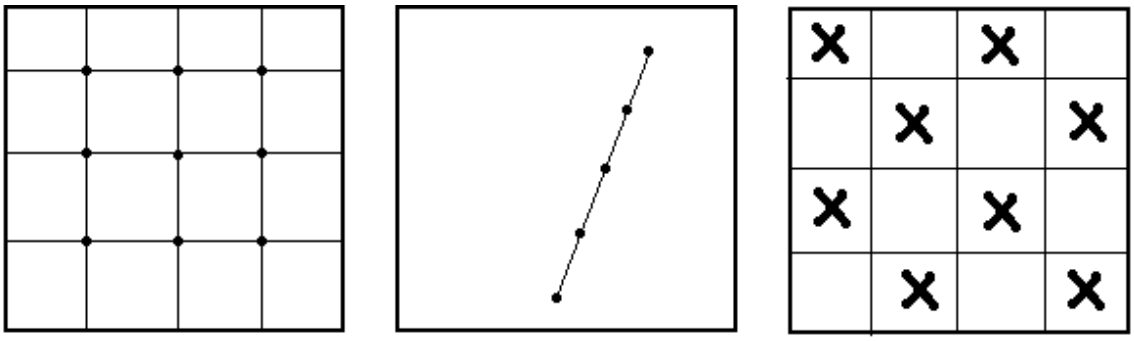


Figure 46: Systematic point, systematic line, and systematic area sampling methods

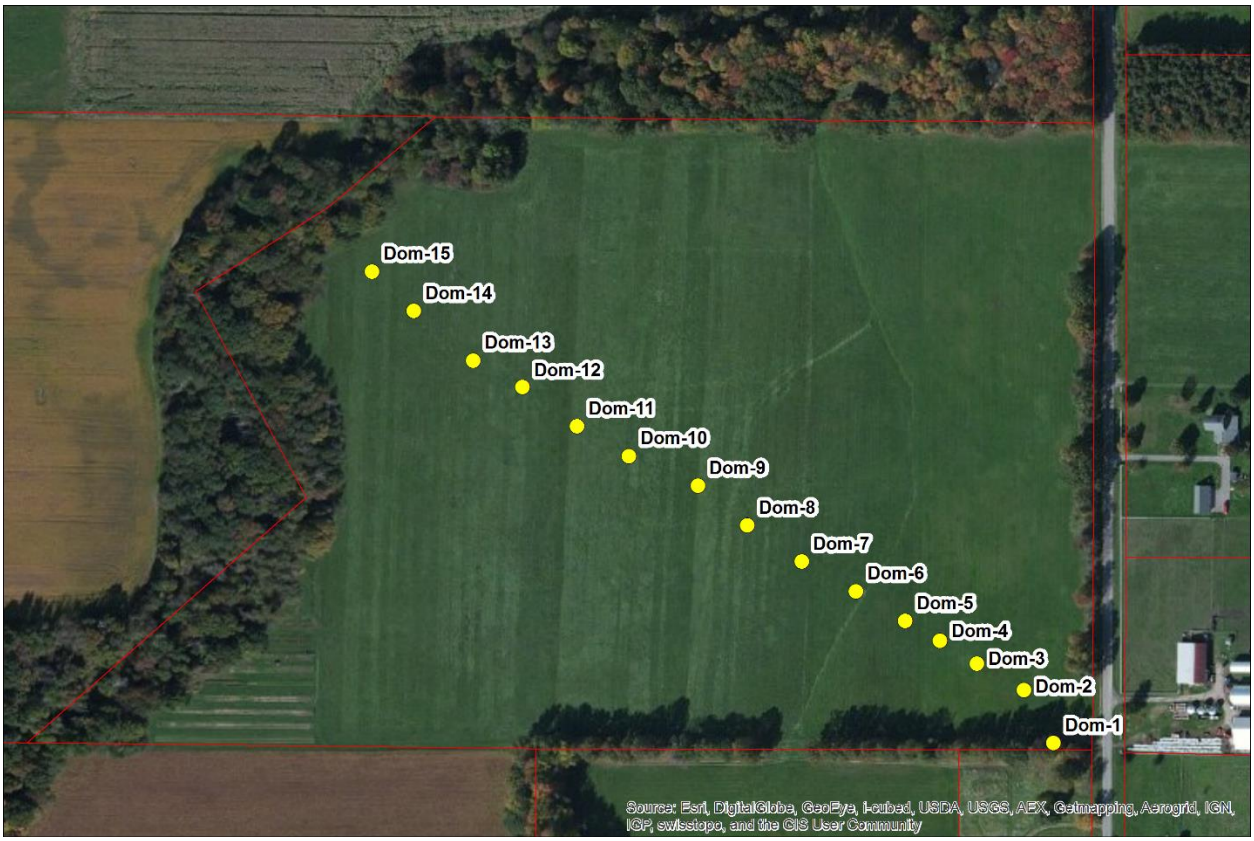


Figure 47: Soil sample locations at Dominic fields.

Spectral Analysis of Ohio Soil Samples

In April of 2015, the 45 soil samples from Ohio (collected during the previous year) were air-dried for three days and then brought to the soil spectroscopy lab at the US Geological Survey (USGS) in Reston, VA. This lab does not have natural lighting or windows, allowing lighting, temperature, humidity, and airflow to be constant, which provided a very controlled atmosphere. The samples were then passed through using a 2 mm sieve. Five spectral measurements were taken for each sample, rotating the sample

plate prior to each consecutive measurement. The five measurements were later averaged to produce more accurate values.

Figure 48: Work space at the USGS Soil Spectroscopy lab in Reston, VA

Soil Sampling Methods in Belize

The soil samples used for the Belize analysis were taken from Ulmer (2015). The samples were collected from five areas in northern Belize which are depicted in the map below: Akab Muclil, Crocodile Lake, Crystal Creek, Laguna Verde, and RB73 (Rio

Bravo). The sampling strategy used in Belize can be described as stratified random sampling since a variety of environments is represented, split between both anthropogenic and natural soils. They were also chemically tested using the Mehlich-II which provides a range of soil nutrients and properties, including total phosphorus values, which can be found in Appendix C (Ulmer 2015).

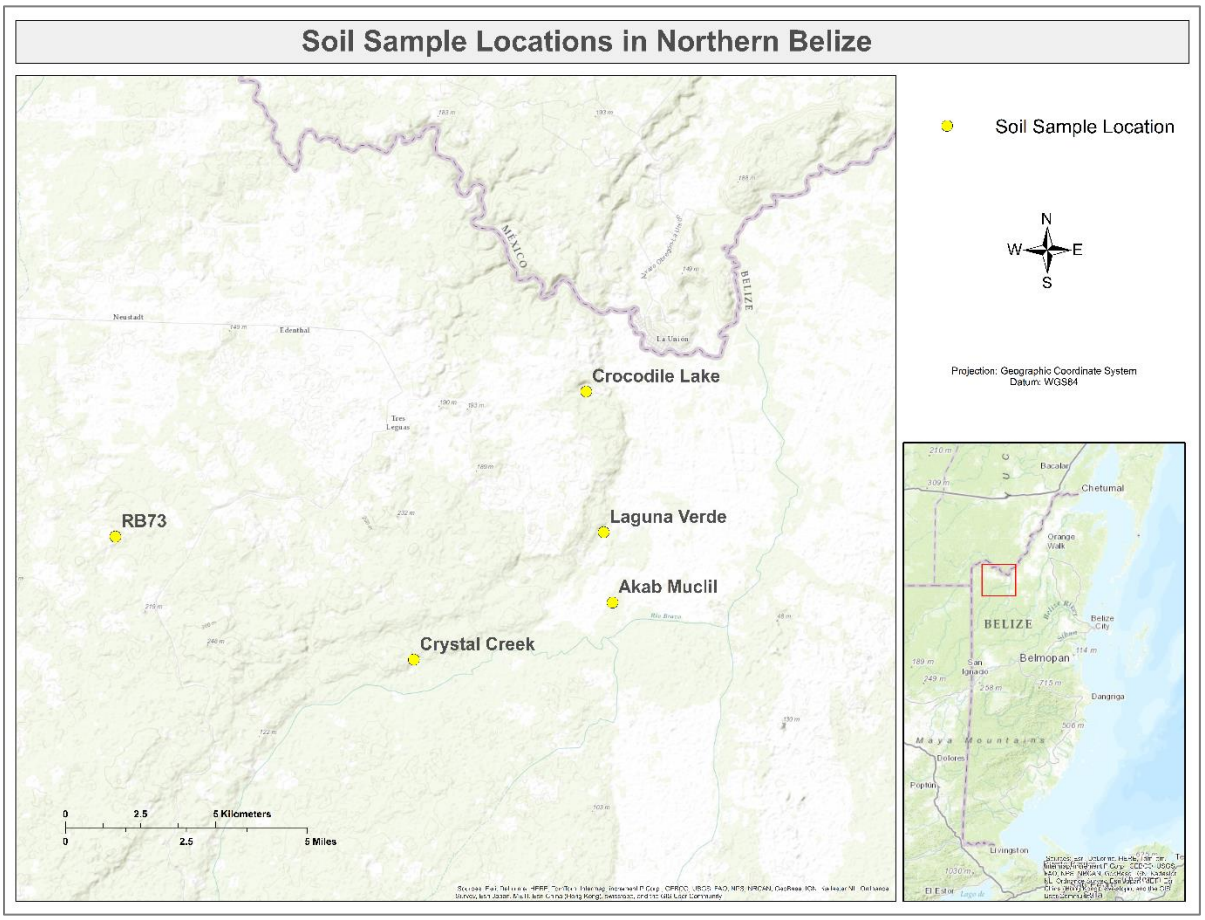


Figure 49: Soil sample field locations in northern Belize.

Akab Muclil – These samples were collected from a wetland area nearby the Maya site of Akab Muclil. Chemical analysis indicated that some samples contained evidence of maize cultivation (Ulmer 2015).

Crocodile Lake – This area is located in a canyon and the immediate area was not likely inhabited by the ancient Maya. However, surrounding areas may have been farmed. The samples collected in this area showed evidence of maize cultivation but it is unknown if they may have been contaminated by sediment from other areas (Ulmer 2015).

Crystal Creek – Samples from this area were collected at the toe slope of the Rio Bravo Escarpment in a valley that is heavily flooded during the rainy season. Some of the samples collected here showed weak evidence of maize cultivation while others showed very strong evidence (Ulmer 2015).

Laguna Verde – This lake is located at the toe slope of the Rio Bravo Escarpment. The toe slope contains no evidence of ancient farming. However, soils collected at the back slope of the escarpment do contain evidence of maize cultivation, likely through terracing (Ulmer 2015).

RB73 – This is a Maya site located near La Milpa, one of the largest Maya sites in Belize. Nearly all of the samples collected here showed no evidence of maize cultivation (Ulmer 2015).

Spectral Analysis of Belize Soil Samples

The Belize soil samples were analyzed in the Maya Research Program's archaeology lab, in July of 2014. The samples were not air-dried or sieved before taking spectral measurements and so the data reflects their natural state. This lab has both

natural and synthetic lighting which creates fluctuations in light, temperature, humidity, and airflow. Although the atmosphere was more controlled than the outdoors, there were significant challenges while using the FieldSpec 4 due to these fluctuations. Compared to working at the USGS, there was significant decrease in the time it took the instrument's white reference measurement to become degraded. The instrument was also much more sensitive to changes in light. In addition, the natural state of the samples, which contained small roots and rock fragments, likely altered the true reflectance spectra of the samples.

Figure 50: Operating the Field Spec 4 in the Maya Research Program's artifact lab.

Melich-III Soil Testing

The Belize soil samples were chemically tested by Ulmer (2015) using the Mehlich-II soil test. The Ohio soil samples were sent in spring, 2015, to the Pennsylvania

State University Agricultural Analytical Services Lab for chemical testing using the Mehlich-III soil test (see Appendix B for results). These tests are commonly used to measure soil micronutrients and determine the amount that should be added to the soil to improve crop production. Phosphorous is measured in one of three ways: inductively coupled plasma (ICP), the ammonium molybdate colormetric method, or the ammonium vanadate colorimetric method. Since ICP is the most efficient of the three, it is used in the majority of labs in the northeast United States (Wolf and Beegle 1995).

Partial Least Squares Regression

PLSR is a quantitative multivariate statistical analysis tool created to handle multiple independent and dependent variables and also employs dimension reduction to allow for analysis of data that are multicollinear and contains noise (Hecker, et al. 2012). It blends principal component analysis (PCA) and multiple linear regression to predict one or more dependent variables from a matrix of independent variables (H. Abdi 2010). This prediction is achieved by extracting from the predictors a set of orthogonal factors called latent variables which have the best predictive power. The resulting principal components are used in a regression to determine coefficient scores instead of the original data. Ultimately, this allows for the extraction of latent variables that are otherwise impossible to detect in multicollinear data (Hecker, et al. 2012). PLSR was specifically selected since phosphorus is not expected to have a strong impact on the reflectance spectra and its detection will require a tool for extracting latent variables.

The PLSR method can be explained as follows. To begin, most general multivariate regression analysis often starts with the equation

$$y = b_1 x_1 + b_2 x_2 + \ldots + \varepsilon$$

Equation 7: Multivariate regression

which can also be written in vector format such as the following

$$\boldsymbol{y} = [x_1, x_2, x_3, \ldots\ x_m] \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \ldots \\ b_m \end{bmatrix} + \varepsilon$$

Equation 8: Multivariate regression in vector format

where **y** is the predicted (dependent) value, **x** is a vector of independent variables, **b** is a vector of regression coefficients, and **ε** represents the residual error (Wehrens 2011). PLSR can be more easily written in vector form as

$$Y = XB + \ \varepsilon$$

Equation 9: Multivariate regression in matrix format, or PLSR

where Y is the matrix (or vector) of independent variables (phosphorus content), X is the matrix of dependent variables (spectra), B is the matrix of regression coefficients and ε is again the residual error. The structure of Eq. (3) is nearly the same as that of Eq. (1) but now will be used to derive PLS components in order to predict Y (here, phosphorus content) instead of relying only on the explained variance of the predictors. In other words, both the dependent and independent variables in PLSR can matrix-derived components (Garcia and Filzmoser 2011).

Principal components are derived from the matrix of spectral reflectance values, i.e. X, which will become a new matrix, T. Each value within T will represent a linear

combination of vectors in the original X matrix. Then, the T matrix will be used to predict values in Y and can be written as

$$Y = \boldsymbol{t}_1\boldsymbol{c}'_1 + \boldsymbol{t}_2\boldsymbol{c}'_2 + \ldots + \boldsymbol{t}_k\boldsymbol{c}'_k + \varepsilon \quad or \quad Y = TC' + \varepsilon$$

Equation 10: Equation of Y matrix, derived from PLSR components

where Y is equal to the sum of all the products of T vectors and the transposed matrix of coefficients vectors, plus the residual ε (Sanchez n.d.).

The PLSR analysis was implemented using packages for the R statistical analysis and programming language environment. This statistical software has a robust set of extensions and graphical capabilities that make it more efficient than Excel or SPSS for this work since it allows for greater customization and does not limit the maximum number of independent or dependent variables (R Development Core Team 2016).

Results

Hypothesis #1

A preliminary analysis of P differences was conducted on the soil samples collected from three agricultural fields in northeast Ohio. An F-test was first run on the samples to determine the equality of variance between datasets. The names of the three farms have been abbreviated in the tables below as CAT, DOM, and BUR.

Table 3: F-Test Two-Sample for Variances: CAT and DOM

	CAT	*DOM*
Mean	5.595	6.214
Variance	0.494	0.341
Observations	19	14
df	18	13

F	1.447
P(F<=f) one-tail	0.252
F Critical one-tail	2.484

Table 4: F-Test Two-Sample for Variances: DOM and BUR

	DOM	*BUR*
Mean	6.131	5.067
Variance	0.264	0.165
Observations	13	9
df	12	8
F	1.600	
P(F<=f) one-tail	0.257	
F Critical one-tail	3.284	

Table 5: F-Test Two-Sample for Variances: CAT and BUR

	CAT	*BUR*
Mean	5.595	5.067
Variance	0.494	0.165
Observations	19	9
df	18	8
F	2.993	
P(F<=f) one-tail	0.059	
F Critical one-tail	3.173	

The results of the three F-tests resulted in p-values that were all above the threshold of 0.05 which indicates the variances among the sample datasets collected at the three farms are not significantly different. Based on these results, a t-test assuming equal variances was used to determine if there was a significant different between P values of at the three Ohio farms.

Table 6: t-Test: Two-Sample Assuming Equal Variances: CAT and DOM

	CAT	*DOM*
Mean	5.595	6.214
Variance	0.494	0.341
Observations	19	14
Pooled Variance	0.430	
Hypothesized Mean Difference	0	
df	31	
t Stat	-2.683	
P(T<=t) one-tail	0.006	
t Critical one-tail	1.696	
P(T<=t) two-tail	0.012	
t Critical two-tail	2.040	

Table 7: t-Test: Two-Sample Assuming Equal Variances: DOM and BUR

	DOM	*BUR*
Mean	6.214	5.067
Variance	0.341	0.165
Observations	14	9
Pooled Variance	0.274	
Hypothesized Mean Difference	0	
df	21	
t Stat	5.130	
P(T<=t) one-tail	0	
t Critical one-tail	1.721	
P(T<=t) two-tail	0	
t Critical two-tail	2.080	

Table 8: t-Test: Two-Sample Assuming Equal Variances: CAT and BUR

	CAT	*BUR*
Mean	5.595	5.067
Variance	0.494	0.165
Observations	19	9

Pooled Variance	0.393
Hypothesized Mean Difference	0
df	26
t Stat	2.083
P(T<=t) one-tail	0.024
t Critical one-tail	1.706
P(T<=t) two-tail	0.047
t Critical two-tail	2.056

The calculated two-tail p-value for all three t-tests were below the threshold of .05, therefore the null hypothesis is rejected. This shows that there is a significant difference in phosphorus levels between all three farms.

In order to determine whether there is a significant difference in P between natural soils and the anthropogenic soils of the ancient wetland agricultural fields of northern Belize, data were acquired from Ulmer (2015). Fifty-two P values were provided for soil samples collected at five locations in northern Belize in 2014. The samples were collected from both areas known to contain wetland fields and areas that are believed to have not been used for this purpose. These values are listed below.

Table 9: Natural versus Anthropogenic P values for Belize soil samples

Natural	Anthropogenic
5.3	15.1
5	5.1
4.8	5.2
9.4	5.7
5.3	4.5
3.9	5.7
5.2	4.7
5.2	4.3

5	3.7
4.3	4
4.5	3.6
5.7	3.7
5.4	3.1
5.3	2.5
10.1	3.7
3.8	3.3
3.7	3.7
3.4	3.6
3.3	5.5
3.5	6.3
2.6	6
4	5.9
3.3	4
3.9	
3.5	
3.4	
3.4	
3.7	
3.5	
2.8	
3	

A preliminary F-test was run on these data in Microsoft Excel to determine the equality of variances. The results showed that the variances of the natural and anthropogenic soil datasets were significantly different, with F=2.43 and p=0.02.

Table 10: F-Test Two-Sample for Variances: Natural and Anthropogenic

	Natural	*Anthropogenic*
Mean	4.463	4.445
Variance	2.809	1.156
Observations	30	22
df	29	21
F	2.430	
P(F<=f) one-tail	0.019	

F Critical one-tail	2.016

Based on these results, a t-test assuming unequal variances was used to determine if there was a significant difference in P values between natural soils and anthropogenic wetland agricultural field soils.

Table 11: t-Test: Two-Sample Assuming Unequal Variances: Natural and Anthropogenic

	Natural	*Anthropogenic*
Mean	4.463	4.445
Variance	2.809	1.156
Observations	30	22
Hypothesized Mean Difference	0	
df	49	
t Stat	0.047	
P(T<=t) one-tail	0.481	
t Critical one-tail	1.677	
P(T<=t) two-tail	0.963	
t Critical two-tail	2.010	

The calculated t-statistic is 0.047 and the two-tail p-value for this test is p=0.96. Since the p-value is greater than 0.05, the null hypothesis is accepted that the two samples do not have significantly different levels of phosphorus.

Hypothesis #2

PLSR of Ohio Samples

Spectral reflectance was measured for the 45 samples from Ohio using the FieldSpec spectrometer. Each sample was measured five times, and the sample dish was rotated after each measurement. Then, the five measurements were averaged. The data

were imported into R. Each spectrum contains 2,151 reflectance measurements from 350 nm to 2500 nm. The 45 spectra and the Mehlich-III derived P values were used to conduct PLSR analysis using the "pls" package (Mevick, Wehrens and Liland 2016), where P is the dependent variable and the spectra are the independent variables. The complete matrix used in this analysis contains 45 rows and 2152 columns, the last being the P values.

Following the methods provided in the package documentation, the PLSR model was created using the following command which uses the P values and VNIR spectra as the Y and X data respectively. It also creates ten components and selects cross-validation as the validation method. The results of this model are provided below.

```
> Ohio_pls = plsr(Oh_P ~ Oh_VNIR, ncomp = 10, data = Oh_Spectra,
validation="LOO")

> summary(Ohio_pls)
Data:   X dimension: 45 2151
        Y dimension: 45 1
Fit method: kernelpls
Number of components considered: 10

VALIDATION: RMSEP
Cross-validated using 45 leave-one-out segments.
      (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
CV          8.804    9.019    8.704    8.755    8.523    8.673    8.658
adjCV       8.804    9.024    8.697    8.740    8.511    8.661    8.647

       7 comps  8 comps  9 comps  10 comps
CV       9.035    11.56    10.77     10.45
adjCV    9.008    11.46    10.69     10.38

TRAINING: % variance explained
     1 comps  2 comps  3 comps  4 comps  5 comps
X    98.4877   99.31    99.47    99.87    99.93
Oh_P 0.4905    12.88    24.97    28.16    32.66

      6 comps  7 comps  8 comps 9 comps  10 comps
X       99.97    99.98    99.98   99.99     99.99
Oh_P    37.22    45.06    65.84   71.02     75.07
```

The model generates two tables of information, the first provides the root mean squared error of prediction (RMSEP), the standard deviation of residuals indicating how far the data points are spread away from the regression line. The RMSEP is calculated for two cross-validation estimates: the cross-validation (CV) estimate which segments the data for evaluation, and the adjusted cross-validation (adjCV) which attempts to correct for the reduced performance of the model on smaller subsets of a larger dataset. (Mevik and Cederkvist 2004) The method of CV used here is Leave One Out (LOO). The second table of information generated by the model contains calculations of variance between the original and predicted y values and each of the components.

The lowest cross-validation values occur at four components. Figure 51 below allows for the visual inspection of the RMSEP values by plotting them as a function of the number of components.

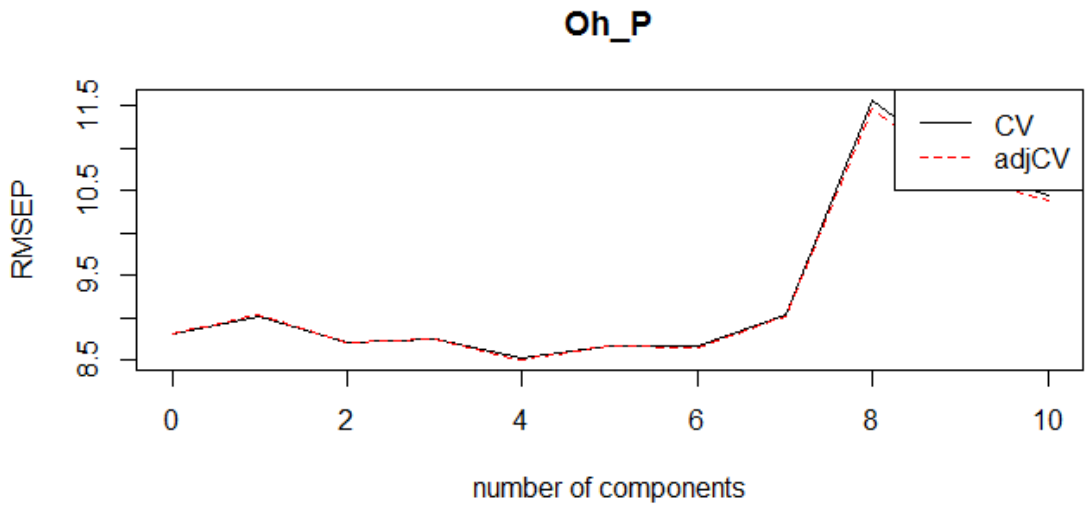


Figure 51: RMSEP as a function of components for Ohio PLSR model

The first four components can be further examined using the style of plot shown in Figure 52 which reveals the amount of variance shared between each of the components.

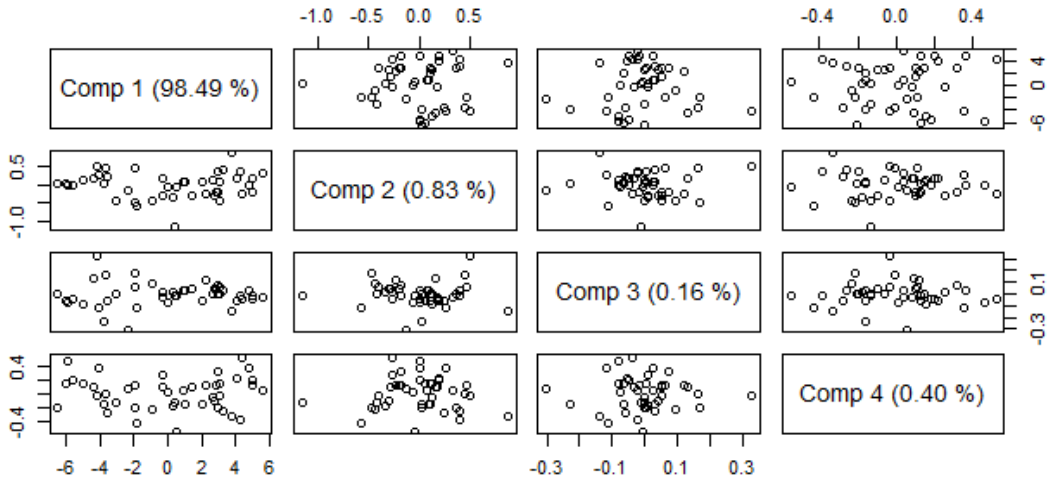


Figure 52: Explained variance for Ohio PLSR model

The fit of the model is displayed in Figure 53, which plots the original versus predicted y values.

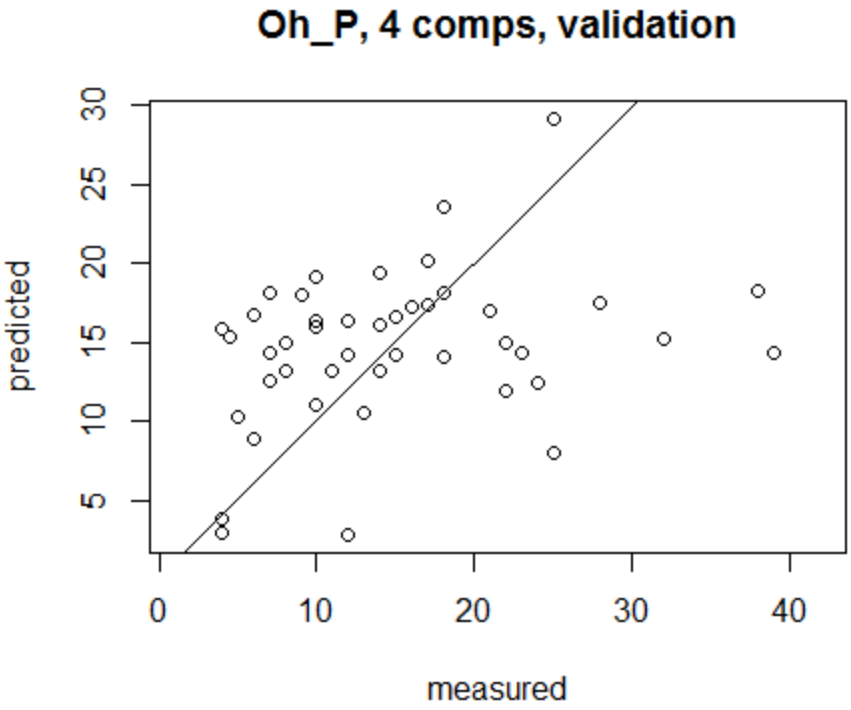


Figure 53: P measured versus predicted for Ohio PLSR model

The final step is to calculate the R^2 for the model, using the original P values measured using the Mehlich-III test and the predicted values. This value represents the amount of variance in the dependent variable, the measured y values, that can be predicted from the independent variables, or components. The R^2 for this model is 0.1524213.

```
> y = Oh_Spectra$Oh_P
> ypred = (Ohio_pls$fitted.values[,,4])
> SST = norm(y-mean(y), type = "2")
> SSE = norm(y-ypred, type="2")
> R2 = 1-(SSE/SST)
> R2
[1] 0.1524213
```

PLSR of Belize Samples

Spectral reflectance was measured for 30 of the samples collected by Ulmer (2015) using the FieldSpec 4 spectrometer. Each sample was measured three times, and the sample dish was rotated after each measurement. Then, the three measurements were averaged in Microsoft Excel to create a more accurate spectral representation of the soil. The data was imported into R and the PLSR was conducted using the final 30 averaged spectra and the Mehlich-II derived P values (Ulmer, 2015). Eight of the wavelengths contained null values, leaving 2143 complete spectra. The complete matrix for this analysis contains 30 rows, or observations, and 2144 columns, or variables. P is once again the dependent variable, and the spectra are the independent variables.

The results of the PLSR model are provided below, which was created using the the P values and VNIR spectra as the Y and X data respectively.

```
> Belize_pls = plsr(Bz_P ~ Bz_VNIR, ncomp = 10, data = Belize_SpectraP,
validation="LOO")

> summary(Belize_pls)
Data:   X dimension: 30 2143
        Y dimension: 30 1
Fit method: kernelpls
Number of components considered: 10

VALIDATION: RMSEP
Cross-validated using 30 leave-one-out segments.
      (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
CV          2.502    2.568    3.017    3.988    4.558    4.403    4.522
adjCV       2.502    2.565    3.003    3.953    4.501    4.357    4.468

       7 comps  8 comps  9 comps  10 comps
CV       4.828    5.199    5.920     6.301
adjCV    4.765    5.131    5.826     6.198

TRAINING: % variance explained
      1 comps  2 comps  3 comps  4 comps  5 comps
X      93.649    98.59    99.44    99.54    99.73
Bz_P    3.915    14.37    19.04    44.92    51.81

      6 comps  7 comps  8 comps 9 comps  10 comps
X       99.79    99.84    99.88   99.89     99.90
```

```
Bz_P    66.71    74.35    79.95    92.95     97.01
```

Figure 54 below is a plot of the Belize RMSEP values as a function of the number of components. For this dataset, the RMSEP continually increases, making it difficult to select the appropriate number of components to use. Since the RMSEP dips slightly at five components, this number was selected for the rest of this dataset's analysis.

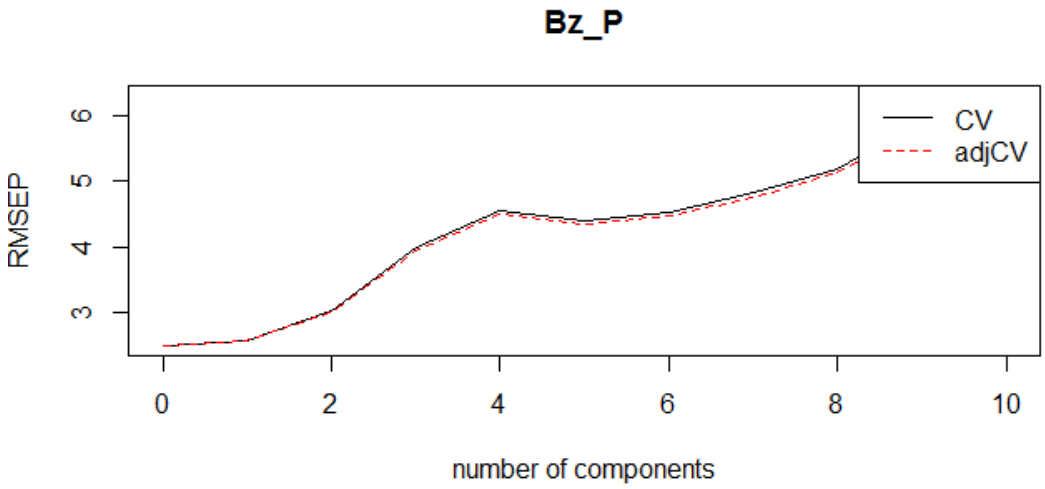


Figure 54: RMSEP as a function of components for Belize PLSR model

The explained variance for each of these components is shown in Figure 55.

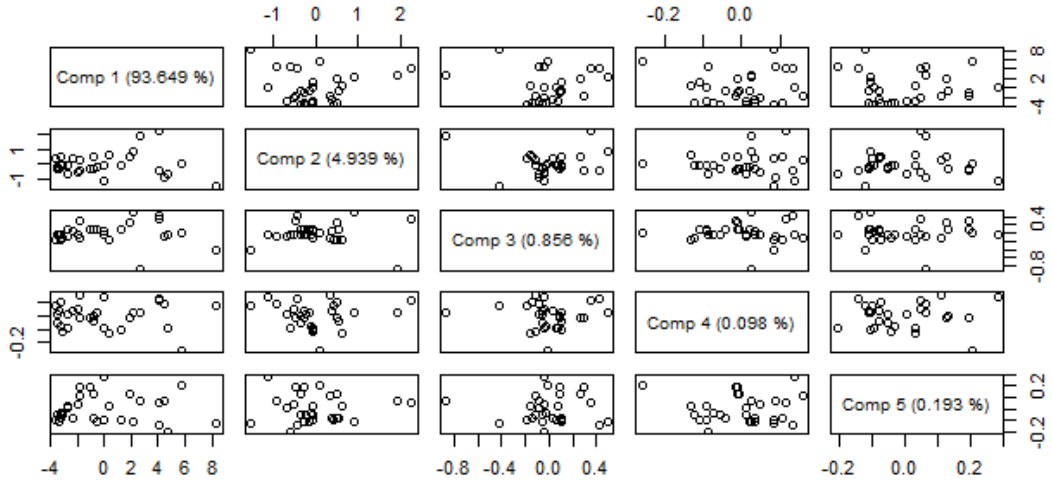


Figure 55: Explained variance for Belize PLSR model

A plot of original versus predicted values, using the first five model components, is shown in Figure 56.

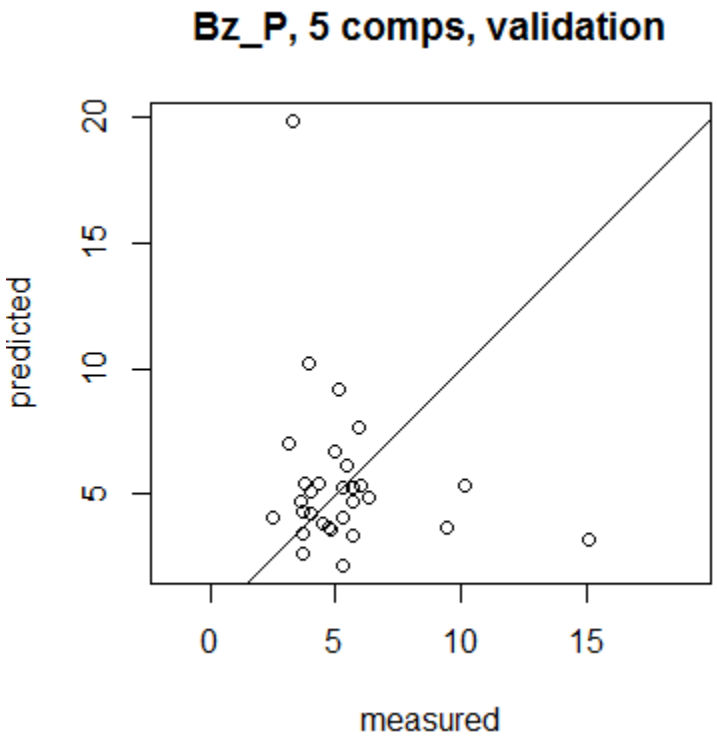


Figure 56: P measured versus predicted for Belize PLSR model

The final step is to use the test dataset to determine the model's performance. The P values for the test dataset are calculated below. The R^2 of the model is 0.305782.

```
> y = Belize_spectraP$Bz_P
> ypred = (Belize_pls$fitted.values[,,5])
> SST = norm(y-mean(y), type = "2")
> SSE = norm(y-ypred, type="2")
> R2 = 1-(SSE/SST)
> R2
[1] 0.305782
```

Discussion

Although the models indicate that over 90% of the variance in each dataset can be explained, the R^2 shows that the overall predictive power is low. It would be possible to increase the number of components used in order to improve the predictive power but it

is necessary to consider the RMSEP of each component used. Therefore, adding too many components in order to improve R^2 would falsely improve the final results. In addition, the datasets used here are relatively small and future work should consider using a more robust dataset to determine the predictive power of PLSR and reflectance spectroscopy for the detection of soil phosphorus.

Conclusions

The results of the Belize and Ohio soil spectroscopy analysis show that it is possible to detect differences in P levels between different areas although there was not a statistically significant difference between anthropogenic and natural soils from Belize. It is likely that having larger sample sizes, and representing more distinct locations, may have produced better results. Also, selecting more samples from key wetland field areas, such as Albion Island, may have also improved the results. Regardless, the use of soil spectroscopy for archaeological analysis is still a new area of research, and should be explored in the future to gain new insights regarding human activity across the ancient Maya landscape. Applying these same methods to the long-wave infrared range of the electromagnetic spectrum may also provide future opportunities to measure soil P in the wetland fields. For decades, Maya archaeologists have appreciated the knowledge that soils offer towards understanding this ancient culture; soil spectroscopy is well suited to complement that area of research and provides a link for additional geospatial and remote sensing research in this region. Much of the remote sensing research in Maya archaeology has focused on lidar in recent years which certainly has improved our understanding of Maya cities and their spatial characteristics and relationships. However,

there are many other tools for geospatial analysis, including reflectance spectroscopy, which can provide new insight on the ancient Maya civilization. The technological advances that are currently on the horizon, such as new satellite imaging sensors and their data being made freely available, will be especially important for challenging landscapes such as in Belize. By branching out and exploring these opportunities, one day the ancient Maya may be far less of a mystery than they are today.

CHAPTER FIVE – COMPARISON OF VNIR, M-LWIR, AND XRF CAPABILITIES IN DETECTING SOIL PHOSPHORUS

Introduction

For decades, archaeologists have been interested in phosphorus in soils as a proxy for locating and understanding ancient human activity. Phosphorus, a relatively stable soil nutrient, has aided in the study of sites because its concentration is often the result of activities such as food preparation or disposal. Agriculture is also known to affect soil phosphorus levels. Fertilization may increase these levels while a continuous cycle of plant growth may remove nutrients from the soil and consequently decrease phosphorus concentration (Holiday and Gartner 2007). This research will evaluate and compare two non-traditional and a third more commonly used method for measuring soil phosphorus in agricultural areas: visible/near- and shortwave-infrared spectroscopy (VNIR/SWIR), mid to long wave infrared spectroscopy (M-LWIR), and x-ray fluorescence (XRF). These will be compared to the more widely used Mehlich-III soil nutrient test. The goal is to show which of these methods can produce the most reliable soil phosphorus measurements in order to avoid costly and laborious chemical-based methods used in archaeology today. In addition, the success of these methods would enable remote sensing based techniques for studying broader areas.

Much of the research on archaeological soil phosphorus has utilized the chemical based Melich-III soil nutrient test and has frequently been applied in ancient Maya research. The cost of such tests can quickly add up, becoming a limiting factor and making research at larger scales less attainable. In addition, it can be logistically

challenging to transport soils that originated from an archaeological site, as many countries place strict regulations on any materials related to their own heritage. The United States also has strict regulations on importing foreign soils that specify designated labs throughout the country where they must be sent, and also the exact procedures for their destruction once no longer needed. Essentially, although soil itself is cheap and easy to procure, there are a whole host of challenging obstacles that archaeologists must endure when incorporating them in their research. Reliable and largely chemistry-free methods for soil analysis are needed in order to advance archaeological research centered, especially on regional-scale studies. This research will attempt to build on the current body of knowledge surrounding this research need.

Background

The Use of VNIR/SWIR Spectroscopy in Detecting Soil Phosphorus

As mentioned in the previous chapter, research in VNIR/SWIR reflectance spectroscopy in recent years has shown progress towards the detection of phosphorus in soil. A handful of studies which used similar methods to those used here, found significant correlations in their analysis. Bogrekci and Lee (2005) detected phosphates from soil samples of the Lake Okeechobee basin in Florida. The results showed that absorbance spectra had distinct bands in the near-infrared and ultraviolet regions of the electromagnetic spectrum and were positively correlated with the results of chemical tests through partial least squares regression (PLSR) analysis (Bogrekci and Lee 2005).

Idowa et al. (2008) used PLSR to detect phosphorus in agricultural soils to determine whether it could be a less expensive and faster option than the Cornell Soil

Health Test. They collected spectra using a Panalytical FieldSpec Pro in the range of 350-2,500 nm and a spectral resolution of 1 nm. In total, 387 soil samples were air-dried and their spectra measured five times each for averaging. The PLSR ultimately showed that phosphorus was only moderately predictable, with an R^2 of 0.63, when using reflectance, as opposed to absorbance, moving averages, or the first derivative of the reflectance. More recently, Abdi et al. (2012) were able to predict total phosphorus in agricultural soils, using the 1100 to 2500 nm spectral range and PLSR, producing an R^2 of 0.78. However, other phosphorus-related properties such as P-uptake and P-budget were not successfully predicted.

Xue-Yu (2013) significantly predicted phosphorus among 448 soil samples using the 350 to 2,500 nm spectral range and PLSR to produce an R^2 of 0.65. The predicted phosphorus values were verified using the Student's T statistical test. In addition, other studies have attempted to model soil phosphorus using multispectral band combinations and indices such as the normalized difference vegetation index (NDVI) and leaf area index (LAI).

To date, reflectance spectroscopy-based soil phosphorus research in archaeology has focused on individual sites and the features they contain (Holiday and Gartner 2007, Matney, et al. 2014), but few if any have specifically applied this method to regional studies such as ancient agricultural areas. A related study used VNIR and midwave-infrared (MWIR) spectroscopy to study Amazonian Dark Earths (ADE), also known as terra preta or terra mulata, in Brazil (Araújo, et al. 2015). It is unclear whether ADE was the result of intentional or unintentional human influences, and whether they were created

through habitation, agriculture, or both (Glaser, 2007). Unfortunately, the research conducted by Araújo, et al. (2015) was not able to accurately predict phosphorus in soils applying PLSR to VNIR/MWIR spectra.

The Use of M-LWIR Spectroscopy in Detecting Soil Phosphorus

Mid to longwave infrared spectroscopy may also be a potential tool for the detection of phosphorus in archaeological contexts. Its use has been less documented but there is growing interest after a handful of studies in recent years on its application. For example, M-LWIR spectroscopy has been used to study the impact of weathering and rainfall on soils (Howington, Ballard Jr. and Wilhelms 2012, Ballard, Howington and Wilhelms 2013, Rozenstein and Karnieli 2015), and for mineral mapping (Riley and Janaskie 2012, Notesco, Ogen and Ben-Dor 2015). Nevertheless, the technology has been rarely used for the detection of phosphates. An early study (Pierzynski, et al. 1990) used Fourier-transform infrared (FTIR) and several other instruments to quantitatively analyze phosphorus-rich particles in soil but found the FTIR results were not successful. A handful of other studies have used FTIR to understand phosphorus sorption and molecular changes, but not specifically for its quantification in soil (Eisazadeh, Kassim and Nur 2012, Xue-Yu 2013).

An infrared spectrometer collects a spectrum of light displaying absorption features in regions where the infrared frequency matches the vibrational frequency of chemical bonds within the sample, in this case a soil sample. Groups of bonds can therefore provide unique information about a material's chemical properties. (Bradley 2016). The Nicolet FTIR spectrometer measures the intensity of an infrared beam after it

has passed through a suitably prepared sample. The resulting acquired signal is referred to as an interferogram which contains information about the material it passed through. The accompanying software called OMNIC inputs the interferogram and applies the fourier transform to recover the IR spectrum of the sample (Thermo Fisher Scientific 2016).

The spectrum, or energy curve, is displayed in arbitrary units on the Y-axis and frequency in the X-axis. The general shape resembles that of an infrared energy source, with dips indicating frequencies that were absorbed by the material. The OMNIC software program divides the sample energy curve by the background energy curve (created using a gold-plated disc) to produce a transmission spectrum, which represents the percentage of energy that passed through the sample at each frequency (Thermo Fisher Scientific 2016).

The Nicolet FTIR collects data in wavenumber, cm^{-1}, which can be converted to wavelength. The equation below converts the wavenumber 700 cm^{-1} to wavelength in µm.

Equation 11: Conversion of wavenumber to wavelength

$$\frac{1cm^{-1}}{700} \times \frac{1m}{100cm} \times \frac{1 \times 10^{9}nm}{1m} \times \frac{0.001\mu m}{1nm} = 14.29\mu m$$

The FTIR instrument used for this research utilizes an integrating sphere, a common accessory for collecting reflectance measurements of diffuse surfaces. Specifically, the reflectance collected by the integrating sphere is biconical (Figure 57), which means the

incident and reflected light are confined to narrow and specific ranges. The integrating sphere is ideal for obtaining accurate quantitative measurements for two reasons: 1) the highly reflective and uniform, or Lambertian, interior coating, and 2) the spherical shape which creates constant uniform irradiance of the sample surface, allowing it to be accurately collected by the detector (Hanssen and Snail 2002).

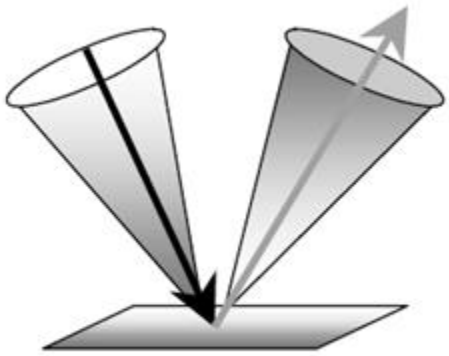


Figure 57: Biconical reflectance (Hanssen and Snail, 2002); Nicolet FTIR integrating sphere, USGS Reston VA.

The Use of X-Ray Fluorescence Spectroscopy in Detecting Soil Phosphorus

X-ray fluorescence (XRF) has been the most prevalent of these three techniques discussed here among archaeological applications. It has been proven to be a fast and cost-effective method for chemical element analysis in the field and has been utilized by archaeologists since the 1960s. The earliest applications focused on analyses of obsidian, ceramic, and other rock-related provenance (Shackley 2011).

In recent years there have been a number of studies that use XRF to understand intra-site patterns and those of specific archaeological soil features (Frahm, et al. 2016, Neff, Voorhies and Umaña 2012). Specifically, Mesoamerican studies incorporating XRF have frequently considered phosphorus among sites and soil features (Ulmer 2015) and have even been able to map linear phosphorus patterns at an ancient Maya plaza that likely served as a marketplace (Coronel, et al. 2015).

XRF requires calibration specific to each instrument as well as the material being analyzed (Shackley 2011). In a recent study, Speakman and Shackley (2013) highlighted a number of pitfalls that archaeologists often ignore and the complexities of XRF analysis that must be addressed prior to any quantitative analysis. Ultimately, they place the onus of responsible science on the user and not the technology itself (Speakman and Shackley 2013).

Ideally, a sample should be homogenous, and for this reason the XRF uses a homogenous metal alloy reference standard to maintain its calibration. Soil samples, however, are never truly uniform, and each reading will generate slight variations. In addition, humidity can cause samples to become denser, making it more difficult for x-rays to pass through the material. Elevation and changes in air density can also affect the consistency of XRF measurements (Bruker 2016).

When taking XRF measurements of soils in-situ, moisture should not exceed 20%. If it is necessary to do so, then additional measurements should be taken after the soil has been allowed to dry. Measurements of semi-prepared samples, where large debris has been removed, can be taken while the soil is in a plastic bag so long as it is not too

thick. Fully prepared soil samples are ones that have been sieved using a large mesh to improve homogeneity, and covered with 4µm polypropylene film which does not attenuate returning x-rays when measuring light elements to include phosphorus (Bruker 2016). For this research, soil samples were dried for over a week and ground using a mortar and pestle to ensure uniformity.

Study Area – Soils of Northeast Trumbull County Ohio

Trumbull County, Ohio was once covered by glaciers and the soils there are formed from glacial deposits. They are part of the Mahoning-Canfield-Rittman-Chili Soil Region, as defined by the Ohio Department of Natural Resources, and range from course to fine-textured. Trumbull County is relatively flat and the soils generally consist of silty loams with 0 to 6 percent slopes (Williams 1992).

One-hundred and thirty soil samples were collected from five locations in northeast Trumbull County, Ohio, during the summer of 2016. Four of them have been described in the previous chapter and they vary in the type of crops grown as well as their frequency of use. These five locations include four agricultural areas and a churchyard that has been in existence for at least 142 years (Everts 1874, US Geological Survey 1907). The first church to be built on this property no longer exists, but the current structure was built in 1899 (Figure 58) and is surrounded by around five acres of land.

Figure 58: Johnston Federated United Methodist Church, current structure built in 1899.

The samples collected around the church will be used as a baseline for the area's natural soil nutrient levels since this area has not been farmed since it was first settled and this location is likely to have experienced the least amount of disruption during that time. It is possible that it may have been used for agriculture prior to being settled by Europeans.

The sampling locations for this study are shown in the map in Figure 59. Twenty-six samples were collected from the fields labeled DOM, SUN, CAT, and Church. Twenty-seven samples were collected from the field labeled BUR for a total of 131 samples (Figure 59).

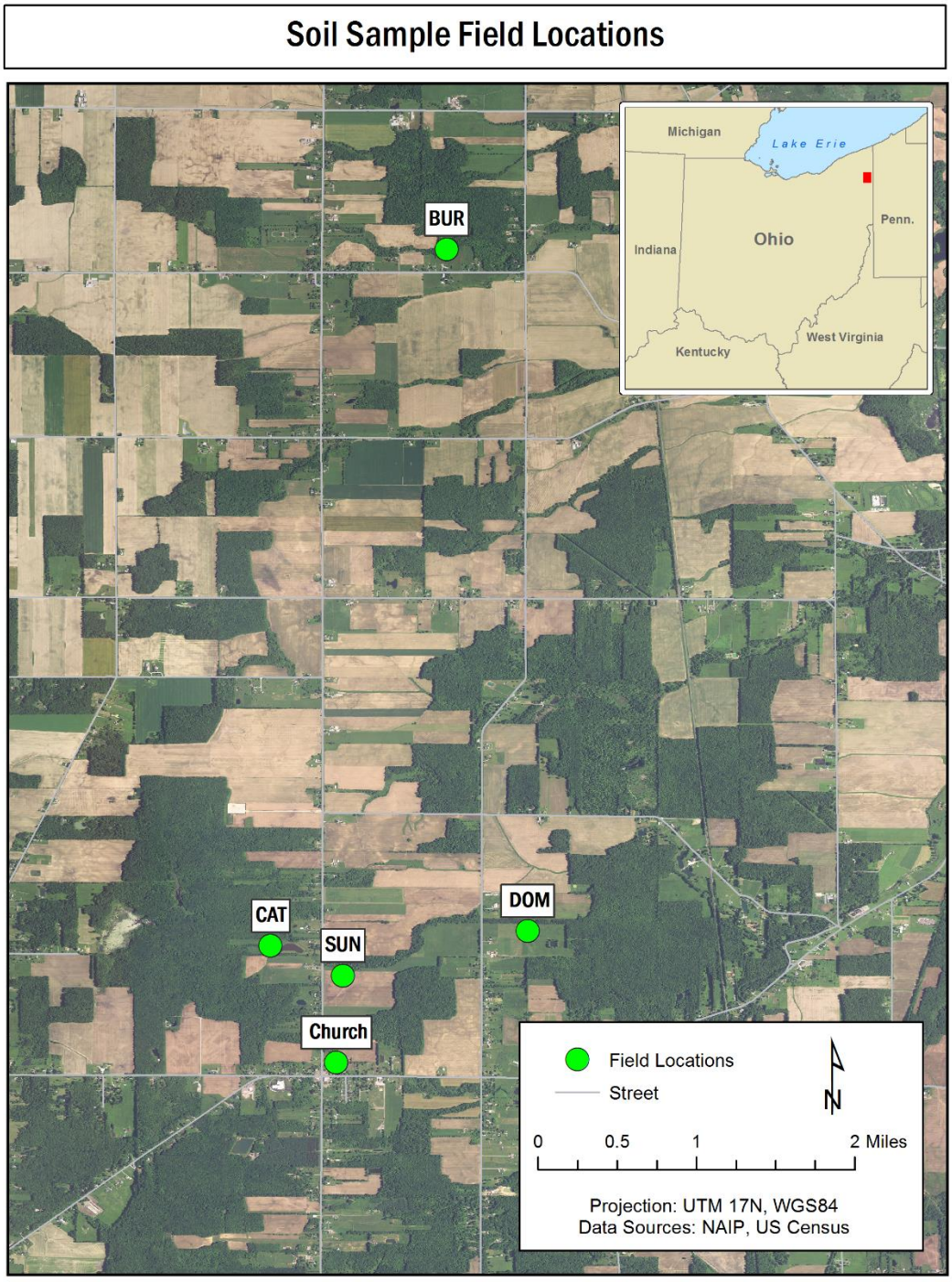


Figure 59: Soil sample field locations.

Methods

Although the goal is to determine which combination of methods would be best for field use, it is important to accurately portray the capabilities of the instruments prior to drawing conclusions about their potential. The soil samples used in this research were minimally prepared before taking measurements. First, large plant debris, insects, and rocks were removed from the samples. Then, the soils were air-dried for roughly one week in a lab at the USGS headquarters in Reston, VA prior to analysis. Each sample was then roughly split in half and one part was sent to the Soil Analysis Lab at Pennsylvania State University in State College, PA, for wet chemistry analysis using the Mehlich-III soil nutrient test. The test provides a range of soil nutrient and property values including phosphorus in parts per million (ppm).

An analysis of variance (ANOVA) was conducted on the resulting Mehlich-derived phosphorus values to determine if a significant difference existed between each of the five locations where soil samples were collected. The null hypothesis of this statistical test is that the amount of soil phosphorus is roughly equal among each of the five fields, or in other words, the difference between each field is not significant. The alternative hypothesis is that the amount of soil phosphorus at each of the five fields is significantly different.

Table 12: Analysis of Variance among five sample locations.

SUMMARY

Groups	*# Samples*	*Sum P ppm*	*Average*	*Variance*
BUR	27	420	15.556	30.487

CAT	26	225	8.654	16.635
Church	26	294	11.308	57.742
DOM	26	426	16.385	118.646
SUN	26	1329	51.115	153.386

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F critical*
Between Groups	31315.301	4	7828.825	104.352	5.06E-39	2.444
Within Groups	9452.897	126	75.023			
Total	40768.198	130				

In the first table, the summary data provide descriptive statistics showing the average Mehlich-derived phosphorus value for each field, as well as the amount of variance at each location. The second table contains the ANOVA results which are based on the descriptive statistics. The F-value represents the ratio of variation between field locations and the variation within each field. If the F-value is larger than the F-critical value, this indicates that the null hypothesis can be rejected and there is a significant difference in soil phosphorus among the five fields. If the results of either the VNIR, M-LWIR, or XRF analysis are significant, it may then be possible to determine if they cannot only be used to detect soil phosphorus, but also identify differences between fields.

VNIR

The 131 air-dried soil samples were first measured using an ASD/PANalytical Field Spec VNIR/SWIR spectrometer which collects reflectance from 350nm to 2500

nm. Each sample was rotated under the sensor and measured three times each, then the measurements were averaged in order to create a more representative dataset. The Field Spec passively collects reflected light using the RS^3 version 6.0 software, which also calibrates the instrument. These reflectance files were exported to text files using the ViewSpec Pro version 6.0 software for further analysis in Microsoft Excel.

M-LWIR

Next, the samples were measured using a Thermo Nicolet fourier transform infrared (FTIR) spectrometer (Figure 60) 5DXB and integrating sphere which collects reflectance spectra from approximately 2,000 nm to 14,000 nm. The Nicolet FTIR is operated by the OMNIC software program. The program was set to collect the full range of data, which includes 1,000 scans of the sample material. Each measurement requires 15 minutes to complete, and the instrument was calibrated every 45 minutes (or after every 3 sample measurements) using a gold-plated calibration disk, placed directly under the instrument's integrating sphere. The final measurement of each sample was saved as a .csv file for processing and further analysis.

Figure 60: Thermo Nicolet 5DXB FTIR Spectrometer at USGS in Reston, VA.

XRF

Finally, 30 of the 131 soils samples were analyzed using a Bruker Tracer-III Handheld x-ray fluorescence (XRF) analyzer. XRF is a non-destructive method for identifying chemical elements in a sample material. It works by firing x-rays at a sample material, knocking electrons out of place, which releases energy that is then read by the detector. Since the energy produced by the movement of electrons for each element is unique, the measurement produced by the XRF is similarly unique. Bruker's ARTAX software interprets the measurements and identifies the elements that are contained in the material (Bruker, 2016). The Bruker Tracer-III is a hand-held device that can be setup in a lab environment, similar to the arrangement seen in Figure 61 and requires calibration in order for the data to be interpreted accurately (Speakman and Shackley 2013).

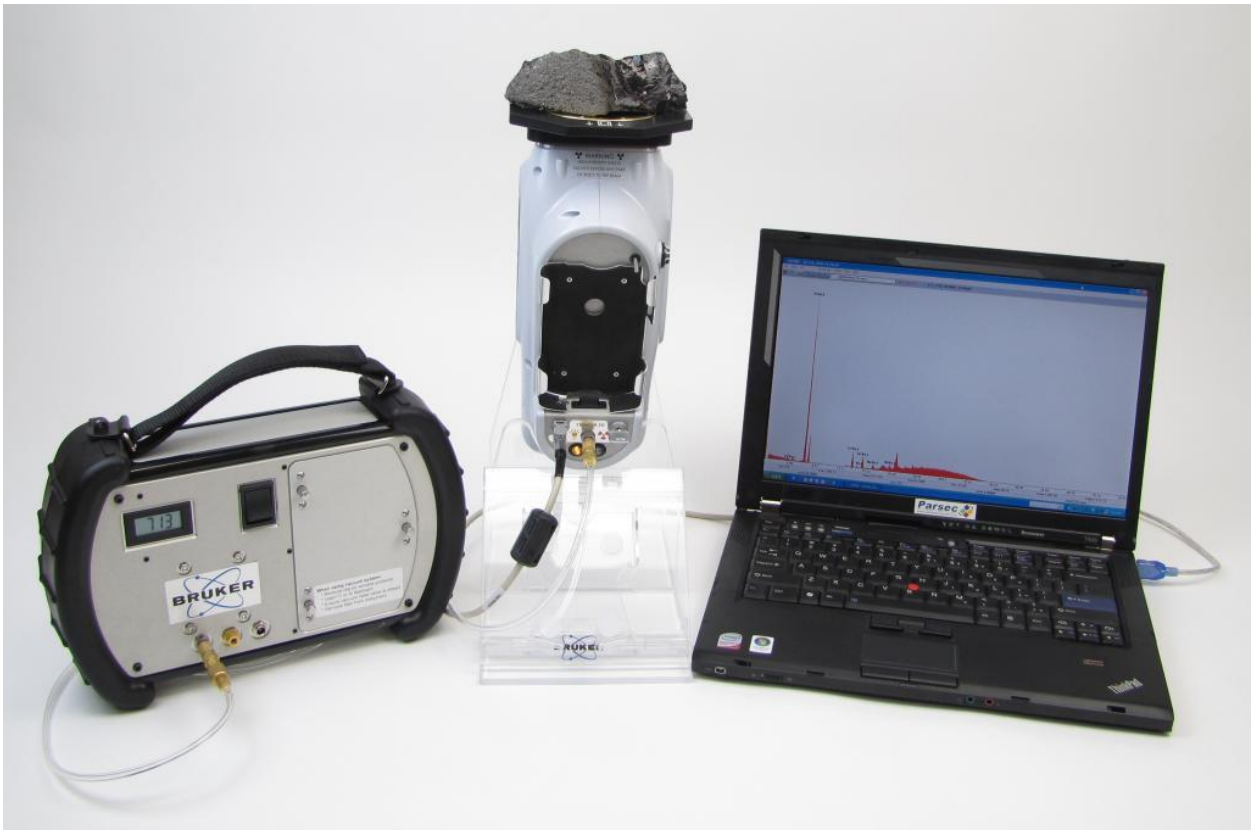


Figure 61: Bruker Tracer-III equipment setup (MSITECH, 2011)

Calibration of the XRF for soil phosphorus analysis was conducted by first creating a P-containing solution with a known quantity of P ppm. The solution was then added to soil samples at known increments which covered their P range, which was found to be as low as 2ppm for some samples. The final mixtures were measured with the XRF and the net count rate was plotted against their known phosphorus ppm values. If an XRF is precise and accurate, then the plotted data points should theoretically form a straight line, the slope of which can be used to calculate P content (in ppm) in subsequent measurements. A detailed list of the calibration steps is described below.

1) Since phosphorus does not exist in nature on its own, it was necessary to use a potassium phosphorus compound ($K_2HPO_4 \bullet 3H_2O$) standard for the initial solution. This compound was measured using a Mettler Toledo AE50 analytical balance (3.3218 g) and added to a 200ml flask. The molecular weight (MW) of the compound was calculated as:

   2K+7H+1P+7O = 2(39.0983) + 7(1.00794) + 1(30.9736) + 7(15.9994) = 228.218

2) The 200 ml flask was then filled with deionized water. Based on the atomic weight and the total volume of the flask, the concentration of stock solution was calculated as: 3.3218 g / (228.218 MW * 0.2) = 0.072846

3) Soil from two samples, CAT-109 which contained the least P (2 ppm) and SUN-319 which contained the most P (91 ppm), were ground with a mortar and pestle and used to create six samples cups containing 2 grams of ground soil, for a total of 12 cups.

4) Solution (V1) and regular deionized water was added to both sets of cups in varying amounts for a total of 2ml each (V2). The P ppm was calculated by first determining the concentration in moles per liter, then eventually converting to ppm as shown in Table 2.

Table 13: Calculation of P ppm for serial dilution calibration.

Sample	V1 (ml)	V2 (ml)	M2 (mole/L)	mM (molex10^3/L)	g/L	P ppm
Cup 0	0.00	2	0.0000	0.0000	0.0000	0.0000
Cup 1	0.10	2	0.0036	3.6423	0.1128	112.8174
Cup 2	0.25	2	0.0091	9.1058	0.2820	282.0434

Cup 3	0.50	2	0.0182	18.2116	0.5641	564.0868
Cup 4	0.75	2	0.0273	27.3174	0.8461	846.1302
Cup 5	1.00	2	0.0364	36.4232	1.1282	1128.1736

5) An XRF measurement was taken for each of these solutions using the following settings selected to enhance the phosphorus since it is at the low end of the detection range. These selected settings follow the recommendations for elements with low atomic numbers (McGlinchey 2012).
    a. Time – 240 seconds
    b. kilovolts (kV) – 15
    c. microamperes (µA) – 55
    d. no filter
    e. vacuum on
6) The results were analyzed in the ARTAX software program and a Bayesian deconvolution was performed to determine which elements were contained in the soil samples. This step allows the software to calculate quantitatively the amount of influence that phosphorus has in the detection and provides an XRF net count rate for each element in the sample. An example of this analysis is shown in Figure 62. The location of the phosphorus within the detection is flagged and labeled “P”. There is a single blue line that does not fit with the rest, this is the Bayesian model fit; fit changes as different elements are added or removed, depending on whether they influenced the final shape of the total XRF spectrum.

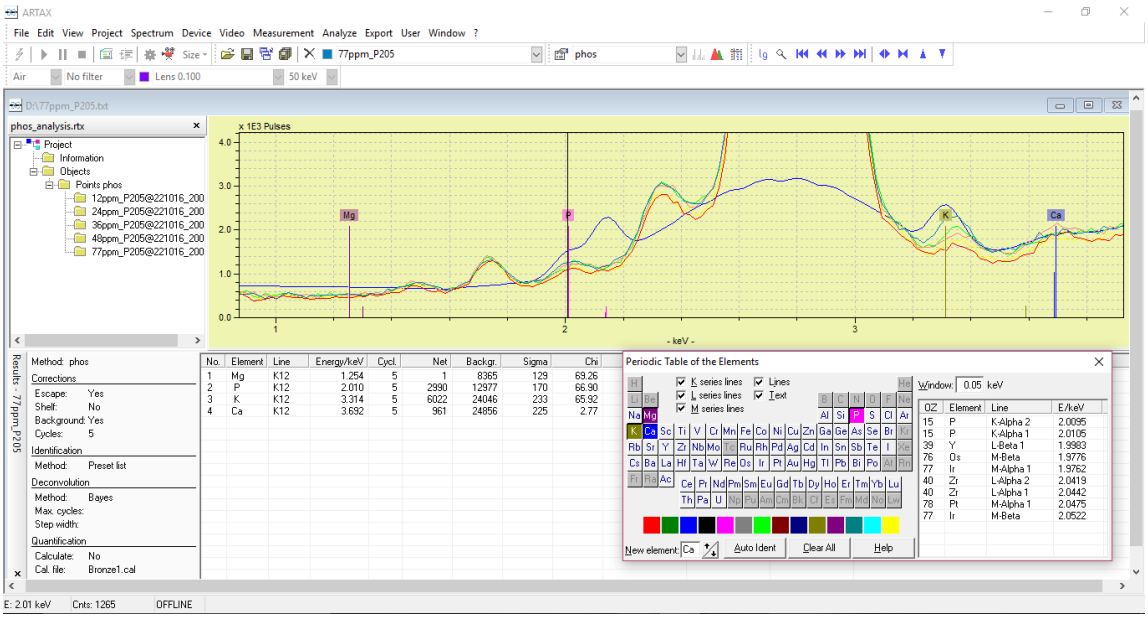


Figure 62: Screen capture showing the ARTAX Bayesian Deconvolution modeling

7) Finally, the XRF counts derived from the Bayesian deconvolution were then plotted against the known P ppm (Figure 63) to determine the relationship between these two sets of data, and to build a calibration curve to be applied to the rest of the soil XRF measurements. The results indicate that the XRF net count rate and the known P ppm are highly correlated and the instrument is properly calibrated.

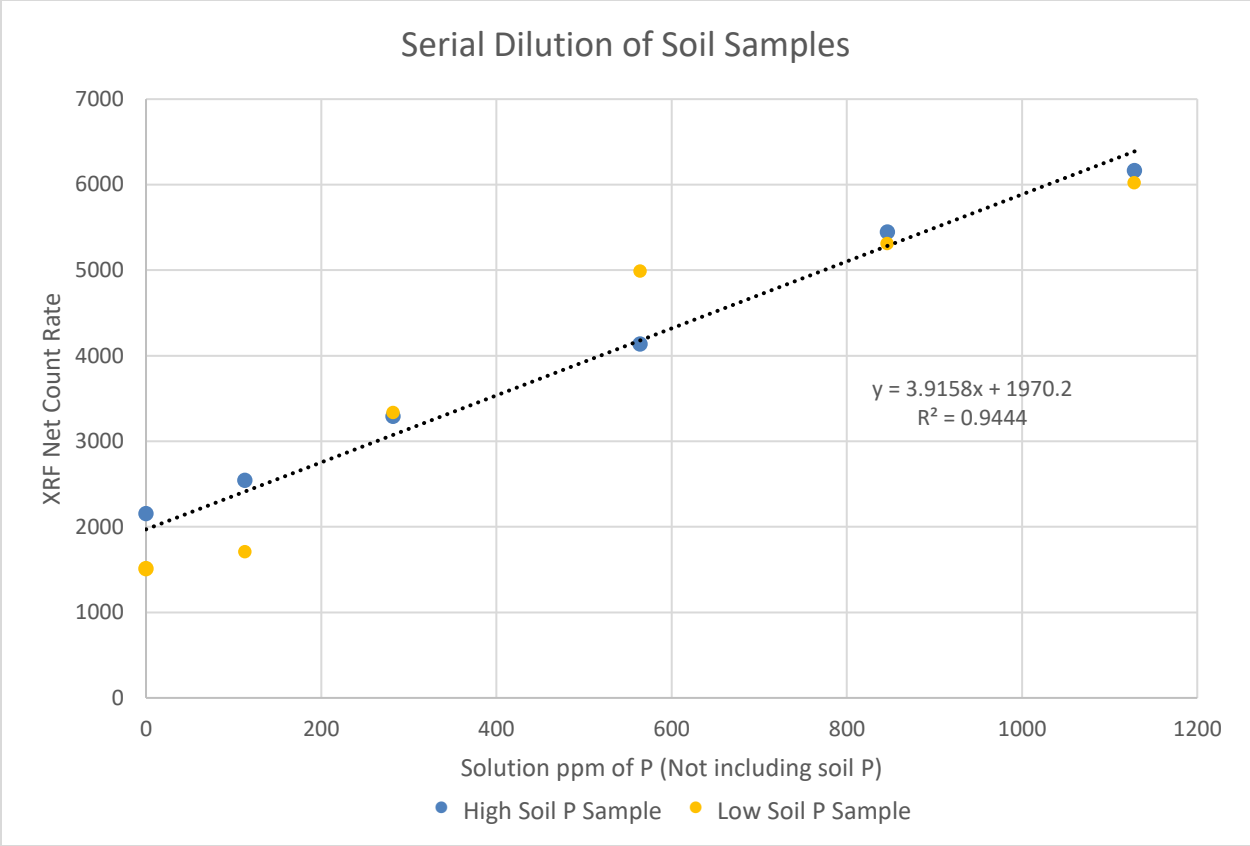


Figure 63: Serial Dilution of Soil Samples for XRF Calibration; the high P sample originally contained 91 ppm before adding P solution and the low P sample originally contained 2 ppm

Once the calibration was performed, the slope of the calibration line was then used to estimate the P ppm for the rest of the soil samples analyzed using XRF. Thirty soil samples were randomly selected, six from each field, and were then ground, and 2 grams were placed in a sample cup and covered with ProleneThin-Film specifically designed for XRF use (Chemplex, 2016).

Table 14: Soil samples used in XRF analysis

BUR-405	CAT-100	Church-501	DOM-201	SUN-300

BUR-409	CAT-102	Church-505	DOM-202	SUN-304
BUR-420	CAT-105	Church-510	DOM-205	SUN-308
BUR-417	CAT-114	Church-521	DOM-210	SUN-311
BUR-420	CAT-116	Church-522	DOM-218	SUN-323
BUR-426	CAT-125	Church-524	DOM-219	SUN-324

These soil samples did not have any liquid added to them since sensor manufacturer (Bruker) advises that this is not necessary, and that liquid may actually lessen the detection abilities. In fact, after drying some of the calibration samples and retesting them, a slightly higher detection was found. However, since the XRF net count rate and known P ppm was high using the wet samples it was not recreated using dried samples. Each of the randomly selected samples were then analyzed using the same settings as the calibration (240 seconds of acquisition time, 15 kV, 55 μA, no filter, vacuum on). They were then analyzed using the ARTAX software and the derived XRF net count rate for each was plotted against the Mehlich-III P ppm.

Results

The VNIR, M-LWIR datasets were each compiled into a single table, along with the Mehlich-III P ppm values. The data was ingested into the R statistical computing software program, and a PLSR was calculated for each dataset in order to evaluate how well each of the three types of measurements were able to model and predict P. The output of the models were then plotted to determine how well they performed.

VNIR Reflectance and PLSR Results

The PLSR model was calculated using the "pls" package in R. The model uses the original P values and VNIR spectra as the Y and X data respectively. The results of this model are provided in Appendix D, which include the LOO cross-validation results, and the calculations of variance between the original and predicted y values and each of the components.

Since the dataset used here is larger than the one used in the previous chapter, it was possible to calculate 50 components when running the model. The cross-validation values plateaued at around 15 components and reached their lowest scores by around 45 components. For this model, 15 components are used. Figure 64 allows for the visual inspection of the RMSEP values by plotting them as a function of the number of components.

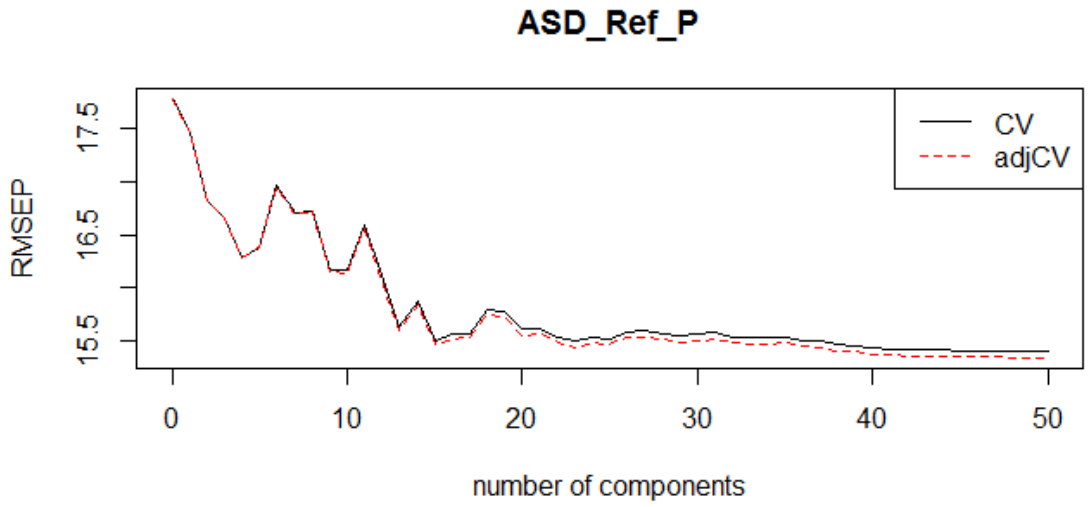


Figure 64: RMSEP as a function of number of components for VNIR/PLSR model

The first five components can be further examined using the plot shown in Figure 65 which reveals the amount of variance shared between each of the components.

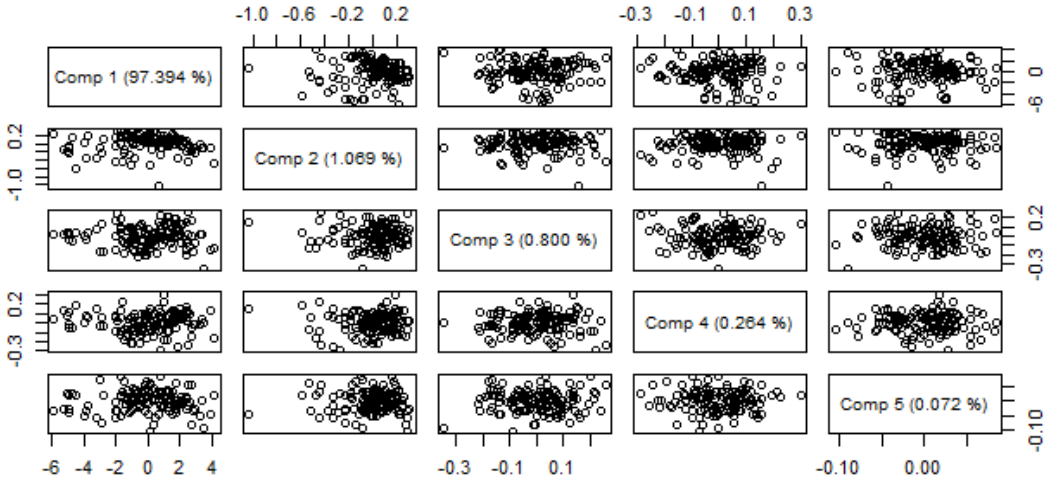


Figure 65: Explained variance for VNIR/PLSR model

The fit of the model is displayed in Figure 66, which plots the original versus predicted y values.

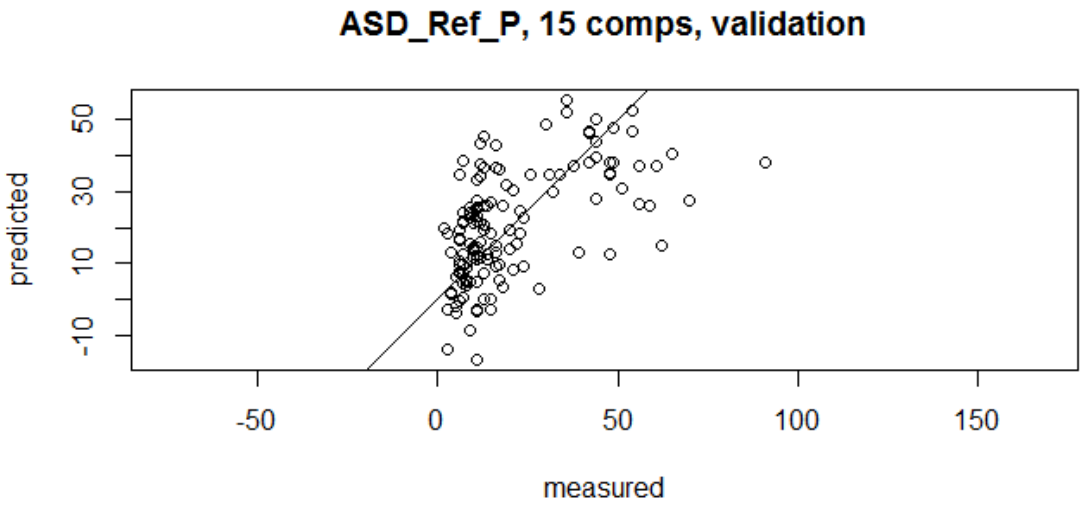


Figure 66: P measured versus predicted for VNIR/PLSR model

The final step is to calculate the R^2 for the model, using the original P values measured using the Mehlich-III test and the predicted values. This value represents the amount of variance in the dependent variable, the measured y values, that can be predicted from the independent variables, or components. The R^2 for this model is 0.6935764.

```
> y = Oh_VNIR_Spectra$ASD_Ref_P
> ypred = (Ohio_VNIR_pls$fitted.values[,,15])
> SST = norm(y-mean(y), type = "2")
> SSE = norm(y-ypred, type="2")
> R2 = 1-(SSE/SST)
> R2
[1] 0.6935764
```

M-LWIR Reflectance and PLSR Results

The cross-validation values for the M-LWIR/PLSR model were lowest at 14 components (See Appendix E). Figure 67 plots the RMSEP values as a function of the number of the components of components calculated.

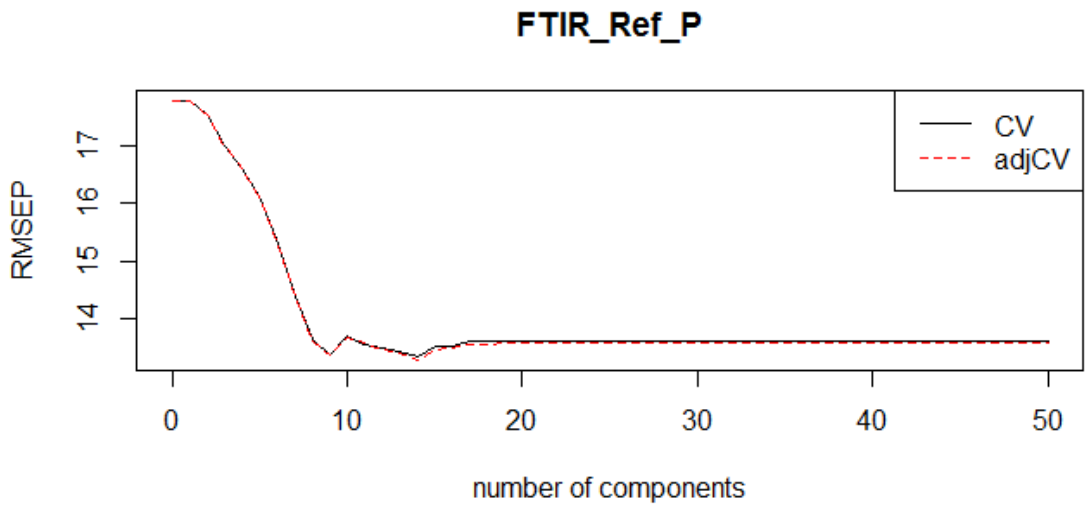


Figure 67: RMSEP as a function of number of components for the M-LWIR/PLSR model

The first five components can be further examined using the plot shown in Figure 68 which reveals the amount of variance shared between each of the components.

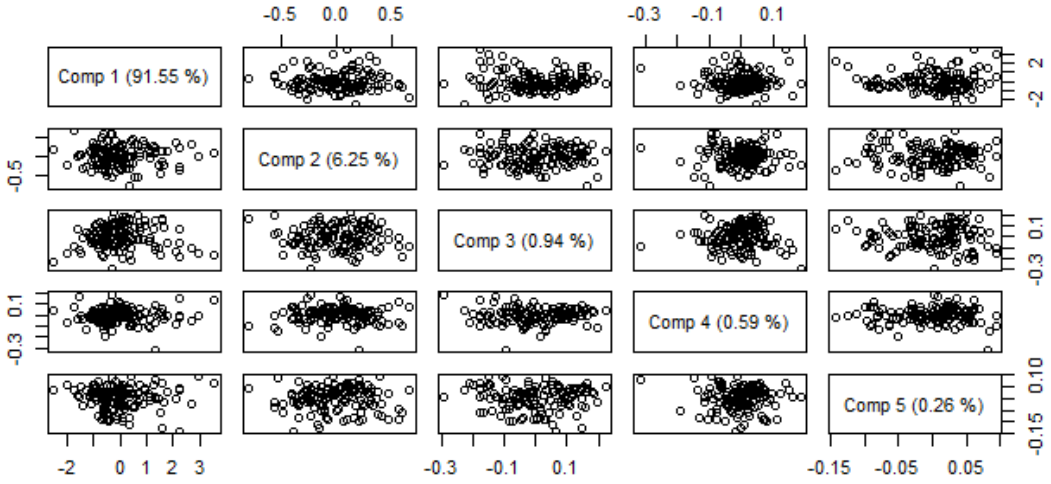


Figure 68: Explained variance of first five components of the M-LWIR/PLSR model

The fit of the model is displayed in Figure 69, which plots the original versus predicted y values.

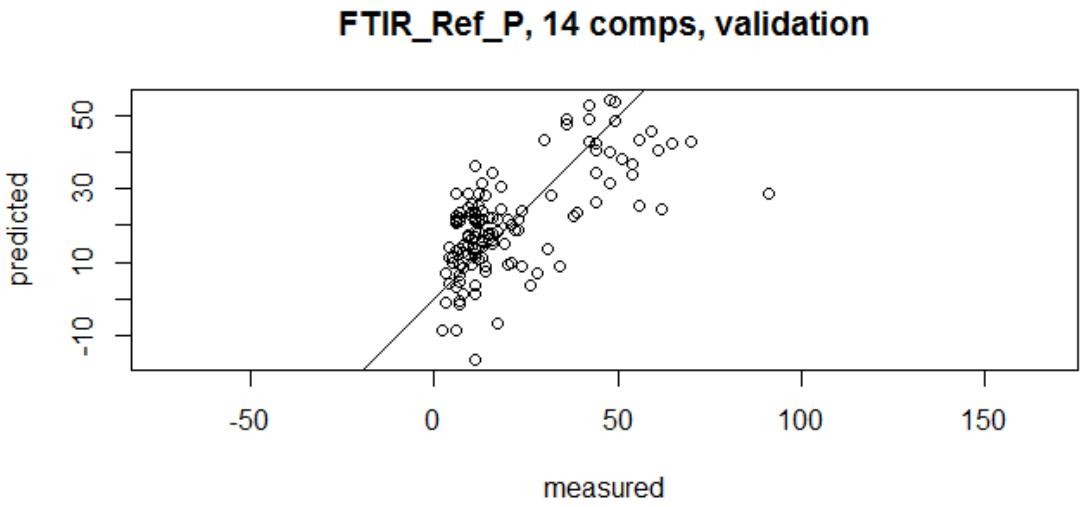


Figure 69: Explained variance for M-LWIR/PLSR model

The R^2 for the model using 14 components is 0.4855194.

```
> y = Oh_LWIR_Spectra$FTIR_Ref_P
> ypred = (Ohio_LWIR_pls$fitted.values[,,9])
> SST = norm(y-mean(y), type = "2")
> SSE = norm(y-ypred, type="2")
> R2 = 1-(SSE/SST)
> R2
[1] 0.4855194
```

XRF Results

Since the XRF data collected for the soil samples is only one value per sample, a PLSR model could not be created. Instead, the net count rates for thirty of the soil samples were plotted against the Mehlich-III P values, resulting in an R^2 of 0.507 (Figure 70).

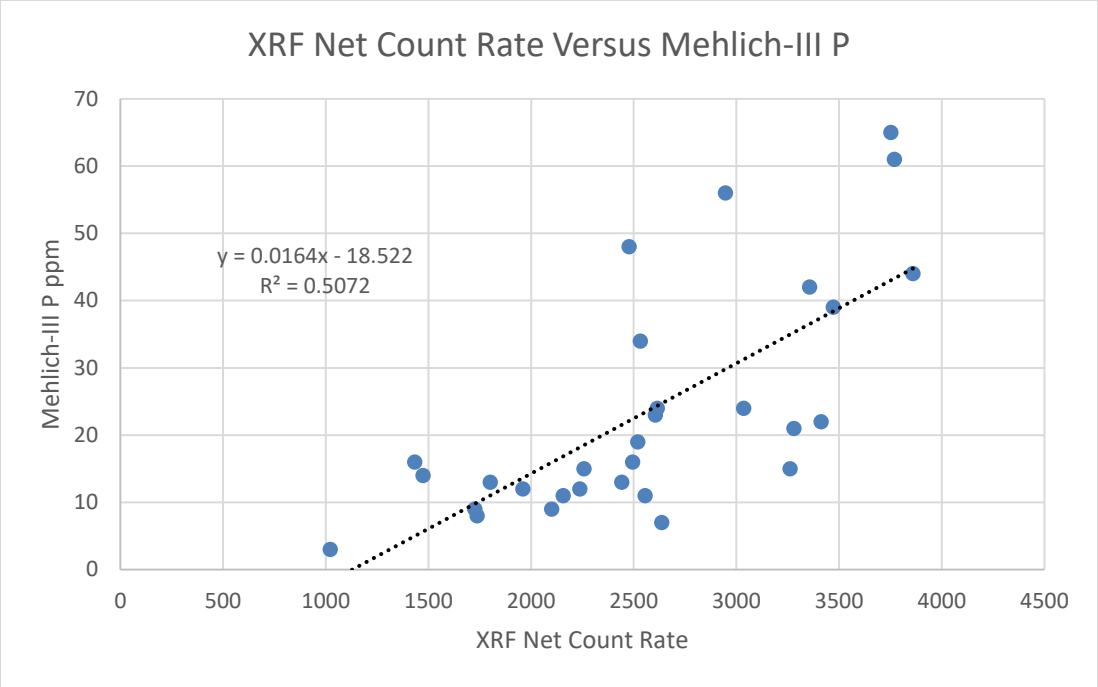


Figure 70: XRF Net Count Rate versus Mehlich-III P

A PLSR model was created for the VNIR and M-LWIR datasets, this time using the XRF net count rate in place of the Mehlich-III. The VNIR/PLSR model results are below.

```
> summary(XRF_VNIR_pls)
Data:   X dimension: 30 2151
        Y dimension: 30 1
Fit method: kernelpls
Number of components considered: 10

VALIDATION: RMSEP
Cross-validated using 30 leave-one-out segments.
    (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
CV        752.8    749.0    766.1    748.9    907.7    922.7    948.1
adjCV     752.8    748.3    763.8    749.7    901.5    911.1    932.5

       7 comps  8 comps  9 comps  10 comps
CV       898.9    919.1    939.4     986.1
adjCV    888.8    908.4    926.8     970.5

TRAINING: % variance explained
         1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps
X         96.390    98.28    99.36    99.56    99.59    99.66    99.75
```

```
XRF_VNIR_P 5.508    20.84    25.48    44.20    70.47    79.52    84.63

            8 comps  9 comps  10 comps
X             99.80    99.84     99.85
XRF_VNIR_P    89.83    93.67     97.40
```

The RMSEP for this model is lowest using three components (Figure 71), however, the RMSEP values are significantly higher than in previous PLSR models.

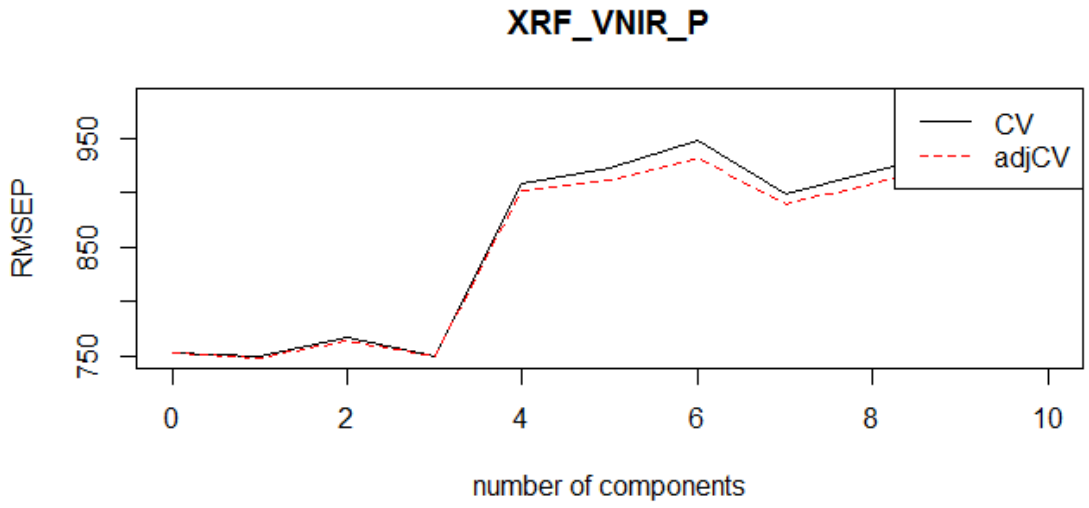


Figure 71: RMSEP as a function of components for the XRF/VNIR PLSR model

The explained variance for the first three components used is shown below in Figure 72. The original versus predicted y values are shown in Figure 73.

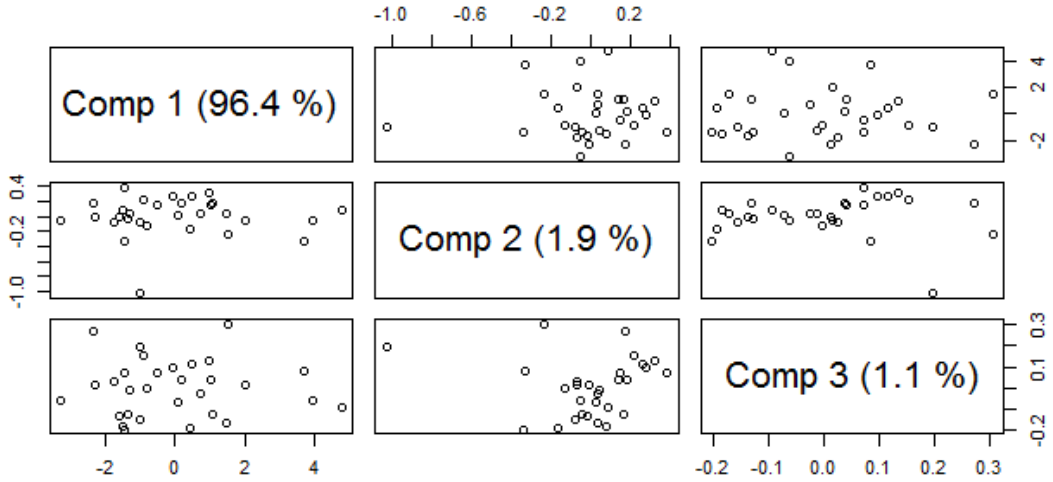


Figure 72: Explained variance of XRF/VNIR PLSR model components

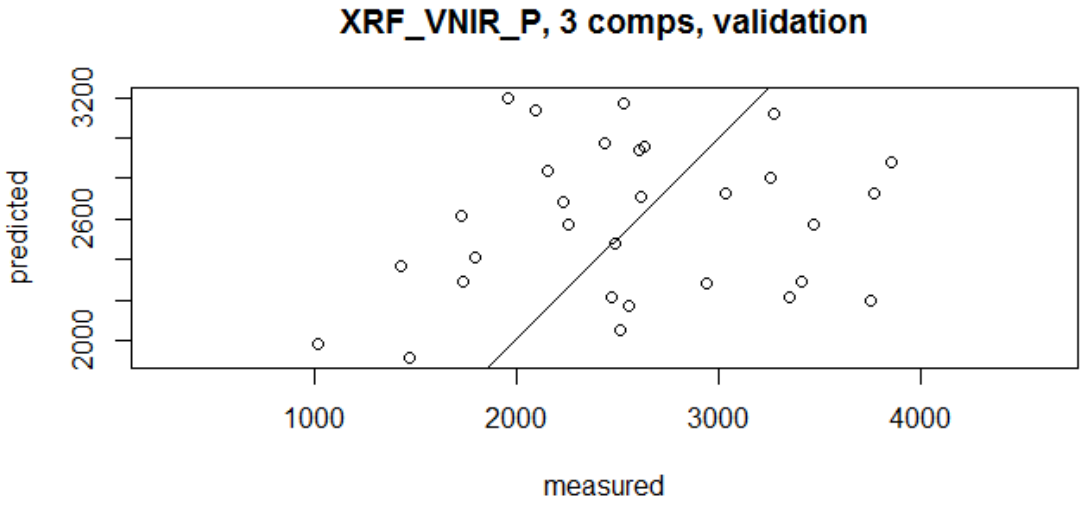


Figure 73: Original y values versus predicted for the XRF/VNIR PLSR model

The calculated R^2 of the XRF VNIR PLSR model is 0.1367354.

```
> y = XRF_VNIR_Spectra$XRF_VNIR_P
> ypred = (XRF_VNIR_pls$fitted.values[,,3])
> SST = norm(y-mean(y), type = "2")
```

```
> SSE = norm(y-ypred, type="2")
> R2 = 1-(SSE/SST)
> R2
[1] 0.1367354
```

The same analysis was conducted using the XRF net count rate as original y values and the M-LWIR data in a PLSR model. The results of this model are shown in Appendix E. Figure 74 shows the RMSEP as a function of components for this model. The lowest value occurs at 15 components.

```
> XRF_LWIR_pls = plsr(XRF_P ~ as.matrix(XRF_LWIR), ncomp = 15, data = X
RF_LWIR_Spectra, validation="LOO")
> summary(XRF_LWIR_pls)
Data:   X dimension: 30 2231
        Y dimension: 30 1
Fit method: kernelpls
Number of components considered: 15

VALIDATION: RMSEP
Cross-validated using 30 leave-one-out segments.
  (Intercept)  1comps  2comps  3comps  4comps  5comps  6 comps  7comps
CV    752.8    686.3   678.8   739.3   891     774.2   721.1    733.8
adjCV 752.8    685.5   677.7   738.4   880     729.0   695.0    725.5

   8comps  9comps  10comps  11comps  12comps  13comps  14comps  15comps
CV    648  641.2   667.4    659.4    664.5    663.2    663.2    663.1
adjCV 641  630.7   656.0    648.4    653.4    652.0    652.0    652.0

TRAINING: % variance explained
       1comps  2comps  3comps  4comps  5comps  6comps  7comps  8comps
X       95.29  98.68   99.20   99.30   99.36   99.56   99.83   99.91
XRF_P   23.71  31.01   37.09   59.36   77.95   85.83   89.78   93.08

      9comps 10 comps  11 comps  12 comps  13 comps  14 comps  15 comps
X      99.92  99.93     99.94     99.95    99.95     99.96     99.96
XRF_P  98.58  99.71     99.94     99.99    100.00    100.00    100.00
```

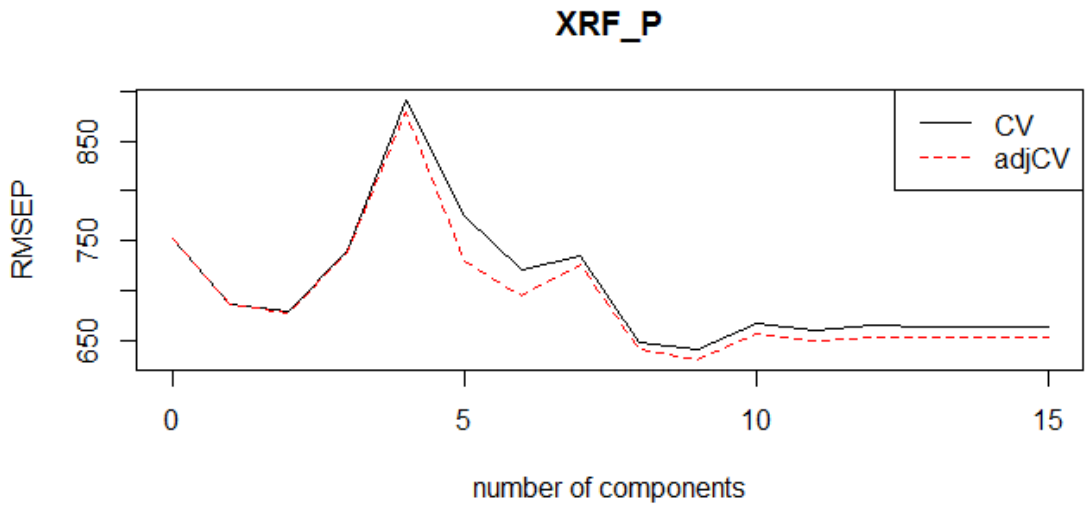


Figure 74: RMSEP as a function of components for the XRF/M-LWIR PLSR model

The explained variance for the first three components used is shown below in Figure 75. The original versus predicted y values are shown in Figure 76.

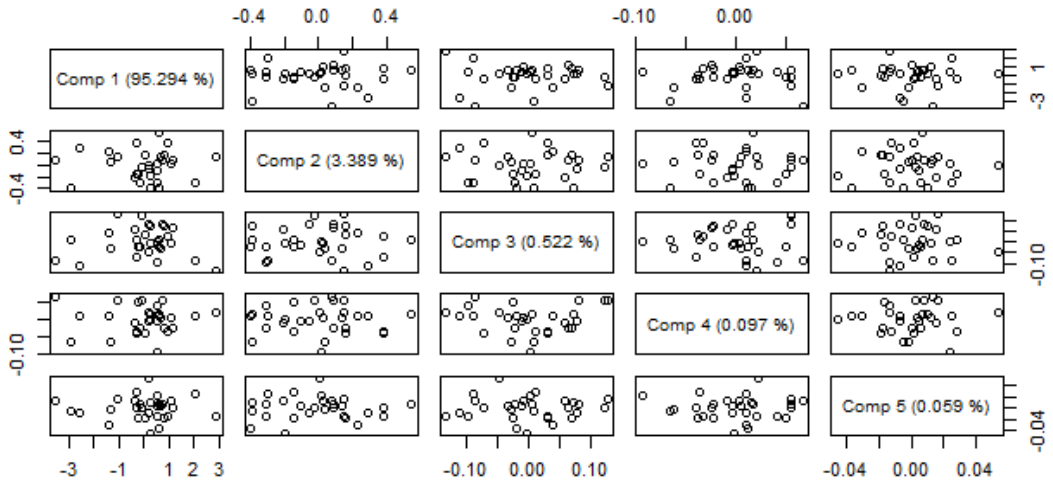


Figure 75: Explained variance of the first five components for the XRF/M-LWIR PLSR model

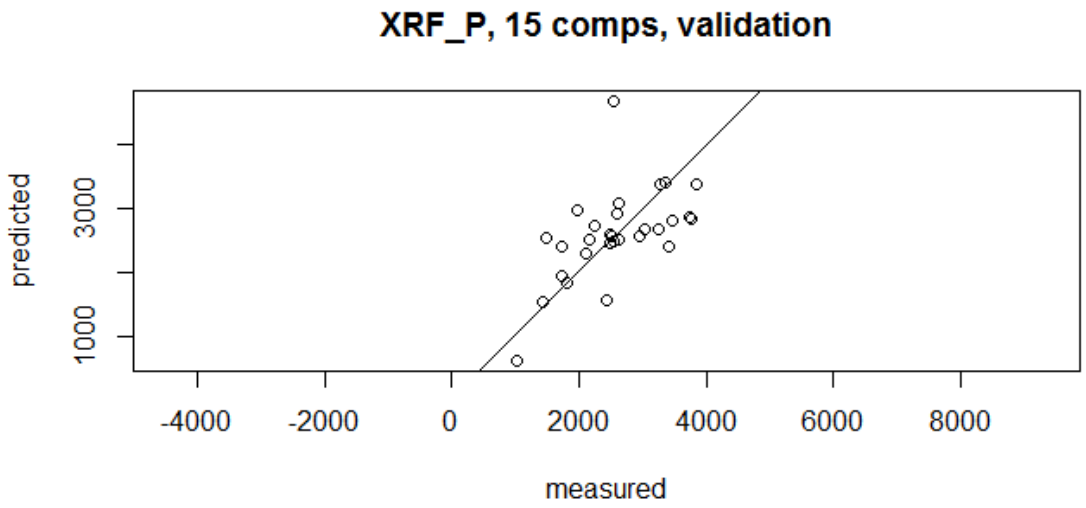


Figure 76: Original y values versus predicted for the XRF/M-LWIR PLSR model

The R^2 for this model is 0.9997253

```
> y = XRF_LWIR_Spectra$XRF_P
> ypred = (XRF_LWIR_pls$fitted.values[,,15])
> SST = norm(y-mean(y), type = "2")
> SSE = norm(y-ypred, type="2")
> R2 = 1-(SSE/SST)
> R2
[1] 0.9997253
```

Discussion

The models generated a range of results, some of which indicate that reflectance and x-ray fluorescence spectroscopy may be useful for archaeological soil P field measurements. The first two models indicate that VNIR reflectance is more highly correlated with Mehlich-III P than M-LWIR reflectance, but both VNIR and M-LWIR models indicate a significant correlation with Mehlich-III that warrants future research. Since XRF-derived P ppm was not closely correlated with the Mehlich-III P ppm, more research is required before the XRF-derived P ppm values can be used in similar PLSR

soil models. However, the XRF calibration showed that the relationship between net count rate and P ppm was highly correlated, producing an R^2 of .94, so it is likely a much better benchmark than Mehlich-III. Mehlich-III measures only the available phosphorus, which may not exist at a constant ratio to total phosphorus in these fields.

Conclusions

This research shows that although VNIR, M-LWIR, and XRF spectroscopy cannot easily replace laboratory-based wet chemistry tests for soil phosphorus, there may be avenues for improvement. For example, future work should consider using XRF as a benchmark instead of Mehlich-III. The XRF may actually be the best choice for collecting measurements in the field due to its accuracy, portability, and relative durability. Also, XRF can take measurements of samples that have had little to no preparation, as long as they are not too wet. Future work attempting to correlate field measurements to satellite or airborne imagery will still require the use of reflectance spectroscopy.

CHAPTER SIX - CONCLUSION

The three articles presented in this dissertation utilized geospatial analysis and remote sensing to identify the location of ancient Maya wetland agricultural fields, model the probability of their spatial extent, and determine the impact of agriculture on wetland soils. The first article explained how high-resolution imagery has become an essential tool in locating and the spatial analysis of wetland fields. Commercial imagery has rapidly improved over the last decade to enable researchers and students to greatly benefit from its use. Google Earth is a freely available imagery resource that also contains historic imagery, providing a way to follow changes on the earth's surface. In Belize, modern agriculture and deforestation is uncovering new fields and ancient cities, and the number of visible archaeological features is increasing all the time. Although this sometimes results in the loss of archaeological information, it is important to record it while it still exists.

The second article showed the location of wetland fields can be predictively modeled in order to determine estimates of their past extent. A number of environment variables were selected based on past research and current knowledge of wetland fields. These variables were used in a binary logistic regression analysis to determine which variables were correlated with wetland field locations. The analysis was run multiple times to find the best combination of variables and the coefficients produced were used to develop a spatial model of field probability. Finally, the probability equation was applied to the entire study area and the results showed the amount of suitable land is far greater

than the 16,000 acres of wetland fields initially identified using high-resolution multispectral imagery. The odds of finding even more fields that has been discovered so far is high. This analysis can also be used to help answer other questions regarding the variety of subsistence techniques and population size.

The third article attempted to use reflectance spectroscopy of visible and near-infrared wavelengths to detect and measure soil phosphorus. Soil samples were collected in northwest Belize and spectra were measured using a Field Spec 4 reflectance spectrometer. A parallel analysis was conducted on samples collected in northeast Ohio for comparison. Unfortunately, the null hypotheses were not rejected but since this area of research is in its early stages, it would be useful to attempt this work again using different algorithms. Phosphorus plays a significant role in archaeology and understanding past cultures. It will continue to be used as an indicator of human activity on the landscape and should therefore continue to be studied using reflectance spectroscopy. The advantage of a field spectrometer is that it can provide information while working in the field. The ability to determine phosphorus levels quickly and efficiently using a field spectrometer would create even more avenues for remote sensing and geospatial analysis in Maya archaeology. This technique would also be extremely valuable to modern day farmers around the globe.

This dissertation builds on the regional knowledge of wetland fields of northern Belize and will help to accelerate the discovery of previously unknown field locations. Ultimately, it is important to understand where the fields once existed to better understand their significance in the lives of the ancient Maya. This will explain the

amount of food they could have produced from year to year and the size of the population they supported. The ancient Maya relied on a variety of agricultural techniques to procure food, and understanding the role that each of them played will lead to a broader understanding of their existence over several millennia and sharp decline around 900 A.D. In addition, the methods used in this research could be applied to other parts of the world where wetland agriculture was use by other ancient societies, and may be valuable to conduct similar research on a much broader scale. The far-reaching impact of studying ancient Maya wetland fields is that modern civilizations can learn from the lessons of the Maya. Beyond the interests of the Maya archaeology community, societies and governments around the world will benefit from understanding how a civilization adept at supporting large numbers of people with minimal resources was unable to avoid a dramatic decline.

The ability to study wetland fields, or any other remnants of ancient Maya life, may have a time limit. In recent years, historical imagery in Google Earth clearly shows the continual destruction of forest to make way for modern agriculture. Stories in the news explain how temples are being demolished and used for building materials. Many temples that are discovered have also been tunneled into and looted. Although archaeologists continue to discover entire cities buried deep within the jungles of Belize and other parts of Mesoamerica, many others may also become lost before they are ever found.

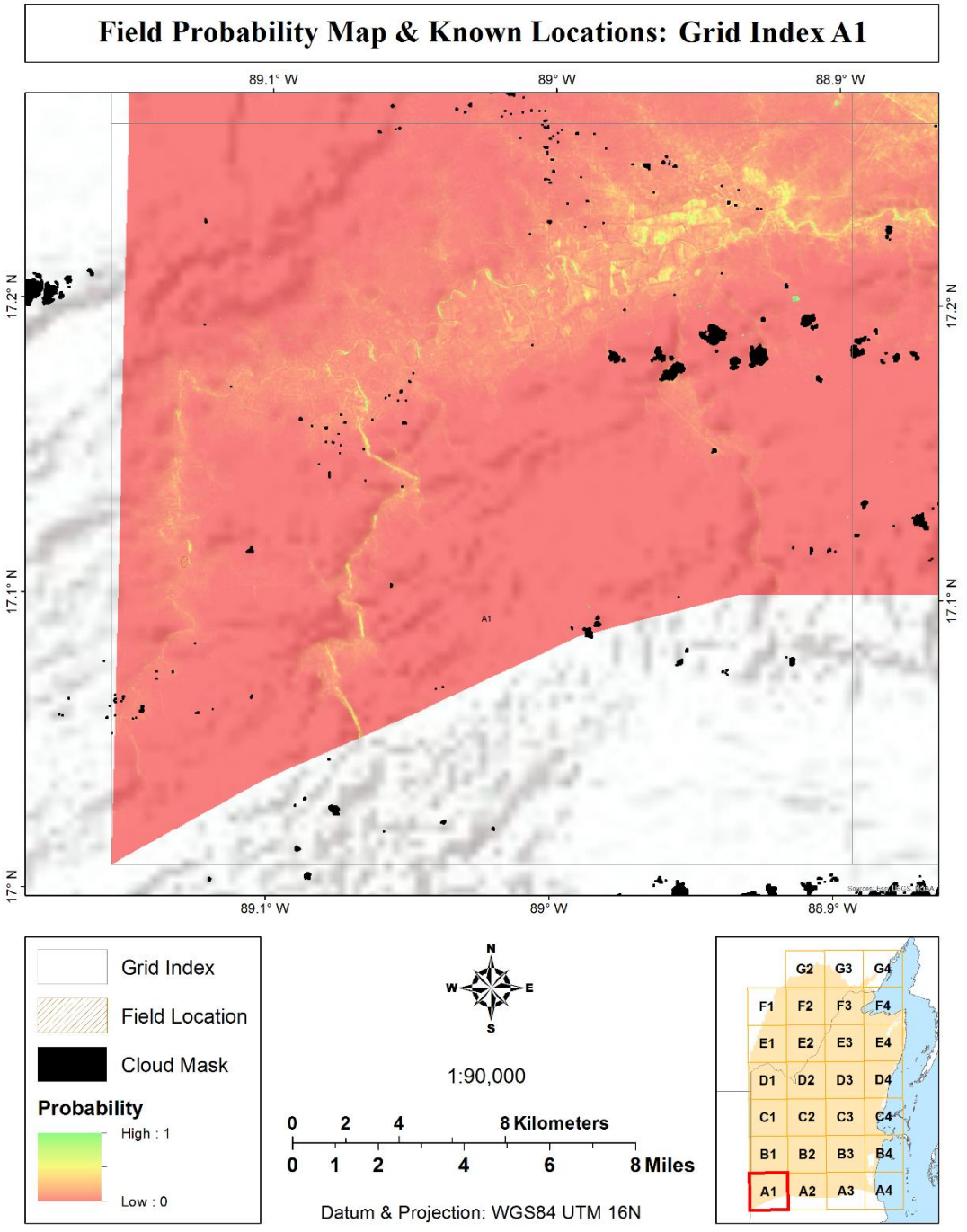
Figure 77: Looter's trench in the side of a temple at the Maya site of La Milpa.

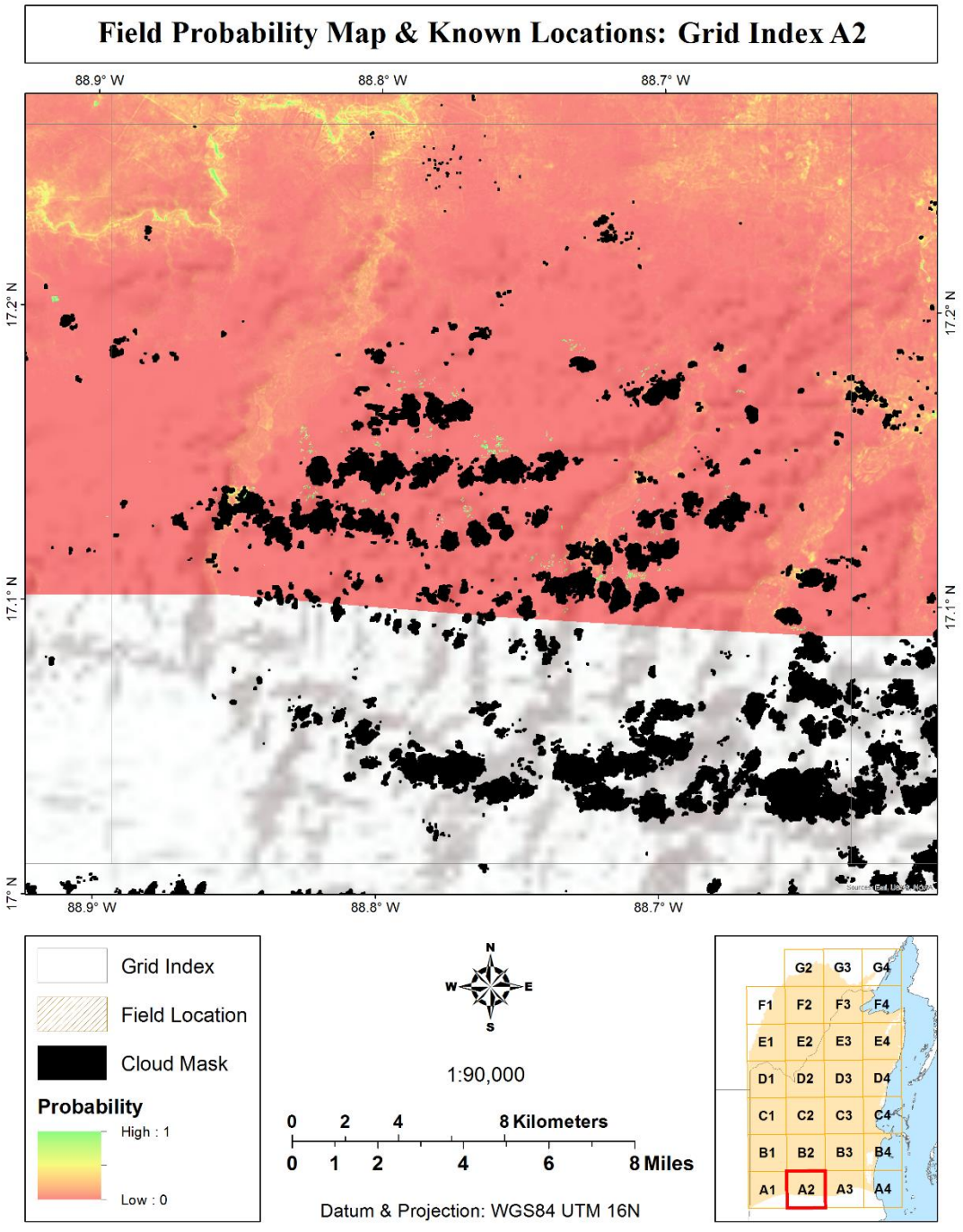
There is no way to know exactly how many more are still undiscovered since nature has skillfully reclaimed them. High-resolution elevation data such as Lidar has quickened the pace of discovery in some areas, but until it is feasible to collect this type of data for all of Mesoamerica, we will only be able to speculate where and how many Maya cities once existed. Even still, the ability to collect this data will require further research to understand aspects such as the timeline of their creation or destruction. Remote sensing and geospatial analysis will therefore play an important role in Maya

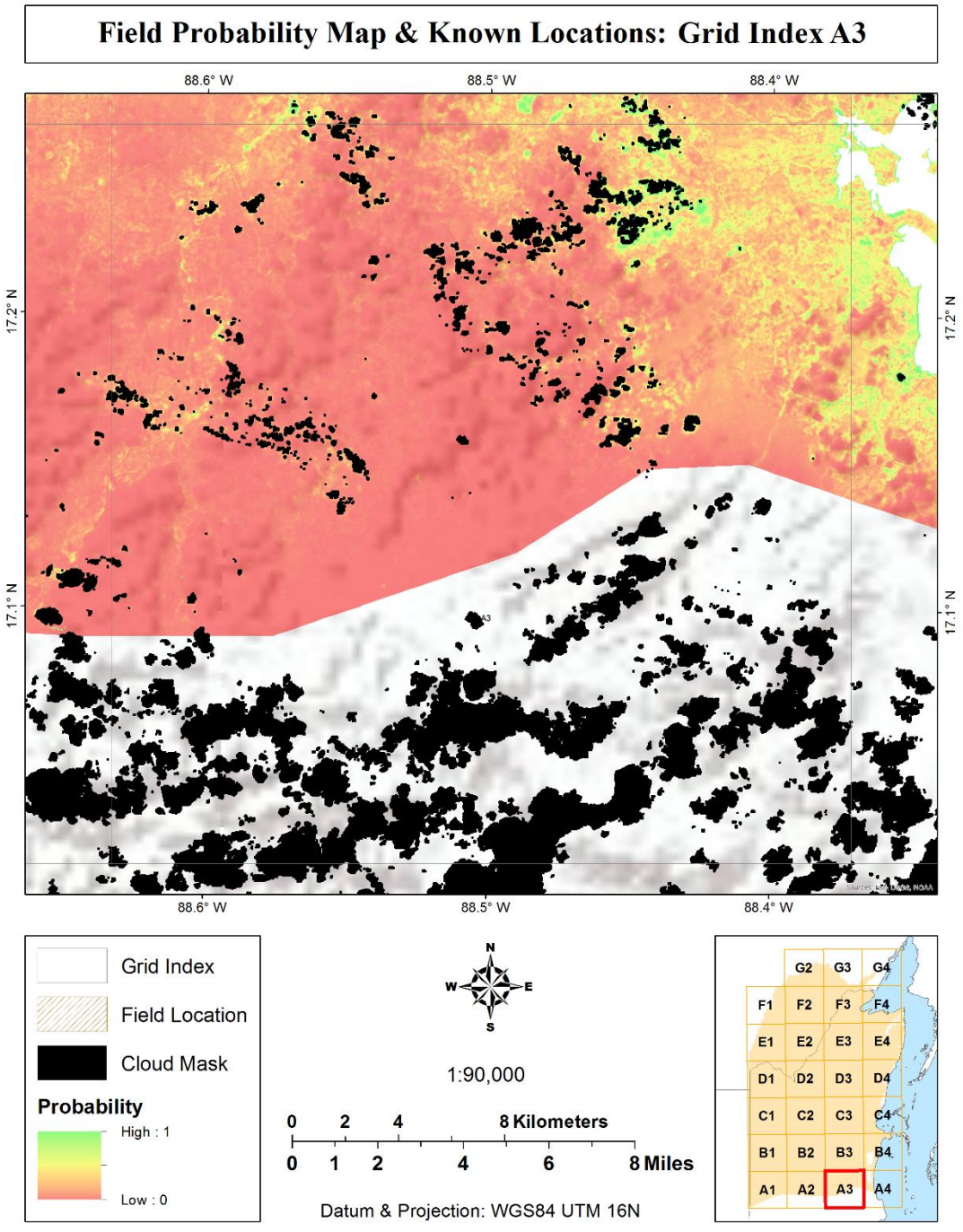
research in the future since historic satellite images may one day be all that is left of some ancient places.

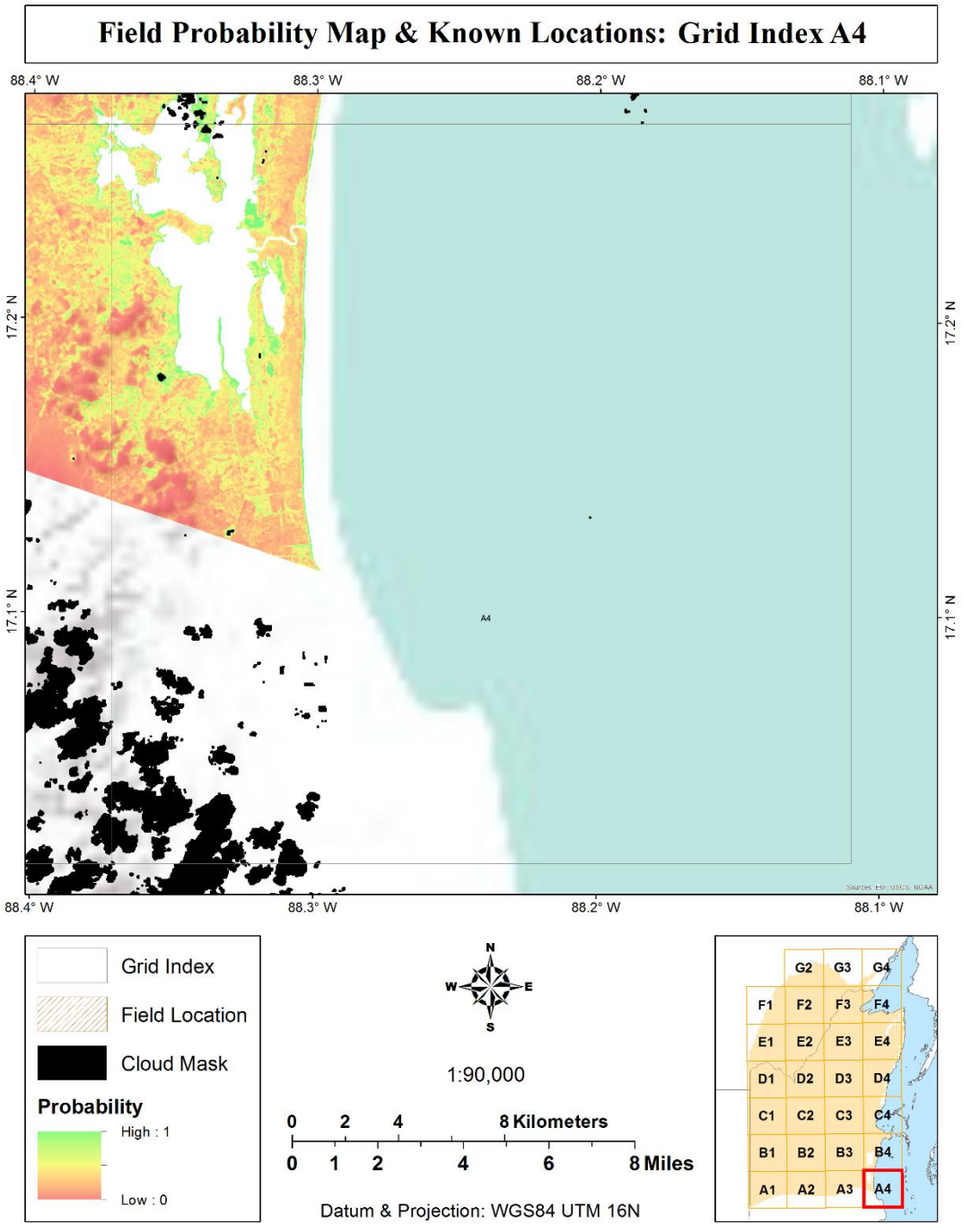
Figure 78: Mask Temple at Lamanai

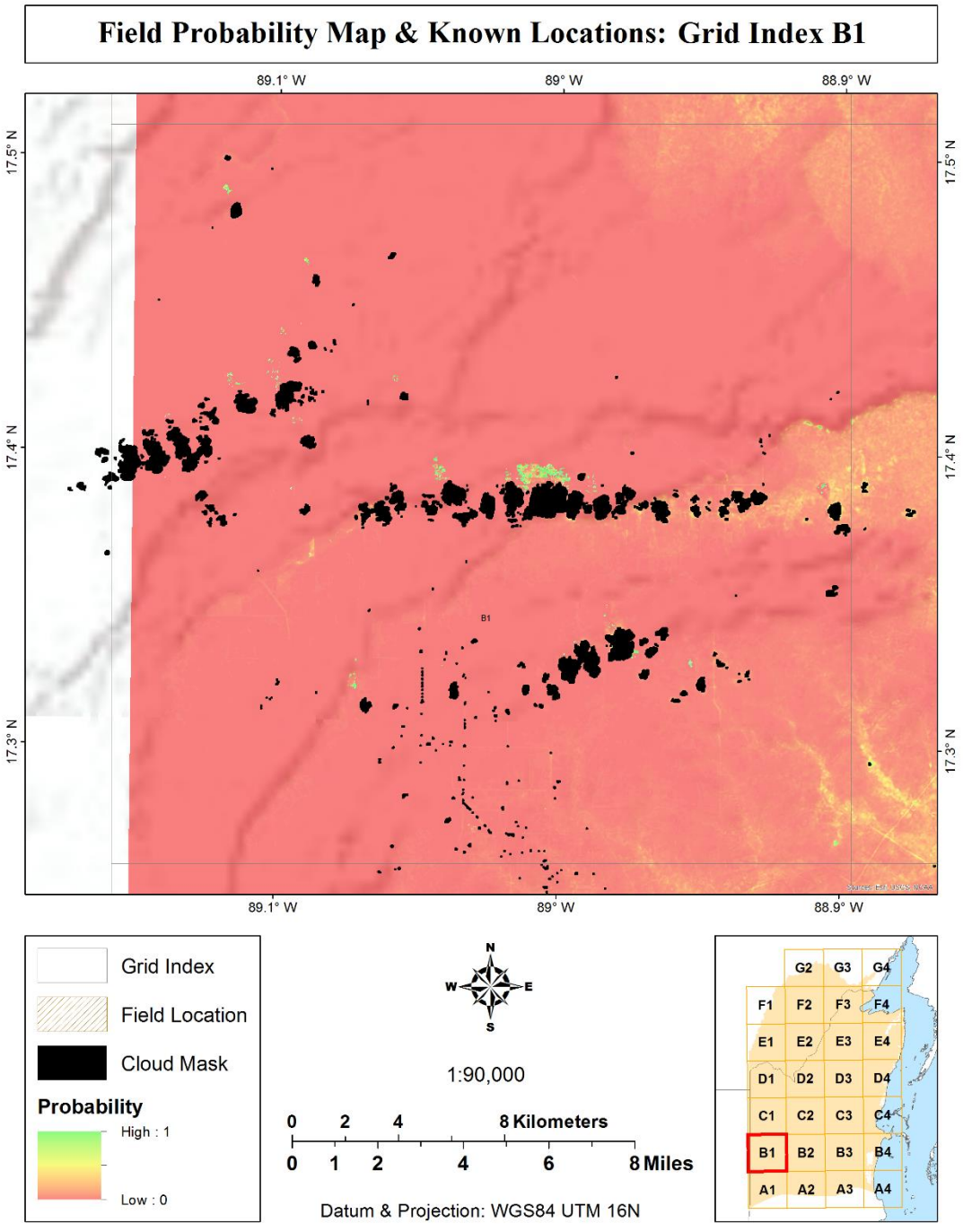
APPENDIX A – PROBABILITY MAP SERIES

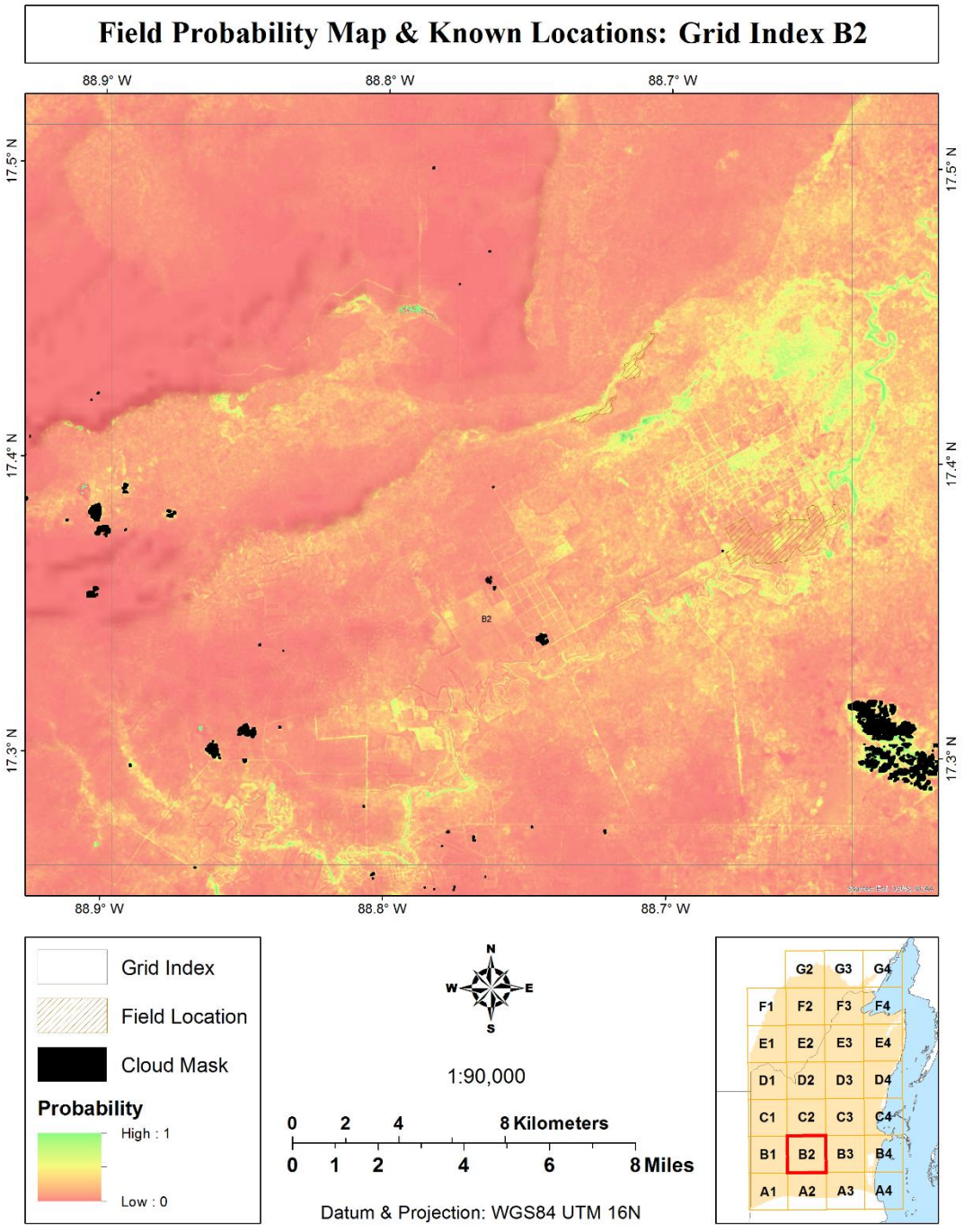


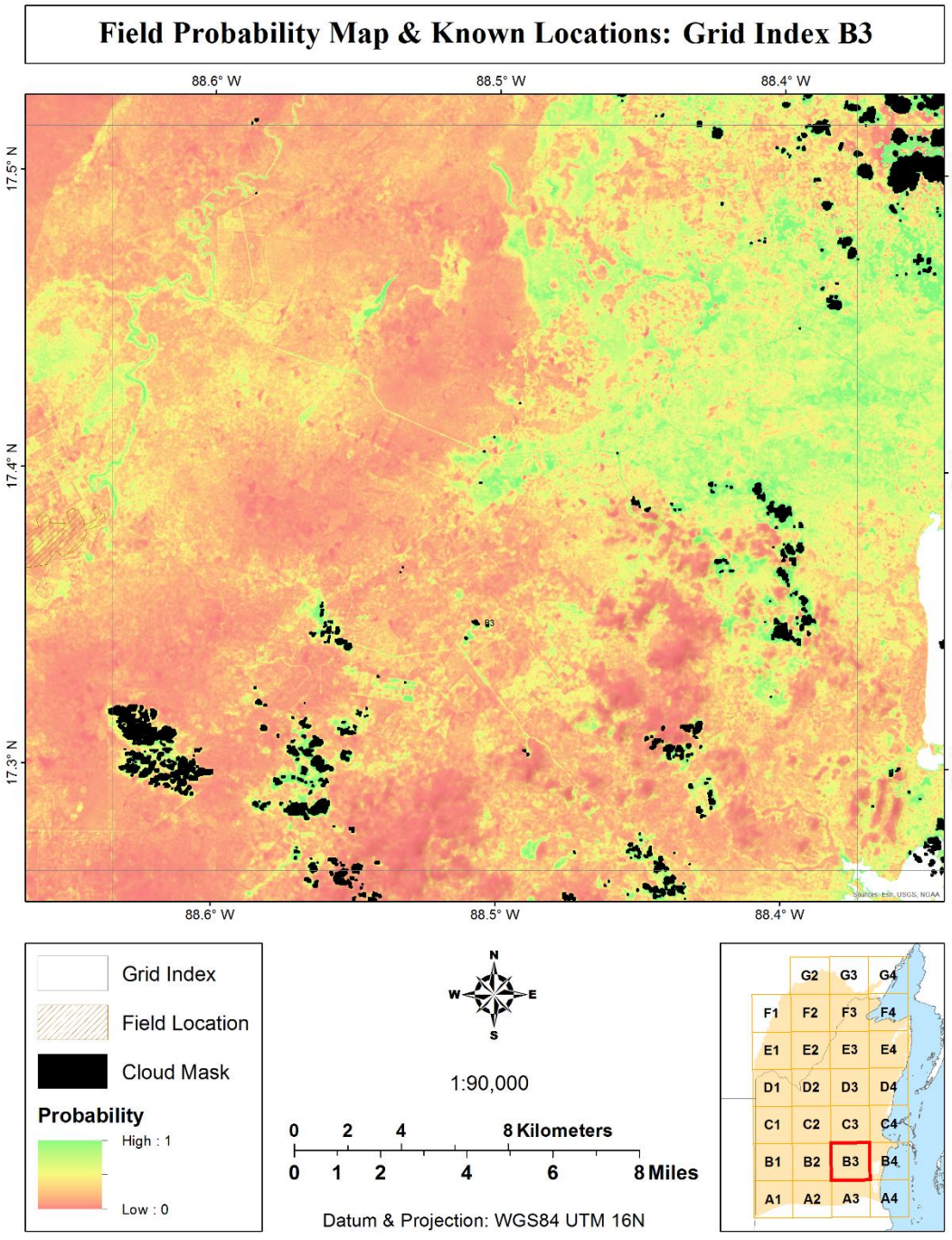

Field Probability Map & Known Locations: Grid Index A2
88.9° W
88.8° W
88.7° W
17.2° N
17.1° N
17° N
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

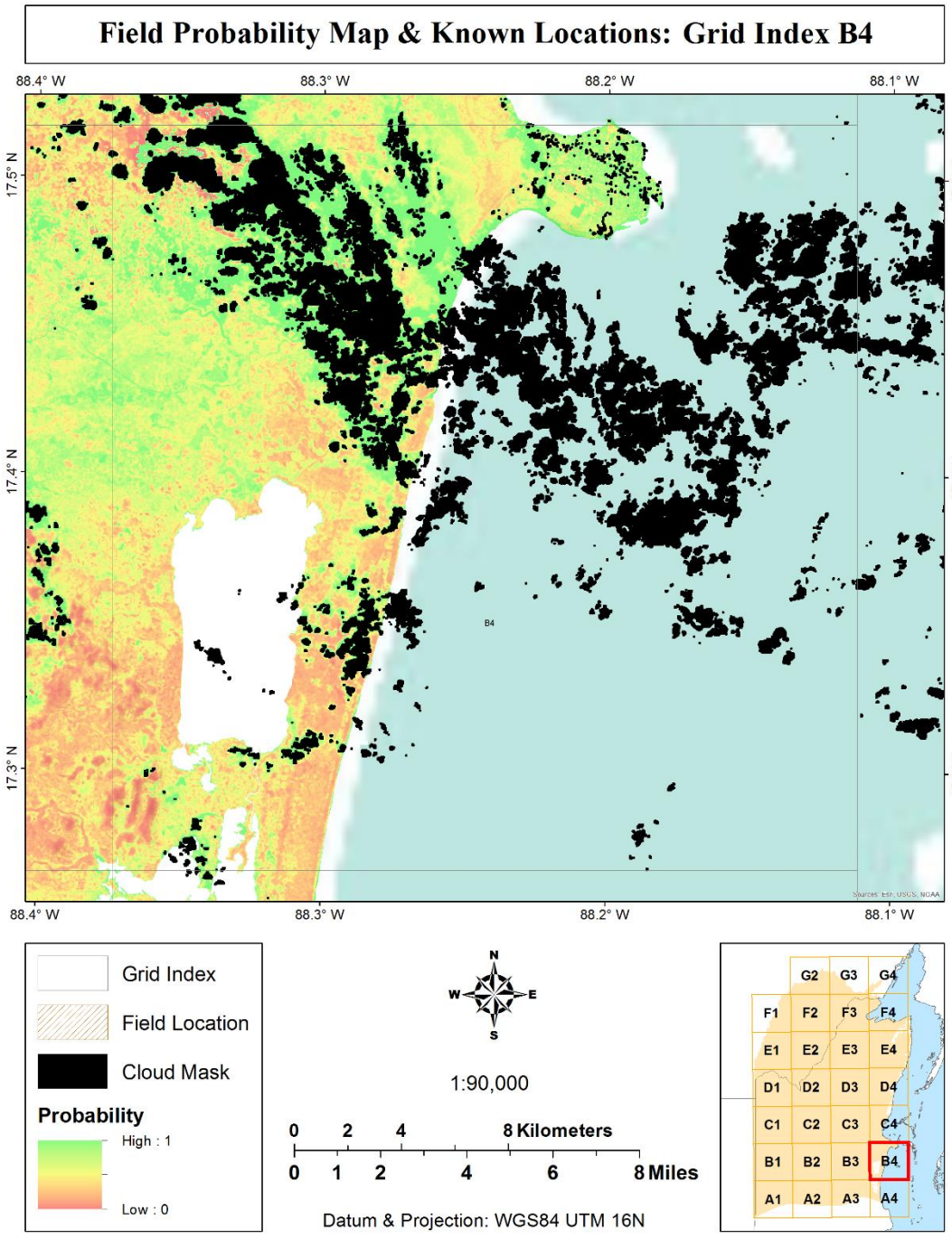

Field Probability Map & Known Locations: Grid Index A3
88.6° W
88.5° W
88.4° W
17.2° N
17.1° N
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

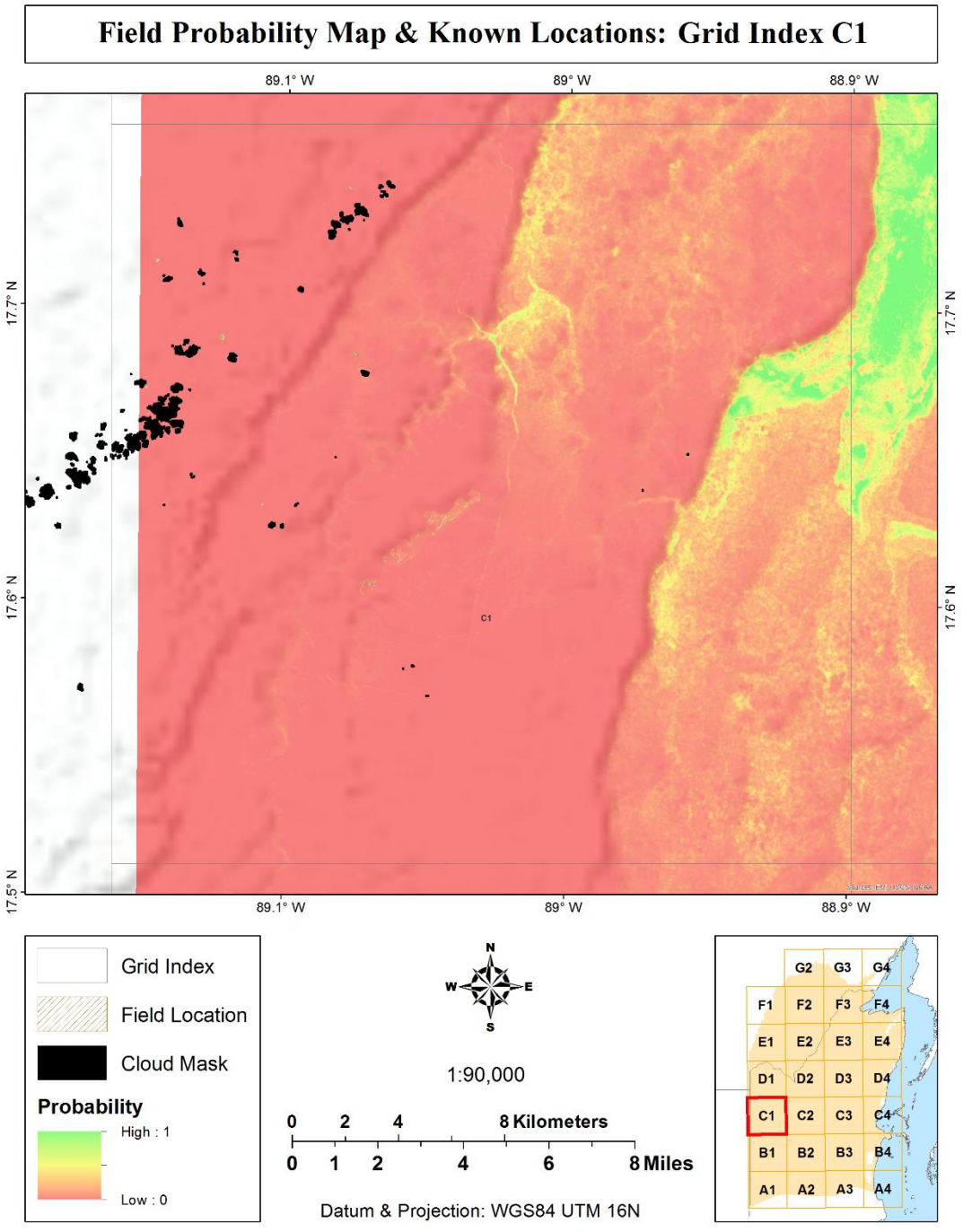

Field Probability Map & Known Locations: Grid Index A4
88.4° W
88.3° W
88.2° W
88.1° W
17.2° N
17.1° N
A4
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

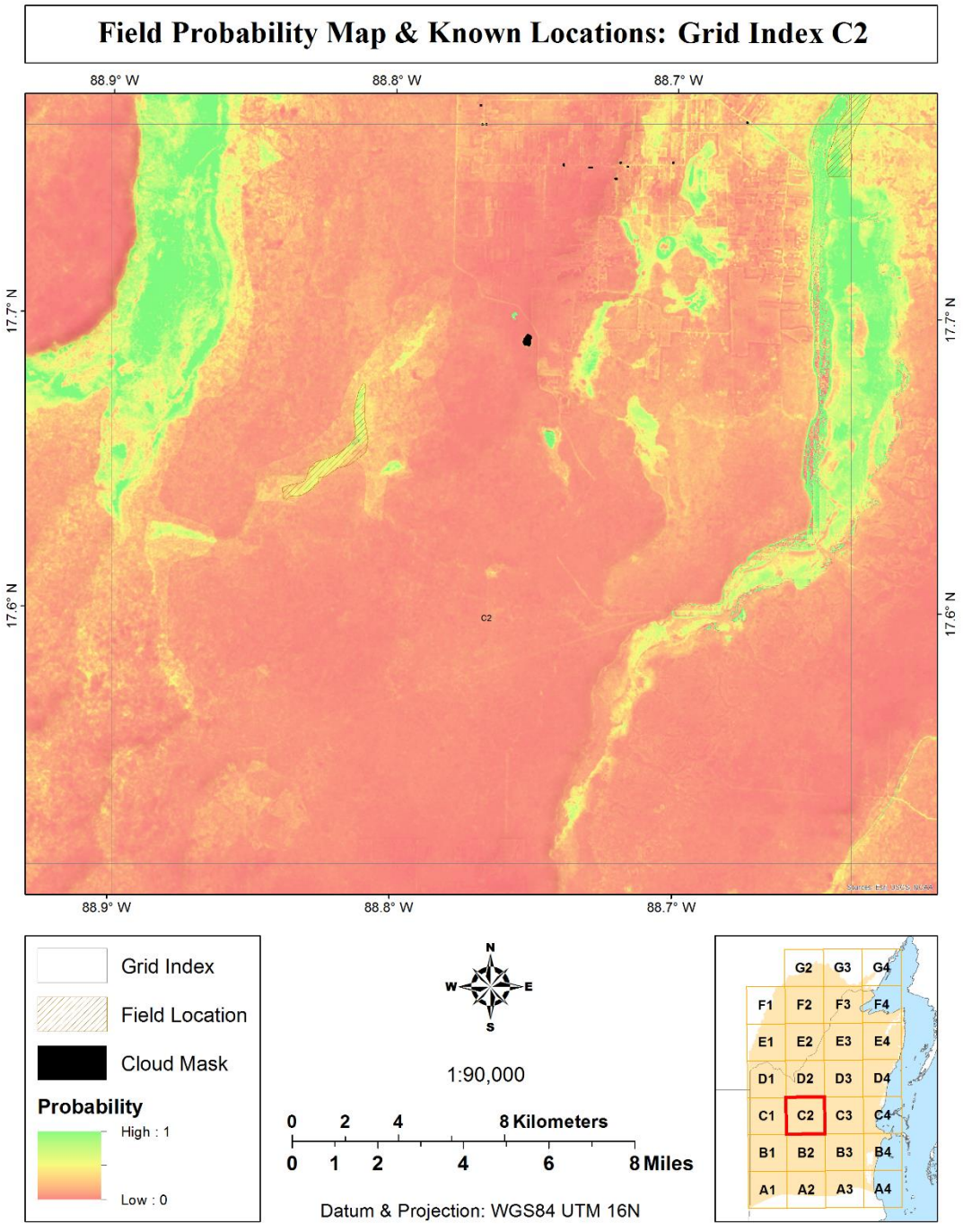

Field Probability Map & Known Locations: Grid Index B1
89.1° W
89° W
88.9° W
17.5° N
17.4° N
17.3° N
B1
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

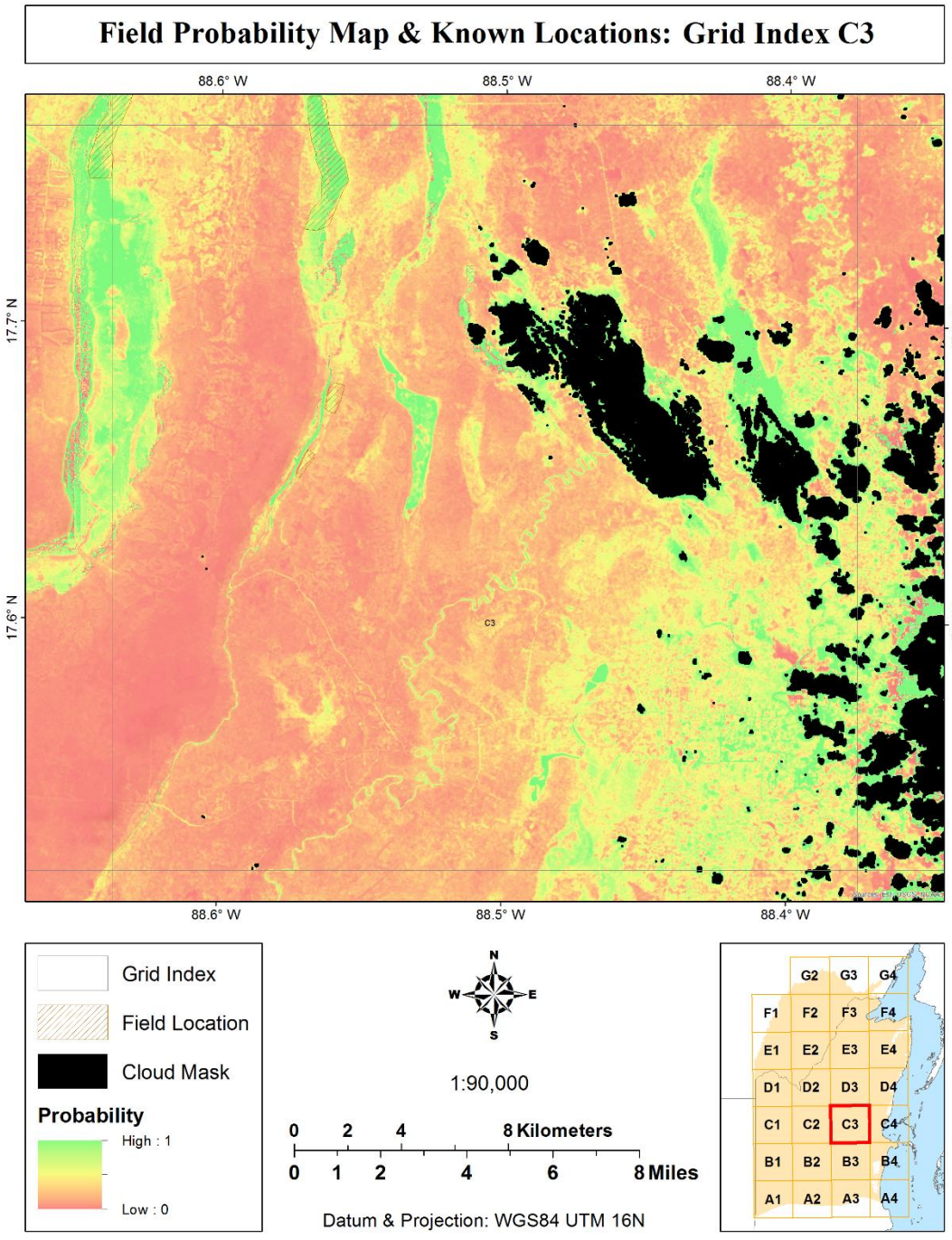

Field Probability Map & Known Locations: Grid Index B2
88.9° W
88.8° W
88.7° W
17.5° N
17.4° N
17.3° N
B2
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

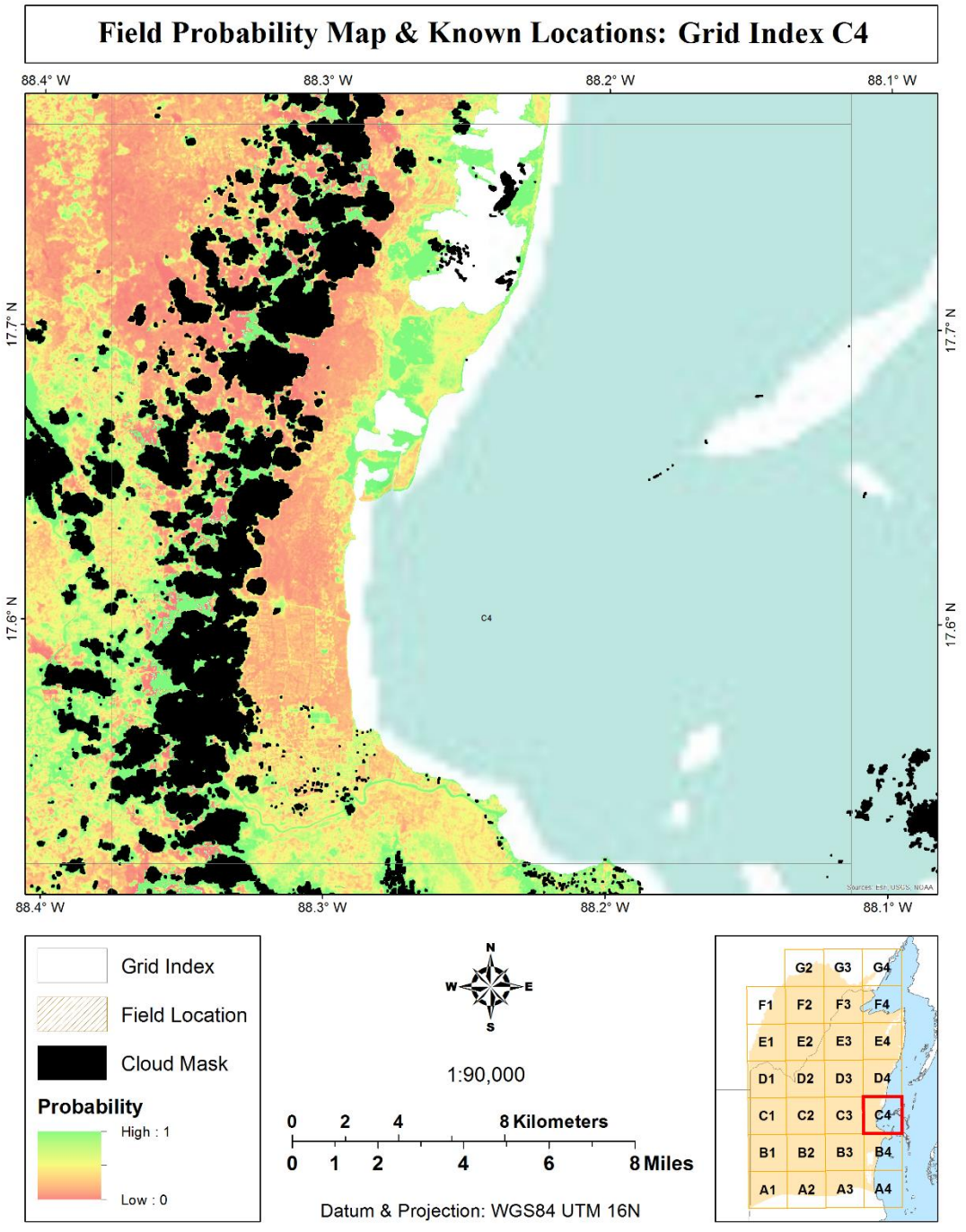

Field Probability Map & Known Locations: Grid Index B3
88.6° W
88.5° W
88.4° W
17.5° N
17.4° N
17.3° N
B3
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

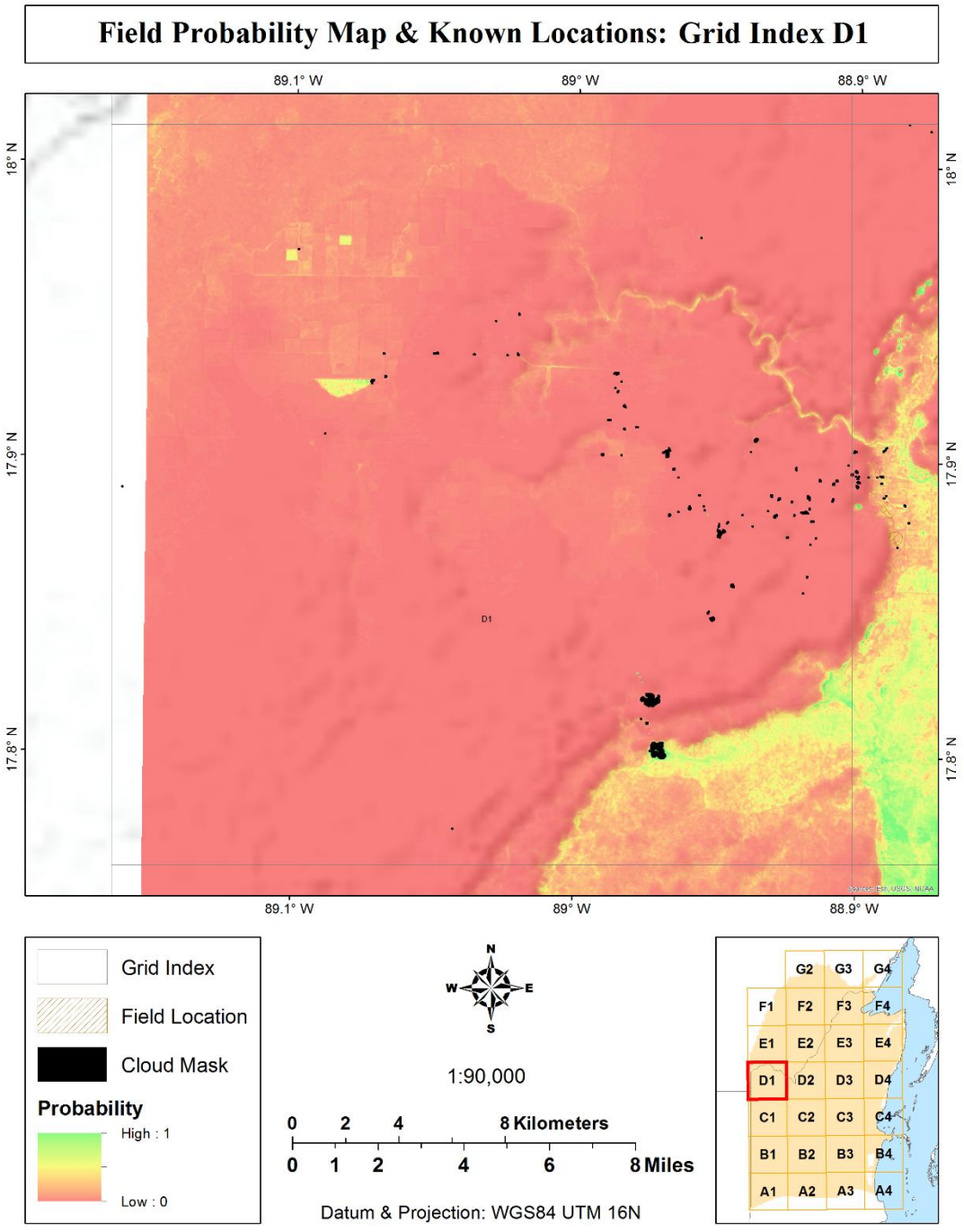

Field Probability Map & Known Locations: Grid Index B4
88.4° W
88.3° W
88.2° W
88.1° W
17.5° N
17.4° N
17.3° N
B4
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

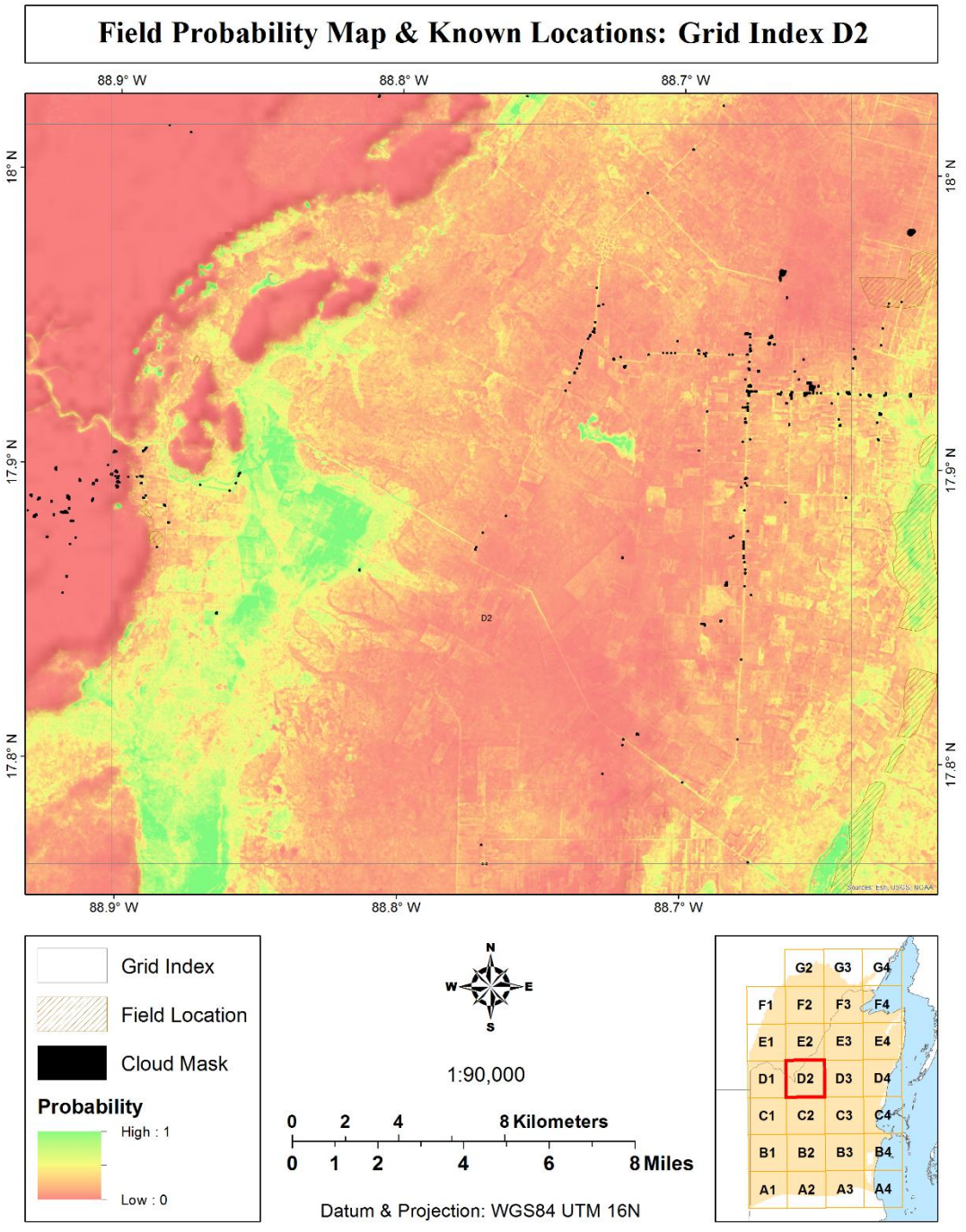

Field Probability Map & Known Locations: Grid Index C1
89.1° W
89° W
88.9° W
17.7° N
17.6° N
17.5° N
C1
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

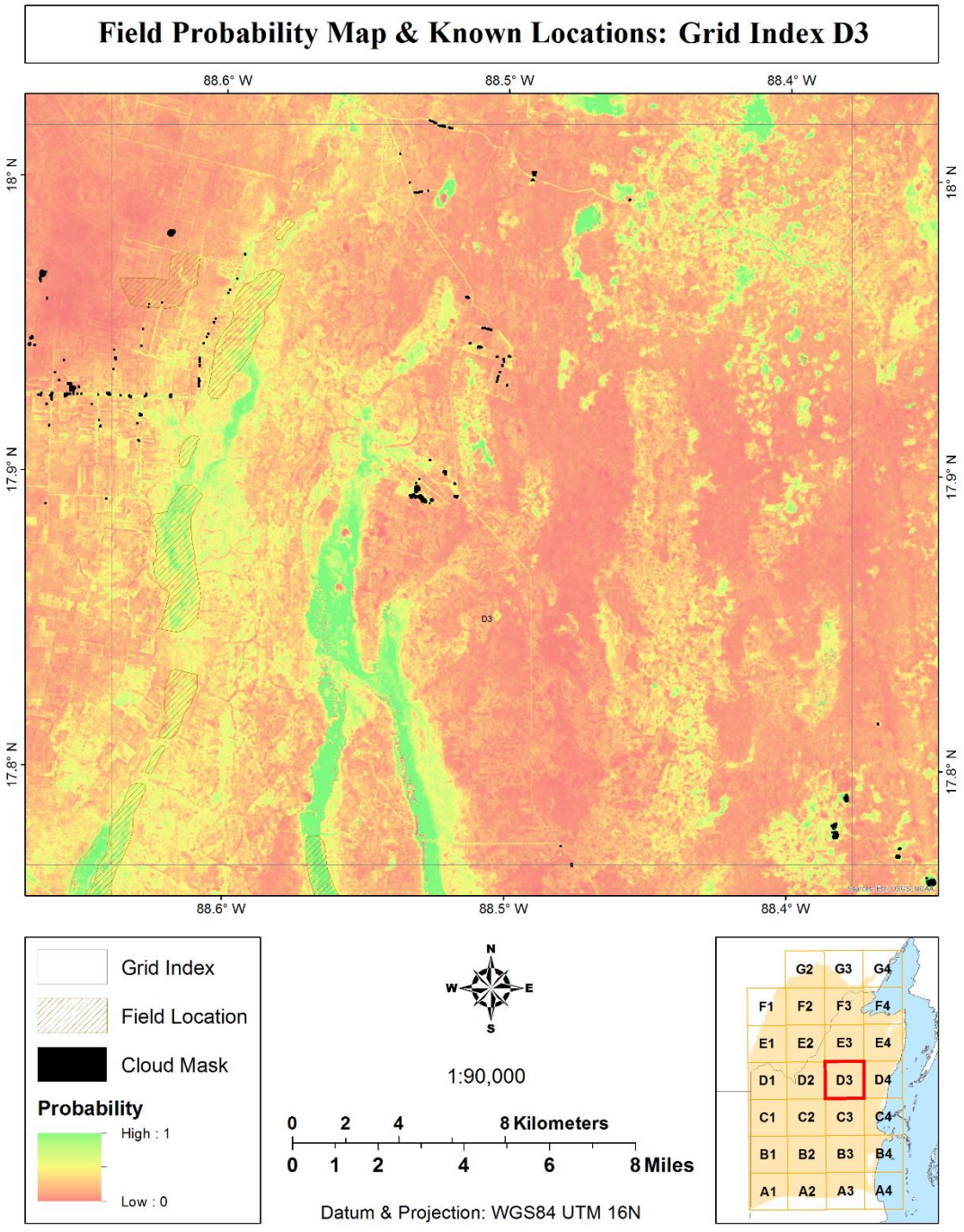

Field Probability Map & Known Locations: Grid Index C2
88.9° W
88.8° W
88.7° W
17.7° N
17.6° N
C2
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

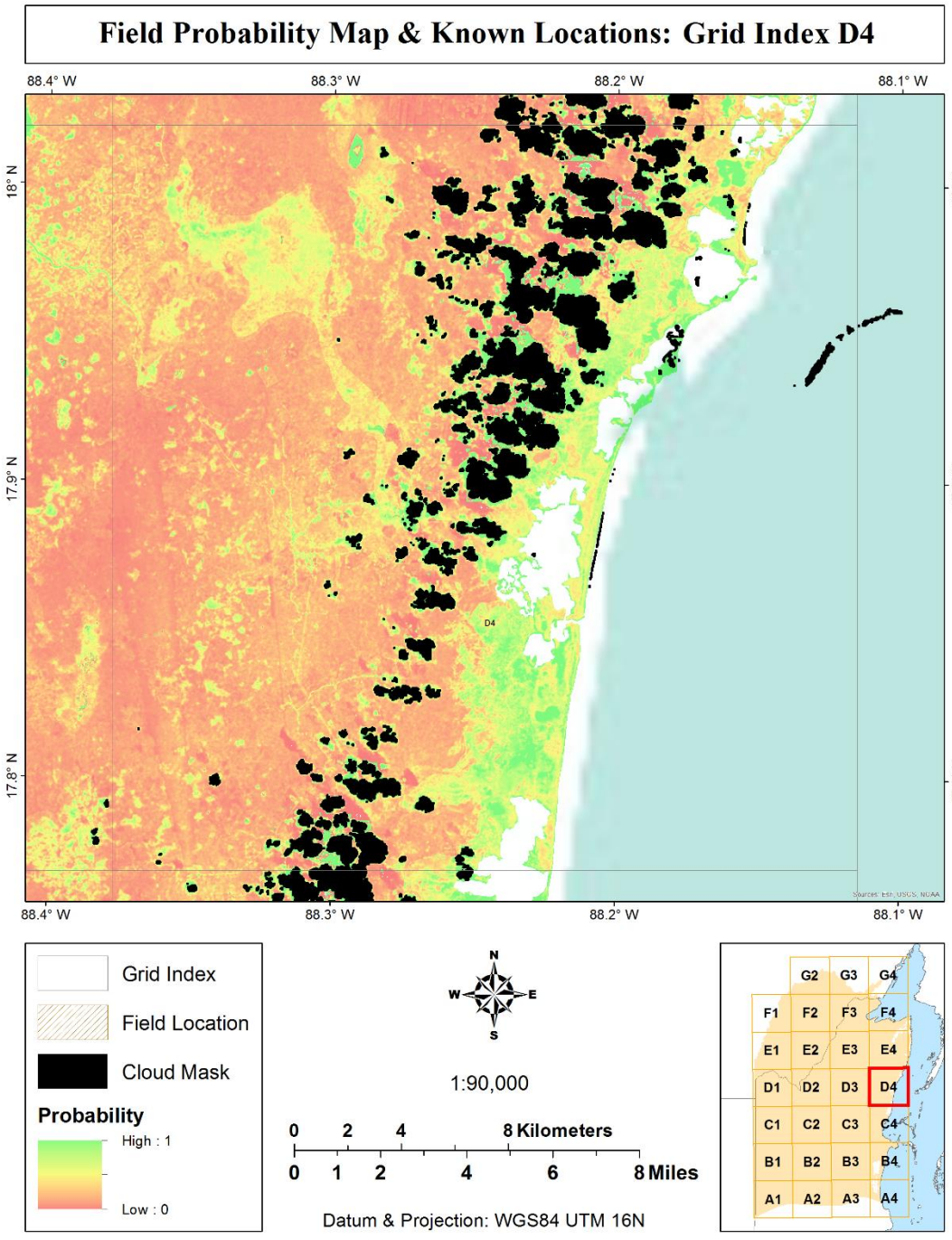

Field Probability Map & Known Locations: Grid Index C3
88.6° W
88.5° W
88.4° W
17.7° N
17.6° N
C3
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

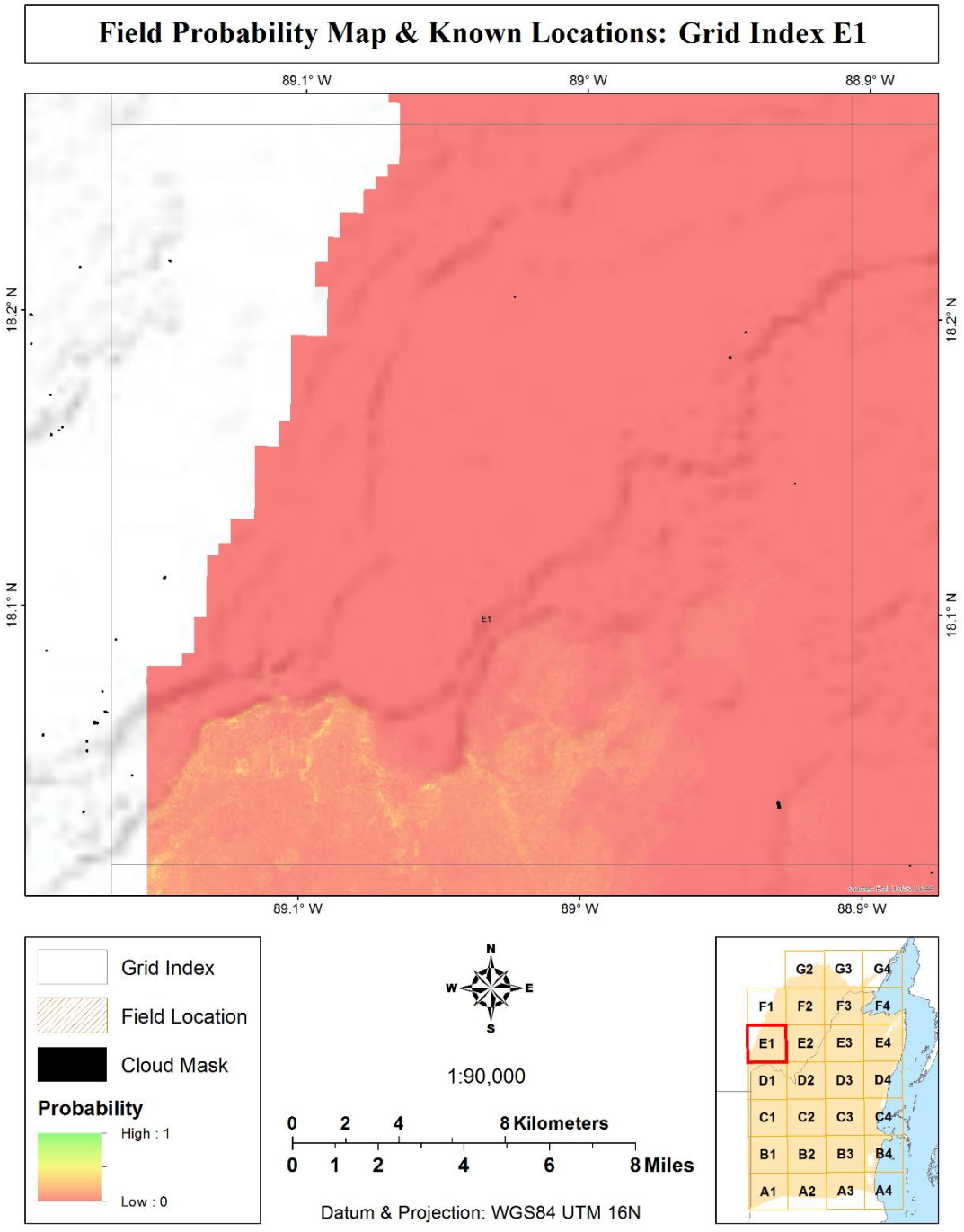

Field Probability Map & Known Locations: Grid Index C4
88.4° W
88.3° W
88.2° W
88.1° W
17.7° N
17.6° N
C4
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

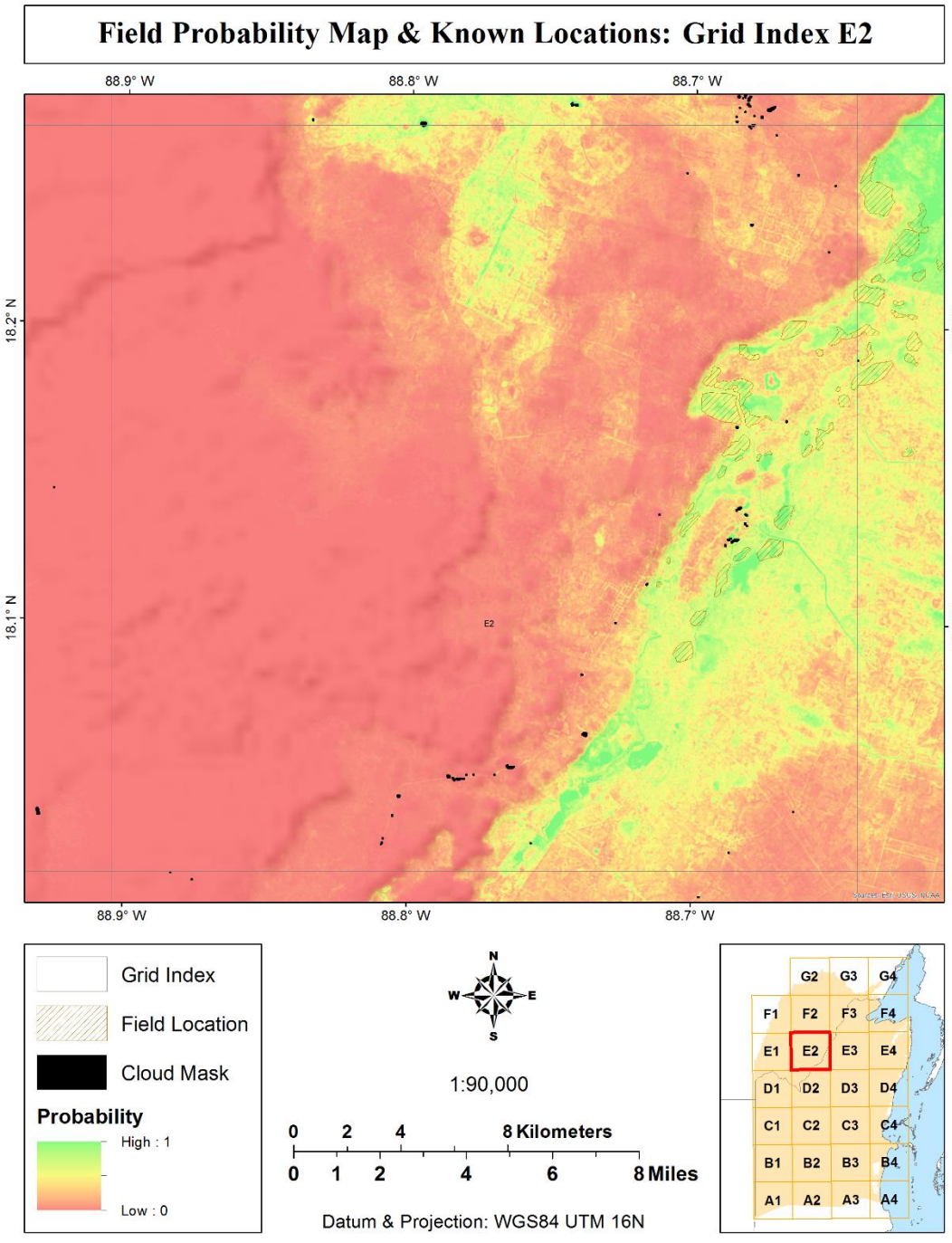

Field Probability Map & Known Locations: Grid Index D1
89.1° W
89° W
88.9° W
18° N
17.9° N
17.8° N
D1
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

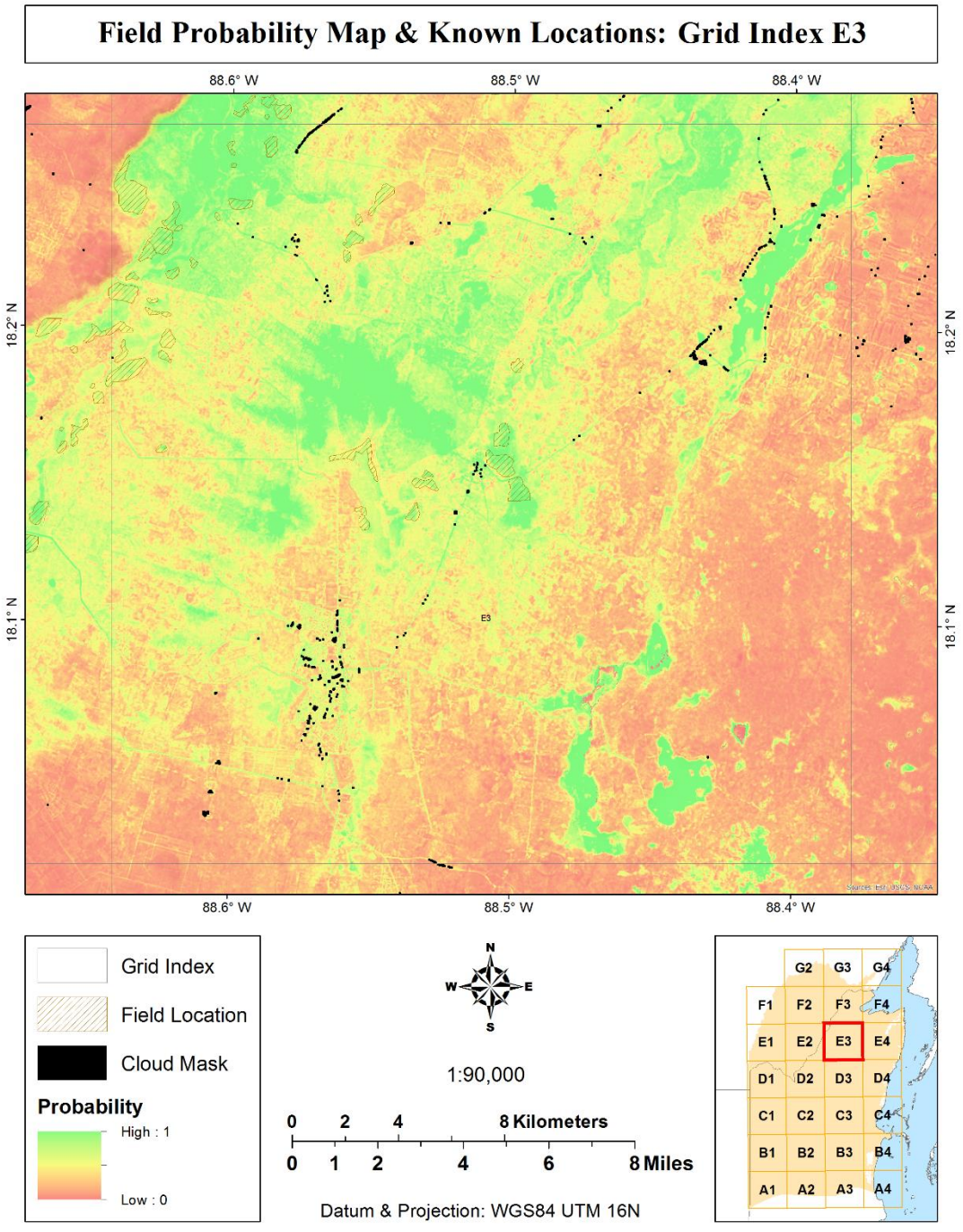

Field Probability Map & Known Locations: Grid Index D2
88.9° W
88.8° W
88.7° W
18° N
17.9° N
17.8° N
D2
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

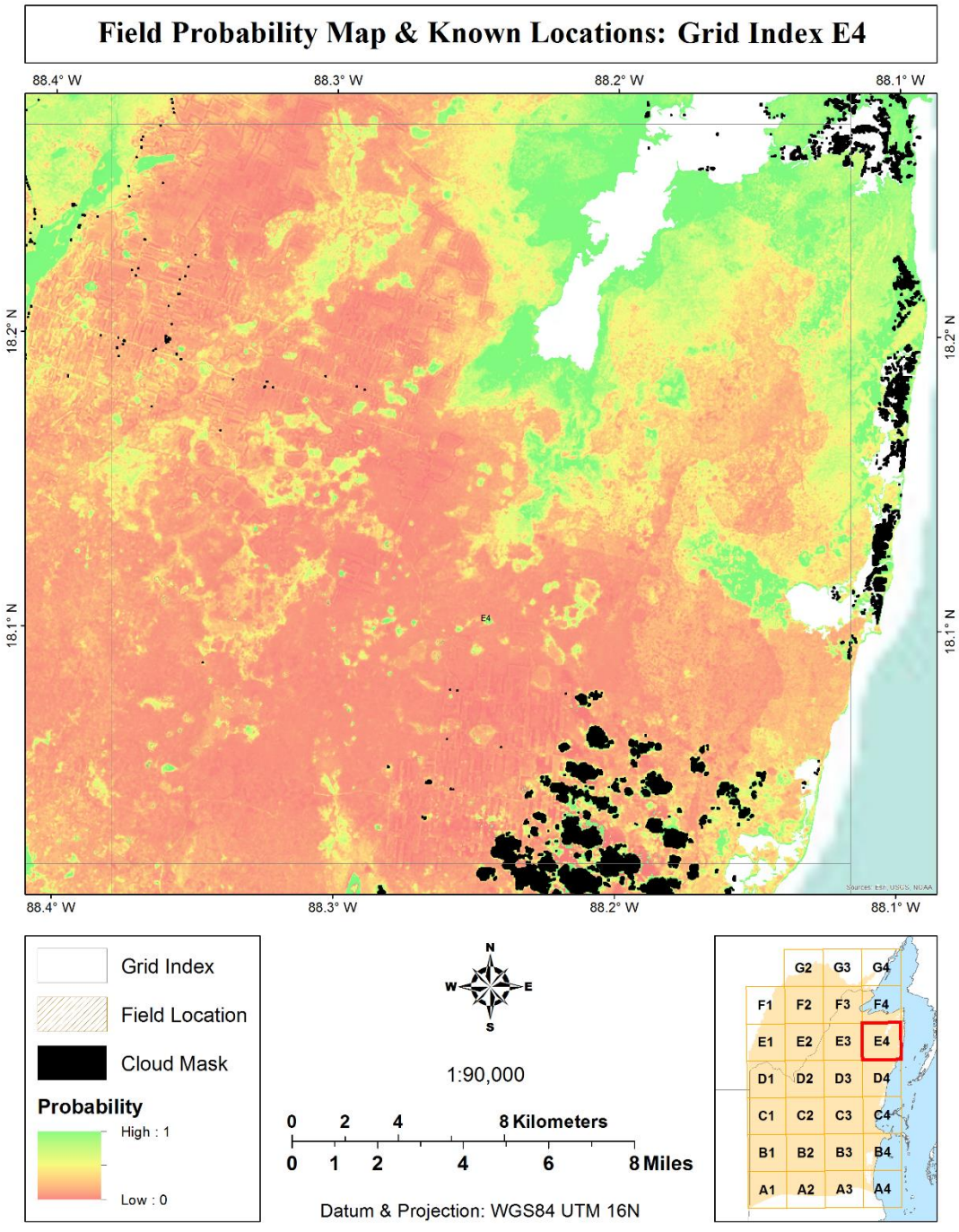

Field Probability Map & Known Locations: Grid Index D3
88.6° W
88.5° W
88.4° W
18° N
17.9° N
17.8° N
D3
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

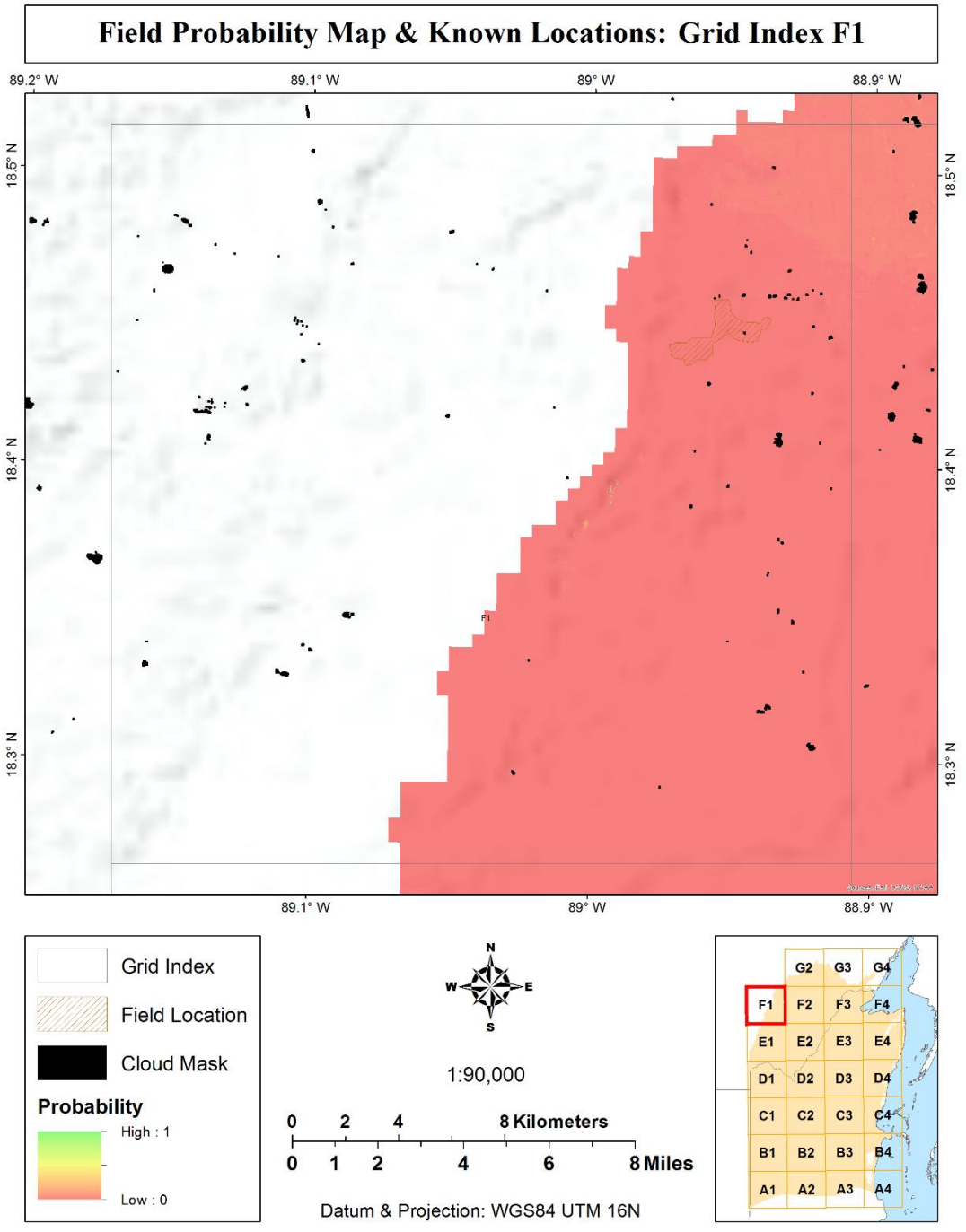

Field Probability Map & Known Locations: Grid Index D4
88.4° W
88.3° W
88.2° W
88.1° W
18° N
17.9° N
17.8° N
D4
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

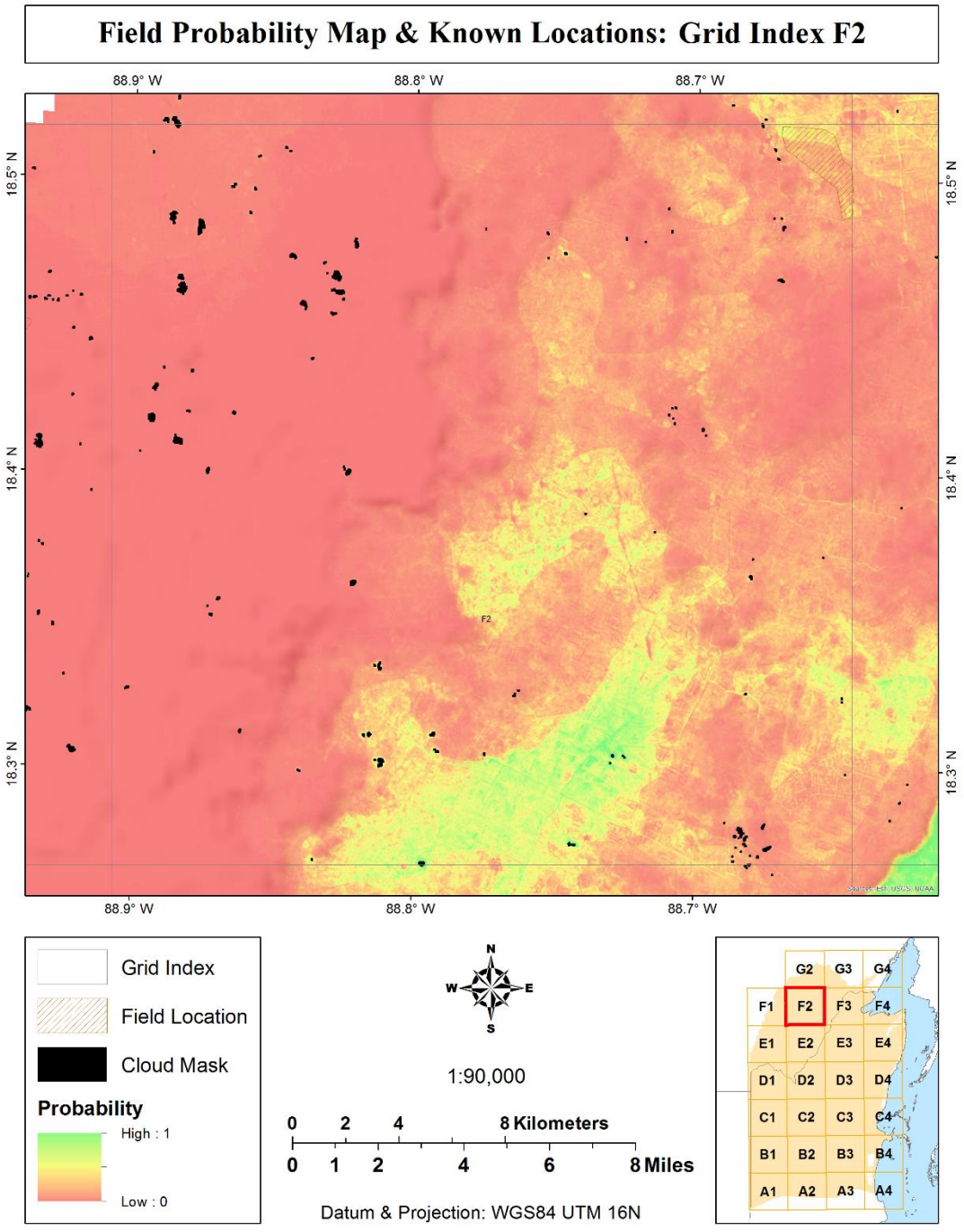

Field Probability Map & Known Locations: Grid Index E1
89.1° W
89° W
88.9° W
18.2° N
18.1° N
E1
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

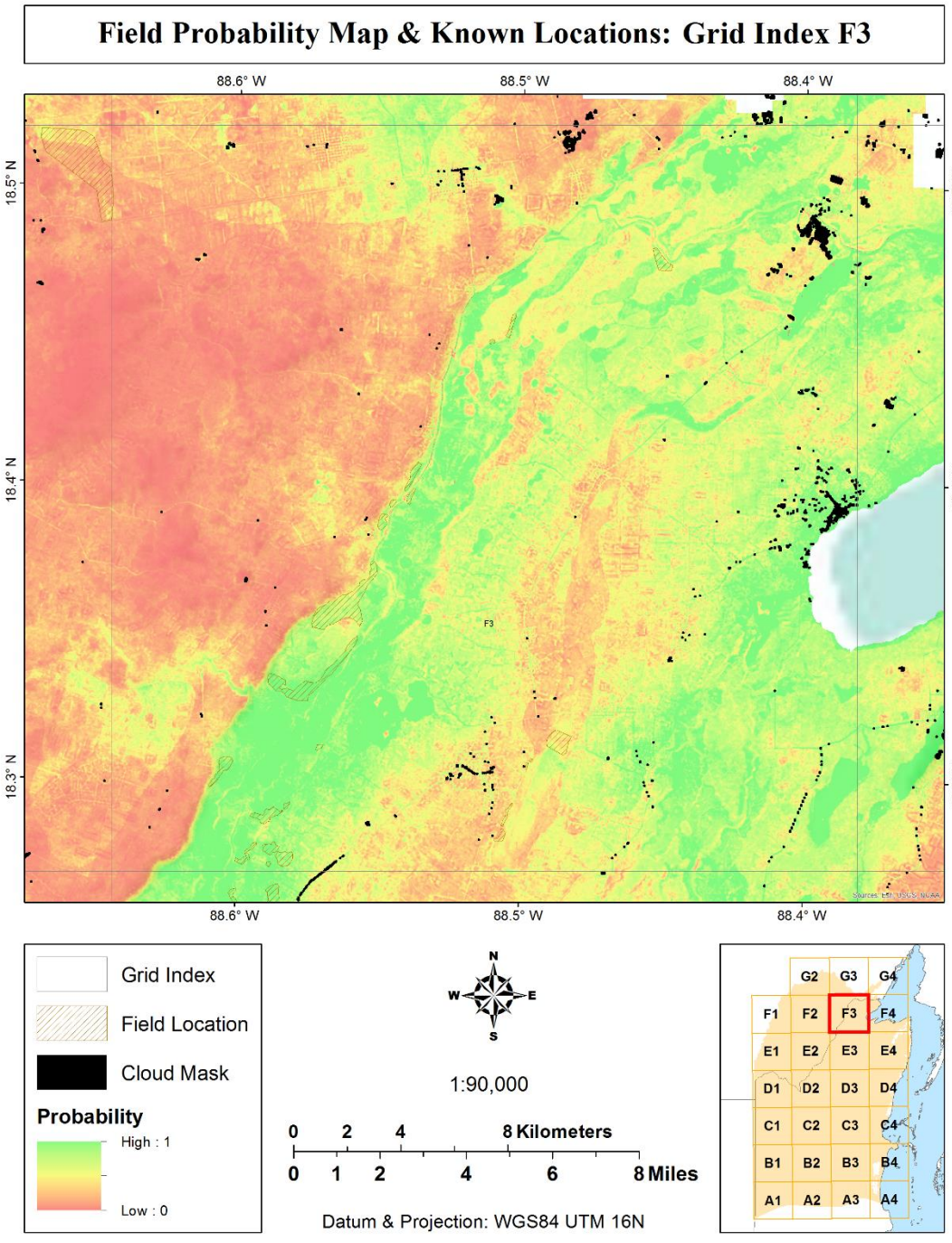

Field Probability Map & Known Locations: Grid Index E2
88.9° W
88.8° W
88.7° W
18.2° N
18.1° N
E2
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

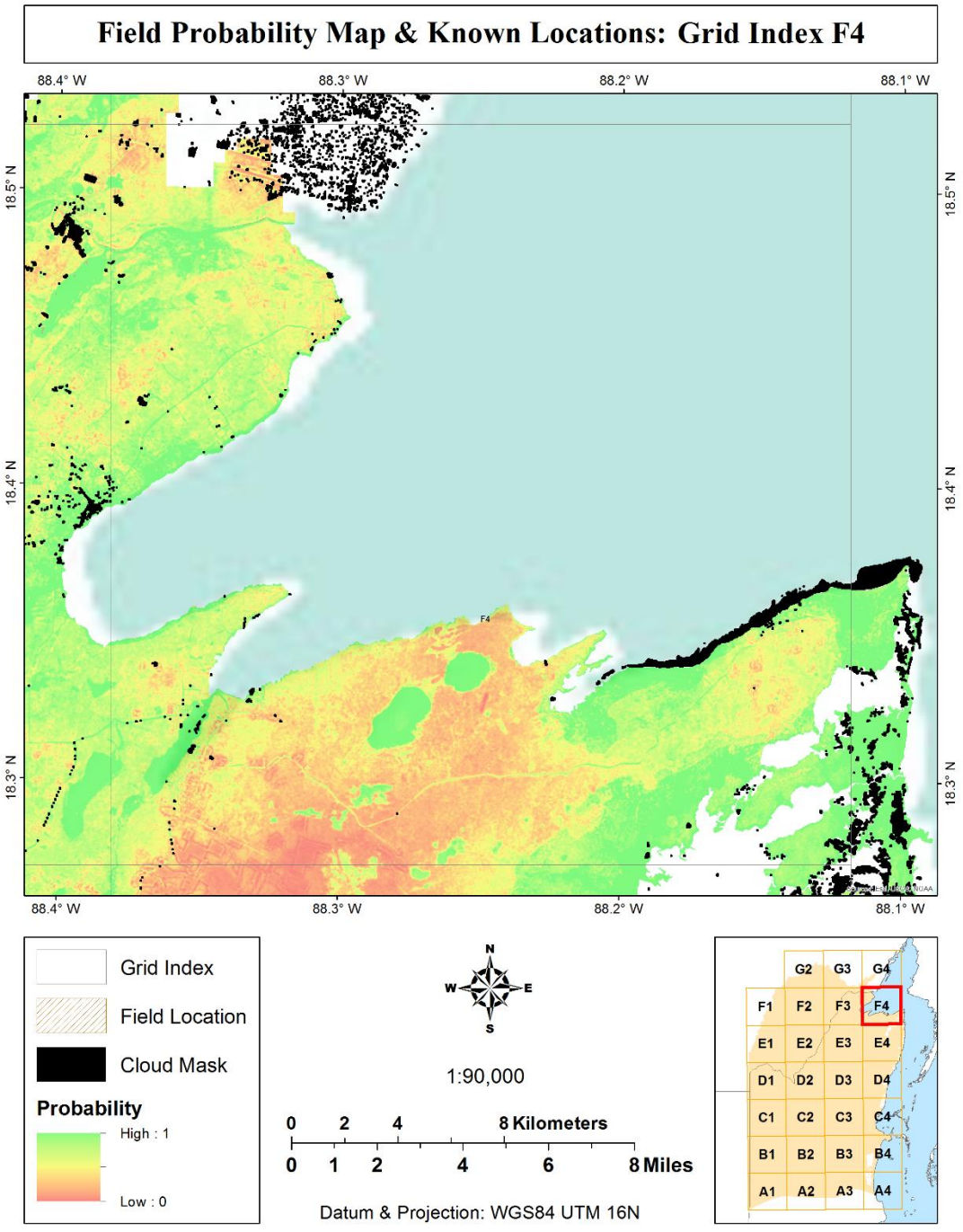

Field Probability Map & Known Locations: Grid Index E3
88.6° W
88.5° W
88.4° W
18.2° N
18.1° N
E3
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

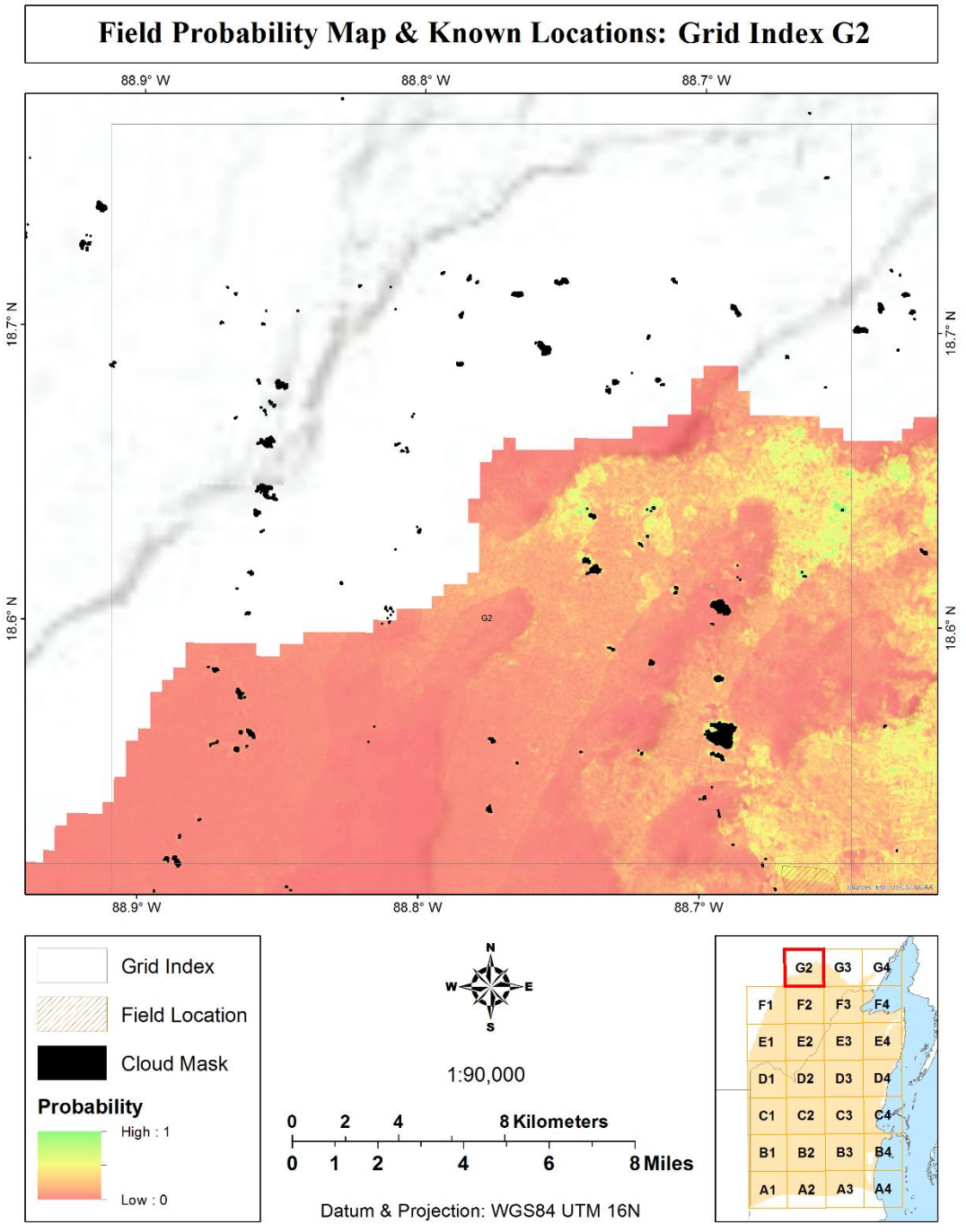

Field Probability Map & Known Locations: Grid Index E4
88.4° W
88.3° W
88.2° W
88.1° W
18.2° N
18.1° N
E4
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

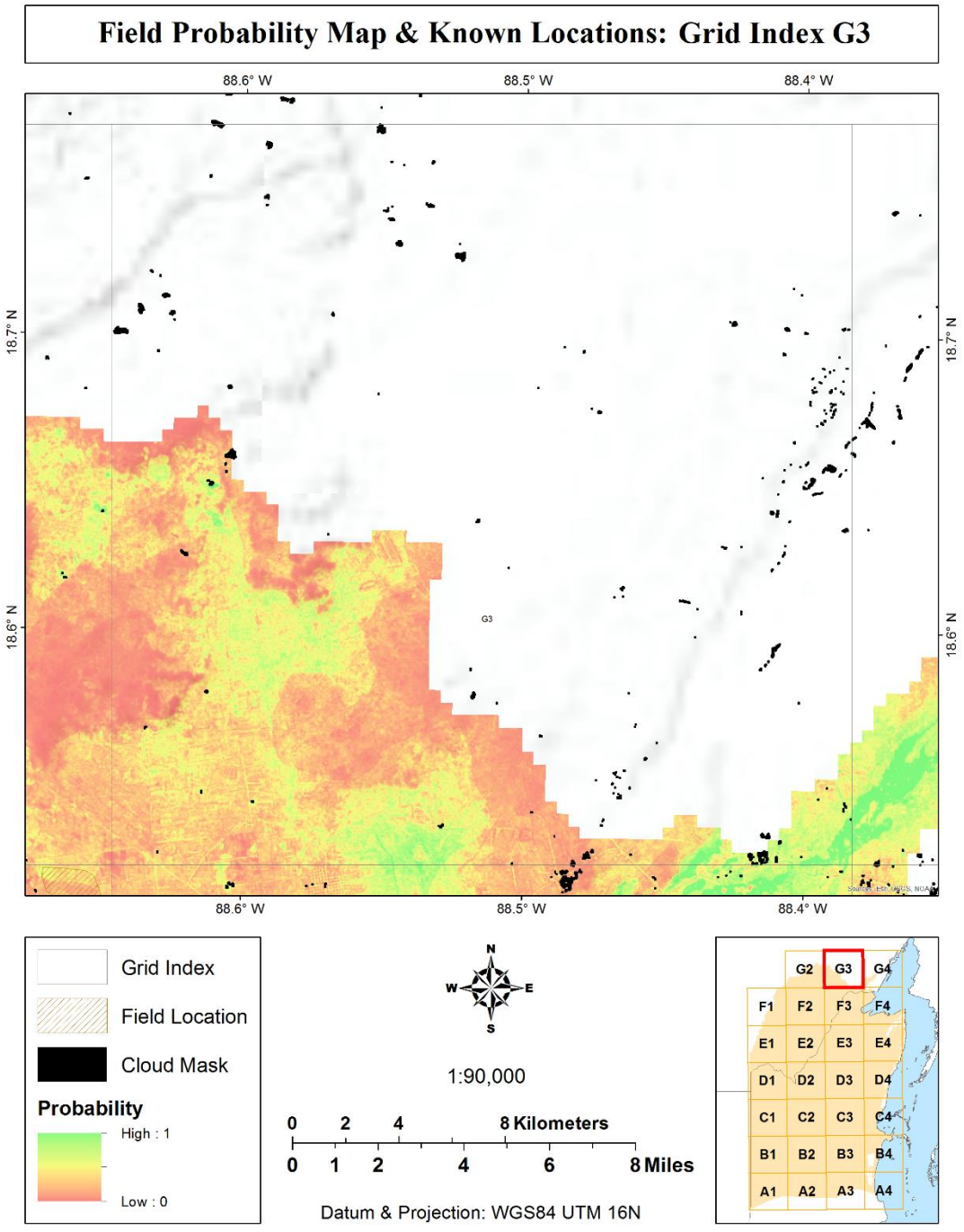

Field Probability Map & Known Locations: Grid Index F1
89.2° W
89.1° W
89° W
88.9° W
18.5° N
18.4° N
18.3° N
F1
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

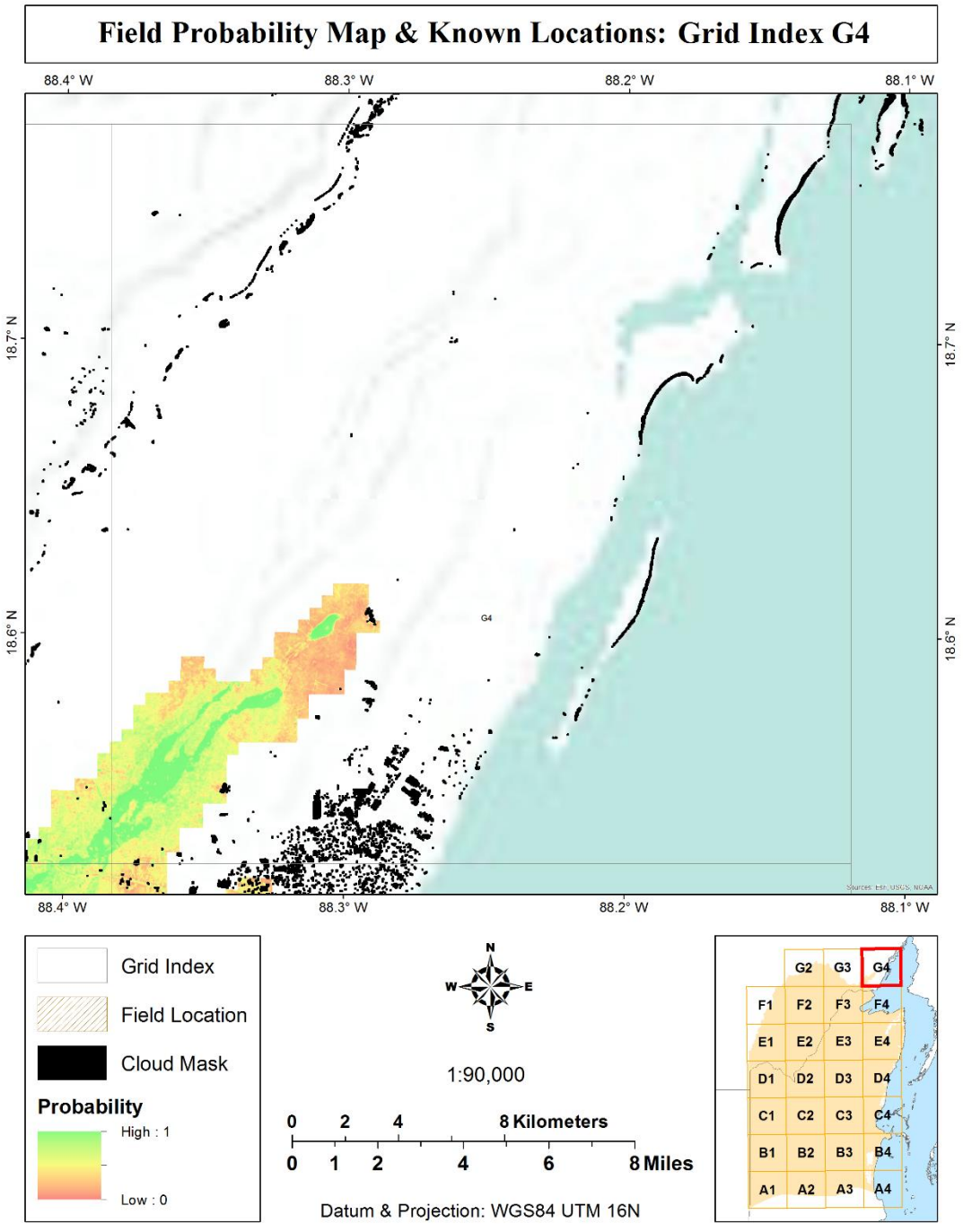

Field Probability Map & Known Locations: Grid Index F2
88.9° W
88.8° W
88.7° W
18.5° N
18.4° N
18.3° N
F2
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4


Field Probability Map & Known Locations: Grid Index F3
88.6° W
88.5° W
88.4° W
18.5° N
18.4° N
18.3° N
F3
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4


Field Probability Map & Known Locations: Grid Index F4
88.4° W
88.3° W
88.2° W
88.1° W
18.5° N
18.4° N
18.3° N
F4
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4


Field Probability Map & Known Locations: Grid Index G2
88.9° W
88.8° W
88.7° W
18.7° N
18.6° N
G2
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4


Field Probability Map & Known Locations: Grid Index G3
88.6° W
88.5° W
88.4° W
18.7° N
18.6° N
G3
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4


Field Probability Map & Known Locations: Grid Index G4
88.4° W
88.3° W
88.2° W
88.1° W
18.7° N
18.6° N
G4
Grid Index
Field Location
Cloud Mask
Probability
High : 1
Low : 0
N
W
E
S
1:90,000
0 2 4 8 Kilometers
0 1 2 4 6 8 Miles
Datum & Projection: WGS84 UTM 16N
G2 G3 G4
F1 F2 F3 F4
E1 E2 E3 E4
D1 D2 D3 D4
C1 C2 C3 C4
B1 B2 B3 B4
A1 A2 A3 A4

APPENDIX B – OHIO SOIL SAMPLE MEHLICH-III P VALUES AND PLSR PREDICTED P VALUES

	Mehlich-III	PLSR Y Predicted
V1	8.0	13.197
V2	38.0	20.958
V3	28.0	18.922
V4	14.0	13.108
V5	14.0	11.137
V6	12.0	11.770
V7	15.0	11.749
V8	6.0	11.219
V9	25.0	11.198
V10	12.0	13.893
V11	25.0	24.450
V12	18.0	19.922
V13	7.0	15.304
V14	17.0	18.922
V15	14.0	15.036
V16	7.0	14.998
V17	10.0	13.233
V18	4.0	5.579
V19	12.0	12.011
V20	4.0	7.336
V21	15.0	14.649
V22	32.0	16.093
V23	18.0	12.004
V24	24.0	13.656
V25	16.0	20.938
V26	21.0	14.368
V27	22.0	15.375
V28	39.0	12.891
V29	22.0	12.139
V30	4.5	13.244
V31	8.0	13.383
V32	6.0	14.981
V33	17.0	18.501
V34	13.0	16.534
V35	10.0	20.036

V36	10.0	18.053
V37	18.0	18.722
V38	9.0	16.167
V39	11.0	14.399
V40	23.0	20.665
V41	10.0	13.448
V42	8.0	10.969
V43	5.0	9.820
V44	7.0	13.726
V45	4.0	13.800

APPENDIX C – BELIZE SOIL SAMPLE MEHLICH-II P VALUES AND PLSR PREDICTED P VALUES

	Mehlich-II	PLSR Y Predicted
V1	5.3	5.070
V2	5.0	5.103
V3	4.8	4.980
V4	9.4	4.448
V5	5.3	6.016
V6	3.9	6.218
V7	15.1	8.042
V8	5.1	5.151
V9	5.7	5.268
V10	5.4	5.367
V11	5.3	5.224
V12	10.1	5.532
V13	3.8	5.643
V14	5.7	4.956
V15	5.7	4.464
V16	4.7	5.113
V17	4.5	5.214
V18	4.3	4.589
V19	4.0	4.288
V20	3.6	4.590
V21	3.7	4.815
V22	3.1	4.628
V23	2.5	4.674
V24	3.7	4.772
V25	3.3	8.155
V26	3.7	4.398
V27	6.3	4.896
V28	6.0	4.988
V29	5.9	6.002
V30	4.0	6.297

APPENDIX D – OHIO VNIR PLSR MODEL SUMMARY

```
> Ohio_VNIR_pls = plsr(ASD_Ref_P ~ as.matrix(ASD_Ref_VNIR), ncomp = 50,
data = Oh_VNIR_Spectra, validation="LOO")
> summary(Ohio_VNIR_pls)
Data:   X dimension: 131 2151
        Y dimension: 131 1
Fit method: kernelpls
Number of components considered: 50

VALIDATION: RMSEP
Cross-validated using 131 leave-one-out segments.
    (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
CV        17.78    17.48    16.82    16.66    16.28    16.39    16.97
adjCV     17.78    17.48    16.82    16.66    16.28    16.38    16.95
    7 comps  8 comps  9 comps  10 comps  11 comps  12 comps  13 comps
CV    16.70    16.73    16.18     16.16     16.60     16.11     15.64
adjCV 16.68    16.73    16.15     16.14     16.56     16.08     15.60
   14 comps  15 comps  16 comps  17 comps  18 comps  19 comps  20 comps
CV    15.87     15.51     15.57     15.57     15.79     15.78     15.62
adjCV 15.84     15.47     15.52     15.53     15.74     15.72     15.56
   21 comps  22 comps  23 comps  24 comps  25 comps  26 comps  27 comps
CV    15.62     15.54     15.50     15.54     15.52     15.59     15.60
adjCV 15.57     15.49     15.45     15.48     15.46     15.53     15.54
   28 comps  29 comps  30 comps  31 comps  32 comps  33 comps  34 comps
CV    15.57     15.55     15.57     15.58     15.54     15.53     15.53
adjCV 15.51     15.49     15.51     15.52     15.48     15.47     15.47
   35 comps  36 comps  37 comps  38 comps  39 comps  40 comps  41 comps
CV    15.54     15.51     15.50     15.47     15.46     15.44     15.43
adjC  15.48     15.45     15.44     15.41     15.40     15.38     15.37
   42 comps  43 comps  44 comps  45 comps  46 comps  47 comps  48 comps
CV    15.42     15.42     15.42     15.41     15.41     15.41     15.41
adjCV 15.36     15.36     15.36     15.35     15.35     15.35     15.35
   49 comps  50 comps
CV    15.41     15.41
adjCV 15.35     15.35

TRAINING: % variance explained
       1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps
X        97.39    98.46    99.26    99.53    99.60    99.65    99.69
ASD_Ref_P 4.59    14.53    18.95    28.67    36.25    46.62    56.44
     8 comps  9 comps  10 comps  11 comps  12 comps  13 comps  14 comps
X      99.79    99.81     99.83     99.84     99.86     99.87     99.88
ASD_Ref_P 58.68  66.58    71.95     79.26     83.22     85.83     87.96
           15 comps  16 comps  17 comps  18 comps  19 comps  20 comps
X             99.89     99.90     99.90     99.91     99.91     99.92
ASD_Ref_P     90.61     92.72     94.66     95.68     96.76     97.72
           21 comps  22 comps  23 comps  24 comps  25 comps  26 comps
X             99.92     99.93     99.93     99.93     99.94     99.94
ASD_Ref_P     98.22     98.59     99.02     99.27     99.42     99.61
           27 comps  28 comps  29 comps  30 comps  31 comps  32 comps
```

X	99.94	99.94	99.95	99.95	99.95	99.95
ASD_Ref_P	99.73	99.84	99.90	99.92	99.95	99.96
	33 comps	34 comps	35 comps	36 comps	37 comps	38 comps
X	99.96	99.96	99.96	99.96	99.96	99.96
ASD_Ref_P	99.97	99.98	99.99	99.99	100.00	100.00
	39 comps	40 comps	41 comps	42 comps	43 comps	44 comps
X	99.97	99.97	99.97	99.97	99.97	99.97
ASD_Ref_P	100.00	100.00	100.00	100.00	100.00	100.00
	45 comps	46 comps	47 comps	48 comps	49 comps	50 comps
X	99.97	99.97	99.98	99.98	99.98	99.98
ASD_Ref_P	100.00	100.00	100.00	100.00	100.00	100.00

APPENDIX E – OHIO LWIR PLSR MODEL SUMMARY

```
> summary(Ohio_LWIR_pls)
Data:   X dimension: 131 2231
        Y dimension: 131 1
Fit method: kernelpls
Number of components considered: 50

VALIDATION: RMSEP
Cross-validated using 131 leave-one-out segments.
    (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
CV        17.78    17.77    17.53    16.99     16.6    16.15    15.38
adjCV     17.78    17.77    17.53    16.99     16.6    16.14    15.38
      7 comps  8 comps  9 comps  10 comps  11 comps  12 comps  13 comps
CV      14.39    13.63    13.37     13.70     13.54     13.48     13.42
adjCV   14.39    13.62    13.36     13.68     13.58     13.45     13.38
   14 comps  15 comps  16 comps  17 comps  18 comps  19 comps  20 comps
CV    13.34     13.51     13.53     13.61     13.60     13.62     13.62
adjCV 13.29     13.46     13.48     13.56     13.55     13.57     13.57
   21 comps  22 comps  23 comps  24 comps  25 comps  26 comps  27 comps
CV    13.62     13.62     13.62     13.62     13.62     13.62     13.62
adjCV 13.57     13.57     13.57     13.56     13.56     13.56     13.56
   28 comps  29 comps  30 comps  31 comps  32 comps  33 comps  34 comps
CV    13.62     13.62     13.62     13.62     13.62     13.62     13.62
adjCV 13.56     13.56     13.56     13.56     13.56     13.56     13.56
   35 comps  36 comps  37 comps  38 comps  39 comps  40 comps  41 comps
CV    13.62     13.62     13.62     13.62     13.62     13.62     13.62
adjCV 13.56     13.56     13.56     13.56     13.56     13.56     13.56
   42 comps  43 comps  44 comps  45 comps  46 comps  47 comps  48 comps
CV    13.62     13.62     13.62     13.62     13.62     13.62     13.62
adjCV 13.56     13.56     13.56     13.56     13.56     13.56     13.56
       49 comps  50 comps
CV        13.62     13.62
adjCV     13.56     13.56

TRAINING: % variance explained
         1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps
X        91.5497   97.798    98.74    99.33    99.59    99.69    99.76
FTIR_Ref_P 0.9998    5.874    14.94    20.12    28.58    40.43    53.40
            8 comps  9 comps  10 comps  11 comps  12 comps  13 comps
X             99.82    99.83     99.85     99.88     99.89     99.90
FTIR_Ref_P    63.55    73.53     80.46     86.85     94.19     96.17
            14 comps  15 comps  16 comps  17 comps  18 comps  19 comps
X              99.90     99.91     99.91     99.91     99.91     99.92
FTIR_Ref_P     98.74     99.50     99.86     99.96     99.99    100.00
            20 comps  21 comps  22 comps  23 comps  24 comps  25 comps
X              99.92     99.92     99.92     99.92     99.92     99.92
FTIR_Ref_P    100.00    100.00    100.00    100.00    100.00    100.00
            26 comps  27 comps  28 comps  29 comps  30 comps  31 comps
X              99.92     99.93     99.93     99.93     99.93     99.93
FTIR_Ref_P    100.00    100.00    100.00    100.00    100.00    100.00
```

	32 comps	33 comps	34 comps	35 comps	36 comps	37 comps
X	99.93	99.93	99.93	99.93	99.93	99.94
FTIR_Ref_P	100.00	100.00	100.00	100.00	100.00	100.00
	38 comps	39 comps	40 comps	41 comps	42 comps	43 comps
X	99.94	99.94	99.94	99.94	99.94	99.94
FTIR_Ref_P	100.00	100.00	100.00	100.00	100.00	100.00
	44 comps	45 comps	46 comps	47 comps	48 comps	49 comps
X	99.94	99.94	99.94	99.94	99.95	99.95
FTIR_Ref_P	100.00	100.00	100.00	100.00	100.00	100.00
	50 comps					
X	99.95					
FTIR_Ref_P	100.00					

REFERENCES

Abdi, Dalel, Gaëtan F. Tremblay, Noura Ziadi, Giles Bélanger, and Léon-Étienne Parent. 2012. "Predicting soil phosphorus-related properties using near-infrared reflectance spectroscopy." *Soil Science Society of America Journal* 76 (6): 2318-2326.

Abdi, Hervé. 2010. "Partial least squares regression and projection on latent structure regression (PLS Regression)." *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (1): 97–106.

Adams, Richard E. W., T. Patrick Culbert, Walter E. Brown, Peter D. Harrison, and Laura J. Levi. 1990. "Commentary: Rebuttal to Pope and Dahlin." *Journal of Field Archaeology 17.2* 17 (2): 241-244.

Adams, Richard E. W., Walter E. Brown Jr., and T. Patrick Culbert. 1981. "Radar mapping, Archaeology, and ancient Maya land use." *Science* 213 (4515): 1457-1463.

Araújo, Suzana Romeiro, Mats Söderström, Jan Eriksson, Christian Isendahl, Per Stenborg, and José M. Demattê. 2015. "Determining soil properties in Amazonian Dark Earths by reflectance spectroscopy." *Geoderma* 237: 308-317.

Authentic Maya. 2005. *Maya Agriculture.* Accessed 2015. http://www.authenticmaya.com/maya_agriculture.htm.

Ballard, Jerrell R., Stacy E. Howington, and Steven C. Wilhelms. 2013. "Laboratory-based rainfall effects on LWIR soil reflectance." *IEEE Geoscience and Remote Sensing Letters* 10 (3): 627-630.

Bauman, Paul R. 2009. *History of Remote Sensing, Satellite Imagery, Part II.* Accessed April 7, 2015. http://www.oneonta.edu/faculty/baumanpr/geosat2/RS%20History%20II/RS-History-Part-2.html.

Beach, Tim, Sheryl Luzzadder-Beach, and Nicholas Dunning. 2006. "A Soils History of Mesoamerica and the Caribbean Islands." In *Soils and Societies*, by J. R. McNeill and Verena Winiwarter, 51-90. The White Horse Press.

Beach, Timothy, Nicholas Dunning, Sheryl Luzzadder-Beach, and Vernon Scarborough. 2003. "Depression Soils in the Lowland Tropics of Northwestern Belize: Anthropogenic and Natural Origins." In *Lowland Maya Area: Three Millennia at the Human-Wildland Interface*, by A. Gomez-Pompa, M. Allen, S. Fedick and J. Jiménez-Osornio, 139-174. Binghamton, NY: Haworth Press.

Beach, Timothy, Nicholas Dunning, Sheryl Luzzadder-Beach, D. E. Cooke, and J. Lohse. 2006. "Impacts of the ancient Maya on soils and soil erosion in the central Maya Lowlands." *Catena* 65 (2): 166-178.

Beach, Timothy, Sheryl Luzzadder-Beach, Nicholas Dunning, and Duncan Cooke. 2008. "Human and natural impacts on fluvial and karst depressions of the Maya Lowlands." *Geomorphology* 101 (1-2): 308-331.

Beach, Timothy, Sheryl Luzzadder-Beach, Nicholas Dunning, John Jones, John Lohse, Tom Guderjan, Steve Bozarth, Sarah Millspaugh, and Tripti Bhattacharya. 2009. "A review of human and natural changes in Maya Lowland wetlands over the Holocene." *Quaternary Science Reviews* 28 (17-18): 1710-1724.

Beach, Timothy, Sheryl Luzzadder-Beach, Richard Terry, Nicholas Dunning, Stephen Houston, and Thomas Garrison. 2011. "Carbon isotopic ratios of wetland and terrace soil sequences in the Maya Lowlands of Belize and Guatemala." *Catena* 85 (2): 109-118.

Biodiversity and Environmental Resource Data System of Belize (BERDS). 2012. *Spatial Data Warehouse.* October 7. Accessed July 1, 2012. http://www.biodiversity.bz/mapping/warehouse/ (now defunct).

Bogrekci, I., and W. S. Lee. 2005. "Spectral soil signatures and sensing phosphorus." *Biosystems Engineering* 527-533.

Bradley, Michael. 2016. *FTIR Basics.* Accessed November 2016. https://www.thermofisher.com/us/en/home/industrial/spectroscopy-elemental-isotope-analysis/spectroscopy-elemental-isotope-analysis-learning-center/molecular-spectroscopy-information/ftir-information/ftir-basics.html.

BRUKER. 2016. *Handheld XRF: How it works.* Accessed September 26, 2016. https://www.bruker.com/products/x-ray-diffraction-and-elemental-analysis/handheld-xrf/how-xrf-works.html.

Bruker. 2016. *Understanding what's happening when you take an assay.* Accessed August 2016. https://www.bruker.com/products/x-ray-diffraction-and-elemental-analysis/handheld-xrf/how-do-i-know-if-my-xrf-analyzer-is-working.html.

Bundschuh, Jochen, and Guillermo E. Alvarado Induni. 2007. *Central America: geology, resources, and hazards.* Taylor & Francis.

Burns, Robert P., and Burns Richard. 2008. *Business Research Methods and Statistics Using SPSS.* London: Sage Publications.

Busman, Lowell, John Lamb, Gyles Randall, George Rehm, and Michael Schmitt. 2002. *The Nature of Phosphorous in Soils.* Accessed October 10, 2013. http://www.extension.umn.edu/distribution/cropsystems/dc6795.html.

Carleton, Chris W., James Conolly, and Gyles Ianonne. 2012. "A locally-adaptive model of archaeological potential (LAMAP)." *Journal of Archaeological Science* 3 (11): 3371-3385.

Chase, Arlen F., Diane Z. Chase, and John F. Weishampel. 2010. "Lasers in the Jungle." *Archaeology* 64 (4): 29-31.

Chase, Arlen F., Diane Z. Chase, and John F. Weishampel. 2013. "The Use of LiDAR at the Maya Site of Caracol, Belize." In *Mapping Archaeological Landscapes from Space*, by Douglas C., and Michael J. Harrower Comer, 187-197. New York: Springer.

Chase, Arlen F., Diane Z. Chase, Christopher T. Fisher, Stephen J. Leisz, and John F. Weishampel. 2012. "Geospatial revolution and remote sensing LiDAR in Mesoamerican archaeology." *Proceedings of the National Academy of Sciences* 109 (32): 12916-12921.

Chase, Arlen F., Diane Z. Chase, Jaime Awe, John Weishampel, Gyles Iannone, Holley Moyes, Jason Yaeger, and M. Kathryn Brown. 2014. "The use of LiDAR in understanding the Ancient Maya landscape." *Advances in Archaeological Practice* 2 (3): 208-221.

Chase, Arlen F., Diane Z. Chase, Jaime J. Awe, John F. Weishampel, Gyles Iannone, Holley Moyes, and Jason Yaeger. 2014. "Ancient Maya Regional Settlement and Inter-Site Analysis: The 2013 West-Central Belize LiDAR Survey." *Remote Sensing* 6 (9): 8671-8695.

Chase, Arlen F., Diane Z. Chase, John F. Weishampel, Jason B. Drake, Ramesh L. Shrestha, K. Clint Slatton, Jaime J. Awe, and William E. Carter. 2011. "Airborne LiDAR, archaeology, and the ancient Maya landscape at Caracol, Belize." *Journal of Archaeological Science* 38 (2): 387-398.

Clark, R. N. 1999. "Chapter 1: Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy." In *Manual of Remote Sensing, Volume 3, Remote Sensing for the Earth Sciences*, by A. N. Rencz, 3-58. New York: John Wiley and Sons.

Clark, Roger N., Gregg A. Swayze, K. Eric Livo, Raymond F. Kokaly, Trude V.V. King, J. Brad Dalton, J. Sam Vance, Barnaby W. Rockwell, Todd Hoefen, and Robert R. McDougal. 2003. *Surface Reflectance Calibration of Terrestrial Imaging Spectroscopy Data: a Tutorial Using AVIRIS.* October 7. Accessed January 2016. https://speclab.cr.usgs.gov/PAPERS.calibration.tutorial/.

Coe, Michael D. 2011. *The Maya.* New York: Thames and Hudson.

Coronel, Eric G., Scott Hutson, Aline Magnoni, Chris Balzotti, Austin Ulmer, and Richard E. Terry. 2015. "Geochemical analysis of Late Classic and Post Classic Maya marketplace activities at the Plazas of Coba." *Journal of Field Archaeology* 40 (1): 89-109.

Curtis, Jason H., Mark Brenner, David A. Hodell, Richard A. Balser, Gerald A. Islebe, and Henry Hooghiemstra. 1998. "A multi-proxy study of Holocene environmental change in the Maya Lowlands of Peten, Guatemala." *Journal of Paleolimnology* 19: 139-15.

Dunning, Nicholas P., and Timothy Beach. 1994. "Soil erosion, slope management, and ancient terracing in the Maya lowlands." *Latin American Antiquity* 5 (1): 51-69.

Dunning, Nicholas P., J. G. Jones, Timothy Beach, and Sheryl Luzzadder-Beach. 2003. "Physiography, habitats, and landscapes of the Three Rivers Region." In *Heterarchy, political economy, and the ancient Maya: the Three Rivers region of the east-central Yucatan peninsula*, by V. L. Scarborough, F. Valdez Jr. and Nicholas P. Dunning, 14-24. Tucson: University of Arizona Press.

Dunning, Nicholas, Sheryl Luzzadder-Beach, Timothy Beach, John G. Jones, Vernon Scarborough, and Patrick T. Culbert. 2002. "Arising from the bajos: The evolution of a neotropical landscape and the rise of Maya Civilization." *Annals of the Association of American Geographers* 92 (2): 267-283.

Dunning, Nicholas, Timothy Beach, and David Rue. 1997. "The paleoecology and ancient settlement of the Petexbatun region, Guatemala." *Ancient Mesoamerica* 8 (2): 255-266.

Dunning, Nicholas, Timothy Beach, Pat Farrell, and Sheryl Luzzadder-Beach. 1998. "Prehispanic agrosystems and adaptive regions in the Maya Lowlands." *Culture and Agriculture* 20 (2-3): 87-101.

Eisazadeh, Amin, Khairul Anuar Kassim, and Hadi Nur. 2012. "Stabilization of tropical kaolin soil with phosphoric acid and lime." *Natural Hazards* 61 (3): 931-942.

Eismann, Michael T. 2012. *Hyperspectral Remote Sensing.* Bellingham, WA: SPIE.

ESRI. 2012. *Extract Values to Points (Spatial Analyst).* November 8. Accessed April 30, 2015. http://resources.arcgis.com/en/help/main/10.1/index.html#//009z0000002t000000.

—. 2016. *Extract Values to Points.* Accessed December 1, 2015. http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/extract-values-to-points.htm.

—. 2013. *Near (Analysis).* ESRI. May 6. Accessed November 16, 2014. http://resources.arcgis.com/en/help/main/10.1/index.html#//00080000001q000000.

Everts, L. H. 1874. "Johnston Township." *Historic Map Works.* Accessed August 2016. http://www.historicmapworks.com/Map/US/20140/Johnston+Township/Trumbull+County+1874/Ohio/.

Fedick, Scott L. 1996. "An Interpretive Kaleidoscope: Alternative Perspectives on Ancient Agricultural Landscapes of the Maya Lowlands." In *The Managed Mosaic: Ancient Maya Agriculture and Resource Use*, by Scott L. Fedick, 107-131. Salt Lake City: University of Utah Press.

Frahm, Ellery, Gilliane F. Monnier, Nicolas A. Jelinski, Edward P. Fleming, Brian L. Barber, and Justice B. Lambon. 2016. "Chemical soil surveys at the Bremer Site (Dakota county, Minnesota, USA): Measuring phosphorous content of sediment by portable XRF and ICP-OES." *Journal of Archaeological Science* 75: 115-138.

Garcia, Heide, and Peter Filzmoser. 2011. "Multivariate Statistical Analysis Using the R Package Chemometrics." *The Comprehensive R Archive Project.* November 7. Accessed 2013. http://cran.r-project.org/web/packages/chemometrics/vignettes/chemometrics-vignette.pdf.

Garrison, Thomas G., Bruce Chapman, Stephen Houston, Edwin Román, Jose Luis, and Garrido López. 2011. "Discovering ancient Maya settlements using airborne radar elevation data." *Journal of Archaeological Science* 38 (7): 1655-1662.

Garrison, Tom G. 2010. "Remote sensing ancient Maya rural populations using QuickBird satellite imagery." *International Journal of Remote Sensing* 31 (1): 213-231.

Garrity, C. P., and D. R. Soller. 2009. *Database of the Geologic Map of North America.* Accessed 2013. http://pubs.usgs.gov/ds/424/.

GISGeography. n.d. *GISGeography.* Accessed April 8 2015, 2015. http://gisgeography.com/landsat-1/.

Google. 2015. "Google Earth Basics." *Google Earth Blog.* Accessed April 8, 2015. http://www.gearthblog.com/basics.

—. 2015. *Understanding Google Earth Imagery.* Accessed April 8, 2015. https://support.google.com/earth/answer/176147?hl=en.

Green, Ernestene L. 1973. "Location analysis of prehistoric Maya sites in Northern British Honduras." *American Antiquity* 38 (3): 279-293.
Hanssen, Leonard M., and Keith A. Snail. 2002. "Integrating Spheres for Mid-and Near-Infrared Reflection Spectroscopy." In *Handbook of Vibrational Spectroscopy*, by John M. Chalmers and Peter R. Griffiths. Chichester: John Wiley & Sons Ltd.
Harris Corporation. 2016. *ENVI.* Accessed May 1, 2016. http://www.harrisgeospatial.com/ProductsandSolutions/GeospatialProducts/ENVI.aspx.
Harris Geospatial Solutions. 2016. *Atmospheric Correction.* Accessed December 2015. http://www.harrisgeospatial.com/docs/AtmosphericCorrection.html.
—. 2016. *Atmospheric Correction.* Accessed January 25, 2016. http://www.harrisgeospatial.com/docs/AtmosphericCorrection.html.
—. 2017. *Convolution and Morphology Filters.* Accessed 1 10, 2017. http://www.harrisgeospatial.com/docs/ConvolutionMorphologyFilters.html.
Harrison, Peter D., and B. L. Turner. 1978. *Pre-Hispanic Maya Agriculture.* Univeristy of New Mexico Press.
Hecker, Cristoph, John H. Dilles, Mark van der Meijde, and Freek D. van der Meer. 2012. "Thermal infrared spectroscopy and partial least squares regression to determine mineral modes of granitoid rocks." *Geochemistry, Geophysics, Geosystems* 13 (3).
Holiday, Vance T., and William G. Gartner. 2007. "Methods of soil P analysis in archaeology." *Journal of Archaeological Science* 34 (2): 301-333.
Howington, S. E., J. Ballard Jr., and S. Wilhelms. 2012. "Laboratory-Measured Rainfall Effects on LWIR Soil Reflectance." *AGU Fall Meeting Abstracts* 1182.
Idowu, Omololu J., Harold M. van Es, George S. Abawi, David W. Wolfe, Judith I. Ball, Beth K. Gugino, Bianca N. Moebius, Robert R. Schindelbeck, and Ali V. Bilgili. 2008. "Farmer-oriented assessment of soil quality using field, laboratory, and VNIR spectroscopy methods." *Plant Soil* 307: 243-253.
Instituto Nacional de Estadistica y Geografia. n.d. Aguascalientes. Accessed July 30, 2015. http://www.inegi.org.mx/default.aspx.
ISRIC. 2016. *Welcome to ISRIC - World Soil Information.* Accessed January 25, 2016. http://www.isric.org/.
ISRIC World Soil Information Service. 2016. Accessed 2016. http://www.isric.org/data/wosis.
Jain, Ramesh, Rangachar Kasturi, and Brian G Schunck. 1995. *Machine Vision.* New York: McGraw-Hill.
Jeffrey, Gillan. 2013. *The Landscape Toolbox.* February 11. Accessed April 3, 2015. http://wiki.landscapetoolbox.org/doku.php/remote_sensor_types:orbview-2.
Johnson, William C. 1983. "The Physical Setting: Northern Belize and Pulltrouser Swamp." In *Pulltrouser Swamp: ancient Maya habitat, agriculture, and settlement in northern Belize*, by Billie Lee Turner and Peter D. Harrison, 8-20. Austin: University of Texas Press.
Jordan, Clif. 2002. *Holocene Sediments of the Belize Shelf.* Accessed June 24, 2013. http://ambergriscaye.com/pages/mayan/holocene.html.

Kehnel, Emil . 2008. *Mexico Cenotes.* Accessed 2015. https://commons.wikimedia.org/wiki/File:Mexico_Cenotes.jpg.

Kennedy, David, and M. C. Bishop. 2011. "Google earth and the archaeology of Saudi Arabia. A case study from the Jedah area." *Journal of Archaeological Science* 38 (6): 1284-1293.

Kidder, Alfred V. 1930. "Five days over the Maya country." *The Scientific Monthly* 30 (3): 193-205.

Kresge Foundation. 2016. *Data Basin.* Accessed February 17, 2013. https://databasin.org/.

Latrou, M., A. Papadopoulos, F. Papadopoulos, O. Dichala, P. Psoma, and A. Bountla. 2014. "Determination of Soil Available Phosphorus using the Olsen and Mehlich 3 Methods for Greek Soils Having Variable Amounts of Calcium Carbonate." *Communications in Soil Science and Plant Analysis* 45 (16): 2207-2214.

Lavers, Christopher. 2015. *Directions Magazine.* Accessed April 7, 2015. http://www.directionsmag.com/entry/the-origins-of-high-resolution-civilian-satellite-imaging-part-2-civil/307714.

Lucero, Lisa J., Joel D. Gunn, and Vernon L. Scarborough. 2011. "Climate change and Classic Maya water management." *Water* 3 (2): 479-494.

Luzzadder-Beach, Sheryl, and Timothy Beach. 2009. "Arising from the wetlands: Mechanisms and chronology of landscape aggradation in the northern coastal plainof Belize." *Annals of the Association of American Geographers* 99 (1): 1-26.

Luzzadder-Beach, Sheryl, Timothy Beach, and Nicholas P. Dunning. 2012. "Wetland fields as mirrors of drought and the Maya abandonment." *Proceedings of the National Academy of Sciences* 109 (10): 3646-3651.

MacRae, Andrew. 1996. *The TalkOrigins Archive.* Accessed 2015. http://www.talkorigins.org/faqs/timescale.html.

Marschner, Ian. 2014. "glm2: Fitting Generalized Linear Models." *The Comprehensive R Archive Network.* May 7. Accessed January 25, 2016. https://cran.r-project.org/web/packages/glm2/glm2.pdf.

Matney, Timothy, Linda R. Barrett, Mahesh B. Dawadi, D. Maki, C. Maxton, David S. Perry, D. C. Roper, L. Somers, and L. G. Whitman. 2014. "In situ shallow subsurface reflectance spectroscopy of archaeological soils and features: a case-study of two Native American settlement sites in Kansas." *Journal of Archaeological Science* 43: 315-324.

McDaniel, Paul. 2013. *The Twelve Soil Orders.* Accessed April 4, 2015. http://www.cals.uidaho.edu/soilorders/vertisols_10.htm.

McGlinchey, Chris. 2012. "Handheld XRF for the examination of paintings: proper use and limitations." In *Handheld XRF for Art and Archaeology*, by Aaron N. Shugar and Jennifer L. Mass, 131-158. Leuven, Belgium: Leuven University Press.

Mellen, Mickey. 2014. *High resolution imagery coming to Google Earth.* June 14. Accessed April 8, 2015. http://www.gearthblog.com/blog/archives/2014/06/higher-resolution-imagery-coming-google-earth.html.

Merlín-Uribe, Yair, Carlos E. González-Esquivel, Armando Contreras-Hernández, Luis Zambrano, Patricia Moreno-Casasola, and Marta Astier. 2013. "Environmental and socio-economic sustainability of chinampas (raised beds) in Xochimilco, Mexico City." *International Journal of Agricultural Sustainability* 11 (3): 216-233.

Mevick, Bjørn-Helge, Ron Wehrens, and Kristian Hovde Liland. 2016. "pls: Partial Least Squares and Principal Component Regression." *The R Project for Statistical Computing.* Accessed January 10, 2017. https://cran.r-project.org/web/packages/pls/pls.pdf.

Mevik, Bjørn-Helge, and Henrick René Cederkvist. 2004. "Mean squared error of prediction (MSEP) estimates for principal component regression (PCR) and partial least squares regression (PLSR)." *Journal of Chemometrics* 18 (9): 422-429.

Miller, Thomas E. 1996. "Geologic and hydrologic controls on karst and cave development in Belize." *Journal of Cave and Karst Studies* 58 (2): 100-120.

Morehart, Christopher T. 2012. "Mapping ancient chinampa landscapes in the Basin of Mexico: a remote sensing and GIS approach." *Journal of Archaeological Science* 39 (7): 2541-2551.

National Aeronautics and Space Administration. 2012. "The Afternoon Constellation." *National Aeronautics and Space Administration.* June 5. Accessed April 12, 2015. http://atrain.nasa.gov/.

Neff, Hector, Barbara Voorhies, and Frederico Paredes Umaña. 2012. "Handheld XRF elemental analysis of archaeological sediments: some examples from Mesoamerica." In *Handheld XRF for Art and Archaeology*, by Aaron N. Shugar and Jennifer L. Mass, 379-400. Leuven, Belgium: Leuven University Press.

Notesco, Gila, Yaron Ogen, and Eyal Ben-Dor. 2015. "Mineral classification of Makhtesh Ramon in Israel using hyperspectral longwave infrared (LWIR) remote-sensing data." *Remote Sensing* 7 (9): 12282-12296.

Ohio Department of Agriculture. n.d. *About Ohio's Soils.* Accessed November 2016. http://www.agri.ohio.gov/divs/SWC/SWC.aspx.

Palacios-Mayorga, Sergio, Ana Luisa Anaya, Eleazar González-Velázquez, Lázaro Huerta-Arcos, and Arturo Gómez-Pompa. 2003. "Periphyton as a potential biofertilizer in intensive agriculture of the ancient Maya." In *The lowland Maya area: Three millennia at the human-wildland interface*, by M. F. Allen, S. L. Fedick and J. J. Jiménez-Osornio, 389-400. New York: Food Products Press.

Parnell, Jacob J., Richard E. Terry, and Payson Sheets. 2002. "Soil chemical analysis of ancient activities in Ceren, El Salvador: a case study of a rapidly abandoned site." *Latin American Antiquity* 331-342.

Perry, E., G. Velazquez-Oliman, and R. A. Socki. 2003. "Hydrogeology of the Yucatan Peninsula." In *The Lowland Maya Area: Three Millenia at the Human-Wildland Interface*, by A. Gomez-Pompa, M. F. Allen, S. L. Fedick and J. J. Jimenez-Osornio, 115-138. New York: Food Products Press.

Pierzynski, G. M., T. J. Logan, S. J. Traina, and J. M. Bigham. 1990. "Phosphorus chemistry and minerology in excessively fertilized soils: Quantitative analysis of

phosphorus-rich particles." *Soil Science Society of America Journal* 54 (6): 1576-1583.

Piperno, R. Dolores, and M. Deborah Pearsall. 1998. *Origins of agriculture in the lowland neotropics.* San Diego: Emerald Group Publishing.

Pohl, Mary D., Kevin O. Pope, John G. Jones, John S. Jacob, Dolores R. Piperno, Susan D. deFrance, David L. Lentz, John A. Gifford, Marie E. Danforth, and J. Kathryn Josserand. 1996. "Early Agriculture in the Maya Lowlands." *Latin American Antiquity* 7 (4): 355-372.

Pohl, Mary. 1990. "The Rio Hondo Project of Northern Belize." In *Ancient Maya Wetland Agriculture: Excavations on Albion Island, Northern Belize*, edited by Mary Pohl. Boulder, CO: Westview Press.

Pope, Kevin O., and Bruce H. Dahlin. 1989. "Ancient Maya wetland agriculture: New insights from ecological and remote sensing research." *Journal of Field Archaeology* 16 (1): 87-106.

Puleston, Dennis E. 1978. "Terracing, raised fields, and tree cropping in the Maya lowlands: A new perspective on the geography of power." In *Pre-Hispanic Maya Agriculture*, by P. D. Harrison and B. L. Turner, 225-245. Austin, TX: University of New Mexico Press.

R Development Core Team. 2016. *The R Project for Statistical Computing.* Accessed July 11, 2015. http://www.R-project.org.

Rapp, George Robert Rip, and Christopher L. Hill. 2006. *Geoarchaeology: the earth-science approach to archaeological interpretation.* Yale University Press.

Renard, D., J. Iriarte, J. J. Birk, S. Rostain, B. Glaser, and D. McKey. 2012. "Ecological engineers ahead of their time: The functioning of pre-Colombian raised-field agriculture and its potential contributions to sustainability today." *Ecological Engineering* 45: 30-44.

Ricketson, Oliver, and Alfred V. Kidder. 1930. "An archeological reconnaissance by air in Central America." *Geographical Review* 177-206.

Riley, D., and J. Janaskie. 2012. "Initial comparison of mineral mapping simultaneously collected hyperspectral short-wave infrared (SWIR) and long-wave infrared (LWIR) data over Cuprite, Nevada." *GRSG AGM 2012 - Monitoring and Managing the Earth's Resources.* London: EarthDoc.

Rozenstein, Offer, and Arnon Karnieli. 2015. "Identification and Characterization of Biological Soil Crusts in a Sand Dune Desert Environment Across Israel-Egypt Border Using LWIR Emittance Spectroscopy." *Journal of Arid Environments* 112: 75-86.

Rusli, Noradila, M. Rafee Majid, and Ami Hassan Md Din. 2014. "Google Earth's derived digital elevation model: A comparative assessment with Aster and SRTM data." *IOP Conference Series: Earth and Environmental Science.* IOP Publishing. 012065.

Sadr, Karim, and Xavier Rodier. 2012. "Google Earth, GIS and stone-walled structures in southern Gauteng, South Africa." *Journal of Archaeological Science* 39 (4): 1034-1042.

Sanchez, Gaston. n.d. "Gaston Sanchez." Accessed June 23, 2013. http://www.gastonsanchez.com/teaching/.

—. 2012. "plsdepot: Partial Least Squares (PLS) Data Analysis Methods." *The R Project for Statistical Computing.* November 13. Accessed July 31, 2015. https://cran.r-project.org/web/packages/plsdepot/.

Sanders, William T. 1979. "The Jolly Green Giant in tenth century Yucatan, or fact and fancy in Classic Maya agriculture." *Reviews in Anthropology* 6 (4): 493-506.

Sandholt, Inge, Kjeld Rasmussen, and Jens Andersen. 2002. "A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status." *Remote Sensing of Environment* 213-224.

Satellite Imaging Corporation. 2014b. *GeoEye-1 Satellite Sensor.* Accessed April 3, 2015. http://www.satimagingcorp.com/satellite-sensors/geoeye-1/.

—. 2014a. *SPOT-5 Satellite Sensor.* Accessed April 6, 2015. http://www.satimagingcorp.com/satellite-sensors/other-satellite-sensors/spot-5/.

—. 2014c. *WorldView-3 Satellite Sensor.* Accessed April 3, 2015. http://www.satimagingcorp.com/satellite-sensors/worldview-3/.

Schaetzl, Randall J., and Sharon Anderson. 2005. *Soils: Genesis and geomorphology.* Cambridge University Press.

Selva Maya Consortium. 2016. *Data Basin.* Accessed 14 2014, April. https://databasin.org/datasets/21a4f58393904edcbdb1ae031a4c6b68.

Sever, Thomas L., and Daniel E. Irwin. 2003. "Landscape archaeology: Remote-sensing investigation of the ancient Maya in the Peten rainforest of northern Guatemala." *Ancient Mesoamerican* 14 (1): 113-122.

Shackley, Steven M. 2011. "An introduction to X-ray fluorescence (XRF) analysis in archaeology." In *X-ray fluorescence spectrometry (XRF) in geoarchaeology*, 7-44. New York: Springer.

Siemens, Alfred H., and Dennis E. Puleston. 1972. "Ridged Fields and Associated Features in Southern Campeche: New Perspectives on the Lowland Maya." *American Antiquity* 37 (2): 228-239.

Sluyter, Andrew. 1994. "Intensive wetland agriculture in Mesoamerica: Space, time, and form." *Annals of the Association of American Geographers* 84 (4): 557-584.

2014. *Spatial Data Warehouse.* Accessed April 14, 2013. http://www.biodiversity.bz/mapping/warehouse/.

Speakman, Robert J., and M. Steven Shackley. 2013. "Silo science and portable XRF in archaeology: a response to Frahm." *Journal of Archaeological Science* 40 (2): 1435-1443.

Stein, Julie K. 1990. "Sedimentation and Maya agriculture along Rio Hondo, Belize." In *Ancient Maya Wetland Agriculture: Excavations on Albion Island, Northern Belize*, edited by Mary D. Pohl, 323-337. Boulder, CO: Westview Press.

Terry, Richard E., Fabian G. Fernández, J Jacob Parnell, and Takeshi Inomata. 2004. "The story in the floors: chemical signatures of ancient and modern Maya activities at Aguateca, Guatemala." *Journal of Archaeological Science* 1237-1250.

Terry, Richard, Sheldon D. Nelson, Jared Carr, Jacob Parnell, Perry J. Hardin, Mark W. Jackson, and Stephen D. Houston. 2000. "Quantitative phosphorus measurement: a field test procedure for archaeological site analysis at Piedras Negras, Guatemala." *Geoarchaeology* 15 (2): 151-166.

Thakuria, Tilok, Tosabanta Padhan, Rabindra Kumar Mohanty, and Monica L. Smith. 2013. "Google Earth As An Archaeological Tool In The Developing World: An Example From India." *The SAA Archaeological Record* 13 (1): 20-24.

Thermo Fisher Scientific. 2016. *Spectroscopy, Elemental & Isotope Analysis Resource Library.* Accessed November 2016. https://www.thermofisher.com/us/en/home/industrial/spectroscopy-elemental-isotope-analysis/spectroscopy-elemental-isotope-analysis-learning-center/spectroscopy-elemental-isotope-analysis-resource-library.html.

Ulmer, Austin M. 2015. *Indications of Ancient Maya Soil Resource Management in Northern Belize.* Master's Thesis, Provo, UT: Brigham Young University.

US Geological Survey. 1907. "Historical Topographic Map Explorer." *USGS.* Accessed 2016. http://historicalmaps.arcgis.com/usgs/.

USGS. 2014. *Land Processes Distributed Active Archive Center.* April 14. Accessed November 12, 2013. https://lpdaac.usgs.gov/data_access/data_pool.

USGS Land Processes Distributed Active Archive Center. 2014. *Data Pool.* April 14. Accessed December 1, 2015. https://lpdaac.usgs.gov/data_access/data_pool.

USGS. 2015. *Using the USGS Landsat 8 Product.* August 27. Accessed December 1, 2015. http://landsat.usgs.gov/Landsat8_Using_Product.php.

—. 2015. *Using the USGS Landsat 8 Product.* August 27. Accessed December 15, 2015. http://landsat.usgs.gov/Landsat8_Using_Product.php.

Vaughn, Sallie, and Tom Crawford. 2009. "A predictive model of archaeological potential: an example from northwestern Belize." *Applied Geography* 24 (9): 542-555.

Wehrens, Ron. 2011. *Chemometrics with R: multivariate data analysis in the natural sciences and life sciences.* Berlin: Springer.

Williams, Norris L. 1992. *Soil Survey of Trumbull County, Ohio.* USDA Soil Conservation Service.

Witschey, Walter R. T., and Clifford T. Brown. 2010. *The Electronic Atlas.* January 31. Accessed July 30, 2015. http://mayagis.smv.org/index.htm.

Wolf, A. M., and D. B. Beegle. 1995. "Recommended soil tests for macronutrients: Phosphorus, potassium, calcium, and magnesium." In *Recommended soil testing procedures for the northeastern United States*, 25-34. Northeast Regional Bull.

Xue-Yu, Hu. 2013. "Application of Visible/Near-Infrared Spectra in Modeling of Soil Total Phosphorus." *Pedosphere* 23 (4): 417–421.

Young, Jaime. 2009. "LiDAR Trends and Developments." *WiLDER LiDAR Blog.* March 6. Accessed April 7, 2015. https://bloglidar.files.wordpress.com/2011/02/2009_whitepaper_lidartrends.pdf.

Zhang, Chunlong. 2007. *Fundamentals of environmental sampling and analysis.* John Wiley & Sons.

Zumel, Nina, and John Mount. 2014. *Practical Data Science with R.* Shelter Island, NY: Manning.

BIOGRAPHY

Tanya Catignani graduated from Maplewood High School, Cortland, Ohio in 1998. She received her Bachelor of Science from Mercyhurst College in 2002 and her Master of Arts from the State University of New York at Buffalo in 2004. She has worked for the federal government as a geographer since 2002.