5-2015

Accuracy of Supervised Classification of Cropland in sub−Saharan Africa

Sarah Lynn Lewis-Gonzales
University of Tennessee - Knoxville, slewisgo@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Sarah Lynn Lewis-Gonzales entitled "Accuracy of Supervised Classification of Cropland in sub–Saharan Africa." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Geography.

Nicholas N. Nagle, Major Professor

We have read this thesis and recommend its acceptance:

Henri D. Grissino-Mayer, Liem T. Tran

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Accuracy of Supervised Classification of Cropland in sub–Saharan Africa

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Sarah Lynn Lewis-Gonzales
May 2015

Acknowledgements

I am extremely thankful to all of the individuals who have helped me throughout my thesis process. First and foremost, I have to thank my advisor, Dr. Nicholas Nagle, for being extremely understanding, compassionate, and a fierce mentor who continually fought on my side. In addition to being an inspiring and encouraging advisor, he also taught me more than I ever thought possible about statistical methods and analysis, and efficient scientific writing. I must also thank my committee members Dr. Liem Tran and Dr. Henri Grissino-Mayer not only for their thoughtful comments and questions, but also for the confidence and clarity that those comments and questions brought to my thesis.

The training data for this research were graciously provided by Dr. Kathryn Grace of the University of Utah, Department of Geography and the Famine Early Warning Systems Network (FEWS Net). I am also very grateful to the University of Tennessee Geography Department for their support and consideration in both good times and bad, especially Dr. Derek Alderman who had faith in me and never turned me away. To my fellow graduate students, I owe special thanks, for if it weren't for their camaraderie, humorous wit, and support I would have lost my wit and more along the way.

Finally, and most importantly, I wish to thank my husband Martin. Over the course of my education, he has shown me more patience, encouragement, and love than anyone could possibly deserve. Thank you Martin, for not only keeping our house in order, but also for holding on to my sanity for me.

Abstract

Mali is a country in sub–Saharan Africa where monitoring of cropped land area would greatly benefit food security initiatives and aid organizations. More importantly village-scale studies on cropped land are fundamental to making a difference in the way we look at cropped land area and food availability in this region of the world. Using Landsat surface reflectance imagery and World View-2 derived labeled data, this study focuses on accuracy of supervised classification methods while addressing various levels of scale. Several classification methods are taken into account to determine the best method possible to produce cropped area estimates using these data. The relationship between classification and scale was addressed by taking into account how distance and proximity affect accuracy. Accuracy is measured by kappa coefficients, and results among the different methods vary. Kappa coefficients generated are very low and results suggest that estimates between labels are more accurate than estimates far from labels.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

Mali is an arid to semi-arid country in sub-Saharan Africa that suffers from extreme

environmental, political, and economic stress and experiences severe food security hazards. In

this and other sub-Saharan countries, the identification and monitoring of cropland is crucial to

the international efforts aimed at monitoring food availability and preparing for food shortages.

Furthermore, land cover classification is a valuable resource for assessing the extent of

environmental degradation and its effect on agricultural productivity. The use of satellite-based

land cover classifications is a safe and low-cost method for assessing changes in agricultural

development and degradation. The accurate and rapid evaluation of agriculturally productive

land is vital to the improvement of land use practices and food security in developing countries

(Barrett, 2010; Grace et al., 2012).

1.1 Research Question and Objectives

Previous land cover classification studies in sub-Saharan Africa have been performed using

coarse resolution MODIS or AVHRR imagery (Lobel, 2004; Wardlow, 2007; Wessels, 2004; Xiao,

2003). However, the concern is that the granularity of the agriculture in this region of the world

is too fine to be detected by coarse resolution imagery. The goal of this research is to determine

whether or not accurate cropped land classifications can be established using WorldView-2

(WV2) and Landsat data. Given the fine granularity of subsistence cropland in sub-Saharan

Africa, I focus this thesis on village-scale cropland monitoring while also asking: Can we

produce an accurate classification using 30-meter Landsat surface reflectance data that are

trained using labeled data from WV2 data? How will classification accuracy vary with scale? How far from labels can we accurately classify data? Can we be confident about classifications far from labels, and if so, how far? Finally, can we be confident in classifications between labels?

1.2 Background and Historical Methods

While remote sensing tools for land cover classification are important, the historical standard for estimating cropped area at large scales is not based on remote sensing data techniques, but on the statistical technique of area-frame studies. In area-frame sampling, an area is equally divided into a grid of squares and (agricultural) information is collected from each using certain field survey techniques (Tsiligirides, 1998). Many agricultural assessments have previously been conducted using area-frame sampling (Fecso, 1985; Pradhan, 2001; Tsiligirides, 1998). These field-based survey methods can take up a great deal of time and resources. Area-frame sampling is considered most effective when considering a 15–20 year time frame, which adds additional expense to already pricey techniques (Pradhan, 2001; Tsiligirides, 1998). This makes area-frame studies less appropriate for producing timely and low cost crop area estimates.

1.3 Physical Environment in Mali

One of the chief concerns in this study is that classification accuracy will depend not only on the distance from the training data, but also on the physical environment, and particularly on the geographic scale of agriculture. Mali represents an interesting case study in this regard because it contains a strong physical gradient. Herrmann et al. (2005) described the environment in northern Mali (approximately 16° N to 24° N) as part of the Saharan desert,

with southern Mali (approximately 10° N to 16° N) ranging from the arid to semiarid zone below the Sahara desert, to the humid-subtropical savannas in the south. Geographical regions based on rainfall time series data include: Sahelo-Sahara (approximately 18° N to 25° N), Sahel (approximately 13–15° N to 18° N), and Soudan (approximately 10° N to 15° N) (Nicholson, 2001). The 2005 Vulnerability Profile of Mali (Simonsson, 2005) summarized that precipitation in Mali varies greatly and considers the existence of four "eco-climate" regions: Sahara, Sahelian, Sudanian, and Sudanian–Guinean, with rainfall throughout these regions averaging from 100 to 1,700 mm/year. To further illustrate the great range in climatic variability in Mali, climate profiles generated for each province by The World Bank characterize the northern-most province of Timbuktu as having maximum temperatures in June of approximately 44 °C, minimum temperatures in January of approximately 11 °C, and maximum rainfall averages reaching 27 mm in August. Conversely, the southern-most province of Mali, Sikasso, has maximum temperatures in March and April of approximately 37 °C, minimum temperatures in December and January of approximately 15 °C, and maximum rainfall averages reaching 282 mm in August (WorldBank 2013).

The size and climatic variability of Mali allow it to have very distinct vegetative and productive zones that are attributed to specific livelihood regions in Mali. Productive regions are influenced by both human–induced and environmental factors. The severe level of desertification of the Sahel region of Mali since at least the 1980s (Nicholson, 2001) is an example of these contributing factors. Following a prominent north-south aridity gradient pattern, livelihood regions in Mali range from nomadic pastoralism in the far northern region just south of the Sahara desert, to herding, to increasingly rain dependent crops as you travel

3

further south. The majority of the population in Mali depends heavily on agro-pastoral resources, however only 25% of Mali is capable of supporting such practices (Simonsson, 2005).

1.4 Brief Review of Classification Methods

Classification is a method of extracting meaningful information from data, with land cover/land use classification of multispectral imagery being one of the most common applications (Jensen, 2004). By combining the pixel values from multispectral imagery with *a priori* knowledge about how measured pixel characteristics correspond to land cover types, each individual pixel of a given study area can be categorized into a specific land cover class (Campbell and Wynne, 2011). Methods that use this *a priori* knowledge are called "supervised classification," as distinct from "unsupervised classification methods," which attempt to make classifications without *a priori* knowledge. Typically, unclassified methods operate by identifying "natural clusters" of data that share similar multispectral values and are therefore easily divided into classes with little to no additional knowledge (Jones and Vaughan, 2010).

Although supervised classification traditionally requires *a priori* knowledge of the region being classified, at a near countrywide scale, knowledge of the entire region to be classified is an impractical requirement. A remedy for the problem of large-scale supervised classification may include the incorporation of training data such as those used in this analysis. Such training data are used in to create this *a priori* knowledge about the relationship between pixel value and land cover type.

1.5 Resolution

Despite the high historical accuracy of area-frame studies for agricultural assessment, remotely sensed data and their classification are regularly used to reduce the costs of land cover estimation (Fecso et al., 1985; Hammond, 1975). These studies have used various resolutions of remotely sensed data for cropped area estimation and monitoring, including those using low resolution (greater than 100 m) imagery such as AVHRR (NOAA) and MODIS (ASTER or TERRA), medium resolution (between 5 m and 100 m) imagery such as Landsat MSS, TM, and ETM+ and SPOT 1, 2, 3 and 4 (Frolking et al., 1999; Gonzalez-Alonso et al., 1997; MacDonald et al., 1975; Pax-Lenney and Woodcock, 1997; Richards and Jia, 1999) and high resolution (less than 5 m) imagery such as WorldView-2, IKONOS, and SPOT 5, 6 and 7 (Grace et al., 2012; Husak et al., 2008; Launay and Guerif, 2005; Ozdogan and Woodcock, 2006; Pradhan, 2001). Comparisons of remote sensing to ground-based methods note that the remote sensing-based crop estimate techniques tend to be both less expensive and time consuming (Frolking et al., 1999; Pradhan, 2001; Tsiligirides, 1998). Some of these studies also note that the estimates resulting from these combinations of ground-based area-frame techniques and remote sensing techniques give higher degrees of accuracy than either as stand alone methods (Gonzalez-Alonso et al., 1997; Tsiligirides, 1998). Studies that incorporate the use of low to medium resolution in their assessments of agriculture using Landsat MSS (MacDonald et al., 1975), TM (Gonzalez-Alonso et al., 1997; Pax-Lenney et al., 1996; Shalaby and Tateishi, 2007; Tsiligirides, 1998), and ETM+ (Lobel and Asnew, 2004; Lobell et al., 2003; Marshall et al., 2011; Shalaby and Tateishi, 2007) are common. Some even attempted to determine crop distribution at a finer local or regional scale (Lobell, 2003 et al.; Moulin et al., 1998; Ozdogan and Asnew, 2006). However,

5

many more studies apply coarse resolution MODIS (Lobel and Asnew, 2004; Lobell et al., 2003; Wardlow et al., 2007; Yuping et al., 2008) or AVHRR (Frolking et al., 1999; Tottrup and Rasmussen, 2004; Xiao et al., 2003) to conduct cropped land classification at the country-wide scale. Lobell and Asner (2004) and Launay and Guerif (2005), among others, even combined different resolutions in their studies, with Lobell and Asner using an unmixing technique with Landsat ETM+ and MODIS, and Launay and Guerif using a combination of SPOT and aerial imagery.

1.6 Accuracy and Uncertainty

Classification methods for crop estimates come in many different forms and there are degrees of uncertainty exist for each. Some studies compared these different methods, or at least specific aspects of different models, in an effort to understand the uncertainties associated with their crop estimates (Lobell et al., 2003; Ozdogan and Woodcock, 2006; Tao et al., 2005; Xiao et al., 2003). Many more studies added satellite-borne data to existing crop models to increase their accuracy and decrease uncertainty (Frolking et al., 1999; Launay and Guerif, 2005; Moulin et al., 1998; Pradhan, 2001; Yuping et al., 2008). Many methods are used to apply imagery resources to agricultural studies. The Maximum Likelihood classification technique stands out, being mentioned frequently as both most accurate and most commonly used (Chen et al., 2008; Shalaby and Tateishi, 2007; Wessels et al., 2004). To increase accuracy and decrease uncertainty, many studies used various multi-temporal techniques, or time-series techniques, to produce crop estimates (Gonzalez-Alonso et al., 1997; Lobel and Asnew, 2004; Pax-Lenney et al., 1996; Shalaby and Tateishi, 2007).

Of central importance to this study is the relationship between the geographic scale (i.e. resolution) of the data and the geographic scale of cropped and uncropped lands. Land cover classification is particularly difficult in developing countries because the geographic scale of the subsistence agriculture is finer than the geographic scale of a MODIS pixel (Ozdogan and Woodcock, 2006; Pax-Lenney et al., 1997). In study regions where subsistence agriculture is the norm, some substantial errors of inclusion and exclusion in the classification process can arise. Omission and commission errors lead to over- or under-estimation of cropped area (Marshall et al., 2011; Ozdogan and Woodcock, 2006). This problem occurs because subsistence agriculture often exhibits a checkerboard pattern with cropped fields interspersed with fallow lands (Ozdogan and Woodcock, 2006). As such, agricultural land delineation can be difficult with low–resolution imagery such as 250- or 500-meter MODIS imagery, from which most national scale land cover classifications are conducted (Lobel and Asner, 2004). This study will evaluate the use of 30–meter Landsat TM imagery combined with relatively cheap training point data derived from WV2 imagery in place of more expensive ground-based truth measurement. While WV2 data have very high resolution (2 m for multispectral, and 0.5 m for panchromatic), enabling a more precise detection of subsistence croplands, they introduce a new problem, owing to the relatively sparse acquisition of images. Usually, imagery is to sparse to cover an entire country within a growing season. This thesis will explore the accuracy of cropped area estimates that combine WV2 images with Landsat imagery, and in particular, the accuracy of producing estimates far from the WV2 images.

Chapter 2
Literature Review

Many different methods of classification have been developed over the years from multiple disciplines. Scientific ideologies, such as those of Charles Booth in the early 20th century, and that of Robert Park and Ernest Burgess in the 1920s, initiated and advocated for the use of clustering in scientific research (Troy, 2008). Cluster analysis is a method of grouping variables of interest into classes based on similarity (Gan et al., 2007). In cases where assumptions are ill-advised and little background or preliminary knowledge have been obtained, clustering is an especially appropriate method for determining and explaining relationships between variables (Jain et al., 1999). Cluster analysis is a vital element to "data mining," which is a method for gathering information from large data sets (Gan et al., 2007). Trends are emerging for using clustering in image analysis, including information retrieval, processing, and pattern recognition. Clustering techniques are used by many different disciplines and research communities including geography, geology, biology, psychiatry, psychology, archaeology, and even marketing (Jain et al., 1999).

Jain et al. (1999) presented an array of different methods for cluster analysis based on hierarchical and partitional techniques. Techniques related to algorithmic structure and operation (agglomerative and divisive), those related to the sequential or simultaneous use of features in the clustering processes (polythetic and monothetic), and methods such as hard versus fuzzy clustering, deterministic, stochastic, incremental and non–incremental make–up the complexity of Jain et al.'s "taxonomy".

2.1 Classification Methods

Classification methods are often used to derive information from satellite imagery. The information obtained from a satellite can come in the form of spectral bands, the combination of which creates a digital image over an area. Each spectral band is made up of thousands of individual pixels which hold numerical information about that pixel and the land type it represents. A single digital image pixel can be associated with multiple bands. By comparing cross-band pixels, especially when associated with known pixels, one can classify all pixels within a digital image into specific regions or groups known as "classes" (Campbell and Wynne, 2011). A class is a compilation of patterns whose spatial distribution is determined by probability density (Jain et al., 1999). The division and reassignment of heterogeneous land cover types into homogenous classes is known as classification.

Detection, recognition, and identification are three of the most fundamental processes in classification of imagery. Detection refers to decision-making processes such as determining the "presence or absence" of landscape features. Recognition concerns holding enough information about detected features to allow broad categorization. Specificity of classes occurs once confidence is high enough to confirm identification (Campbell and Wynne, 2011). Though classification is often used for the delineation of ground cover into class types, the value of using these classes to determine land cover areas (i.e. determining quantities for hectares of agricultural land cover) is questionable unless conducted through computer interpretation of a digital image in which pixel level analyses can be carried out (Richards and Jia, 1999). Unsupervised, or unlabeled, classification of variables into meaningful homogenous groups is a method of cluster analysis (Gan et al., 2007; Jain et al., 1999). One of the primary differences

between unsupervised and supervised, or labeled vs. unlabeled, analyses is that supervised

analysis is based on categorical labels derived specifically from the data. (Campbell and Wynne,

2011; Jain et al., 1999).

2.2 Clustering

Carmichael et al. (1968) demonstrated the importance of "finding natural clusters" when

performing cluster analyses. Recognition of natural clusters expedites understanding the

relationships between cluster members in that information can be extrapolated from these

"mutually exclusive" subdivisions of the data to make generalizations and predictions based on

their un-relatedness to non-members (Carmichael et al., 1968). Implementation of clustering

follows the condition that data points must exist in a continuous region of space with a

relatively dense population compared to the surrounding continuous regions of space

(Carmichael et al., 1968).

2.3 Unsupervised Classification

Unsupervised classification is related to the idea of "natural groups" or "natural

clusters" within an image being grouped together based strictly on spectral similarities within

the data (Campbell and Wynne, 2011). Minimal analyst input is required apart from inputting

the number of classes or the minimum and maximum to be assigned to the data. The

fundamentals of the process itself involve the separation of values into the specified number of

classes, or clusters, based on some measure of distance, such as Euclidean distance (Jones and

Vaughan, 2010). Due to the limited classification-analyst interaction little opportunity exists for

analyst bias to increase error in the classification prior to interpretation of the results (Campbell

and Wynne, 2011). Some examples of unsupervised methods of classification include k–means and maximum likelihood classification (Jones and Vaughan, 2010). One of the primary limitations to unsupervised classification is that the classification is based on natural relationships between spectral values. Often a significant difference will be found between the ideal classes of interest to the interpreter and the classes actually created by the classification. The analyst has less control over the resultant classes in unsupervised classification than in the case of supervised classification, as there are no training fields or informational classes involved (Campbell and Wynne, 2011).

Classification of pixels based on spectral values only becomes significant when correlated with informational classes. Informational classes are determined by the analyst, or whoever the classification is being produced for. When a match between spectral and information classes arises, it is common for them to be assigned based on recognizable patterns in an image. Accurate assignment of classes can be inhibited by the occurrence of mixed pixels, in which case identification and knowledge from the analyst is heavily relied upon (Campbell and Wynne, 2011). An example of a simple method for the allocation of informational classes in a multispectral image is by plotting spectral image bands against each other (i.e. x-axis being band 4, y-axis being band 7, a simple example) using brightness values of pixels. This demarcates landcover types based upon textural, moisture, soil, and topographic differences (Richards and Jia, 1999). The resultant plot can be referred to as a multispectral space or pattern space. These plots aid in the discrimination between landcover types, pattern recognition, and in the assignment of informational classes (Richards and Jia, 1999). Groups of pixels that have been assigned to informational classes can then be used as training fields. Training fields that

clearly represent the spectral values of the region and are suitably distributed without can be

applied to analysis such as supervised classification. Mixed pixels would not be ideally

included in a training field (Campbell and Wynne, 2011).

2.4 Supervised Classification

Supervised classification is the classification of pixels using training data (Campbell and

Wynne, 2011). The availability of accurate training fields is essential to supervised classification.

Supervised classification is especially suitable in such cases where a specific product is expected

from the analysis, or when conducting comparisons of landscapes that have different temporal

or spatial identities (bi-image comparison, time series, nearest neighbor, etc.) (Campbell and

Wynne, 2011). Limitations and ineffectiveness primarily arise in cases where training fields are

ineptly generated, usually through the inclusion of mixed pixels or shadows (Campbell and

Wynne, 2011). A primary aspect of supervised classification is that the analyst has sufficient

knowledge of the area in which the study is being conducted. The added control in the hands of

the analyst adds some degree of detail or precision to the assessment, but can also introduce

unwanted bias due to human subjectivity and opinion (Guo and Mason, 2009).

ERDAS IMAGINE® is a versatile and powerful software package for creating

information from geospatial data (Intergraph Corporation ©, 2013). IMAGINE's supervised

classification function has a maximum likelihood option that incorporates a Bayesian classifier

that takes added variability within a single class into account:

$$D = \ln(a_c) - [0.5 \ln(|Cov_c|)] - [0.5 \, (X{-}M_c)T \, (Cov_c{}^{-1}) \, (X{-}M_c)]$$

D = weighted distance (likelihood)
C = a particular class

X	= measurement vector of the candidate pixel
M_c	= mean vector of the sample of class c
a_c	= percent probability that any candidate pixel is a member of class c (defaults to 1.0, or is entered from a priori knowledge)
Cov_c	= covariance matrix of the pixels in the sample of class c
$|Cov_c|$	= determinant of Covc (matrix algebra)
Cov_c^{-1}	= inverse of Covc (matrix algebra)
ln	= natural logarithm function
T	= transposition function (matrix algebra)

The ERDAS IMAGINE® supervised classification function does not allow the inclusion of point data into the process. Training samples are traditionally acquired by creating a polygon over a specific area of the image that represents a class type. ERDAS IMAGINE® refers to one of these delineated class areas as an AOI, "area of interest" (Intergraph Corporation ©, 2013). As such, I decided that a polygon file representing the data affiliated with each training point should be created. The prime method for creating the AOI training samples from point data was to create a buffer region around each training point.

Guo and Mason (2009) discussed a hybrid classification approach, one in which the initial step was to conduct an unsupervised classification where classes were assigned based on supplementary data, such as aerial imagery or ground truth data. Statistical results of the initial classification were then used as training data, and the image was reclassified using a supervised method. In terms of accuracy, the subjectivity of the analyst must be considered with ground truth and aerial data. Human observation is variable, and its impracticable to expect to measure and extract the exact truth from the field or aerial data when opinions vary with each human observer (Adams and Gillespie, 2006).

Classification and regression tree analysis (CART) is another method applied to the classification of digital imagery. Defined training classes or fields, such as those required for supervised classification, are also necessary for CART. However, ancillary data are the primary source for the parameters required for CART, so less prior knowledge of the study site and variables is needed for the analysis. Ancillary data aid in the delineation of relevant data to put to use to the successful classification of the image. Problems associated with the CART method include the fact that pixels are frequently mislabeled due the requirement that they can only be appointed to a single discrete class. This can be remedied by using a fuzzy clustering method. Where other classification methods strictly require the assignment of pixels into specific classes like forest or water, fuzzy clustering allows a single pixel to be classified into more than one class based on its unsuitability for any single class. Fuzzy clustering is also an obvious solution that can be applied to the problem of mixed pixels. Fuzzy clustering is based on the idea of membership functions. Membership functions deal with relationships between data and classes. In regards to remote sensing, brightness values and spectral characteristics are the relationship variables for a membership function. For example, fuzzy classification takes a pixel and gives it a membership grade for each class between 0 and 1 (nonmembership and membership) instead of simply classifying a pixel as forest or water. This method specifies "partial membership" by assigning a ratio to the pixel (water = 0.3, forest = 0.7) where the total value of the pixel is equal to 1 (Campbell and Wynne, 2011).

Classification can also be applied to studies with interests in time series and/or change detection. Many different techniques can be applied in image comparison. Comparisons across multiple bands are possible but can add a significant level of difficulty to analysis of the change.

14

Easier interpretation can be achieved by using images that have undergone some method of

band combination, such as in vegetation indices (Campbell and Wynne, 2011).

Chapter 3
Accuracy of Supervised Classification of Cropland in sub-Saharan Africa

3.1 Introduction

Satellite-based land cover classification is a safe and low-cost method to assess changes in agricultural development and degradation. The accurate and rapid evaluation of agriculturally productive land is vital to the improvement of land-use practices and food security in developing countries (Barrett, 2010; Grace et al., 2012). Digital imagery provides information about the earth's surface, and classification is an ideal way to extract that information and turn it into a coherent and substantial product. Desired knowledge of urban areas, forests, waterways, and agriculture are just a few of the reasons one might want to extract information from satellite imagery. Over both large geographical areas and small areas at very high resolutions, the knowledge desired is often land cover types within the given area. Land cover types can then be classified into land cover classes.

In regions of the world where food insecurity is a major concern, classification can be used in very beneficial ways, especially in parts of the world that aren't easily accessible. Whether it is because of climate, terrain, or a precarious political environment, classification methods can be used to classify and analyze the environment. This is especially important in analyses involving crop production and food insecurity, as these things can be affected most in regions with stressed environments. REWORD

3.1.1 Research Question and Objectives

Previous land cover classification studies in sub-Saharan Africa have been performed using coarse resolution MODIS or AVHRR imagery (Lobel and Asner, 2004; Wardlow et al., 2007; Wessels et al., 2004; Xiao et al., 2003). However, the concern is that the granularity of the agriculture in this region of the world is too fine to be detected by coarse resolution imagery. The goal of this research is to determine whether or not accurate cropped land classifications can be established using WV2 and Landsat data. Given the fine granularity of subsistence cropland in sub–Saharan Africa, my study focuses more on village-scale cropland monitoring while also asking: can we produce an accurate classification using 30-meter Landsat surface reflectance data that are trained using labeled data from WV2 data? How will classification accuracy vary with scale? How far from labels can we accurately classify data? Can we be confident about classifications far from labels, and if so, how far? Finally, can we be confident in classifications between labels?

3.1.2 Physical Environment:

Herrmann et al. (2005) described the environment in northern Mali (approximately 16° N to 24° N) as part of the Saharan desert, with southern Mali (approximately 10° N to 16° N) ranging from the arid to semi–arid zone below the Sahara desert, to the humid-subtropical savannas in the south. Geographical regions based on rainfall time series data include: Sahelo-Sahara (approximately 18 °N to 25 °N), Sahel (approximately 13-15 °N to 18 °N), and Soudan (approximately 10 °N to 15 °N) (Nicholson, 2001). The 2005 Vulnerability Profile of Mali (Simonsson, 2005) summarized that precipitation in Mali varies greatly and considers the

existence of four "eco-climate" regions: Sahara, Sahelian, Sudanian, and Sudanian–Guinean; with rainfall throughout these regions averaging from 100 to 1,700 mm/year. To further illustrate the aridity gradient in Mali, climate profiles generated for each province by The World Bank characterize the northern-most province of Timbuktu as having maximum temperatures in June of approximately 44 °C, minimum temperatures in January of approximately 11 °C, and maximum rainfall averages reaching 27 mm in August. Conversely, the southern–most province of Mali, Sikasso, has maximum temperatures in March and April of approximately 37 °C, minimum temperatures in December and January of approximately 15 °C, and maximum rainfall averages reaching 282 mm in August (WorldBank 2013). ADD TO REFERENCES

The size and climatic variability of Mali allow it to have very distinct vegetative and productive zones that are attributed to specific livelihood regions. Productive regions are influenced by human-induced and environmental factors. The severe level of desertification of the Sahel region of Mali since at least the 1980s (Nicholson, 2001) is an example of these contributing factors. Following a prominent south to north aridity gradient, livelihood regions in Mali range from nomadic pastoralism in the far northern region just south of the Sahara desert, to herding, to increasingly rain dependent crops as you travel further south. The majority of the population in Mali depends heavily on agro-pastoral resources, however only 25% of Mali is capable of supporting such practices (Simonsson, 2005).

3.2 Data

Because a distinct aridity gradient exists from south to north, we blocked the study region into a more arid north study region, and a more humid south region. We selected two

Landsat TM (at 30x30 m pixel resolution) scenes in each region for a total of four Landsat TM scenes. The scenes were selected based on their spatial and temporal agreement with the labeled training data (Figure 1), the accuracy of which is the premise of this study. The Landsat surface reflectance product was used for this study in the hope it would facilitate comparability across the landscape.

The training data for this study were produced by the USAID Famine Early Warning System Network using the USGS Rapid Land Cover Mapper (RLCM) tool. The RLCM tool overlays a grid of points on top of high-resolution imagery (Figure 1). For this research, the grid of points were set to be 500 m apart. The RLCM displays an interface by which the user is zoomed into each individual point and prompted to label that point as crop or non-crop.

The data used in this thesis were labeled by African workers with familiarity of the sub-Saharan Mali landscape. The high-resolution imagery used was WorldView-2 data. Only small, area-frame scale sections of the WorldView-2 data were used for the labeling process, so the entirety of southern Mali is not covered with these labeled data (Figure 2).

The coverage of the labeled data in southern Mali is limited, owing to the scarcity of acquired imagery within any growing season. The FEWS NET team also labeled the grid points with: elevation (not used), livelihood zone, and latitude and longitude. The swaths of labeled data were selected based on environmental extremes (two swaths in the arid north, two swaths in the humid subtropical south) and on livelihood regions. As the FEWS NET livelihood zone data designated specific crop types and agricultural practices for different regions of Mali, it seemed most logical to use swaths of training data within the same livelihood region. This method of selection was successful for the southern training data, as two swaths were easily

identified that lie within the same livelihood. However, when selecting the two swaths for the

northern training data, we found it difficult to find two swaths that were related both spatially

and according to livelihood (Figure 2). Ultimately, the two swaths at the northern extent of the

training data that were closest to one another (within the same Landsat scene) were selected for

use in cropland classifications. However, these two swaths are distributed among four different

livelihood regions, primarily growing millet, shallots, rice, and sorghum (FEWSNET, 2013).

3.3 Classification Methods

A single digital image pixel can be associated with multiple spectral bands. By

comparing cross-band pixels, especially when associating them with known pixels, one can

classify all pixels within a digital image into specific regions or groups known as "classes"

(Campbell and Wynne, 2011). A class is a compilation of patterns whose spatial distribution is

determined by probability density (Jain et al., 1999). The division and re-assignment of

heterogeneous land cover types into homogenous classes is known as classification. Though

classification is often used for the delineation of ground cover into class types, the value of

using these classes to determine land cover areas is questionable unless conducted through

computer interpretation of a digital image in which pixel level analyses can be carried out

(Richards and Jia, 1999). Unsupervised, or unlabeled, classification of variables into meaningful

homogenous groups is a method of cluster analysis (Gan et al., 2007; Jain et al., 1999). One of the

primary differences between unsupervised and supervised, or unlabeled versus labeled,

analyses is that supervised analysis is based on categorical labels derived specifically from the

data (Campbell and Wynne, 2011; Jain et al., 1999).

Unsupervised classification is related to the idea of "natural groups" or "natural clusters" within an image being grouped together based strictly on spectral similarities within the data (Campbell and Wynne, 2011). Minimal analyst input is required apart from inputting the number of classes or the minimum and maximum to be assigned to the data. The fundamentals of the process itself involve the separation of values into the specified number of classes, or clusters, based on some measure of distance, such as Euclidean distance (Jones and Vaughan, 2010). Due to the limited classification-analyst interaction is little opportunity for analyst bias exists prior to interpretation of the results to increase error in the classification (Campbell, 2011). One of the primary limitations in unsupervised classification is that the classification is based on natural relationships between spectral values. Often, a significant difference will be observed between the ideal classes of interest to the interpreter and the classes actually created by the classification. The analyst has less control over the resultant classes in unsupervised classification than in the case of supervised classification, as no training fields or informational classes are involved (Campbell and Wynne, 2011).

Supervised classification is the classification of pixels using training fields. The availability of accurate training fields is essential to supervised classification. Supervised classification is especially suitable in such cases where a specific product is expected from the analysis, or when conducting comparisons of landscapes that have different temporal or spatial identities (bi-image comparison, time series, or nearest neighbor) (Campbell and Wynne, 2011). Another aspect of supervised classification is that the analyst has sufficient knowledge of the area in which the study is being conducted. The added control in the hands of the analyst adds some degree of detail or precision to the assessment, but can also introduce unwanted bias due

to the subjectivity and opinions of the analyst (Guo and Mason, 2009). A "hybrid classification" approach is one in which the initial step is to conduct a unsupervised classification, assigning classes based on supplementary data such as aerial imagery or ground truth data. Afterward, using the statistical results of the initial classification as training data, the image is reclassified using a supervised method (Guo and Mason, 2009). In terms of accuracy, the subjectivity of the analyst may affect ground truth and aerial data. Human observation is variable, so it is impracticable to expect to measure and extract the exact "truth" from the field (or from aerial data) because it will differ with each observer (Adams and Gillespie, 2006).

3.3.1 Methods

We first organize the bands of imagery, truth data, and auxiliary information that make up the data so that optimal analysis may occur. Scenes with the most training point swaths were considered for the northern and southern extents of the training data. The northern extent scene included two swaths of training points while the southern extent has one complete swath and a partial swath that it shares with the scene directly north of itself. The south scene and the scene directly north of it where mosaicked together to create two complete training swaths. This was to ensure more comparable results between the northern and southern analyses.

The north scene represents the arid part of the study region, while the south scene represents the humid subtropical portion of the study region. Together, they represent the two most extreme environments of the study region's aridity gradient. Entire Landsat TM scenes were not used because this study is interested in the accuracy that can be obtained in a classification using the labeled data and 30 m Landsat. Therefore, the northern and southern

22

classification regions were clipped so that only the areas where swaths of training point data exist for the classification processes.

ERDAS IMAGINE's supervised classification function has a maximum likelihood option that incorporates a Bayesian classifier that takes added variability within a single class into account. However, the supervised classification function does not allow for the inclusion of point data into the process. Training samples are traditionally acquired by creating a polygon over a specific area of the image that represents a class type. As such, we decided that a polygon file representing the data affiliated with each training point should be created. The prime method for creating the AOI training samples from point data is to create a buffer region around each training point.

To assess the best method of supervised classification using these training point data, classifications were conducted at different spatial scales, progressively moving away from the points, and incorporating more pixel values for each. This method of pixel inclusion incorporates increasing numbers of pixels depending on proximity to the training points. The question behind this reasoning is: when using point data to classify an image, instead of the traditional areas of interest, how many pixels need to be included to produce a efficient training space?

As all classification algorithms require the input of some sort of signature information (i.e. parametric, nonparametric, etc). The AOI feature spaces delineated had to be attributed with spectral signature information, resulting in training samples for the classification analysis. The idea behind training samples is to pick out a group of spectral signatures, the spectral values associated with each pixel of the image, that represent a single class and then repeating

the processes multiple times depending on the number of classes desired. Doing this requires the overlaying of the AOI regions onto an image. A signature editing tool takes the values associated with the pixels in the AOI region averages and adds them to the editor. The result of the signature process is the training samples to be used in the maximum likelihood classification. Parametric decision rules within the maximum likelihood classifier in IMAGINE use the training signatures to assign every pixel within the continuous surface of the image being classified into the designated classes.

To answer specific questions generated throughout this research, various methods of dividing labels into training and test data were employed. The three label division schemes (Figure 3) include: Thin, which will be a test accuracy between labels; Corner, which will test the accuracy of predictions just outside of labels; and Swaths A and B, which test how accurate predictions are far away from labels. Swath methods include predicting from swath A to B and also from B to A. For each method, at least 50–80% of the training points will end up being incorporated as training data for the classification process.

After all classification methods were completed, confusion matrices were generated to display the true and predicted data and the overall performance of the classification schema. Kappa coefficients included in these confusion matrix tables are a conservative measure of agreement, one that takes into account agreement occurring by chance. Also included in these tables were user accuracies, which measure the probability that a pixel is a certain class given that the classifier has labeled the pixel into that class; and producer accuracies, which measure the probability that the classifier has labeled and image pixel into a certain class given the ground truth is that class in reality.

3.4 Results

Results for this research included kappa coefficients generated from confusion matrices with error of omission (Type II error), where pixels are not reported as crop when they should be; and error of commission (Type I error), where pixels are reported into crop when they should not be; are taken into account for three analyses. The 7 band supervised classification for the wet season, 2 band NDVI for wet and dry, and the 14 band supervised classification for wet and dry. The kappa coefficients were generated from supervised classifications of cropped land within areas where labeled data existed. The initial 32 classifications, using only the growing season data, were generated and confusion matrices were made showing results of classifications from all methods. User and producer accuracies, along with kappa coefficients were calculated from these confusion matrices (Table 1 (North) and Table 2 (South)). User accuracies for the crop class range from 0.188 to 0.462 in the southern region, and from 0 to 0.497 in the north region. User accuracies measure how good the model is, or, how well the true data can estimate crop. Producer accuracies range from 0.285 to 0.780 in the southern region, and 0 to 0.456 in the northern region. Producer accuracy measures how good the prediction map is, or, how accurate the prediction is.

NDVI were generated for the leaf-on and leaf-off datasets in both the North and the South. The leaf-on and leaf-off NDVI for each region were stacked together to create one 2-band image. A classification using the thinning method was then performed on the 2-band NDVI images for the two regions. The thinning method was used due to its higher accuracy compared to the other methods. The kappa coefficients for the NDVI classifications (Table 3) were very low, the highest being only 0.24, suggesting that more information is needed to classify crop vs.

non-crop than just what is included in the NDVI. However, information was lacking in the NDVI to classify cropland, legitimizing the use of the original 7-band imagery in the classification. However, the 7-band classifications proved to have low kappa, no higher than 0.32 in the most productive region. Therefore, the analysis to determine the difference between the two leaf-on and leaf-off time periods is most promising.

As most of the kappa coefficients were initially rather low, additional analysis was needed to determine just how much information is being included in the solely crop or non-crop classifications. Also, dry season data were collected to determine if the difference between the wet season and the dry season, also known as leaf-on vs. leaf-off analysis, could more accurately discern crop. The 7-band raster used in the initial leaf-on analysis was combined with the 7-bands of the leaf-off analysis, generating one 14-band image. The kappa coefficients from these classifications (Table 4 and 5) were not dramatically higher than the original analysis, but they did increase. The highest kappa was a 0.488 in the South, where it had been 0.32. Though less dramatic, the kappa did increase in the North as well, the highest now being 0.443. The kappa coefficients showed an overall increase across all methods, though marginally.

3.5 Discussion

The results of this research vary between different classification methods and different buffer sizes. Mali livelihood zones explain some of the variation between classification results. As far as environment, Mali has everything from desert in the north, to lush fertile zones in the south. The Niger river plays a huge role as well, with at least 3 livelihood zones depending on it completely. Cultural factors such as dependence on remittance and commercial dominance also

play a role in zone delineation. In the north, swath A falls into two different livelihood zones: zone 4, millet and transhumant livestock rearing, and zone 6, the Niger delta lakes where agro-pastoral rice and livestock rearing is prominent. Zone 4 has an annual precipitation of 300 to 500-mm, and is considered a food difficult zone where external aid is common. Zone 6 has an annual precipitation of 300 to 600-mm, and is considered a surplus-producing zone, as it is more food secure than the zones surrounding it. Swath B in the north also falls into two different livelihood zones: zone 9, west and central rain-fed millet and sorghum, and zone 5, the Dogon plateau where tourism, millet, shallot, and other wild food make up the livelihoods in the region. Zone 9 has an annual precipitation of 600 to 800-mm, and is in the middle of trade routes bringing livestock to the south and crops to the north. Zone 5 has an annual precipitation of 400 to 600-mm, and is considered a food deficit zone primarily due to its poor quality soils. The north results seem to show more variation between classification methods than the south results. This is most likely due to the inclusion of four agriculturally distinct livelihood regions.

In the south, both swaths A and B fell within a single livelihood zone, zone 11, which hosts southern maize, cotton, and fruits as its primary agricultural products. Zone 11 has an annual precipitation of 1000 to 1300-mm, and is considered a highly productive zone. The south displayed higher kappa accuracies and less, though still present, variation between classification methods overall. We believe that these patterns are due to the south having better comparability between swaths, and the north having very little comparability based on the livelihood zones. The higher kappa by buffer size is different for each classification method, including the 14-band and 2-band analyses. No pattern suggests that buffer size has any influence on improving accuracy or kappa coefficients.

Overall, the classification accuracies generated in the results of this study are not great based on kappa coefficients. Even in what was considered the best case, using two swaths within the same livelihood zone in the most agriculturally productive region, the highest kappa coefficients were still below 0.5 and you can see examples of these classifications in Figures 4, 5, 6, and 7 (Figure 4, 5, 6, and 7 shows the resultant classifications for the 90-m south scene, overlaying WV-2 imagery and overlayed with crop/non-crop point data. These figures are from the southern half of Swath B in the south). However, upon improving on the original methods, kappa coefficients did increase for every method, though only moderate change was seen (Figure 4, 5, 6, and 7).

Variation between different buffer sizes is inconsequential, though differences between different classification testing methods were observed. In the original 7-band methods, kappa coefficients between the thinning and corner methods dropped by approximately half. Also, we observed no similarity between swath A and swath B, suggesting that the swaths are not symmetric. The swaths also have no similarity with the thinning or corner methods. In the 14-band image classifications the highest kappa coefficient for the south increased from 0.320 to 0.488. Unlike in the 7-band method, we saw little to no difference between the thinning and corner methods in the south, and very slight difference between the thinning and corner methods in the north. However, the swath methods still show remarkably low kappa accuracy in both the north and south ranging from negative kappas to kappas less than 0.1.

3.6 Conclusions

Several questions can be answered considering the results of these classifications: 1) How does classification accuracy vary with scale? 2) How far from labels can we classify data? 3) Can we be confident about classifications far from labels, if so, how far? Finally, can we be confident of classifications between labels? The degree of accuracy far from labels was measured by classifying one swath with the data from another. In this case there appears to be no real ability to predict accurately between swaths. Accuracy quickly drops off when predictions begin to be made outside of the training area, and the consideration of different buffer sizes doesn't seem to matter at all. There was hope that classification between labels would produce high, reliable accuracy, however even in the best–case scenario kappa coefficients were still less than 0.5. So, if accounting for crop type and agricultural practices isn't enough to ensure comparability between swaths, are there other ways to define landscapes that would make accuracies using these data better? The results suggest that there is a spectral landscape that is changing rapidly from one place to another. Though the original assumption was that individual livelihood zones would be similar spectrally, there is likely to be gradation in aridity even within the same livelihood zone, causing this assumption to be incorrect. There is a trend of under classification of crop in the north and over classification in the south. As the majority of the over classification of cropland in the south was due to all types of vegetation being classified as crop, there was the assumption that the combination of WorldView–2 imagery from different seasons using leaf–off pairs might increase accuracy. This assumption turned out to be correct, though not to the degree that was hoped.

Overall, kappa coefficients were low, with none reaching above 0.5. This suggests that we cannot produce significantly accurate classifications using this method. We do illustrate, however, that livelihood zone differences play a significant role in low agreement for the swath methods in the north. Though, even in the south classifications where the swaths were in the same livelihood zone the kappas were still very low with the methods used. This is most likely due to the aridity gradient in the country, which is explained by the annual precipitation for the different livelihood zones. This suggests that classification far from lables is not possible. Finally, the thin method has the best kappas which suggests that we can be more confident in classifications between labels, though accuracy quickly declines as we move away from labels.

References

Adams, J.B., Gillespie A.R. (2006). *Remote Sensing of Landscapes with Spectral Images: A Physical Modeling Approach*: Cambridge University Press. ISBN: 9780521662215

Barrett, C.B. (2010). Measuring food insecurity. *Science, 327*, 825-828.

Campbell, J.B., Wynne, R.H. (2011). *Introduction to Remote Sensing* (Fifth ed.). New York, NY: The Guilford Press.

Carmichael, J.W., George, J. Alan, & Julius, R. S. (1968). Finding Natural Clusters. *Systematic Biology/Zoology, 17*(2), 144 - 150.

Chen, Z., Li, S., Ren, J., Gong, P., Zhang, M., Wang, L., Xiao, S., Jiang, D. (2008). Chapter 15: Monitoring and Management of Agriculture with Remote Sensing. In S. Liang (Ed.), *Advance in Land Remote Sensing: System, Modeling, Inversion and Application*. College Park, MD: Springer Science+Business Media B.V.

Fecso, R., Tortora, R.D., Vogel, F.A. (1985). Recent and future activities to improve sampling frames for agriculture. *Proceedings of the Survey Resaerch Method Section, American Statistical Association*, 132-141.

FEWSNET. (2013). Famine Early Warning Systems Network. 2013, from http://www.fews.net/pages/livelihoods-country.aspx?gb=ml&l=en

Frolking, S., Xiao, X., Zhuang, Y., Salas, W., Li, C. (1999). Agricultural land-use in China: a comparison of area estimates from ground-based census and satellite-borne remote sensing. *Global Ecology and Biogeography, 8*, 407-416.

Gan, G., Ma, C., & Wu, J. (2007). *Data Clustering: Theory, Algorithms, and Applications ASA-SIAM Series on Statistics and Applied Probability*

Gonzalez-Alonso, F., Cuevas, J.M., Arbiol, R., Baulies, X. (1997). Remote sensing of agricultural statistics: crop area estimation in north-eastern Spain through diachronic Landsat TM and ground sample data. *International Journal of Remote Sensing, 18*(2), 467-470.

Grace, K., Husak, G.J., Pedreros, D., Michaelsen, J. (2012). Using high resolution satellite imagery to estimate cropped area in Guatemala and Haiti. *Applied Geography, 32*, 433-440.

Guo, J., & Mason, P.J. (2009). *Essential Image Processing and GIS for Remote Sensing*: Wiley-Blackwell.

Hammond, A.L. (1975). Crop forecasting from space: toward a global food watch. *Science, 188*(4187), 434-436.

Husak, G.J., Marshall, M.T., Michaelsen, J., Pedreros, D., Funk, C., Galu, G. (2008). Crop area estimation using high and medium resolution satellite imagery in areas with complex topography. *Journal of Geophysical Research, 113*, D14. doi: 10.1029/2007JD009175

Jain, A.K., Murty, M.N., & Flynn, P.J. (1999). Data Clustering: A Review. *ACM Computing Surveys, 31*(3), 264 - 323.

Jensen, J.R. (2004). *Introductory Digital Image Processing: A Remote Sensing Perspective* (3rd ed.). Upper Saddle River: Prentice-Hall, Inc.

Jones, H.G., Vaughan, R.A. (2010). *Remote sensing of vegetation: principles, techniques, and applications*. Great Clarendon Street, Oxford: Oxford University Press.

Launay, M., Guerif, M. (2005). Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. *Agriculture, Ecosystems and Environment, 111*, 321-339.

Lobel, D.B., Asner, G.P. (2004). Cropland distributions from temporal unmixing of MODIS data. *Remote Sensing of the Environment, 93*, 412-422.

Lobel, D.B, Asner, G.P., Oritz-Monasterio, J.I., Benning, T.L. (2003). Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. *Agriculture, Ecosystems and Environment, 94*, 205-220.

MacDonald, R.B., Hall, F.G., Erb, R.B. . (1975). The use of LANDSAT data in a large area crop inventory experiment (LACIE) *LARS Symposia* (Vol. Paper 46).

Marshall, M.T., Husak, G.J., Michaelsen, J., Funk, C., Pedreros, D., Adoum, A. (2011). Testing a high-resolution satellite interpretation technique for crop area monitoring in developing countries. *International Journal of Remote Sensing, 32*(23), 7997-8012. doi: 10.1080/01431161.2010.532168

Moulin, S., Bondeau, A., Delecolle, R. (1998). Combining agricultural crop models and satellite observations: From field to regional scales. *International Journal of Remote Sensing, 19*(6), 1021-1036.

Nicholson, S.E. (2001). Climatic and Environmental change in Africa during the last two centuries. *Climatic Research, 17*, 123-144.

Ozdogan, M., Woodcock, C.E. (2006). Resolution Dependent Errors in Remote Sensing of Cultivated Areas. *Remote Sensing of the Environment, 103*, 203-217.

Pax-Lenney, M., Woodcock, C.E. (1997). The effect of spatial resolution on the ability to monitor the status of agricultural lands. *Remote Sensing of the Environment, 61*, 210-220.

Pax-Lenney, M., Woodcock, C.E., Collins, J.B., Hamdi, H. (1996). The Status of Agricultural Lands in Egypt: The Use of Multitemporal NDVI Features Derived from Landsat TM. *Remote Sensing of the Environment, 56*, 8-22.

Pradhan, S. (2001). Crop area estimation using GIS, remote sensing and area frame sampling. *Journal of Applied Earth Observation and Geoinformation, 3*(1), 86-92.

Richards, J.A., Jia, X. (1999). *Remote sensing digital image analysis: an introduction* (Third ed.). Berlin Heidelberg: Springer-Verlag.

Shalaby, A., Tateishi, R. (2007). Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the northwestern coastal zone of Egypt. *Applied Geography, 27*, 28-41.

Simonsson, L. (2005). Vulnerability profile of Mali *Poverty and Vulnerability Programme*. Stockholm, Sweden: Stockholm Environment Institute (SEI).

Herrmann, S.M., Anyamba, A., Tucker, C.J.. (2005). Recent trends in vegetation dynamics in the African Sahel and their relationship to climate. *Global Environmental Change, 15*, 394-404.

Tao, F., Yokozawa, M., Zhang, Z., Xu, Y., Hayashi, Y. (2005). Remote sensing of crop production in China by production efficiency models: models comparisons, estimates and uncertainties. *Ecological Modelling, 183*, 385-396.

Tottrup, C., Rasmussen, M.S. (2004). Mapping long-term changes in savannah crop productivity in Senegal through trend analysis of time series of remote sensing data. *Agriculture, Ecosystems and Environment, 103*, 545-560.

Troy, A. (2008). Geodemographic Segmentation. In S. Shekhar & H. Xiong (Eds.), *Encyclopedia of GIS* (pp. 348): Springer.

Tsiligirides, T.A. (1998). Remote sensing as a tool for agricultural statistics: a case study of area frame sampling methodology in Hellas. *Computers and Electronics in Agriculture, 20*, 45-77.

Wardlow, B.D., Egbert, S.L., Kastens, J.H. (2007). Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. central Great Plains. *Remote Sensing of the Environment, 108*, 290-310.

Wessels, K.J., DeFries, R.S., Dempewolf, J., Anderson, L.O., Hansen, A.J., Powell, S.L., Moran, E.F. (2004). Mapping regional land cover with MODIS data for biological conservation: Examples from the Greater Yellowstone Ecosystem, USA and Para´ State, Brazil. *Remote Sensing of the Environment, 92*, 67-83.

Xiao, X., Liu, J., Zhuang D., Frolking, S., Boles, S., Xu. B., Liu, M., Salas, W., Moore, B., Li, C. (2003). Uncertainties in estimates of cropland area in China: a comparison between an AVHRR-derived dataset and a Landsat TM-derived dataset. *Global and Planetary Change, 37*, 297-306.

Yuping, M., Shili, W., Li, Z., Yingyu, H., Liwei, Z., Yanbo, H., Futang, W. (2008). Monitoring winter wheat growth in North China by combining a crop model and remote sensing data. *International Journal of Applied Earth Observation and Geoinformation, 10*, 426-437.

Appendices

Table 1: Results for 7–band North wet season classification.

NORTH														
THIN						CORNER			S_SWATH			N_SWATH		
Classification (30-meter)			Classification (120-meter)			Classification (90-meter)			Classification (90-meter)			Classification (90-meter)		
	Prediction			Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	115	245	Crop	164	196	Crop	117	227	Crop	26	200	Crop	1	474
Non-Crop	253	3701	Non-Crop	204	3750	Non-Crop	169	1425	Non-Crop	587	3696	Non-Crop	58	3587
theta1	0.885		theta1	0.907		theta1	0.796		theta1	0.825		theta1	0.871	
theta2	0.845		theta2	0.845		theta2	0.727		theta2	0.828		theta2	0.874	
kappa	0.253		kappa	0.400		kappa	0.251		kappa	-0.012		kappa	-0.022	
user	0.313		user	0.446		user	0.409		user	0.042		user	0.017	
producer	0.319		producer	0.456		producer	0.340		producer	0.115		producer	0.002	
Classification (60-meter)			Classification (150-meter)			Classification (120-meter)			Classification (120-meter)			Classification (120-meter)		
	Prediction			Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	155	205	Crop	154	206	Crop	130	214	Crop	30	196	Crop	0	475
Non-Crop	197	3757	Non-Crop	162	3792	Non-Crop	179	1415	Non-Crop	709	3574	Non-Crop	53	3592
theta1	0.907		theta1	0.915		theta1	0.797		theta1	0.799		theta1	0.872	
theta2	0.849		theta2	0.856		theta2	0.720		theta2	0.802		theta2	0.875	
kappa	0.385		kappa	0.410		kappa	0.277		kappa	-0.016		kappa	-0.024	
user	0.440		user	0.487		user	0.421		user	0.041		user	0.000	
producer	0.431		producer	0.428		producer	0.378		producer	0.133		producer	0.000	
Classification (90-meter)			Classification (180-meter)			Classification (150-meter)			Classification (150-meter)			Classification (150-meter)		
	Prediction			Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	149	211	Crop	155	205	Crop	123	221	Crop	29	197	Crop	0	475
Non-Crop	180	3774	Non-Crop	157	3797	Non-Crop	173	1421	Non-Crop	639	3644	Non-Crop	23	3622
theta1	0.909		theta1	0.916		theta1	0.797		theta1	0.815		theta1	0.879	
theta2	0.853		theta2	0.856		theta2	0.724		theta2	0.817		theta2	0.880	
kappa	0.383		kappa	0.416		kappa	0.263		kappa	-0.011		kappa	-0.011	
user	0.453		user	0.497		user	0.416		user	0.043		user	0.000	
producer	0.414		producer	0.431		producer	0.358		producer	0.128		producer	0.000	
						Classification (180-meter)			Classification (180-meter)			Classification (180-meter)		
							Prediction			Prediction			Prediction	
						Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
						Crop	115	229	Crop	31	195	Crop	0	475
						Non-Crop	162	1432	Non-Crop	705	3578	Non-Crop	32	3613
						theta1	0.798		theta1	0.800		theta1	0.877	
						theta2	0.730		theta2	0.803		theta2	0.879	
						kappa	0.252		kappa	-0.013		kappa	-0.015	
						user	0.415		user	0.042		user	0.000	
						producer	0.334		producer	0.137		producer	0.000	

Table 2: Results for 7–band South wet season classification.

SOUTH														
	THIN						CORNER		S_SWATH			N_SWATH		
Classification (30-meter)			Classification (120-meter)			Classification (90-meter)			Classification (90-meter)			Classification (90-meter)		
	Prediction			Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	348	609	Crop	448	509	Crop	238	313	Crop	468	748	Crop	200	501
Non-Crop	559	2989	Non-Crop	521	3027	Non-Crop	540	1556	Non-Crop	728	2733	Non-Crop	325	3308
theta1	0.741		theta1	0.771		theta1	0.678		theta1	0.684		theta1	0.809	
theta2	0.672		theta2	0.664		theta2	0.620		theta2	0.617		theta2	0.756	
kappa	0.210		kappa	0.320		kappa	0.151		kappa	0.175		kappa	0.218	
user	0.384		user	0.462		user	0.306		user	0.391		user	0.381	
producer	0.364		producer	0.468		producer	0.432		producer	0.385		producer	0.285	
Classification (60-meter)			Classification (150-meter)			Classification (120-meter)			Classification (120-meter)			Classification (120-meter)		
	Prediction			Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	435	522	Crop	446	511	Crop	239	312	Crop	574	642	Crop	547	154
Non-Crop	545	3003	Non-Crop	531	3017	Non-Crop	544	1552	Non-Crop	1161	2300	Non-Crop	2317	1316
theta1	0.763		theta1	0.769		theta1	0.677		theta1	0.614		theta1	0.430	
theta2	0.662		theta2	0.663		theta2	0.619		theta2	0.562		theta2	0.391	
kappa	0.298		kappa	0.314		kappa	0.151		kappa	0.120		kappa	0.063	
user	0.444		user	0.456		user	0.305		user	0.331		user	0.191	
producer	0.455		producer	0.466		producer	0.434		producer	0.472		producer	0.780	
Classification (90-meter)			Classification (180-meter)			Classification (150-meter)			Classification (150-meter)			Classification (150-meter)		
	Prediction			Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	438	519	Crop	428	529	Crop	229	322	Crop	577	639	Crop	522	179
Non-Crop	526	3022	Non-Crop	513	3035	Non-Crop	543	1553	Non-Crop	1174	2287	Non-Crop	2262	1371
theta1	0.768		theta1	0.769		theta1	0.673		theta1	0.612		theta1	0.437	
theta2	0.664		theta2	0.667		theta2	0.622		theta2	0.560		theta2	0.404	
kappa	0.309		kappa	0.305		kappa	0.136		kappa	0.118		kappa	0.055	
user	0.454		user	0.455		user	0.297		user	0.330		user	0.188	
producer	0.458		producer	0.447		producer	0.416		producer	0.475		producer	0.745	
						Classification (180-meter)			Classification (180-meter)			Classification (180-meter)		
							Prediction			Prediction			Prediction	
						Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
						Crop	227	324	Crop	457	759	Crop	470	231
						Non-Crop	514	1582	Non-Crop	823	2638	Non-Crop	2022	1611
						theta1	0.683		theta1	0.662		theta1	0.480	
						theta2	0.6284		theta2	0.609		theta2	0.4493	
						kappa	0.148		kappa	0.136		kappa	0.056	
						user	0.306		user	0.357		user	0.189	
						producer	0.412		producer	0.376		producer	0.670	

Table 3: Results for NDVI analysis on wet and dry season data.

NORTH						SOUTH					
THIN						THIN					
Classification (30-meter)			Classification (120-meter)			Classification (30-meter)			Classification (120-meter)		
	Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	822	135	Crop	814	143	Crop	1874	135	Crop	1839	143
Non-Crop	1675	1874	Non-Crop	1709	1839	Non-Crop	1674	822	Non-Crop	1709	814
theta1	0.598		theta1	0.589		theta1	0.598		theta1	0.589	
theta2	0.469		theta2	0.465		theta2	0.469		theta2	0.465	
kappa	0.244		kappa	0.231		kappa	0.244		kappa	0.231	
user	0.329		user	0.323		user	0.528		user	0.518	
producer	0.859		producer	0.851		producer	0.933		producer	0.928	
Classification (60-meter)			Classification (150-meter)			Classification (60-meter)			Classification (150-meter)		
	Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	822	135	Crop	808	149	Crop	1843	135	Crop	1858	149
Non-Crop	1705	1843	Non-Crop	1690	1858	Non-Crop	1705	822	Non-Crop	1690	808
theta1	0.592		theta1	0.592		theta1	0.592		theta1	0.592	
theta2	0.465		theta2	0.469		theta2	0.465		theta2	0.469	
kappa	0.237		kappa	0.232		kappa	0.237		kappa	0.232	
user	0.325		user	0.323		user	0.519		user	0.524	
producer	0.859		producer	0.844		producer	0.932		producer	0.926	
Classification (90-meter)			Classification (180-meter)			Classification (90-meter)			Classification (180-meter)		
	Prediction			Prediction			Prediction			Prediction	
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	817	140	Crop	807	150	Crop	1874	140	Crop	1896	150
Non-Crop	1701	1847	Non-Crop	1652	1896	Non-Crop	1701	817	Non-Crop	1652	807
theta1	0.591		theta1	0.600		theta1	0.594		theta1	0.600	
theta2	0.466		theta2	0.474		theta2	0.468		theta2	0.474	
kappa	0.235		kappa	0.240		kappa	0.237		kappa	0.240	
user	0.324		user	0.328		user	0.524		user	0.534	
producer	0.854		producer	0.843		producer	0.930		producer	0.927	

Table 4: Results for 14–band North wet and dry season classification.

NORTH

THIN — Classification (30-meter)	Prediction Crop	Non-Crop	THIN — Classification (120-meter)	Prediction Crop	Non-Crop	CORNER (A & B) — Classification (90-meter)	Prediction Crop	Non-Crop	SWATH (B to A) — Classification (90-meter)	Prediction Crop	Non-Crop	SWATH (A to B) — Classification (90-meter)	Prediction Crop	Non-Crop
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	245	115	Crop	249	111	Crop	276	68	Crop	8	218	Crop	0	475
Non-Crop	372	3582	Non-Crop	421	3533	Non-Crop	443	1151	Non-Crop	20	4263	Non-Crop	6	3639
theta1	0.887		theta1	0.877		theta1	0.736		theta1	0.947		theta1	0.883	
theta2	0.797		theta2	0.787		theta2	0.583		theta2	0.944		theta2	0.884	
kappa	0.443		kappa	0.421		kappa	0.367		kappa	0.053		kappa	-0.003	
user	0.397		user	0.372		user	0.384		user	0.286		user	0.000	
producer	0.681		producer	0.692		producer	0.802		producer	0.035		producer	0.000	

THIN — Classification (60-meter)	Prediction Crop	Non-Crop	THIN — Classification (150-meter)	Prediction Crop	Non-Crop	CORNER (A & B) — Classification (120-meter)	Prediction Crop	Non-Crop	SWATH (B to A) — Classification (120-meter)	Prediction Crop	Non-Crop	SWATH (A to B) — Classification (120-meter)	Prediction Crop	Non-Crop
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	251	109	Crop	250	110	Crop	278	66	Crop	10	216	Crop	0	475
Non-Crop	394	3560	Non-Crop	430	3524	Non-Crop	447	1147	Non-Crop	22	4261	Non-Crop	6	3639
theta1	0.883		theta1	0.875		theta1	0.735		theta1	0.947		theta1	0.883	
theta2	0.792		theta2	0.785		theta2	0.581		theta2	0.943		theta2	0.884	
kappa	0.439		kappa	0.417		kappa	0.368		kappa	0.066		kappa	-0.003	
user	0.389		user	0.368		user	0.383		user	0.313		user	0.000	
producer	0.697		producer	0.694		producer	0.808		producer	0.044		producer	0.000	

THIN — Classification (90-meter)	Prediction Crop	Non-Crop	THIN — Classification (180-meter)	Prediction Crop	Non-Crop	CORNER (A & B) — Classification (150-meter)	Prediction Crop	Non-Crop	SWATH (B to A) — Classification (150-meter)	Prediction Crop	Non-Crop	SWATH (A to B) — Classification (150-meter)	Prediction Crop	Non-Crop
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	250	110	Crop	252	108	Crop	276	68	Crop	11	215	Crop	0	475
Non-Crop	407	3547	Non-Crop	436	3518	Non-Crop	451	1143	Non-Crop	29	4254	Non-Crop	6	3639
theta1	0.880		theta1	0.874		theta1	0.732		theta1	0.946		theta1	0.883	
theta2	0.790		theta2	0.784		theta2	0.581		theta2	0.942		theta2	0.884	
kappa	0.430		kappa	0.417		kappa	0.362		kappa	0.069		kappa	-0.003	
user	0.381		user	0.366		user	0.380		user	0.275		user	0.000	
producer	0.694		producer	0.700		producer	0.802		producer	0.049		producer	0.000	

CORNER (A & B) — Classification (180-meter)	Prediction Crop	Non-Crop	SWATH (B to A) — Classification (180-meter)	Prediction Crop	Non-Crop	SWATH (A to B) — Classification (180-meter)	Prediction Crop	Non-Crop
Test	Crop	Non-Crop	Test	Crop	Non-Crop	Test	Crop	Non-Crop
Crop	227	67	Crop	11	215	Crop	0	475
Non-Crop	455	1139	Non-Crop	32	4251	Non-Crop	6	3639
theta1	0.724		theta1	0.945		theta1	0.883	
theta2	0.596		theta2	0.941		theta2	0.884	
kappa	0.316		kappa	0.067		kappa	-0.003	
user	0.333		user	0.256		user	0.000	
producer	0.772		producer	0.049		producer	0.000	

Table 5: Results for 14–band South wet and dry season classification.

SOUTH

	THIN					CORNER (A & B)			B — SWATH (B to A)			A — SWATH (A to B)		

THIN — Classification (30-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	649	308
Non-Crop	545	3003
theta1	0.811	
theta2	0.635	
kappa	0.481	
user	0.544	
producer	0.678	

THIN — Classification (120-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	683	274
Non-Crop	603	2945
theta1	0.805	
theta2	0.623	
kappa	0.483	
user	0.531	
producer	0.714	

CORNER (A & B) — Classification (90-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	326	225
Non-Crop	309	1787
theta1	0.798	
theta2	0.652	
kappa	0.421	
user	0.513	
producer	0.592	

SWATH (B to A) — Classification (90-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	1211	5
Non-Crop	3370	91
theta1	0.278	
theta2	0.270	
kappa	0.012	
user	0.264	
producer	0.996	

SWATH (A to B) — Classification (90-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	415	286
Non-Crop	856	2777
theta1	0.737	
theta2	0.640	
kappa	0.268	
user	0.327	
producer	0.592	

THIN — Classification (60-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	683	274
Non-Crop	592	2956
theta1	0.808	
theta2	0.625	
kappa	0.488	
user	0.536	
producer	0.714	

THIN — Classification (150-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	680	277
Non-Crop	605	2943
theta1	0.804	
theta2	0.624	
kappa	0.480	
user	0.529	
producer	0.711	

CORNER (A & B) — Classification (120-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	324	227
Non-Crop	314	1782
theta1	0.796	
theta2	0.651	
kappa	0.414	
user	0.508	
producer	0.588	

SWATH (B to A) — Classification (120-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	1211	5
Non-Crop	3360	101
theta1	0.281	
theta2	0.271	
kappa	0.013	
user	0.265	
producer	0.996	

SWATH (A to B) — Classification (120-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	433	268
Non-Crop	1007	2626
theta1	0.706	
theta2	0.613	
kappa	0.239	
user	0.301	
producer	0.618	

THIN — Classification (90-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	685	272
Non-Crop	609	2939
theta1	0.804	
theta2	0.622	
kappa	0.482	
user	0.529	
producer	0.716	

THIN — Classification (180-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	680	277
Non-Crop	604	2944
theta1	0.804	
theta2	0.624	
kappa	0.480	
user	0.530	
producer	0.711	

CORNER (A & B) — Classification (150-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	321	230
Non-Crop	316	1780
theta1	0.794	
theta2	0.651	
kappa	0.408	
user	0.504	
producer	0.583	

SWATH (B to A) — Classification (150-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	1211	5
Non-Crop	3372	89
theta1	0.278	
theta2	0.270	
kappa	0.011	
user	0.264	
producer	0.996	

SWATH (A to B) — Classification (150-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	450	251
Non-Crop	1171	2462
theta1	0.672	
theta2	0.585	
kappa	0.209	
user	0.278	
producer	0.642	

CORNER (A & B) — Classification (180-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	324	227
Non-Crop	316	1780
theta1	0.795	
theta2	0.651	
kappa	0.413	
user	0.506	
producer	0.588	

SWATH (B to A) — Classification (180-meter)

Test	Crop	Non-Crop
	Prediction	
Crop	1211	5
Non-Crop	3369	92
theta1	0.279	
theta2	0.270	
kappa	0.012	
user	0.264	
producer	0.996	

SWATH (A to B) — Classification (180-meter)

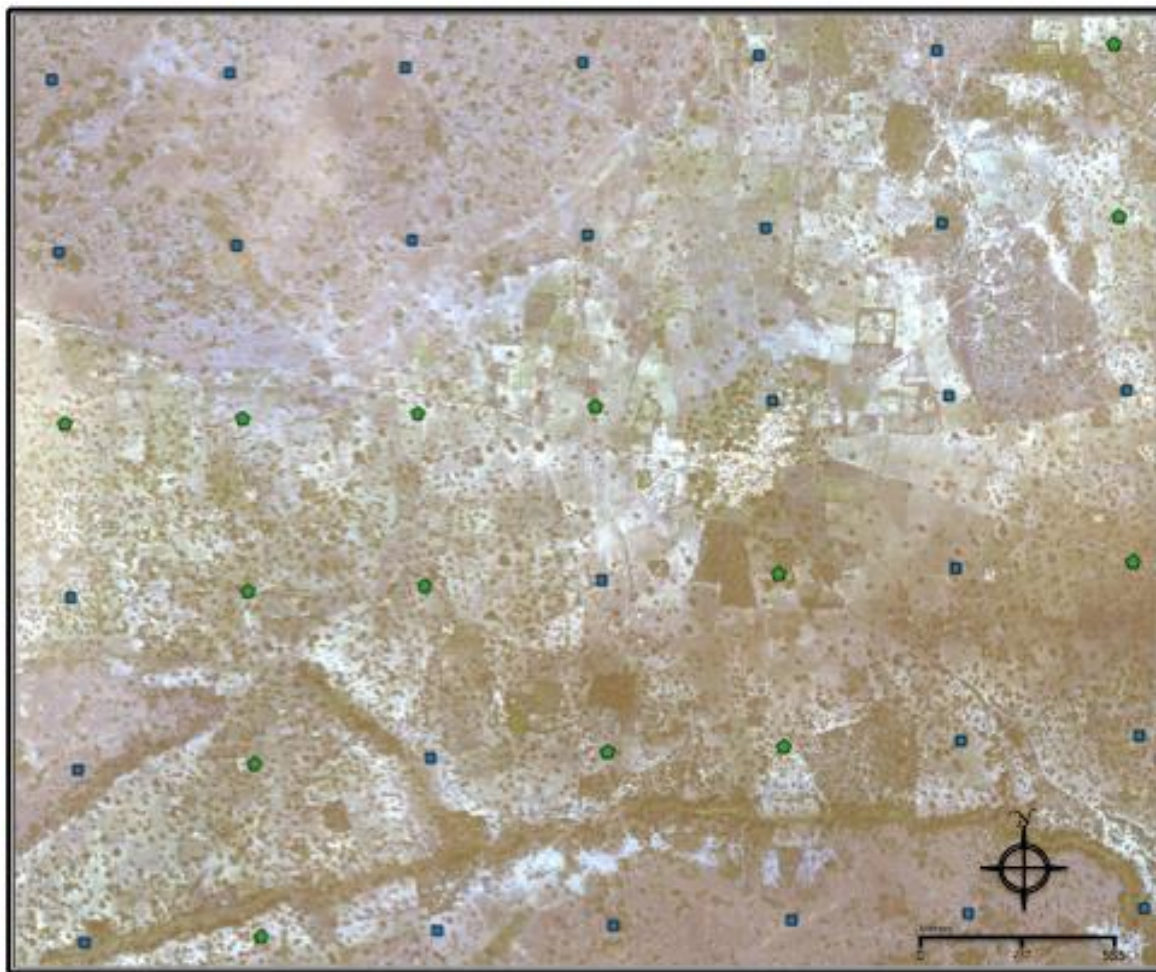| Test	Crop	Non-Crop
	Prediction	
Crop	472	229
Non-Crop	1438	2195
theta1	0.615	
theta2	0.540	
kappa	0.164	
user	0.247	
producer	0.673	

Figure 1: An example of the training data points and WV2 imagery. The points are on a 500 m grid. Green circles indicate points that were classified as crop, while blue squares were classified as not-crop.
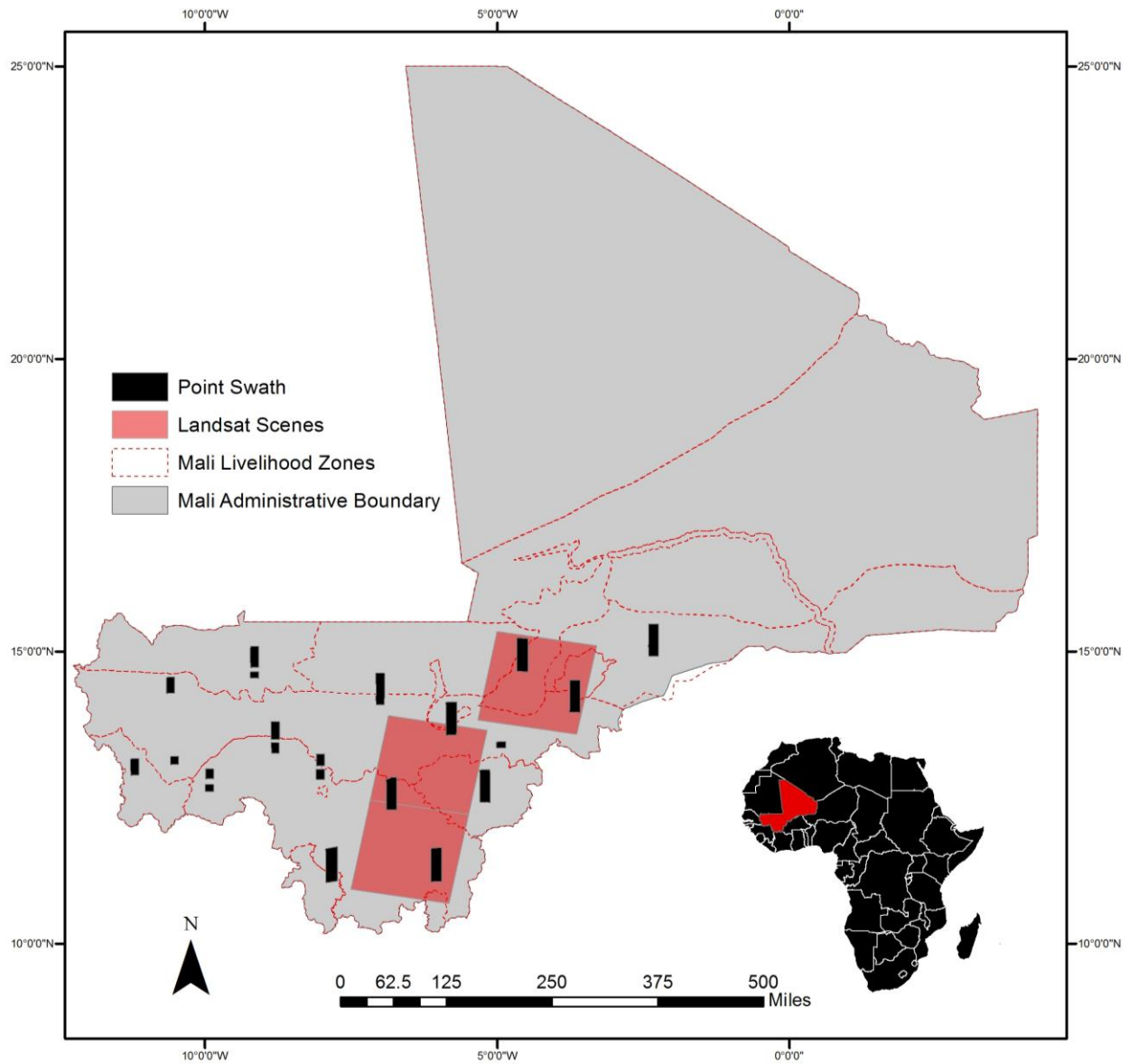
Figure 2: Illustrates the distribution of data. Red squares represent the distribution of the Landsat scenes. The black rectangles represent the distribution of the World View-2 derived point data. The red dotted lines show the boundaries for the different livelihood zones in Mali.
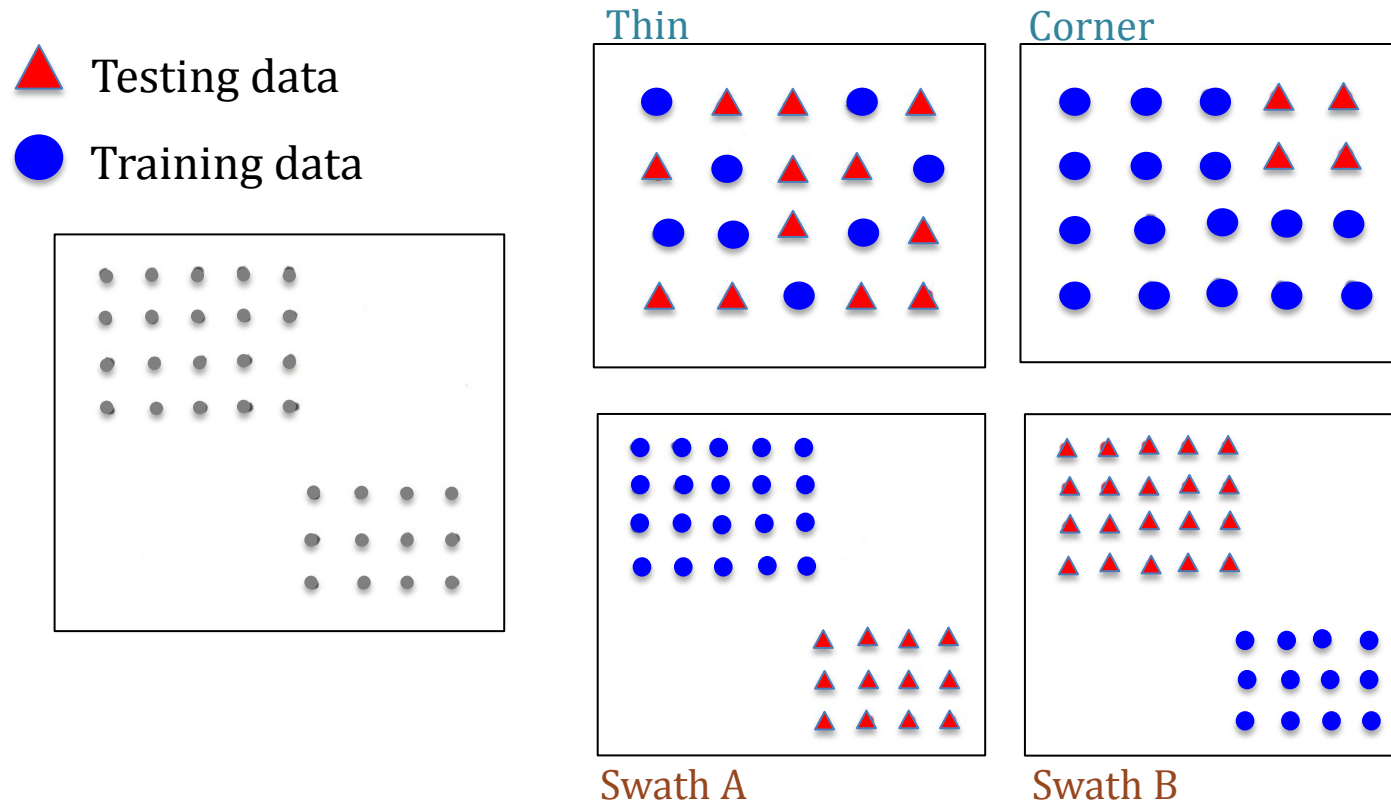
Figure 3: Examples of point exclusion methods of test and training data. The box with grey points is an example of the distribution of the two swaths, from each environment, at an offset. Blue points represent those used for training and red triangles for testing.
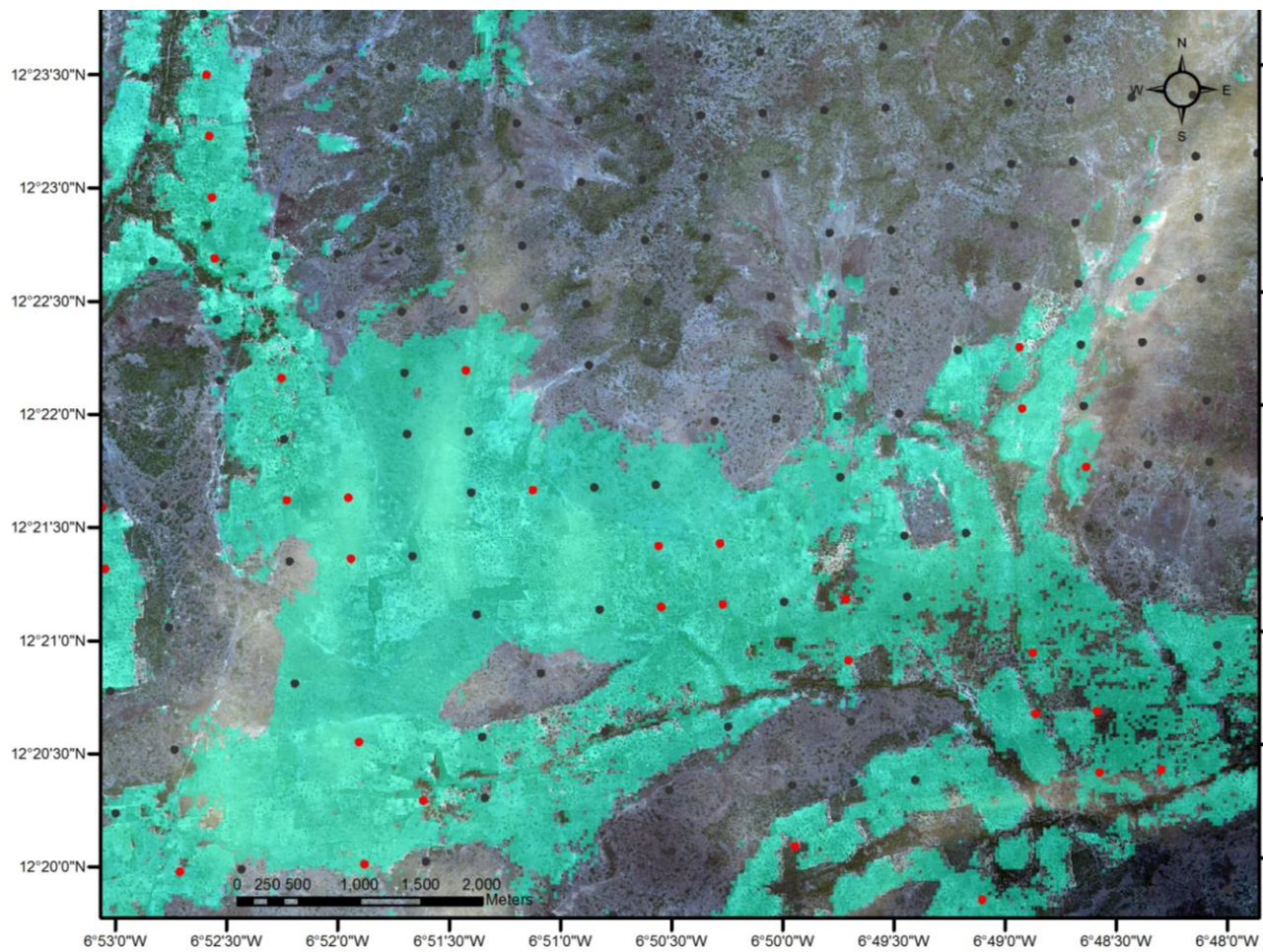
Figure 4: Corner 90-meter, red points represent crop pixels while black points represent non-crop pixels. Blue is the land area classified as crop for this method.
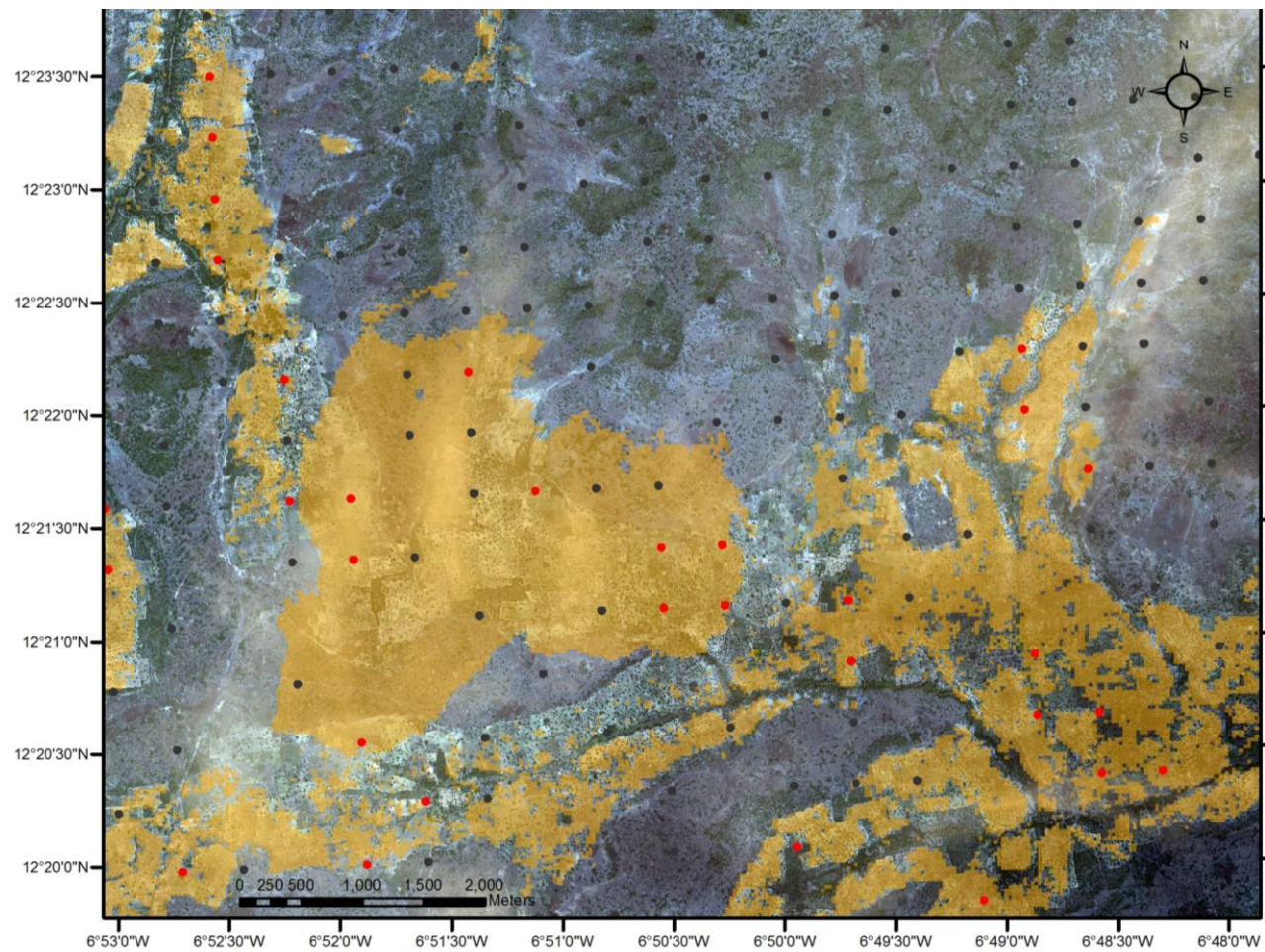
Figure 5: SwathA 90-meter, red points represent crop pixels while black points represent non-crop pixels. Orange is the land area classified as crop for this method.
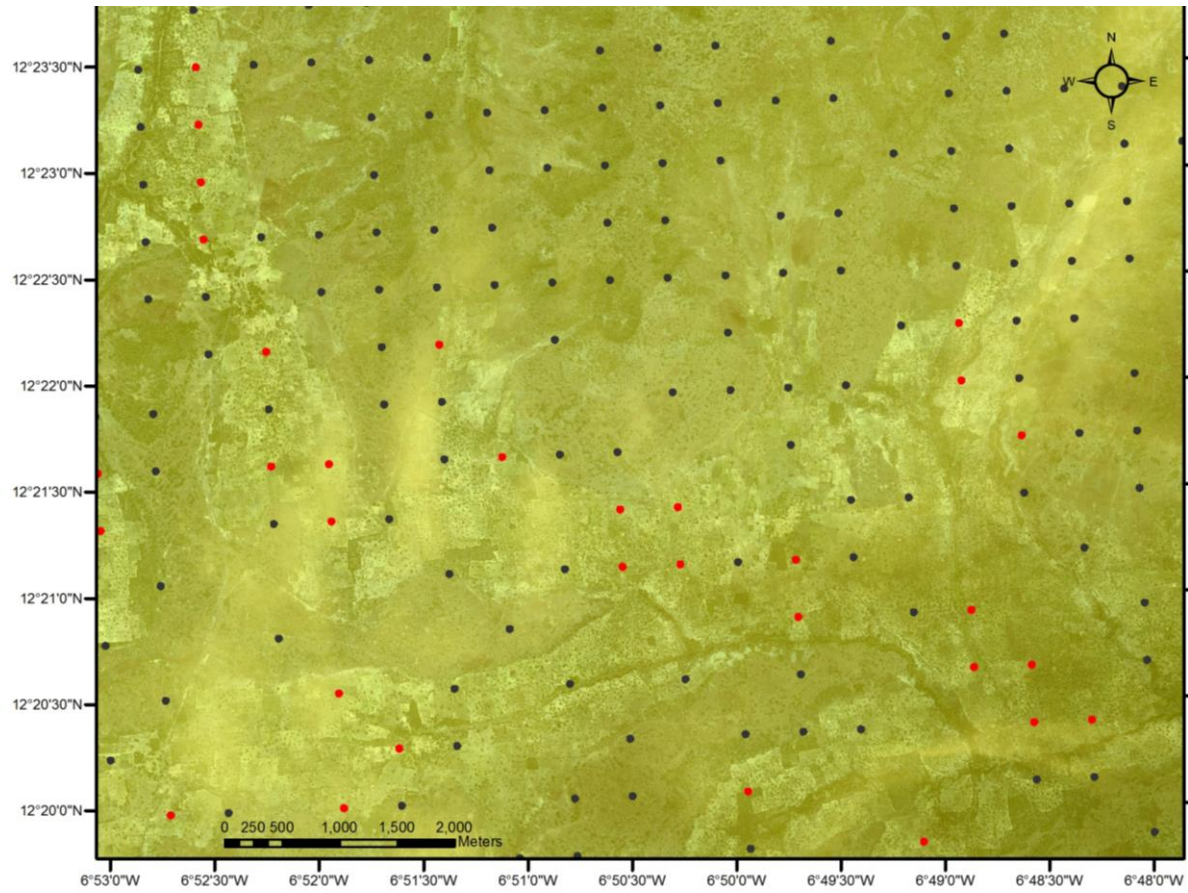
46

Figure 6: SwathB 90-meter, red points represent crop pixels while black points represent non-crop pixels. Yellow is the land area classified as crop for this method.
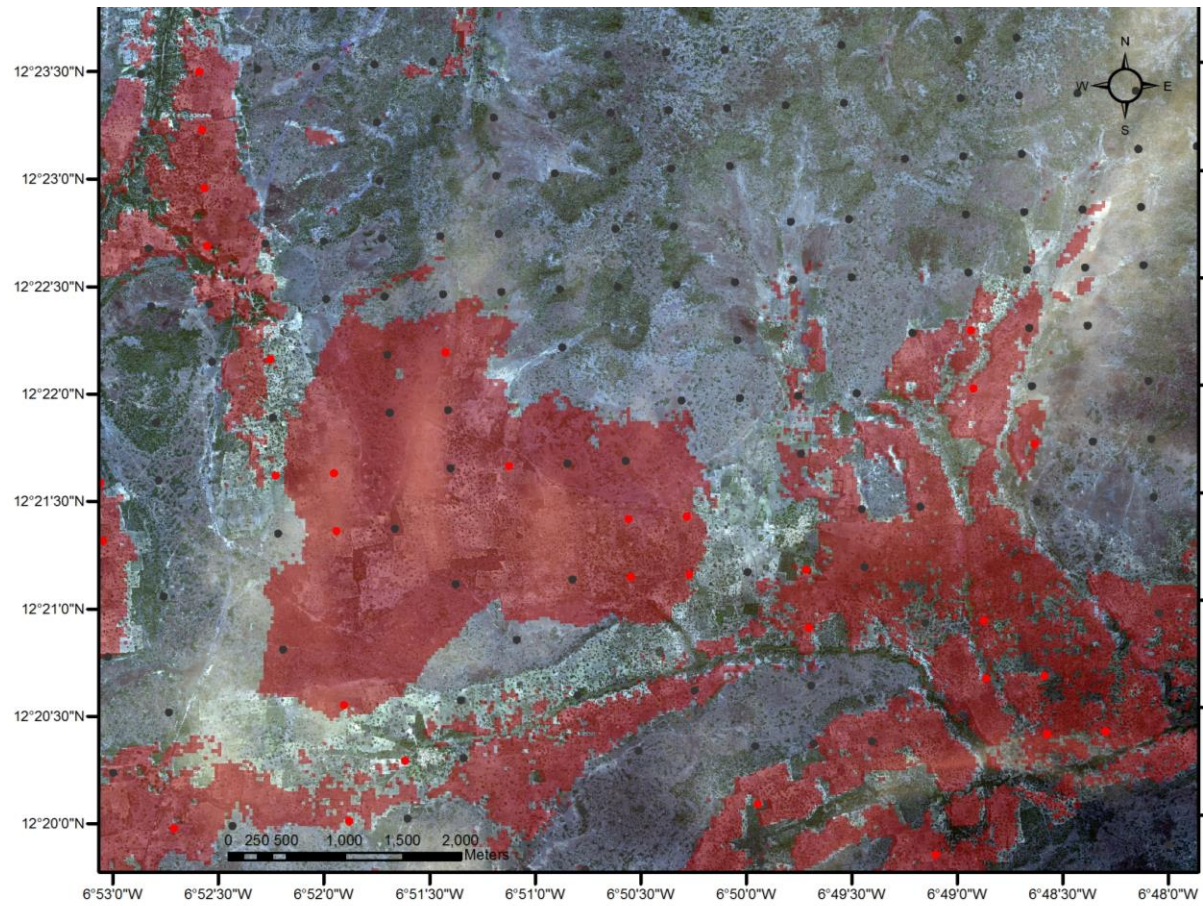
Figure 7: Thin 90-meter, red points represent crop pixels while black points represent non-crop pixels. Red is the land area classified as crop for this method.

VITA

Sarah L. Lewis-Gonzales earned a Bachelor of Arts degree in Geography from Texas Tech University of Lubbock, Texas in May of 2012. Succeeding graduation, Sarah began a three-month internship with the NASA DEVELOP program, as a team lead. Her project involved the development of products for partners in Africa and Washington D. C., generated from satellite imagery, and the remote sensing and GIS techniques. Her team won the opportunity to participate in the highlight talks at NASA headquarters in Washington, D. C. Following her internship Sarah began her graduate degree and Teaching Assistant fellowship at the University of Tennessee of Knoxville, Tennessee in August of 2012. In 2014 she accepted a position at the University of Tennessee as a Research Associate, working under the Department of Forestry, Fisheries, and Wildlife, at the Center for Renewable Carbon.