5-2015

# Time Series Analysis of MODIS NDVI data with Cloudy Pixels: Frequency-domain and SiZer analyses of vegetation change in Western Rwanda

Ephraim Robert Love
*University of Tennessee - Knoxville*, elove4@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Ephraim Robert Love entitled "Time Series Analysis of MODIS NDVI data with Cloudy Pixels: Frequency-domain and SiZer analyses of vegetation change in Western Rwanda." 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:

Liem Tran, Yingkui Li

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

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

# Time Series Analysis of MODIS NDVI data with Cloudy Pixels:

*Frequency-domain and SiZer analyses of vegetation change in Western Rwanda*

A Thesis Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Ephraim Robert Love

May 2015

## Dedication

This thesis is dedicated to my mother and father.

*'Parents who have no equals [in goodness] rear children unlike themselves'*

(Cant. R. ch. i. § 6 to i. 1; D. 483),

but I'll still try.

## Acknowledgements

I would like to first thank Dr. Nicholas Nagle, my major professor, advisor in matters of work and life, for his invaluable help and advice and for allowing me to complete this research. I would also like to thank Dr. Liem Tran and Dr. Yingkui Li, for their membership in my thesis committee and for their advice. Dr. Tran and Dr. Li have also given me their invaluable perspectives throughout my work and I am deeply grateful.

Thanks to my family and friends, who supported me fully, even when I was distant and stressed. I am grateful in particular to my mother, who has kept me on track academically since I began my bachelor's degree and continued to advise me during this research.

Finally, I wish to acknowledge the crucial role of the R open source community. My research relied heavily on the work and accomplishments of statisticians, engineers and geographers alike, working to create software and documentation as well as to provide a comprehensive knowledgebase for developers. My work could not have been completed, nor even begun, without the help of the R community (http://www.r-project.org/).

## Abstract

Remote sensing is a valuable source of data for the study of human ecology in rural areas. In this thesis, I attempt to analyze the presence of a long-term trend indicative of post-resettlement adaptation in the vegetation signals of Western Rwanda. There is a dearth of research utilizing medium resolution imagery to study difficult environments, such as tropical-montane regions, where complex topography and cloud cover diminish image accuracy. I attempt to add to the extant literature on frequency-domain smoothing methods as well as the literature on human-environment interaction in tropical-montane regions by applying a *harmonic filtering and smoothing algorithm* to the 'MOD13Q1', 16-day composite, 250m, NDVI, MODIS imagery. To create a more robust time-series, I combine Gaussian generalized additive models and discrete Fourier analysis of the residuals to impute values to a filtered time series, based on MODIS's own pixel reliability data. These methods significantly improve the quality of the time-series being analyzed, compared with the raw data, or imputation of the mean signal. To control for conflating variables, I take a *difference-in-differences* (DD) approach (Abadie, 2005) comparing resettled regions to older regions, identified in Google Earth. Harmonic filtering and smoothing shows a definite long-term trend of post-resettlement changes in the vegetation signal, demonstrated by the DD approach, analyzed in SiZer maps (Chaudhuri & Marron, 1999). Further research will be needed to determine whether this is indicative of cropping changes, or other impacts of post-resettlement adaptation.

# Table of Contents

# List of Figures

# Introduction

The demand for research on human-environment interaction, driven by changes in the Earth's climate, has led to a new focus on rural regions (Pender, 1999). Low density populations are thought to have relatively higher impacts on their surrounding environment, as they put greater pressure on the land and its natural resources. Additionally, geographic remoteness and weak institutions for environmental protection allow for greater exploitation. Remote sensing (RS) is a promising tool for studying human ecology in these broad regions; however, RS is not without its challenges. A vast array of RS platforms exists, often representing the inborn compromises between spatial, spectral and temporal resolution. The drive for improved inference in each of these dimensions has motivated researchers to explore new advances in feature extraction — to derive project-tailored datasets that maximize information extraction from raw data.

The study of rural regions inherently involves the study of broad-scale spatial data. In the case of image processing applications, such as computations on RS data, spatial resolution is related exponentially to computational complexity (Yin, Wu, & Liu, 2000). Therefore, it is often advantageous in RS studies of rural regions to use moderate to low resolution data to reduce computational steps and memory requirements. In this paper I will explore the suitability of the Moderate Resolution Imaging Spectroradiometer's (MODIS') 250m, scientific-band data for the study of rural, tropical-montane regions with a case study in Western Rwanda.

The tropics include many important regions of interest, as they contain diverse, critical ecosystems, and are often subject to little environmental regulation due to access difficulties (Mittermeier, Myers, Thomsen, Da Fonseca, & Olivieri, 1998). Unfortunately, these regions are not only difficult to study by land, but pose significant challenges to research utilizing RS data.

Most RS platforms available to scientists rely on the passive absorption of sunlight reflected from the Earth's surface. These passive sensors' images are obscured by the light reflected by clouds. Mountainous regions of the tropics are more difficult yet, because of irregular topography, resulting in varied angles of incidence and thus inter-pixel mixing. Development of systematic algorithms to improve data quality in tropical-montane regions is a necessary step to improving the feasibility of broad-scale, rural, RS studies. This in turn may provide insight to developers creating decision tree methods to make flexible, global RS algorithms, that can accommodate these difficult regions.

The primary goal of this thesis is to test quantitative, time-series methods for improvement of raw RS data in tropical-montane regions. The motivation for this research, however, comes from a desire to study the qualitative effects of the Rwandan Imidugudu resettlement policy in multi-temporal RS data. I will apply time-series and frequency-domain methods to study MODIS 16-day composite data. MODIS composites take the best quality observation from a given period, individually for each pixel in the raster. Especially in tropical-montane regions, a 16-day period may still not have a single suitable quality pixel. For this reason, MODIS provides a pixel reliability mask that allows for easy identification of obscured or low quality pixels. The first major step of this research is to filter and smooth these erroneous observations. Rather than simply imputing the mean of the time-series, or a windowed mean, I will use the Fast Fourier Transform (FFT) to retain the frequency-domain characteristics of the raw dataset (Kalacska, Bell, Arturo Sanchez-Azofeifa, & Caelli, 2009). This step aims to test my first hypothesis: *FFT-filtered, 16-day composite, 250m NDVI images provide a more robust time-series for further analysis.*

Finally, after smoothing the data, I will test the applicability of MODIS data to the study of tropical-montane human ecology with a case in Western Rwanda. After the Rwandan Genocide of 1994, the new authorities saw fit to institute broad policies to resettle the rural regions of Western Rwanda (M. van Leeuwen, 2001). I attempt to show long term, post-resettlement adaptation in Western Rwanda, evidenced by long-memory trends in vegetation signal data from MODIS. I take a difference-in-differences (DD) approach (Abadie, 2005) to study the changes in seasonal vegetation signals between areas surrounding new 'Imidugudu' settlements, and older villages to test the my second hypothesis that *there is a significant long-memory/low-frequency trend in the differenced camp and village data.*

Inferences from RS data are promising resources for social scientists, policy scholars, and aid organizations. The applications of RS data seem infinite as new methods for improved feature extraction are continually developed. I attempt to add to the extant RS and human-environment interaction literatures, by developing a smoothing algorithm for MODIS, and by showing the suitability of 250m 16-day composite data for the study of rural, tropical-montane regions. This thesis should be treated as a pilot for researchers interested in studying these difficult regions.

# Chapter 1: Remote Sensing Background

## 1.1 Why use remotely sensed data?

Remote sensing (RS) has quickly become a vital tool for scholars of human ecology (human-environment interaction) (Turner, 2003). Satellite imaging has been shown to be a valuable source of data, as it provides regular, open and objective data. While early RS work was devoted to military and meteorological applications, by the late 1970's researchers of the physical environment had begun to successfully relate RS image features[1] to physical phenomena on the Earth's surface. By 1985, studies including Justice, Townshend, Holben, & Tucker's (1985), had shown that a Normalized Difference Vegetation Index (NDVI), comparable over time and across space, could be generated from multispectral RS data. Subsequently, vast accomplishments have been made in the study of the physical environment and human ecology through remotely sensed feature extraction (Agarwal, Green, Grove, Evans, & Schweik, 2002).

I chose remotely sensed data as the focus of my thesis research because of its public accessibility, ease of scaling, and a wealth of research demonstrating its utility in the natural and social sciences. Over the past half-century, remote sensing has become an integral tool for the study of physical and human environments and is a promising data source for the study of complex phenomena such as post-resettlement adaptation. Remotely sensed data have been leveraged both in the study of humans' impact on the environment (Turner, 2003), as well as the reciprocal effect of environmental changes on humans (McGranahan, Balk, & Anderson, 2007a). Remote sensing platforms vary widely: from ground penetrating radar apertures mounted on

[1] In this paper I will refer to image 'features', which is a simplified means of describing a product of multiple image characteristics, with *a priori* knowledge of their relationship, in order to reduce the dimensional complexity of image analysis.

wheelbarrows, to hyperspectral radiometers mounted on planes or satellites. With dozens of open access data repositories, choosing an appropriate platform should be guided by the combined *radiometric*, *spectral*, *temporal* and *spatial* characteristics of the data gathered (Hamlyn & Vaughn, 2010b). These features integrated with the project's framework can greatly constrain or add to prospective analyses.

Derivation of land cover classes from RS data is basic to many methods of observing and measuring changes in the physical environment. While remotely sensed data have been used to classify land cover for decades (Sellers et al., 1994), recent advances in computational and analytic power have enabled geographers to test far more nuanced theories with satellite imagery (Galford et al., 2008; Kalacska et al., 2009; Ruiz et al., 2004). Analyses leveraging simple measurement of static spectral characteristics of the Earth's surface have been replaced by multi-temporal and multi-resolution approaches that maximize the usefulness of all available data. For example, contemporary techniques observe phenology, texture, distribution and other more abstract indices (Peña-Barragán et al., 2011). Improvements in spatial statistical techniques greatly extend the interpretation of such data (Galford et al., 2008; Sakamoto et al., 2005). Due to the physical constraints on improving the fidelity of RS data, and the exhaustive work that has already taken place on machine learning in this environment (Longbotham et al., 2012), the development of new, sophisticated data inputs and analysis frameworks is most likely to garner significant discoveries. Therefore, *the objective of this thesis is to develop models that better leverage moderate resolution, time-series imagery to study changes of vegetation signals in tropical-montane regions.*

### 1.2 A Brief History of RS

While many geoscientists share a view of RS as a relatively narrow group of airborne and satellite apertures, the term in fact encompasses a wide variety of methods and instrumentation. Remote sensing is "…the measurement or acquisition of information of some property of an object or phenomenon, by a recording device that is not in physical or intimate contact with the object or phenomenon under study" (Levin, 1999). Therefore, one could say that the earliest complete RS device was in fact Nicéphore Niépce's *camera obscura* with a pewter plate for the exposure, invented in 1826 ("History of Photography," 2014). However, RS could not become a critical tool for analysis without great scientific achievement in the understanding of the physics of light. James Clerk Maxwell and Heinrich Hertz theorized and confirmed that visible light is in fact an electromagnetic wave, on a spectrum of such waves (Hamlyn & Vaughn, 2010b). Subsequently, we began to understand that objects in fact radiated electromagnetic waves, across the entire spectrum of possible waves at different amplitudes. In the visible range, these relative differences are what create our perception of color.

Modern RS platforms range greatly in spectral, radiometric, spatial and temporal resolutions. RS instruments usually collect data in *bands* containing data from one segment of the electromagnetic spectrum. In RS, spectral resolution refers to the range of spectra an instrument is able to gather as well as the number of bands, or cross sections that are gathered within that operating spectral range. Radiometric resolution refers to the scale of digital numbers used when converting the incidence of light into recorded data. Radiometric resolution is given as the number of bits allocated to store a single digital number (DN) representing the intensity of incident radiation on a given pixel. The spatial resolution of a sensor encompasses the size of its field of view as well as the finest level at which it records discrete observations. However, in

satellite RS, we typically treat spatial resolution as the true size of a single pixel on the ground. Finally, temporal resolution can refer to several parameters, depending on sensor type, but most commonly refers to the frequency with which a particular area is revisited by an RS platform.

When comparing satellite RS platforms, it is important to keep in mind that many platforms rely not on a single satellite, but on a group of satellites. By employing several satellites, researchers achieve greater spatial or temporal resolution and/or some degree of data redundancy. Platforms vary in the number of satellites used, from the Advanced Very High Resolution Radiometer (AVHRR)[2] currently carried by 5 satellites, to the Defense Meteorological Satellite Program (DMSP)[3], currently carried by 4 satellites. Our platform of interest, the Moderate Resolution Imaging Spectroradiometer (MODIS)[4], is currently carried by only 2 satellites, but this still allows for MODIS to image the entire world's land surface, twice daily.

Multimodal and time-series analyses of RS data add information to a study by combining a temporal dimension into the data. This temporal dimension can be used to examine change, phenology and noise in data (Sakamoto et al., 2005). When effectively employed, high temporal resolution imagery has been found to contain more information than high spectral or spatial resolution imagery (Key, Warner, McGraw, & Fajvan, 2001). As with all RS studies, platform choice should be based on the constraints of the study; however, Key et al.'s (2001) work suggests that temporal resolution is likely the most powerful constraint to RS studies of ecology.

[2] More information about AVHRR can be found at: http://noaasis.noaa.gov/NOAASIS/ml/avhrr.html
[3] More information and data for DMSP: http://ngdc.noaa.gov/eog/dmsp.html
[4] More information and data for MODIS: http://modis.gsfc.nasa.gov/

### 1.3 The Normalized Difference Vegetation Index

The key remote sensing feature, and the dependent variable throughout this thesis, is NDVI. NDVI capitalizes on relative differences in vegetation's reflectivity at different electromagnetic spectra. NDVI is derived by a ratio transform of one band of data with a spectral range from .58 - .68 $\mu$m ('red light'), where chlorophyll absorbs the majority of incident radiation, to another band of data, ranging from .725 – 1.1 $\mu$m (near-infrared), where leaves' spongy structures reflect most incident radiation (Justice, Townshend, Holben, & Tucker, 1985). Figures 1.1 and 1.2 show samples of Near Infrared (NIR) and red bands, from NASA's distributed active archive center (DAAC) for MODIS. Figure 1.3 shows the ratio transform of the data contained in Figures 1.1 and 1.2, as provided in the MODIS DAAC.

NDVI is a transform of the surface reflectance values from the near-infrared (NIR) and red (R) bands of a remote sensing platform. Possible NDVI values range from -1 to 1, and can be used to predict presence and vigor of vegetation:

*Equation 1)* $$NDVI = \frac{NIR-R}{NIR+R}$$

Since the development of the NDVI, other band combinations, capitalizing on other known relationships, have been studied. One such feature is the Normalized Difference Water Index (NDWI). NDWI is an excellent feature for delineating open bodies of water. Another development has been the Enhanced Vegetation Index (EVI), which sought to correct for the effect of atmospheric attenuation in NDVI signal, by adding data from a 'blue light' data band (.46 - .48 $\mu$m). EVI has been shown as a powerful alternative to NDVI when atmospheric correction is necessary (W. J. D. van Leeuwen, Huete, & Laing, 1999). Unfortunately, due to the relatively high energy level of blue light, far less total light can be captured in this portion of the spectrum, the result being that RS platforms generally gather blue band data at a relatively low

spatial resolution. This is the case for MODIS, and will therefore be an important constraint in the choice of dependent variable for this study.

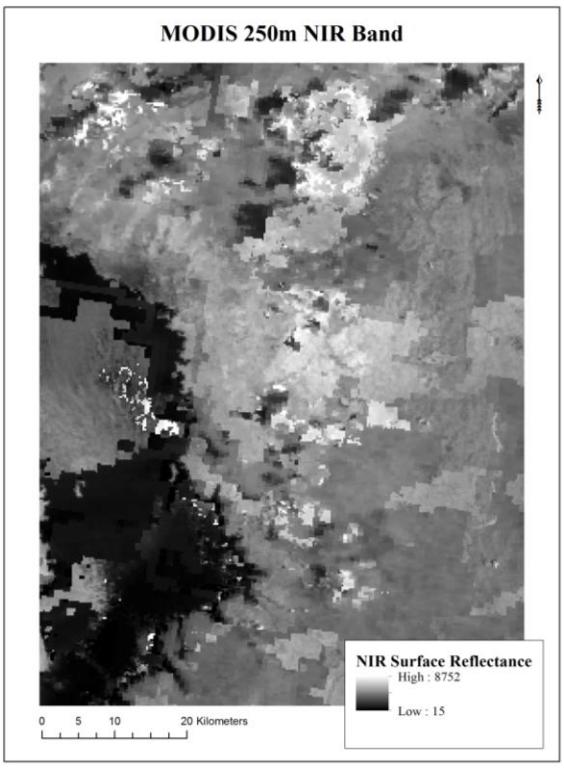


**Figure 1.1.** MODIS 250m NIR band image for the Kigali region of Western Rwanda

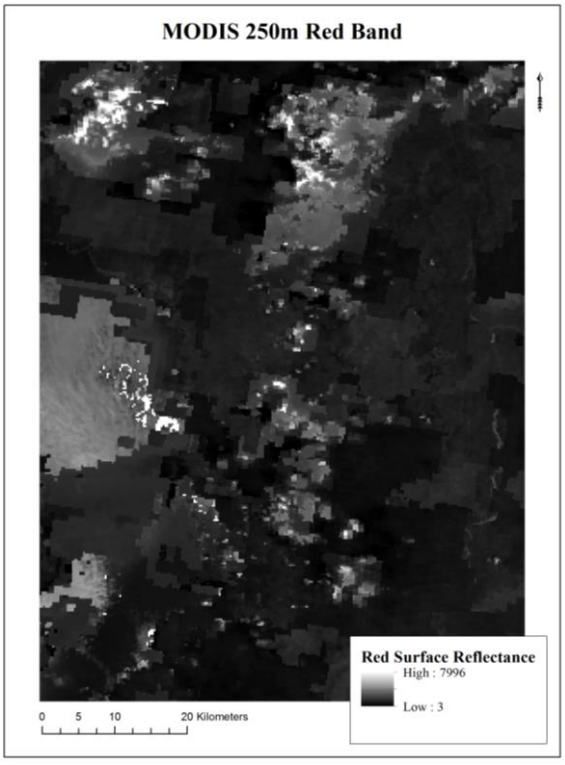


**Figure 1.2.** MODIS 250m red band image of the Kigali.

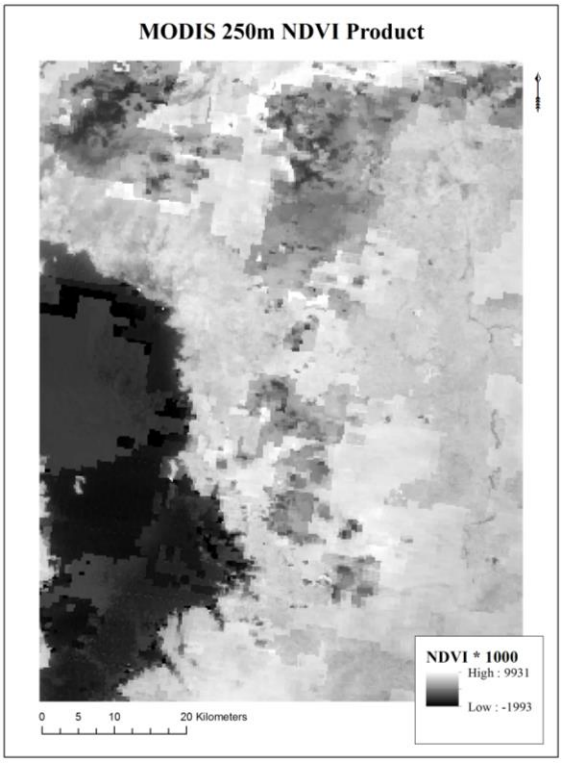


**Figure 1.3.** MODIS 250m NDVI image, generated by ratio transform of the data shown in Figure 1.2 & Figure 1.3.

### 1.4 Noisy Data

In all statistical experiments, analysts use mathematic tools to aggregate and interpret collected data. The common characteristic of these tools is that they attempt to distinguish meaningful information from stochastic or conflating trends in data. In the case of RS, a variety of erroneous signals can be expected and must be accounted for in a nuanced statistical analysis. Common sources of error introduced into RS data include atmospheric attenuation and satellite aperture design. These erroneous signals make up the *noise* component of RS data. To accurately interpret RS data and make valuable inferences we must first use appropriate statistical tools to account for this noise component.

The concept of noise in data suggests that in a group of sampled data, there is a component of 'true' signal and a component of 'noise', erroneous deviations from that true signal. Methods for investigating signal noise range from simple assumptions of perfectly random errors, to autoregressive moving average (ARMA) models, that attempt to model the underlying patterns driving the noise component in the data (Greene, 2003a). Depending on the dataset and research objectives, noise may be a nuisance to contend with, or it may be the critical variable of interest in the study. Therefore, when examining or correcting for signal noise, great care must be taken to choose appropriate methods and to diagnose the efficacy of those chosen methods.

Noise is introduced to RS images through a diverse variety of processes, physical (obscuring objects or phenomena) and digital (issues with an instrument's ability to capture the true signal). Physical processes introducing noise to RS data include atmospheric attenuation, irregular reflectance patterns, cloud cover, shadows, ice and more. Digital noise includes artifacts of an aperture's design, washout from brightness, and if present, is often systemic. If digital noise is present in imagery, care must be taken to evaluate the severity, diagnose, and potentially correct erroneous datas. Most RS image archives provide a mask of values indicating potential problems, and/or an overall metric for the quality of each pixel-observation. These masks are useful, but they do not eliminate noise in the data. More importantly, a simple application of this type of mask throws out observations of poor quality, without replacement, creating a discrete dataset, with irregular spatial and temporal dimensions. Analysis of a discrete spatial time-series is far more complex and computationally intensive than datasets where continuity can be assumed.

Another important, but conceptually separate, source of noise in geographic data comes as a result of the *modifiable areal unit problem* (MAUP). Fotheringham and Wong (1991) first

demonstrated quantitative evidence for an MAUP through multivariate linear and logit regressions. They showed that choices in the level of spatial aggregation (a similar concept to spatial resolution), could introduce a '… rather depressing…' (1004) degree of inaccuracy in analyses. They concluded that a key element to any successful Geographic Information System (GIS) is the adoption of an appropriate scale (the suitability of a GIS' scale can be evaluated through simple spatial analysis). With the knowledge that spatial aggregation of data introduces error, we should be aware of the inborn inaccuracies of satellite data, introduced by uniform resolution. One pixel of satellite data can never capture the spatial distribution of light within that pixel, and this problem grows as resolution decreases. Therefore, the MAUP is a crucial component of choosing the appropriate source of RS data for a study.

While airborne sensors (sensors mounted beneath airplanes) may produce extremely high resolution images, even a pixel of this high resolution data includes a mixture of light, reflected from various objects within the pixel. High resolution imagery also has several important drawbacks. First, the higher the spatial resolution of airborne or satellite imagery, the lower the temporal resolution, due simply to constraints on the size of the aperture, flight height, and thusly the platform's field of view (holding the number of platforms constant). Second, high resolution imagery has much less total luminescence, or brightness, per pixel, diminishing the accuracy of observations as spatial resolution increases. This is of particular concern in multispectral imagery, where the total luminescence is divided irregularly across discrete bands. This phenomenon is why the highest resolution data available from any multispectral sensor tends to be given as a panchromatic or *pan-sharpened* image. Due to tradeoffs between spatial, temporal and spectral resolution, it is often best to choose a lower spatial resolution sensor in order to capture more frequent observations with a reliable spectral signature.

Noise in RS data can be controlled for and modeled with spatial and temporal methods. Common methods for smoothing noise from data include spatial, time and frequency-domain smoothing techniques. Choosing an appropriate smoothing methodology for an RS dataset is dependent on the resolution of the data, the features being examined and whether the analysis is multimodal. Some cutting edge methods for smoothing RS data include Fourier analysis and wavelet decomposition, which transform data to the frequency domain to be clustered and filtered based on sinusoidal behaviors. This class of methods will be discussed further in Chapter 4.

# Chapter 2: Background on Rwanda

## 2.1 Motivation

Land cover (LC) classification of RS data is a common technique for studying the natural environment (Hamlyn & Vaughn, 2010a), as well as human-environment interaction (Agarwal et al., 2002; McGranahan, Balk, & Anderson, 2007b; Turner, 2003; Wasige, Groen, Smaling, & Jetten, 2013). These studies static and dynamic approaches to relate land cover characteristics to human activity. In this thesis, I attempt to demonstrate the informational capacity of medium spatial resolution satellite imagery to build models of land- cover change in Western Rwanda, a difficult region for remote sensors due to perennial cloud cover. Medium resolution data saves on computational complexity and memory requirements, and may still be rich in information when taken in a time-series framework.

Is it possible to study human ecology in extremely cloudy regions, by leveraging temporal resolution over spatial resolution? The motivation for this research comes from the desire to make social inference from RS data. There is a call throughout the social sciences for more comprehensive, and objective sources of data. As discussed, RS will be an important source of such data as improved apertures, satellite coverage and statistical/machine learning tools simplify experimental design and inference for scientists. With this work, I intend to demonstrate the potential for medium-resolution, multispectral, time-series RS data to be applied to the study of human-ecology in tropical-montane regions. Tropical-montane regions are an important general case for remote sensors due to the intersection of ecological, atmospheric and human processes present in these regions.

An appropriate pilot for the study of tropical-montane human ecology exists in Western Rwanda. For remote sensors, Western Rwanda is plagued with difficulties related to climate and topography. Additionally, it presents a plethora of opportunities for the examination of the effects of conflict, resettlement and various national and international policies on the environment. Rwanda has experienced some of the most brutal ethnic conflict in human history, in a period that can be studied, indirectly, with a variety of data from RS apertures. This section will discuss the conflict, resettlement and policy in Western Rwanda, as well as how these may be observable in RS data.

**2.2 Genocide in Rwanda**

The Rwandan genocide and civil war of 1994 were the result of a complex and precarious combination of socio-political institutions. Long standing ethnic and tribal conflict, greatly exacerbated by German and Belgian colonial rule, led to one of the fastest and most brutal mass murders in human history. When the violence ended, '…at least 800,000…' (2), or 10% of the Rwandan population, had perished. The death toll does not begin to capture the entirety of this tragedy, because when the violence ended, for those who lived, the struggle to carry on was only beginning (Akresh, Walque, & Damien, 2008). This Genocide deserves far more attention than can be devoted in this thesis, but I will attempt to summarize the key events leading up to the resettlement of refugees.

The Genocide in Rwanda is often referred to interchangeably as the Rwandan Civil War. This second title may more precisely describe the conflict, as it was not only an ethnic cleansing, but also part of a long standing struggle over power in Rwanda. While tribal and ethnic tensions were part of the conflict, these tensions were inflamed by a short, fraught, colonial rule. During a period of Belgian governance, from 1945 to 1961, the three main ethnic groups were given

identification cards with *Hutu*, *Twa*, or *Tutsi* clearly labeled. While Hutus were a clear majority in the region, power laid with the minority Tutsis. While Belgium attempted to quell the rising violence in Rwanda by splitting it into the nations of Rwanda and Burundi, this did not end the conflict. Rwanda became independent in 1961, but attacks from Burundi and frequent civil skirmishes within Rwanda continued until the tragedy that unfolded in 1994. Despite the long standing history of conflict in Rwanda, the Genocide of 1994 seemed to begin with an incredible abruptness to many in the international community (Akresh et al., 2008).

The Rwandan episode was particularly unique in its speed and efficacy. The period of violence lasted roughly 100 days and has been noted as the most efficient slaughter of human life since the bombing of Hiroshima. What is more astounding, however, is the speed with which Rwandans began to reconstruct their society and institutions. The largest crash crop of Rwanda in the mid 90's was coffee, which declined in production by 40% during the conflict, but returned to pre-war levels by the end of 1995 (see Figure 2.2). In this time an interim government was formed, and reconciliatory policy almost immediately enacted (Akresh et al., 2008).

While the Rwandan Genocide ended as abruptly as it seemed to begin, resettlement, reconciliation and rehabilitation would take much more time. In fact, government sponsored reconciliation events and projects continue to this day ("Outreach Programme on the Rwanda Genocide and the United Nations," 2014). The bulk of resettlement occurred over a period of roughly 15 years. Social, political and economic institutions in Rwanda were decimated not only by genocide, but by the massive reshuffling of power and geography within the nation. One such observable impact has been the conflict's impact on education. Despite an inpouring of international aid, and massive improvements in education policy from 1992 – 2000, research has

still shown an surfeit of evidence that those exposed to the violence had lasting reductions in educational achievement as of 2006 (Akresh et al., 2008). Despite Rwanda's resilience, there is robust evidence of a long period of institutional rehabilitation in the nation.

### 2.3 Imidugudu Policies

After the Rwandan Genocide and Civil war, the size of the refugee population was unclear to the Rwandan authorities and to the international community. Rwanda has the highest population density in Sub-Saharan Africa, with a population of over 12 million in an area (26,338 $km^2$) roughly the size of West Virginia (The World Bank, 2014). The settlement pattern of Rwanda is similar to that of the United States, having a relatively high degree of dispersion. Prior to the Genocide, the vast majority of Rwandans were small land holder, subsistence farmers. Farming was a year-round activity, with 'rainy season' and 'dry season' crops. Crops in the region are stratified mostly by their suitability to given elevations, with more water-tolerant crops such as rice being grown in the marshy lowlands and wheat in the dryer, high altitudinal regions (see Figure 2.1). Prior to the genocide, most rural Rwandans were subsistence farmers, who farmed a wide variety of crops in a tiered system, leveraging the different altitudinal zones. However, after the Genocide, the new governing body saw fit to push the country towards centralized living, large-holder land shares and cash crop farming through the 'Imidugudu' redevelopment plan (W. J. D. van Leeuwen et al., 1999), which has likely also had a significant impact on the dynamics of agriculture in Rwanda.

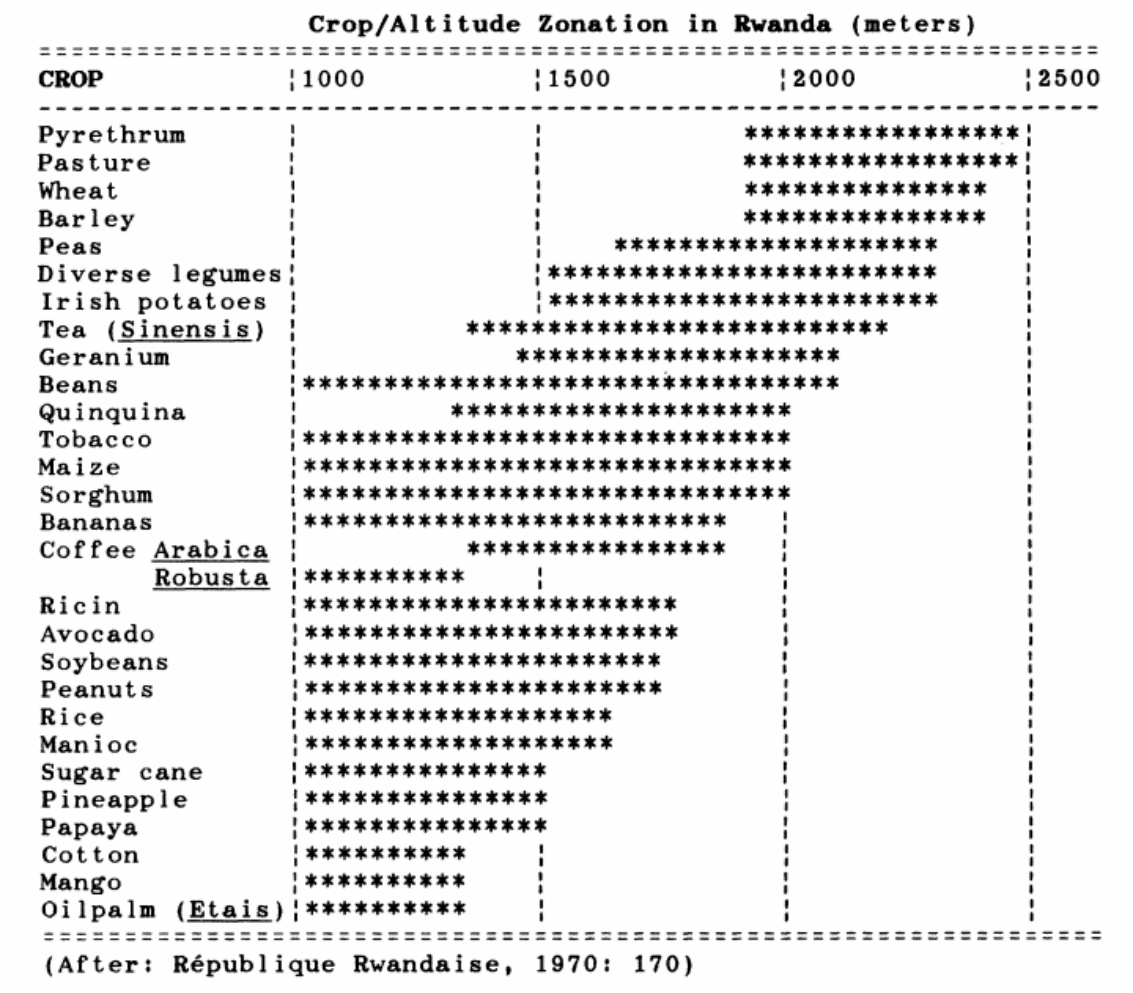


**Figure 2.1.** Crops and altitudinal zonation in Rwanda. Reprinted from *The Dynamics of Human-Environment Interactions in the Tropical Montane Agrosystems of Rwanda* (pg. 45) by R.E. Ford, 1990, JSTOR. Reprinted with permission.

The Imidugudu policies have been met with a mixture of skepticism and mistrust, but thus far seem to have been successful in redeveloping and stabilizing the rural areas of Rwanda (M. van Leeuwen, 2001). The Imidugudu policies were criticized for two main reasons: their similarity to other failed efforts at villagization in other nations in East Africa, and the possibility that greater centralization of power in Rwanda (here, with regards to centralization of agricultural policy) could lead to another deterioration of ethnic relations and a return to authoritarianism.

Three major elements of the Imidugudu policies are likely to have had lasting impacts on the shape of agriculture in Rwanda. The first is the family of aforementioned villagization and resettlement policies. These policies moved hundreds of thousands of farmers to new land, often at drastically different elevations, with different soil conditions, rainfall patterns, etc. Villagization was designed, in part, to incentivize cooperative farming, a major adjustment from farmers' previous experiences with family based, subsistence farming (M. van Leeuwen, 2001). Figure 2.2 shows the total and rural populations of Rwanda from 1990 to 2013, and demonstrates the impacts of villagization policies. Rwanda's population has rebounded, but the vast majority of that growth has been in urban centers and drawn in part from rural areas. Farmers' adaptations to new environments and increasing centralization have likely led to a slow but distinct shift in the agricultural dynamics of Rwanda.

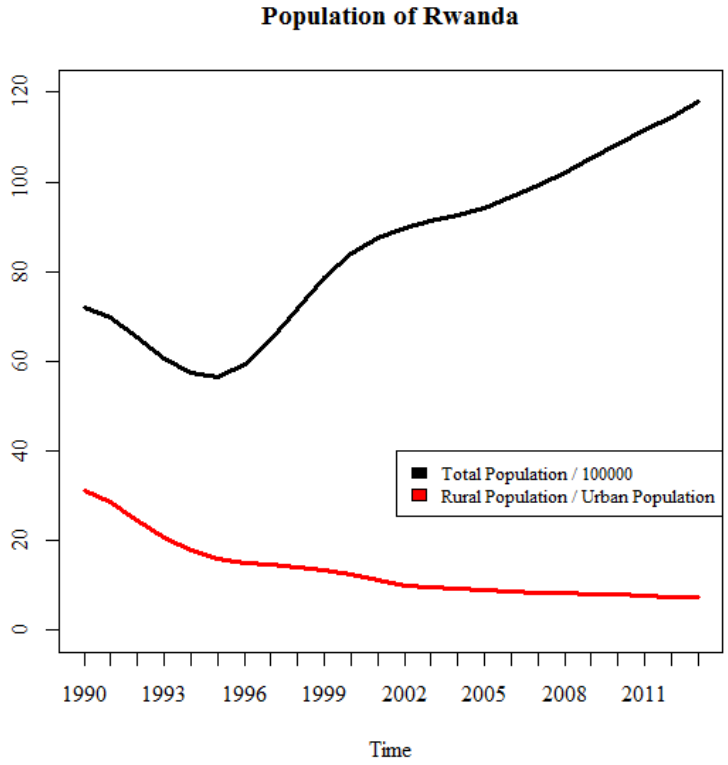


**Figure 2.2.** Total population and rural population in Rwanda (The World Bank, 2014).

The second major element driving change is the direct influence of new agricultural policies. While the Rwandan government wished to improve food security, following the conflict, they were also motivated to reduce the portion of society working directly in agriculture, particularly in subsistence farms. The new authorities responded to scholars' claims that the incongruity of a lack of opportunity outside of agriculture, and a simultaneous lack of opportunity to purchase land for agriculture, was a prominent piece of the conflict of the 1990's. Prior to the Imidugudu policies there was no systematic means of purchasing land. To combat food insecurity while diminishing reliance on subsistence farms, the government incentivized large-holder cropping, the planting of cash crops and streamlined the process of buying and selling land (Havugimana, 2009).

The third major element driving change in Rwanda's agriculture has simply been a rapid increase in population. During the Civil War, the Rwandan population dropped precipitously, mostly due to refugee migration. In the period following the Civil War, many of these refugees returned, and the replacement rate continued to grow Rwanda's population at greater than pre-war levels. The black line in Figure 2.2 shows the rapid growth of Rwanda's total population and illustrates the need for greater agricultural yields. The need to feed Rwanda's increasing population, and it's already fully utilized land, is likely driving a shift towards more efficient agricultural practices.

### 2.4 Mapping Imidugudu Settlements

To analyze LC change in Western Rwanda, we must first map the locations of Imidugudu settlements as well as older remaining villages. Locating the Imidugudu refugee camps of Western Rwanda from RS imagery is relatively straightforward, as the new structures all have the same white-painted roofs. This is likely a product of the international aid influence on

rebuilding in the region. An example of an aerial view of a typical refugee camp, with white roofs clearly visible, is shown in the leftmost panel of Figure 2.3. Google Earth imagery of Western Rwanda from 2009 was analyzed and 11 refugee camps and villages were outlined and exported as shapefiles. The camps and villages come from a geographically disperse sample of the Western Rwandan highlands, and each camp and village is paired based on geographic proximity.


**Figure 2.3.** Google Earth images of the Kigali Refugee Camp, one of several camps in this study, and the surrounding ROI. Google Earth. 2014.

For this study, the key to observing the impact of the Imidugudu settlements is to take a DD approach. In the DD framework, we subtract the observed signal of a 'control' group, in this case older villages, not built in the Imidugudu program, from the signal observed in Imidugudu villages, or the 'treatment' group. Figure 2.4 shows the spatial polygons representing 11 pairs of camps and villages, with the red lines denoting the choice of the pairing step. The village, or control group is subtracted from the camp, or treatment group. This is a simple means of accounting for variability caused by conflating factors, such as climatic patterns, that should be present in the Imidugudu villages as well as their close neighbors.

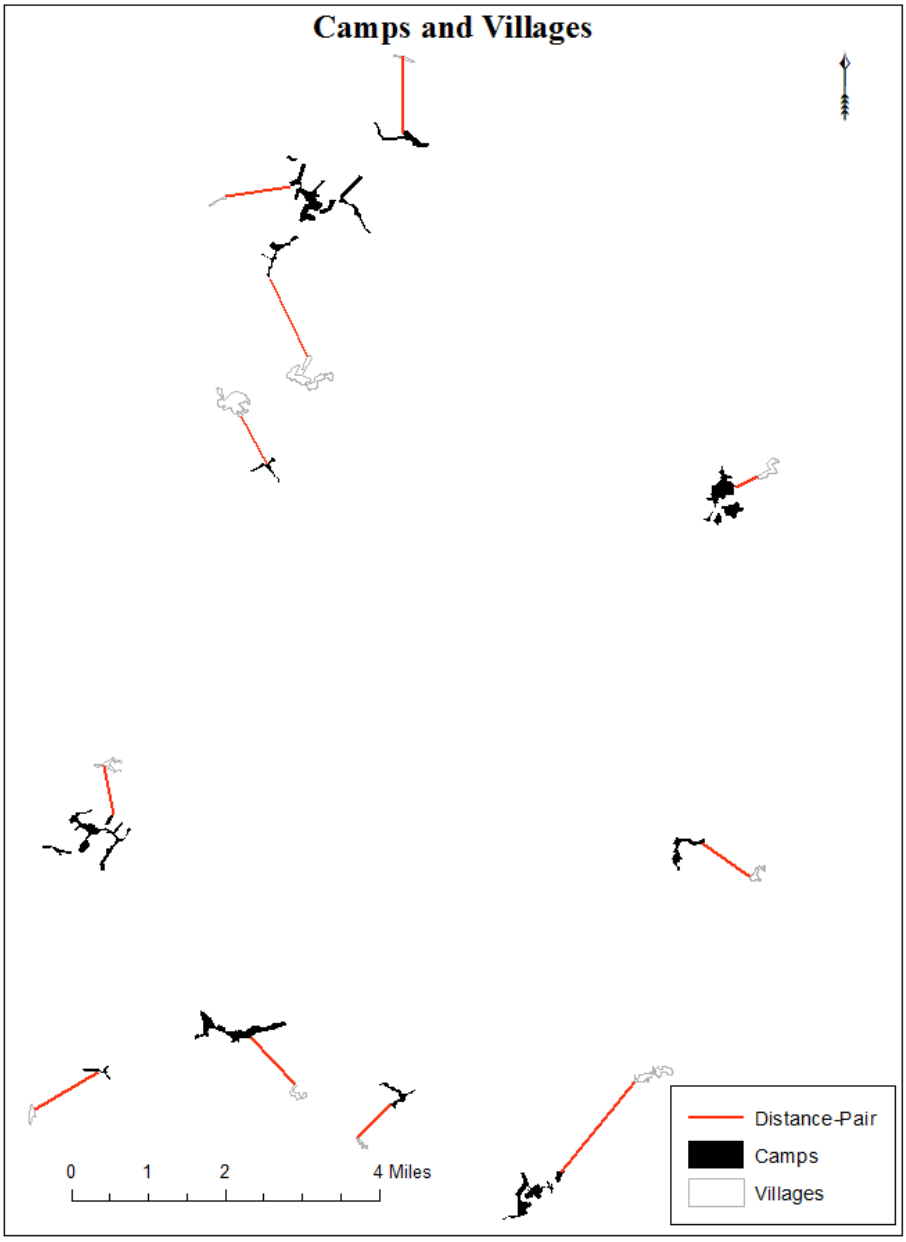


**Figure 2.4.** Camp and village shapefiles for 22 sites, as gathered from Google Earth imagery.

In a more statistical language, we would say that DD preserves variance at the expense of introducing bias (i.e. we sacrifice precision for accuracy). To mitigate introduced bias, a highly systematic approach is taken to the generation of training data. By taking equal sized buffers from each camp and village and pairing according to geographic proximity, the study design attempts minimizes bias, not having ground truth training data.

# Chapter 3: Tropical-Montane RS

## 3.1 Broad-scale RS

When studying a relatively large region, such as the *Kigali*, or Western Rwanda, for most projects it is too computationally expensive to use high spatial resolution data. As spatial resolution increases, everything from simple data cleaning processes to matrix operations grow exponentially in computational space (Yin et al., 2000). RS studies often have very small input datasets but massive amounts of intermediate and output data that can become extremely memory intensive (e.g. gray level co-occurrence matrices, computed for every raster in a high-resolution study, and stored in active memory would promptly outstrip a personal machine's capacity). While it is feasible to complete a study of the Kigali in high resolution data, the computations would likely need to be performed on an extremely powerful machine or run for weeks on end on a typical research machine (certain individual functions written for this thesis required 3 hours to complete on a medium-to-low resolution dataset). Broad-scale RS presents an inborn challenge of memory consumption, that can be addressed either by using lower resolution datasets, or with enormous computational power.

## 3.2 Tropical-Montane Noise

Tropical-montane regions produce particularly noisy observations in multispectral RS data, especially in the NDVI, which has no inborn correction for atmospheric conditions. The tropics are especially obfuscating, due to persistent cloud cover. The exact *rho*, or fraction of total luminescence reflected by cloud cover, is dependent on the density of the cloud. For this reason, it is sometimes possible to have intermediate quality observations, where a correcting factor can be derived from blue band data. This RS product is known as the enhanced vegetation index

(EVI). Raw EVI data is more accurate than NDVI data (Galford et al., 2008), when taking the raw time series. In our case, blue band data is not available at 250 m resolution and so NDVI will be used rather than EVI in a compromise to reduce spectral error introduced by partial atmospheric attenuation with superior spatial resolution (reducing spatial error).

Mountainous regions are of themselves more difficult to study with RS data. In landscapes where there are dramatic changes of slope, in magnitude and sign, errors may be introduced into the data through pixel mixing. Pixel mixing is less of a concern, however, in medium resolution data, where a single pixel contains an array of topographic angles, land cover types, etc. and methods of analysis are appropriately flexible. For instance, if classifying pixels, *fuzzy classification*, where pixel membership can be split, would likely improve classification accuracy. Mountains also cast shadows, which can cause erroneous NDVI observations, though this is also less prominent in medium-resolution data. In sum, mountainous topography adds to the challenge of RS studies in tropical-montane regions, but much less so than cloud cover, particularly in medium-resolution data.

### 3.3 The Moderate Resolution Imaging Spectroradiometer

All of the aforementioned factors, regarding RS principles and the challenges of tropical-montane remote sensing, motivated the decision to use data from NASA's Moderate Resolution Imaging Spectrometer (MODIS). MODIS has several advantages over other RS platforms when studying equatorial regions. First, light in the visible range of the electromagnetic spectrum is reflected by clouds and attenuated by aerosols (Kaufman et al., 1997), making regular remote sensing of the tropics a challenging proposition. MODIS overcomes this challenge simply, by gathering images twice daily and averaging the clear pixels collected over days or weeks, creating a clear, regular output. Our region of interest, Rwanda, experiences two seasons of

heavy rain (c. October and March), but in fact experiences some cloud cover nearly year round. Kigali, the Rwandan capital east of our study area, experiences highs of 100% relative humidity nearly every day outside of summer (see figures 3.1 & 3.2).

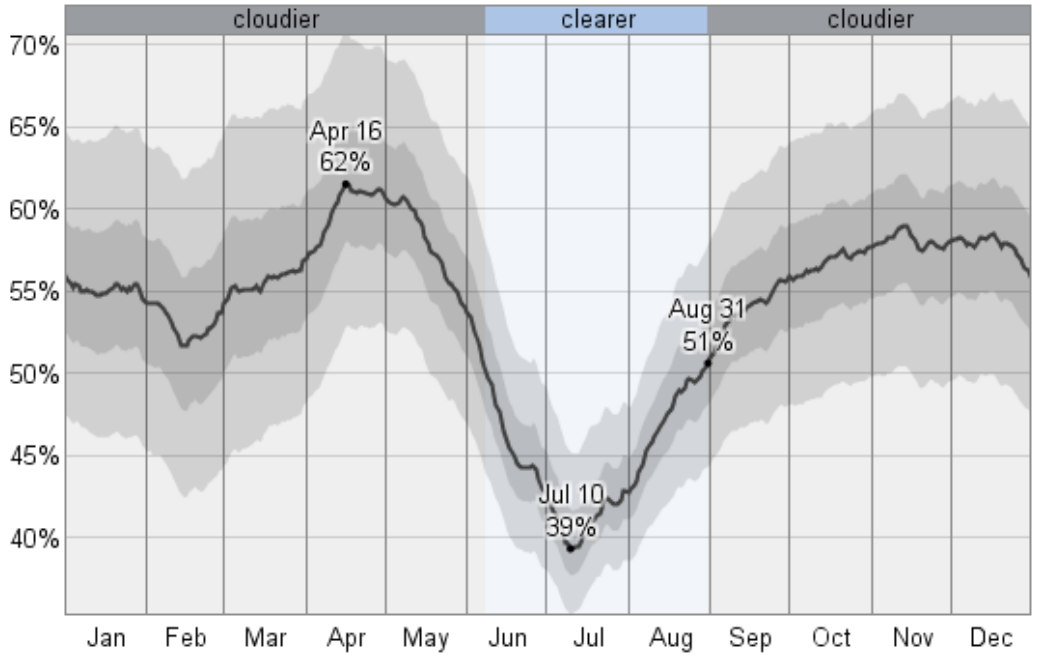


**Figure 3.1.** Median cloud cover for Kigali, Rwanda; outer band represents 1st and 3rd Quartiles, inner bands represent 10% above and below median daily cover (Figure published by Weatherspark).

The second advantage of MODIS is its ability to take reliable imagery in topographically irregular environments. Mountainous terrain can often obscure the interpretation of raw digital numbers from satellite instruments. Mountain topography can alter angles of incidence and reflection for light being reflected to the sensor's aperture, create broad shadows obscuring the landscape, and usually lead to complex topologies of interspersed human activity as well as high ecosystem variability across elevations. For all of these reasons attaining reliable remotely sensed imagery of tropical montane regions has remained a hurdle for researchers using broad scale environmental phenomena. MODIS collects imagery at a resolution more appropriate for studying broad scale phenomena in such a complex topography. Since MODIS employs several

methods, such as a 500m resolution blue band and several measures of aerosol content to control for these climatological limitations, it is a logical data source for this project, focusing on tropical-montane regions.

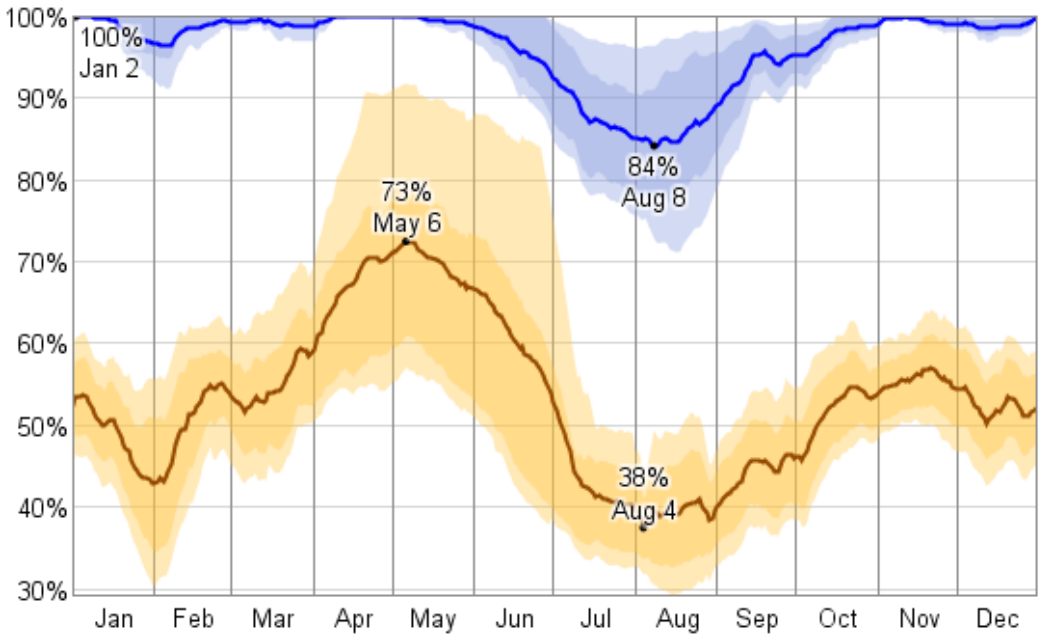


**Figure 3.2.** Average high (blue) and low (orange) relative humidity; Shows both the persistent daily humidity of Rwanda, but also the high degree of daily variability in conditions (i.e. in August where highs reach nearly regularly reach 90% humidity and lows 35%). Bands represent 1st and 3rd Quartiles and 10 and 90% boundaries (Figure published by Weatherspark).

Corrective strategies have been developed to improve data fidelity in these mountainous regions. One common method for topographic correction (correcting for errors driven by off-nadir viewing angles) is the bidirectional reflectance and distribution function (BRDF). BRDF models are applied to time-series imagery to rectify off-nadir viewing angles (Roujean, Leroy, & Deschamps, 1992). MODIS provides a BRDF adjusted product ('MCD43A4'). However, in the tropics, MODIS is often unable to obtain sufficient 'looks,' or repeated observations of a pixel region, without cloud cover. As such, BRDF imagery in this area of interest is regularly missing vast swaths of data. Figure 3.3 shows a typical coverage pattern in our area of interest.

Reflectance can rarely be calculated in areas of high forestation and mixed topography, even in 16 day intervals.

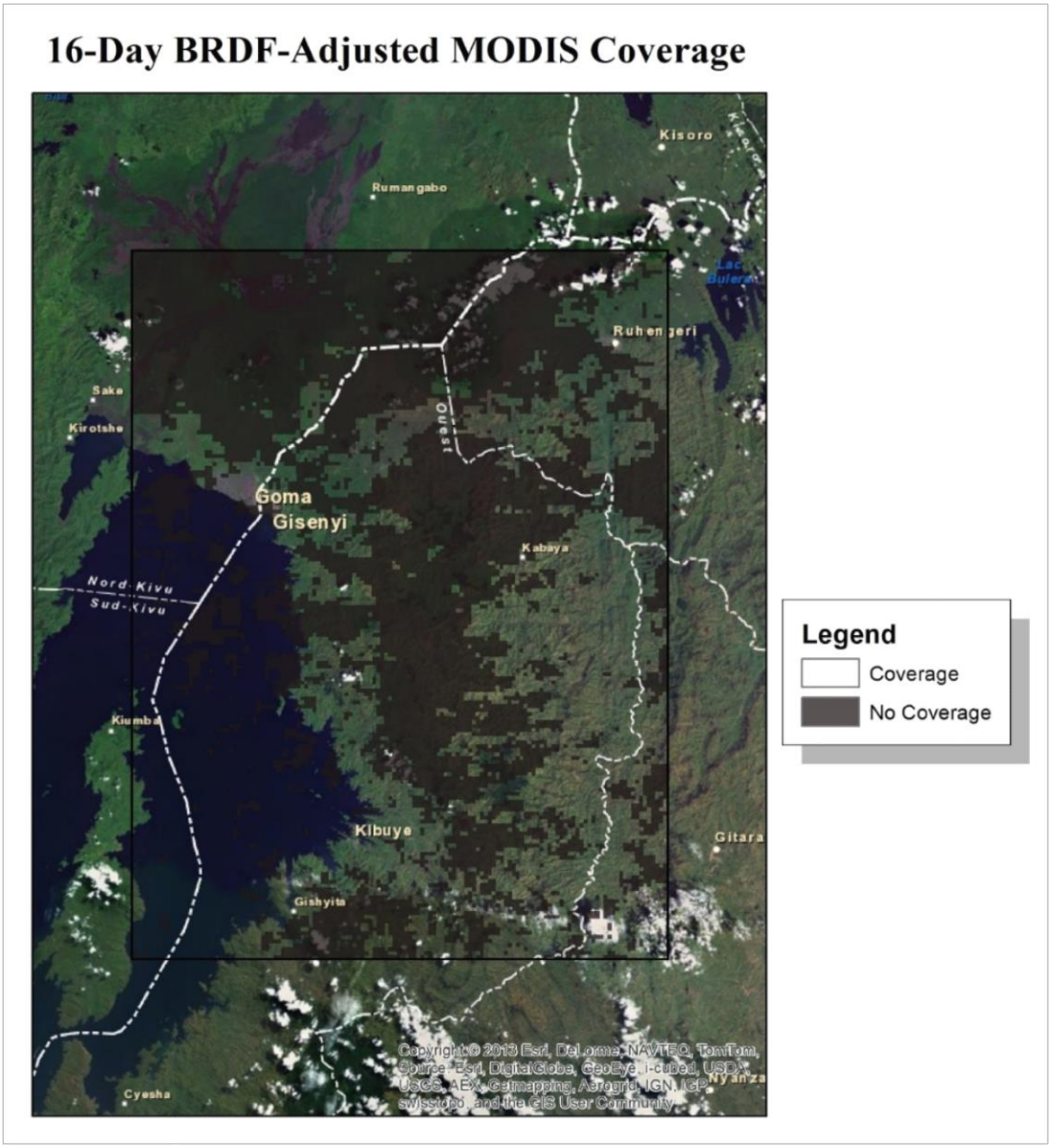


**Figure 3.3.** A typical example of 16-day, 500m MODIS BRDF-adjusted product in Northeast Rwanda (taken 03/12/2013); red regions are pixels for which reflectance could not be computed.

# Chapter 4: Time-Series Analysis

## 4.1 Background

Time-series analysis is a powerful tool for examining multimodal RS imagery; however, frequency-domain analyses are promising new methods for extracting information from multimodal data. Frequency-domain methods are a broad suite of methods adapted from the signal processing and control systems literatures (Bremaud, 2002). Analyses in the frequency-domain examine a series of data points as a function of frequency, amplitude and phase, rather than as a function of time (i.e. we examine how much of the total signal, or a window of the total signal, falls within certain frequencies).

An important assumption for this study is the relationship between long-memory and rapid changes, expressed in the frequency-domain. Long-memory trends, or relatively smooth, deterministic trends in the time-domain, are expressed as extremely low-frequency waves in the frequency-domain. Extremely short-memory trends, such as the expected stochastic noise in the signal described in Section 1.4, should appear as very high-frequency waves in the frequency-domain (Bremaud, 2002). These relative differences in the expression of trends in the frequency-domain allow for the intelligent analysis of noisy time-series data, such as the raw NDVI used in this research.

## 4.1 Frequency-Domain Analysis

Given the known relationships between signal/noise and low/high frequency waves, analysis and filtering in the frequency domain should be an effective means for analyzing seasonal vegetation data. Frequency-domain analysis has classically been used to create filters, such as the high-band-pass filter, for signal processing. A low-band-pass filter is generally used to select and

attenuate high frequency data-points. The purpose of this is to remove noise from a signal, or as in the case of audio systems, redirect the high frequency component of a signal to a speaker designed for higher notes. This type of filtering can in fact be done within any band of frequencies through combination of a high-pass filter. Frequency-domain methods such as the low-band-pass filter can be applied to RS in data in the spatial or temporal dimension. In a spatial application a high-pass filter would pick out features such as edges, where sudden changes occur, whereas a low-pass filter smooths the data, as it picks up low-frequency trends (see Figure 4.1). The same is true in the temporal dimension, but rather than picking out edges, the filter would pick out rapid changes in the value of a single pixel.

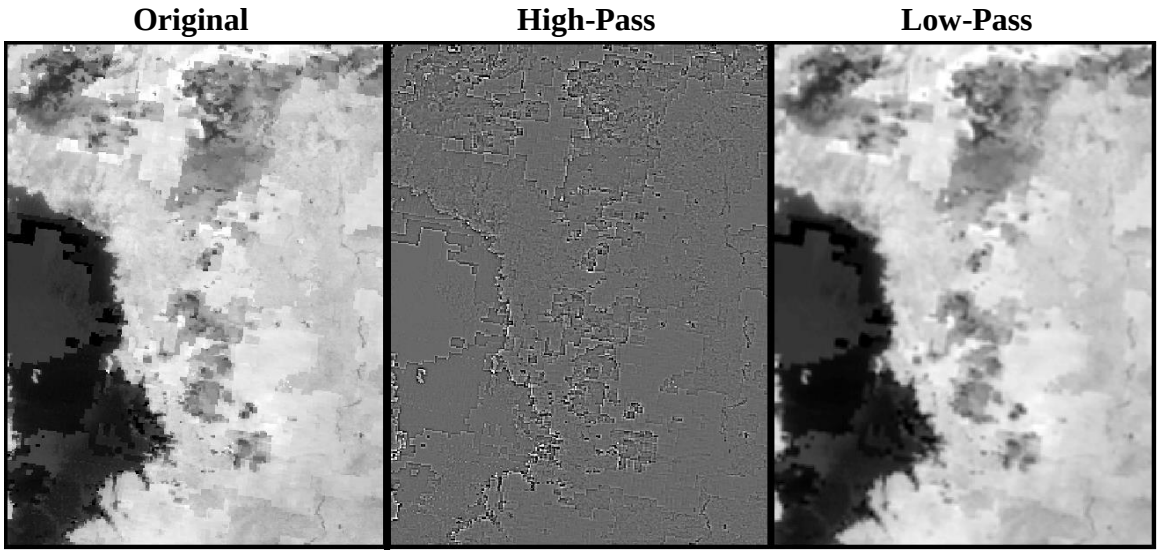


**Figure 4.1.** Example of a high and low-pass 3x3 filter applied to a MODIS NDVI image in the spatial dimension.

### 4.1 Fast Fourier Transform

In this thesis I apply frequency-domain analysis by filtering MODIS images in the temporal dimension with the *Fast Fourier Transform* (FFT). FFT is simply an algorithm that computes the *discrete Fourier transform* (DFT) of a time-series, a special case of the *continuous*

*Fourier Transform*. FFT is particularly useful in that it can be simply reapplied to find the *inverse Fourier transform* (IFT), in order to return data from the frequency-domain to the time-domain. The continuous Fourier transform can expressed as follows ( "Discrete Fourier Transform," 2014):

*Equation 2)* $$f(v) = \int_{-\infty}^{\infty} F(t)e^{-2\pi ivt}dt$$

If we now consider an approximation, where F(t) becomes $F(t_k)$ by allowing $F_k \equiv F(t_k)$, where $t_k \equiv k\Delta$, when k = 0, …, N-1. This will give us the DFT:

*Equation 3)* $$f(n) = \int_{k=0}^{N-1} dF(k)e^{-2\pi ikn/N}$$

The inverse of the DFT, the IFT is as follows:

*Equation 4)* $$f(k) = \frac{1}{N}\int_{n=0}^{N-1} dF(n)e^{2\pi ikn}$$

Frequency domain analysis has been shown to be a powerful method of improving data quality in RS studies (Andres, Salas, & Skole, 1994; Galford et al., 2008; Lunetta, Knight, Ediriwickrema, Lyon, & Worthy, 2006; Sakamoto et al., 2005). Frequency smoothing can occur

in the spatial or temporal domain, depending on the resolution and the intended use of the data, and covers a broad category of methods. Galford et al. (2008) used Morlet wavelets to smooth crop phenology signals, while Lunetta et al. (2006) used a simple FFT-based method that will serve as the basis for my analysis. The common thread in all of these studies is the utilization of a time-series' frequency-domain information to improve the quality of the data being analyzed. I will utilize this same approach, to be described in Section 5.3.

# Chapter 5: Case Study

## 5.1 Objectives

The primary objective of this research is quantitative, but the motivation for the objective is qualitative. The primary target is to test the applicability of frequency-domain and time-frequency domain smoothing to medium resolution RS studies. The motivation for this objective is to test the presence of a long-memory trend in the vegetation signals of resettled refugee villages of Rwanda compared with older, established villages. It is likely that due to noise in the raw NDVI signal, a smoothing algorithm using the FFT will improve the ratio of 'signal' to 'noise,' and provide a more robust time-series for analysis. By improving data-quality, the primary objective of this research, I hope to make possible inference relevant to human ecology in tropical-montane regions. I will follow a similar filtering process to Lunetta, Knight, Ediriwickrema, Lyon, and Worthy (2006) and visually inspect the data for outliers to assess Hypothesis 1:

*$H1_o$: The original 16-day composite, 250m NDVI images will provide a more robust time-series for further analysis.*

*$H1_a$: FFT-filtered, 16-day composite, 250m NDVI images will provide a more robust time-series for further analysis.*

The second objective, to analyze the time-series for the presence of a long-memory trend is a vastly more difficult test to pin down. The first simple, time-series test would be to check the difference between paired villages and refugee camps for linear stationarity using the Augmented Dickey-Fuller (ADF) test (Greene, 2003). The ADF tests the assumption that the time series is a non-mean-stationary process with a probability distribution that significantly changes when lagged in the time dimension, i.e. contains a *unit root*. In this case, the ADFs' test statistics were

highly significant for time-series of mean NDVI values for all villages, camps and the differences in the two series, as shown in Table 5.1. Taking these tests alone, we would reject the assumption that there is a unit root, or a long-term, deterministic trend in the data, and conclude that the vegetation signals of the camps, villages, and the differences between the two are all mean-stationary processes.

**Table 5.1.** ADF test for stationarity. The ADF function in R chooses an appropriate lag based on a typical ARMA(p,q) model. All of the ADF's showed significant evidence to accept that the time series are stationary.

| **Augmented Dickey-Fuller Test of Mean Time-Series** | | | |
|---|---|---|---|
| | Lag Order | Test Statistic | p-Value |
| Villages | 6 | -8.568358 | $p < .01$ |
| Camps | 6 | -8.298578 | $p < .01$ |
| (Camps – Villages) | 6 | -8.395433 | $p < .01$ |

The significance of the ADF for the differenced could demonstrate that there is no low frequency component to the data. However, I would posit that a more nuanced, seasonal approach to this study, combined with frequency-domain filtering will show significant evidence to the contrary. Therefore, I would posit:

*$H2_o$: There is no statistically significant long memory/low frequency trend in the differenced camp and village data.*

*$H2_a$: There is a statistically significant long memory/low frequency trend in the difference camp and village data.*

### 5.2 Data & Methods

The first key method to this thesis was the generation of paired training data with the locations of Imidugudu 'camps' as well as older, established villages, for the DD approach described earlier. DD is a simple means of reducing the potential for conflating variables that

might cause biased results. The camps and villages were identified by their architecture and settlement patterns, and the polygons generated are shown in Figure 2.4. The camps and villages have similar buffer sizes, and are paired on geographic proximity to further avoid biased results.

Google Earth exports polygon features in the *keyhole markup language* (KML) format. I first imported the KML files to ArcGIS 10.2 to convert them to a more useful shapefile format, in a projection consistent with that chosen for the MODIS imagery ('WGS 1984 UTM Zone 35'). I then generated 2500 m buffers (roughly 1.5 miles), which translates to a one half hour of walking at an average pace of 3 mph, around the centroids of each camp and village. I experimented with 5000 m, 2500 m, and 1000 m buffers, and found little variation in outcome through buffer-size modification. Initially, the most encompassing (largest) buffer seemed most appropriate, however, visual inspection revealed that, due to the proximity of settlements in Rwanda, the paired buffers overlapped by nearly 50%. Therefore, I chose the middle-ground, 2500 m buffer. Buffers were created around the centroids, rather than the original polygons, because agriculture is often present in the settled areas. Villages and camps were paired based on geographic proximity, and so many of the camp and village buffers overlapped. This was dealt with using a maximum likelihood (euclidean distance) classifier, which assigned pixels to the nearest camp or village centroid. These buffered and divided regions were used as the final sampling areas from which to gather and analyze MODIS data.

Two layers were taken from the MODIS DAAC: a derived NDVI layer and a summary layer of 'pixel reliability'. Both of these features are available from the MOD13Q1 product, which is used to distribute 250m, 16-day composite imagery of MODIS's scientific band data for vegetation indices (i.e red, near-infrared, mid-infrared and blue bands). NDVI is a simple computation as described in Equation 1. The MODIS pixel reliability mask in Figure 5.1 is a

simple, 2-bit summary of quality indicators, given at the pixel level. I retain 'good' and 'acceptable' pixels, but filtered (described in methods section) all other values. In the case of Western Rwanda, this is nearly always due to cloud cover.

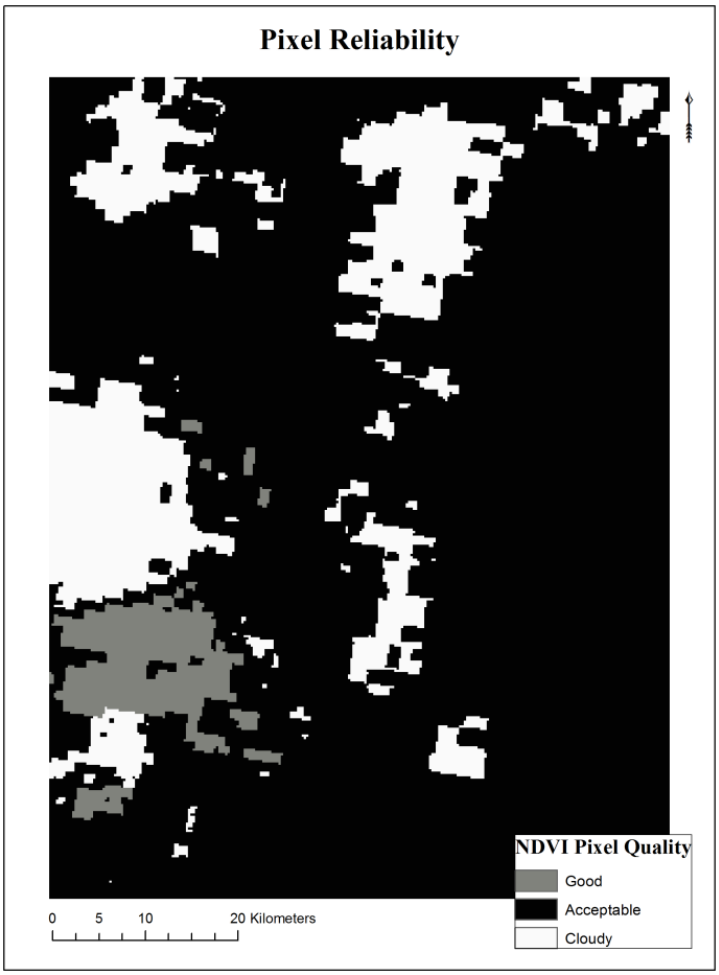


**Figure 5.1.** 250m NDVI pixel quality assessment layer from MOD13Q1 product of the MODIS sensor

The second key method is the harmonic smoothing algorithm used in the analysis of the MODIS 250m NDVI data. This algorithm can be roughly subset into three sections. The first section is the filtering stage, where a generalized additive model and FFT of the raw data are used to impute smoothed observations in place of cloudy or otherwise poor quality observations.

The second major step is to refactor the dataset into annual-average waves (AAWs), which allowed for a more simple analysis of seasonal trends and also reduced the proportion of noise in the ultimate time-series to be analyzed. Finally, a DD approach, analyzed for effect-size and robustness with *SiZer maps*, is taken to test H1 and H2.

The first key step of this research is the filtering algorithm. The goal of this step is to create a complete MODIS record by replacing poor quality data with data including seasonal and long-term trends evident in the series. This algorithm is modeled after Lunetta et al. (2006), where the authors used the FFT in the spatial domain to smooth MODIS images. In this case, I smooth in the time dimension and use g*eneralized additive models* (GAM). GAM's are used for *harmonic filtering*, or removing the signal driven by the harmonics at each data point. The GAM is simply a *generalized linear model* with an additional link function, where the linear predictor depends on a smoothed variable, in this case time. By first removing the known true components of the signal, the FFT can more effectively pick up on erroneous, high-frequency signal components.

GAM's are fit to each camp and village, with a factor for each pixel, i.e. both the intercept and the slope are allowed to vary for each pixel, in each model expressed by Equation 5. A Gaussian family is used with a linear link function, regressing harmonics with periods of one year, half-year, and quarter-year, and with sine and cosine waves. The residuals of this regression should contain only higher frequency components of the signal, including the 'noise' we are removing, as well as any long-memory (greater than 1 year in period) trends.

A GAM regression removes known (annual, semiannual & quarterly) as well as low frequency signal from MODIS NDVI pixels. $a$ is an intercept varying with each pixel ($p$), $\lambda$ is the period of $k$ harmonic waves at time $t$. Each harmonic wave has a linear coefficient $\beta$, which

varies by village/camp ($v$), $p$ & $k$. The GAM regression is smoothed over time, to remove ultra-low-frequency/deterministic trends, $S_v(t)$. This regression will leave noise as well as some higher-frequency signal in the residuals, $\hat{\varepsilon}_{pt}$.

*Equation 5)* $$NDVI_{vpt} = a_p + \sum_k (\sin \lambda_{tk} \beta_{vpk} + \cos \lambda_{tk} \beta_{vpk}) + S_v(t) + \hat{\varepsilon}_{pt}$$

Next, to attenuate the purely stochastic noise and leave unknown, higher-frequency signal (signal not captured in the harmonic components of the GAM), an FFT is performed on the residuals of the GAM at each pixel. Remember, FFT is simply an algorithm for approximating the DFT. $X_{p\gamma}$ are the residuals from Equation 4, expressed as frequency-domain pulses at $\gamma$ frequencies, at $p$ pixels.

*Equation 6)* $$X_{p\gamma} = \sum_{t=0}^{N} \hat{\varepsilon}_{pt} \, e^{-i2\pi\gamma \frac{t}{N}}$$

The FFT is then convolved with a simple probability density function, Gaussian filter, centered at 0 with standard deviation σ:

*Equation 7)* $$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

The one-dimensional Gaussian filter smooths the frequency-domain GAM residuals when convolved. Here, the Fourier transform acts much like high-pass filter, by attenuating the extreme high frequency components of each time-series. See Figure 5.2 for a representation of

the Gaussian density function used for the FFT convolution. The filtered residuals $\hat{X}_{py}$ are converted back to the time-domain using the IFT from Equation 4 as $\hat{\varepsilon}_{pt}$ and added back to the predicted values from the GAM harmonics, $H_{vpkt}$, intercepts and smoothers, and we are left with our final smoothed time-series, $\widehat{NDVI}_{pt}$.

*Equation 8)* $$\widehat{NDVI}_{pt} = a_p + H_{vpkt} + \mathrm{S}_v(t) + \hat{\varepsilon}_{pt}$$

An example of these smoothed results, for a single pixel, is shown in green in Figure 5.4.

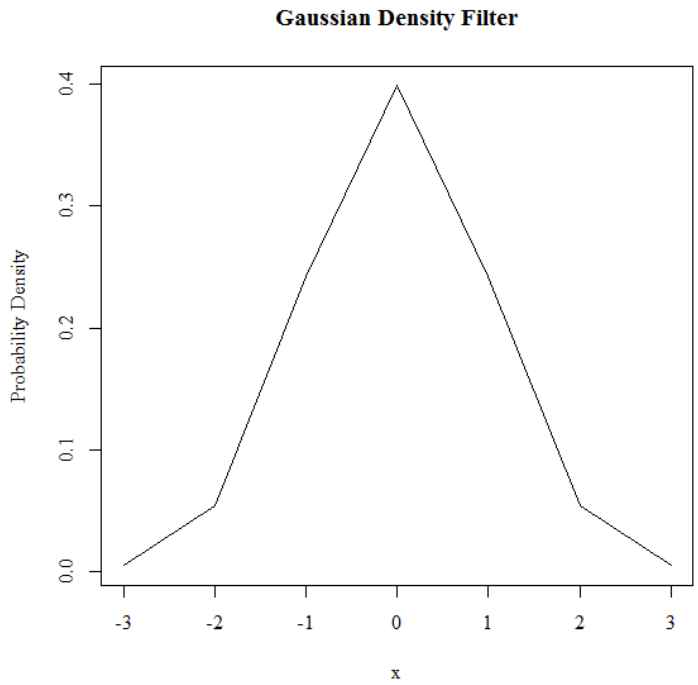


**Figure 5.2.** Shows the probability density function for a Gaussian density filter to be used in convolved with the products of the FFT's of the GAM residuals.

Finally, the raw NDVI time-series for each pixel is filtered using the accompanying pixel reliability masks, where observations are retained only if pixel reliability is 0 or 1 (good or

acceptable). Imagine Figure 5.1 to be a 3-dimensional array of images, stacked by time. Each pixel's raw NDVI time-series is filtered by the pixel's value at the same point in the time-stacked dimension. This process leaves missing values (shown in the red graph in Figure 5.4) which are then imputed with the values from the FFT smoothed time-series. Figure 5.3 provides a graphical depiction of the key steps in the harmonic smoothing algorithm described above.

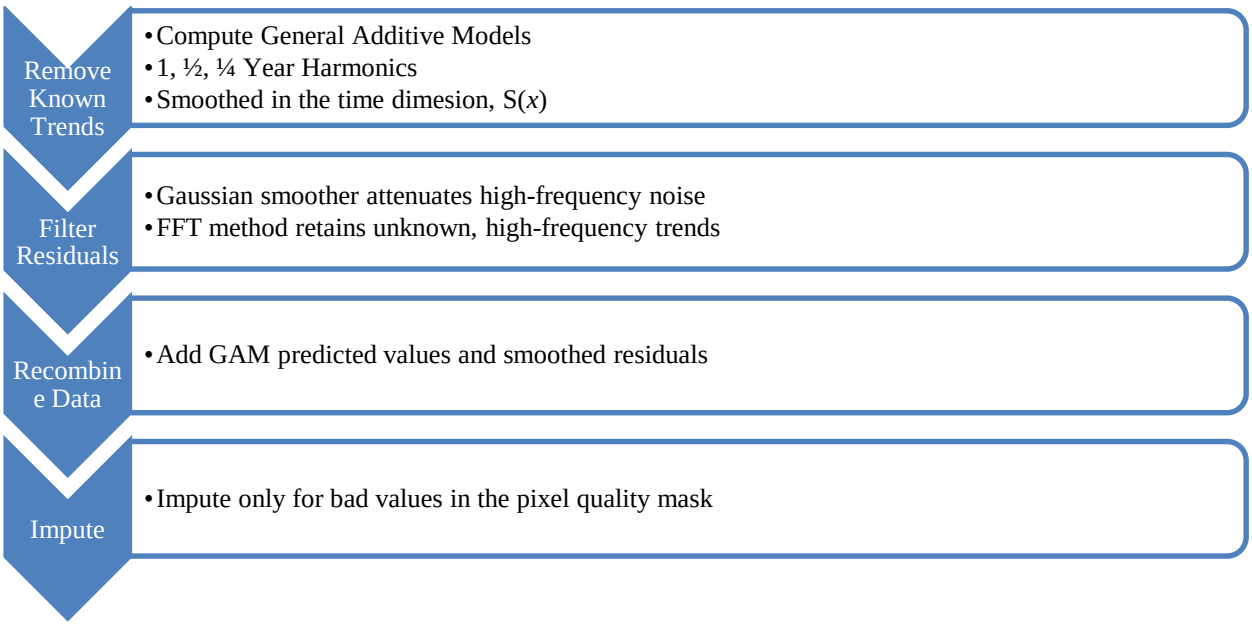


**Figure 5.3** *Harmonic Smoothing Algorithm*: **f**low chart illustrating key steps of the algorithm.

Comparing the objective quality of the smoothed and unsmoothed data is difficult because we have no knowledge of the 'truth.' There is no certain way to state that the smoothed data is an improvement over the unsmoothed data. We can be fairly certain that the imputed values improve upon the raw NDVI, because MODIS provides the accompanying masking data. The pixel reliability mask seems consistent with wave theory, where water vapor (in this case clouds), reflect much more in the infrared and visible spectrum and thus create erroneous, low

NDVI points. This can be seen in Figure 5.5, where the dropped pixels are almost entirely below the mean. The imputed points do not simply conform to the mean or local mean, but instead are consistent with time and frequency.

The second important step to improving the efficacy of this analysis is to refactor the entire 12 year time-series into a series of 23 mean values, one point for each 16-day period (i.e. each village and camp buffer would have 23 NDVI data points). Examining all 12 years of data, from 2000 to 2013, as a single time series makes it much more difficult to model the cyclical nature of vegetation-signal data. By refactoring the data into an Average Annual Wave (AAW), one can achieve more certain results by increasing the observations to be averaged by a factor of roughly 23. This process will not only streamline the analysis, but should also diminish the proportion of noise to true signal in the data.

To compare the time-series of Rwandan refugee camps to that of neighboring villages, one may simply subtract their observations at each time-period to create a differenced dataset. However, because of the refactored, seasonal AAW, this differenced time-series does not retain any of the original time-dimension (i.e. what year an observation occurred in), excluding each data-point's relative position in the AAW. To analyze changes in the AAW over time, one must first dissect the wave into some logical time-frames. For this study, I simply split the differenced dataset (camps minus villages) into two time-series, prior to refactoring into the AAW, the first from 2000 to 2006 and the second from 2006 to 2013. These results are again differenced, creating a wave that shows how the average change between villages and camps has changed between the first and second periods, or the two periods may simply be examined visually in SiZer maps.

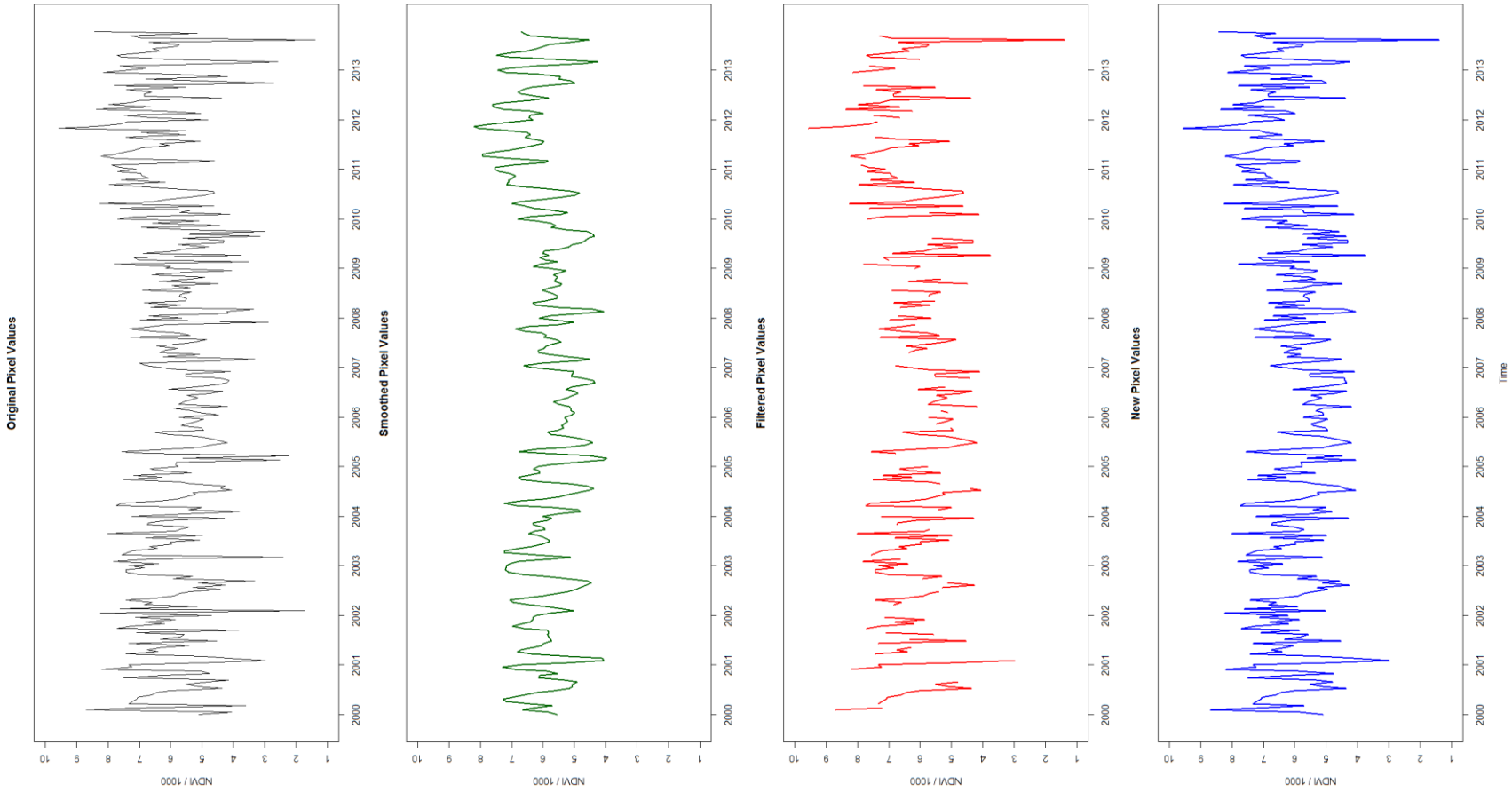


**Figure 5.4.** Graphs showing the four stages of data in the FFT filtering algorithm. The first graph, in black, shows the raw time-series of NDVI values for a single pixel. The second graph, in green, shows the harmonic and FFT filtered time-series for the pixel ($\widehat{\boldsymbol{NDVI}}$). The third graph, in red, shows the masked time series with missing values. The fourth graph shows the final, filtered time series

SiZer is commonly held abbreviation of "**Si**gnificant **Zer**o crossings of the derivative," and is a tool originally proposed by Chaudhuri and Marron (1999). In our exploration of time-series RS data, accurately modeling the data again requires fitting of GAMs, smoothed along the time dimension. Once one has attained the predicted values of this GAM, the next logical step would be to take the first derivative of this model, to observe where changes between camps and villages are diverging most (i.e. the derivative shows the rate of change at a given time-period, between camps and villages). A SiZer map allows us to be more certain of the nature and significance of features of the GAM, with respect to different degrees of smoothing.

Chaudhuri and Marron (1999) define the basic SiZer algorithm, from a "Scale-Space Viewpoint" (810) as follows:

*Equation 9)* $$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - X_i)$$

Where *h* is the bandwidth, or width of the smoothing kernel (e.g. Figure 5.2), and $K_h$ represents the kernel function at each bandwidth. $\hat{f}_h$ becomes flat as *h* increases, due to the mean-zero kernel. Plotting SiZer results is made simple through the R package, 'SiZer' (http://cran.r-project.org/web/packages/SiZer/SiZer.pdf), which includes a custom function ('plot.SiZer') for creating simple color-coded charts.

### 5.3 Results & Discussion

The results of the FFT and harmonic smoothing algorithm, applied to the 250m NDVI MODIS data were consistent with expected outcomes. As discussed in Chapter 3, cloud cover is likely the most common source of noise in RS data of Western Rwanda, and is likely to cause erroneously low observations. The MODIS pixel-quality mask flagged mostly low-value NDVI

observations, as seen in the example mean-village time-series in Figure 5.5, where nearly all dropped pixels are significantly below the FFT smoothed (black) line.

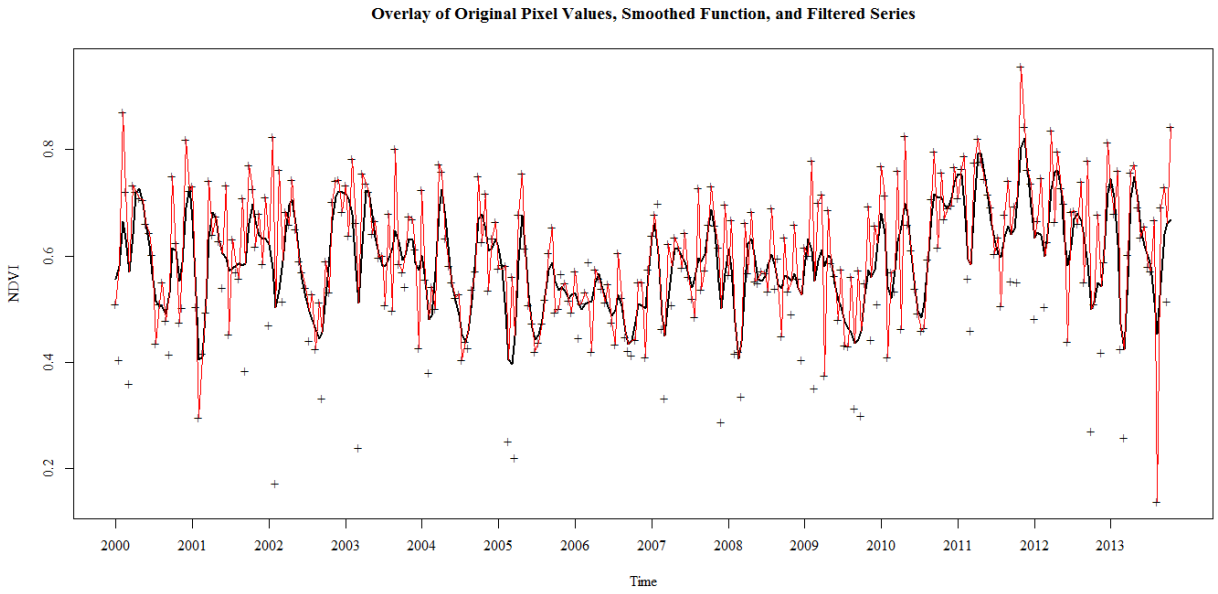


**Figure 5.5.** Overlay of original, smoothed and filtered pixel values for one village in Western Rwanda. The original values are shown as '+', the red line is the black lined is the smoothed function, and the new, filtered time-series is in red.

To evaluate the impact of the FFT and harmonic smoothing algorithm on the raw NDVI data, I produced two SiZer maps. The first map, shown in Figure 5.8, is a refactored AAW for the entire original, raw time-series. The second map, shown in Figure 5.9, is the SiZer of the same refactored AAW, but using the FFT and harmonic smoothed data. One can clearly see the difference made by smoothing in the changes insignificance of features of a fitted GAM across smoothing bandwidths. The major features of the GAM, fitted to the AAW, are significant at nearly all bandwidths. This implies that the patterns expressed in the GAM predictions are real, irrespective of the level of time-dimension smoothing.

For the smoothed term (time) of the GAMs used in the smoothing algorithm, the f-statistics ranged from 49.7 to 238.1, and the corresponding p-values had a *maximum* of $1.89e^{-90}$. The $R^2$ values for the models ranged from 0.12 to 0.23. The $R^2$ values were lower than expected, but this is likely symptomatic of the large, inclusive buffer size, which likely includes more bias (less noise) than a smaller-radius buffer would have.

A disconcerting feature of the SiZer results is the correlation between the presence of cloudy pixels in the data and the significant regions of the SiZer maps. A graph of percent-cloudy pixels across the entire MODIS NDVI scene for Western Rwanda, refactored to the same AAW, is shown in Figure 5.6. It appears that the regions of highest significance in our GAM model predictions are also periods where there are large numbers of cloudy pixels, raising concern that the results are being driven by the FFT and harmonic smoothing algorithm.

However, it is unlikely that the results are being driven by the smoothing algorithm. First, the algorithm is unlikely to create differentiated results for camps and villages, as each camp and each village is fit with an individual GAM. The SiZer map in Figure 5.9 is the result of the differenced time-series of the camps and villages, and any systematic impact of the smoothing algorithm, which is fit to each pixel and each buffer, should either be removed by this differencing or should be simply insignificant.

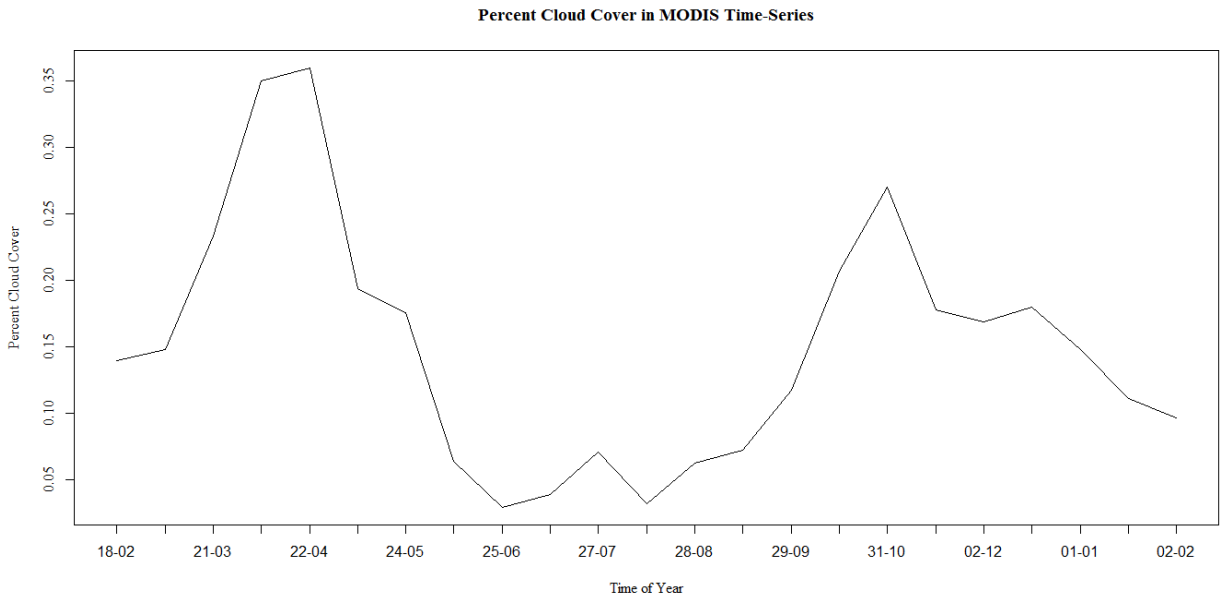


**Figure 5.6.** Average percent cloud cover by 16-day period (*dd-mm*) in MODIS time-series for Western Rwanda.

After affirming the results of the smoothing algorithm, attention was turned to assessing H2, that post-resettlement adaptation has long-term consequences on vegetation-signals within the camp and village buffers. As aforementioned, the time-series was split into two time frames, one from 2000-2006 and another from 2006-2013. These were then refactored into two AAWs for comparison. This was done with both the original and smoothed datasets, so as to create a second comparison for the efficacy of the smoothing algorithm. The GAM predictions, derivatives and SiZer maps are shown in Figures 5.8–13.

Upon visual inspection, it is clear that a different pattern exists in both the unsmoothed and smoothed datasets, between the two timeframes. In both cases, the pattern observed in the second-half SiZers more closely represents the patterns in the AAWs of the entire time-series. This suggests that the magnitude of the difference between camps and villages is increasing between the two time periods, because the second-half pattern is dominating the entire series-average results.

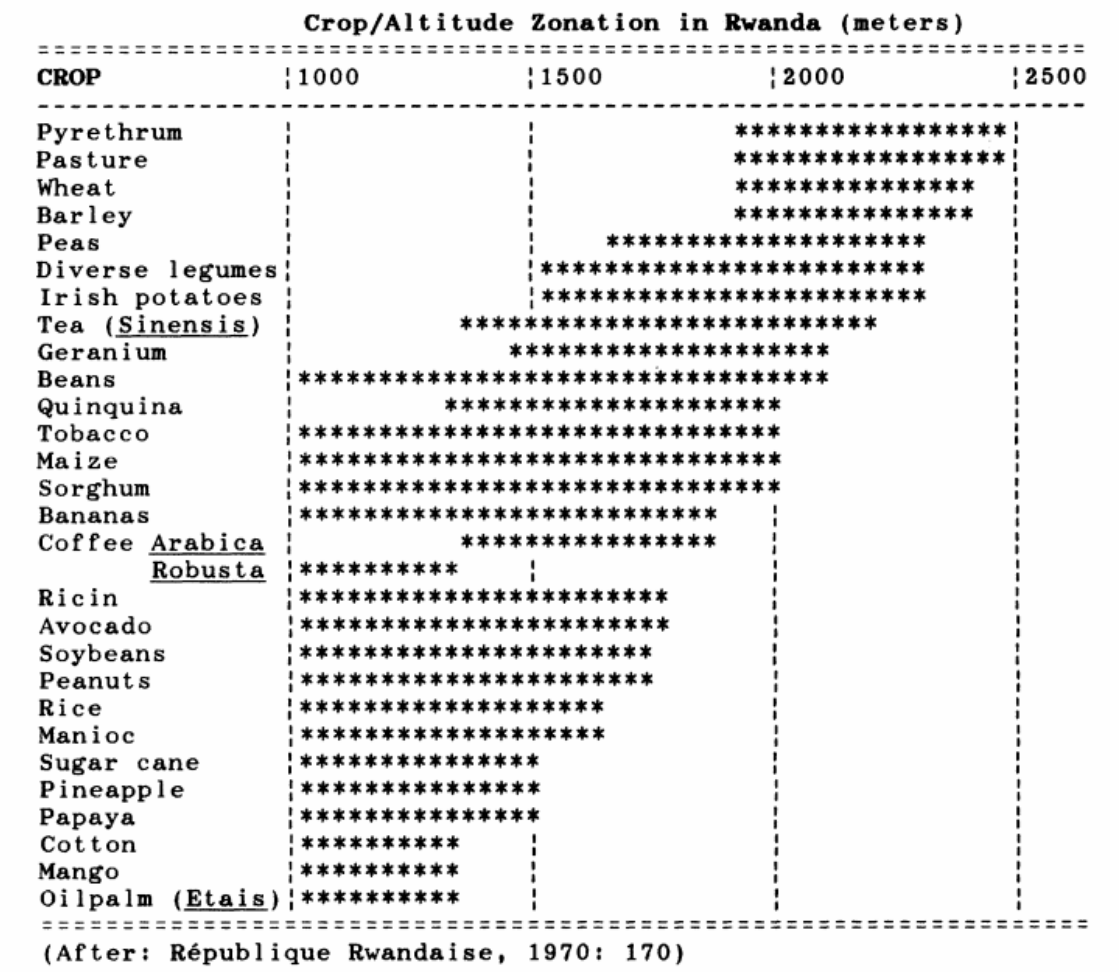


**Figure 5.7.** Crops and altitudinal zonation in Rwanda. Reprinted from *The Dynamics of Human-Environment Interactions in the Tropical Montane Agrosystems of Rwanda* (pg. 45) by R.E. Ford, 1990, JSTOR. Reprinted with permission.

Several theories could explain the pattern observed in the final time-frame comparison of the smoothed data. We observe a significantly lower vegetation signal in March, and likely a significantly higher signal in June. When considering potential causes of this trend, we can diminish the number of potential explanatory crops, due to the high elevation of the Western Rwanda and the elevation stratification of crops in the region. Only a small subset of crops shown again in Figure 5.7 are grown in the Rwandan Highlands. Based on simple crop calendars, and the medium-to-high elevation of the region, this likely represents an increase in bean, pea

and potato farming. This is a sensible outcome, due to the government's subsidization of staple crop and cash crop production, as well as new, more efficient farming practices to accommodate the growing population. However, this analysis of the results leaves much to chance, and could be dramatically improved with better training data, or by someone specializing in agriculture of this region.

There are several statistical approaches that might shed light on the robustness of the results above. One potential approach would be to create synthetic data, with a known spatial and temporal distribution, degrade this data with artificial 'clouds,' and then use our same smoothing and SiZer analyses. If this results in a similar outcome, we could much more confidently validate H1 and H2. A simpler approach to assessing H1 in particular, would be to degrade our current data further, adding extra clouds, or growing the existing clouds in a Bayesian framework. If the smoothing algorithm is able to recreate the degraded pixels, this is a simple means of validating the harmonic smoothing algorithm. Either of these approaches could add significant power to the analysis, but simply fall outside the scope of this thesis.

Original SiZer

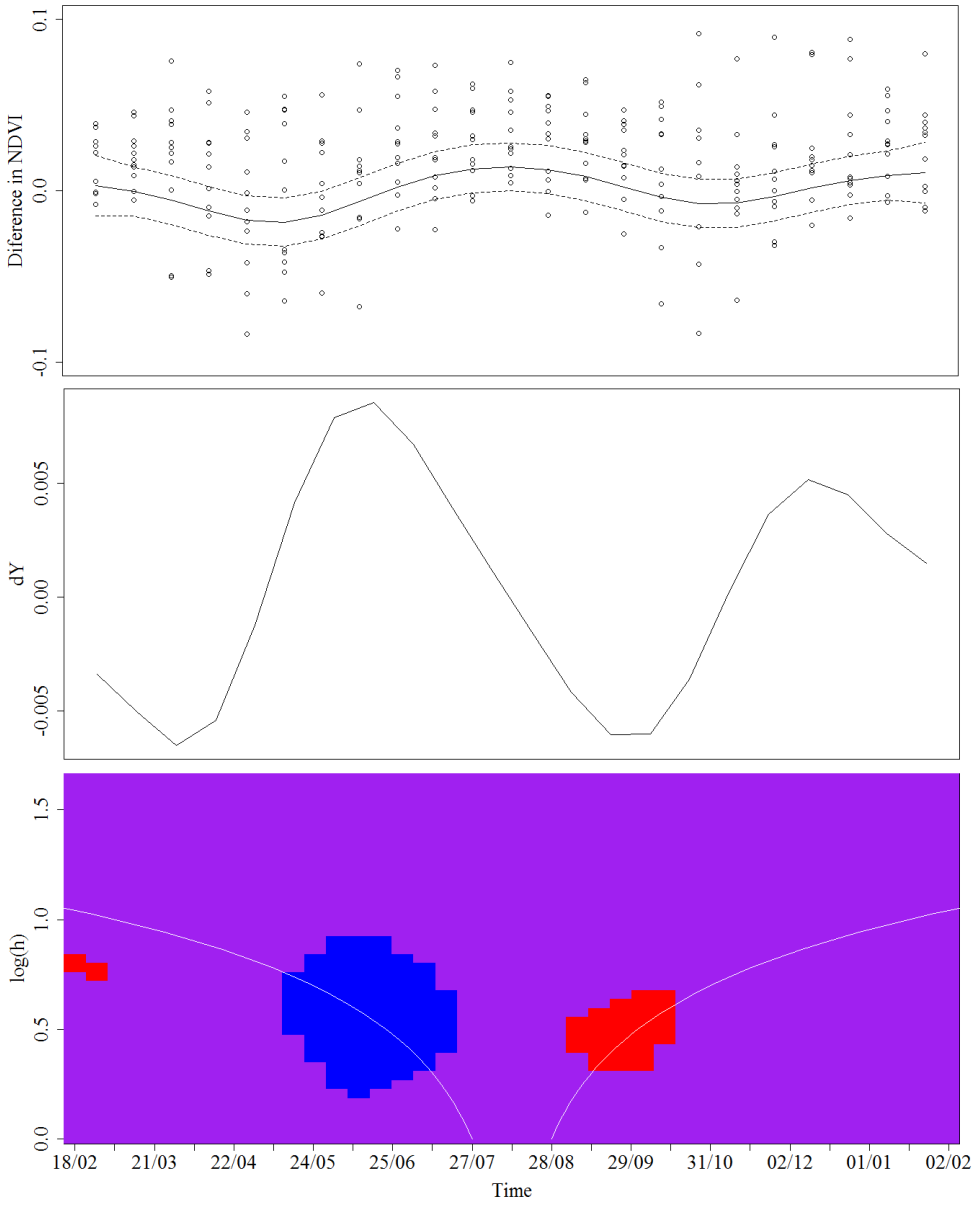


**Figure 5.8.** Shows the SiZer map of the original, unsmoothed, difference in NDVI values (camps – villages) and

Smoothed SiZer

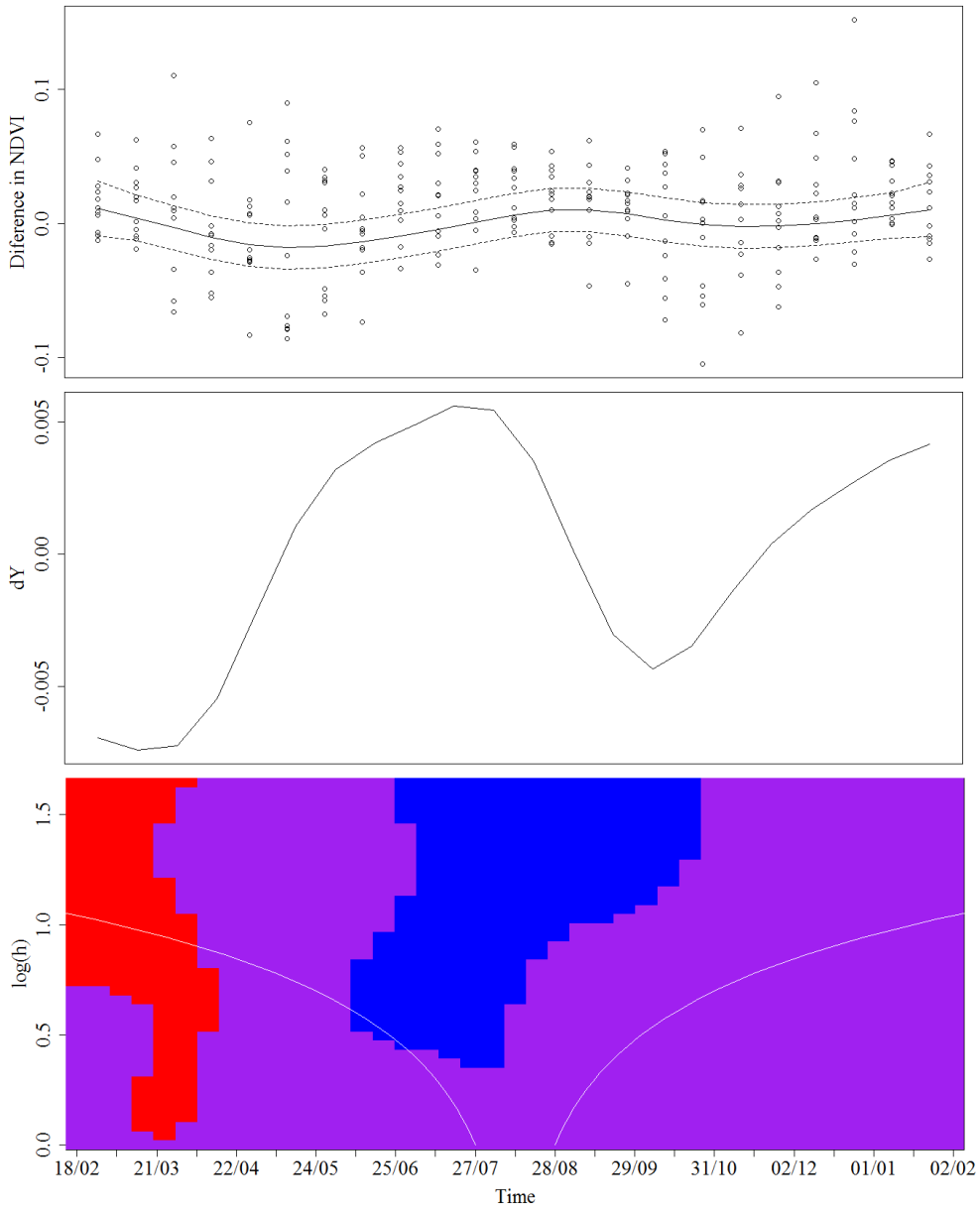


**Figure 5.9.** shows the same SiZer map, but with FFT smoothed differences. Purple indicates the derivative is possibly 0, blue indicates a significant positive derivative and red indicates a significant negative derivative. There is a dramatic improvement in the robustness of the GAM models fitted to the FFT smoothed time-series as compared with the original data.

Original First ½

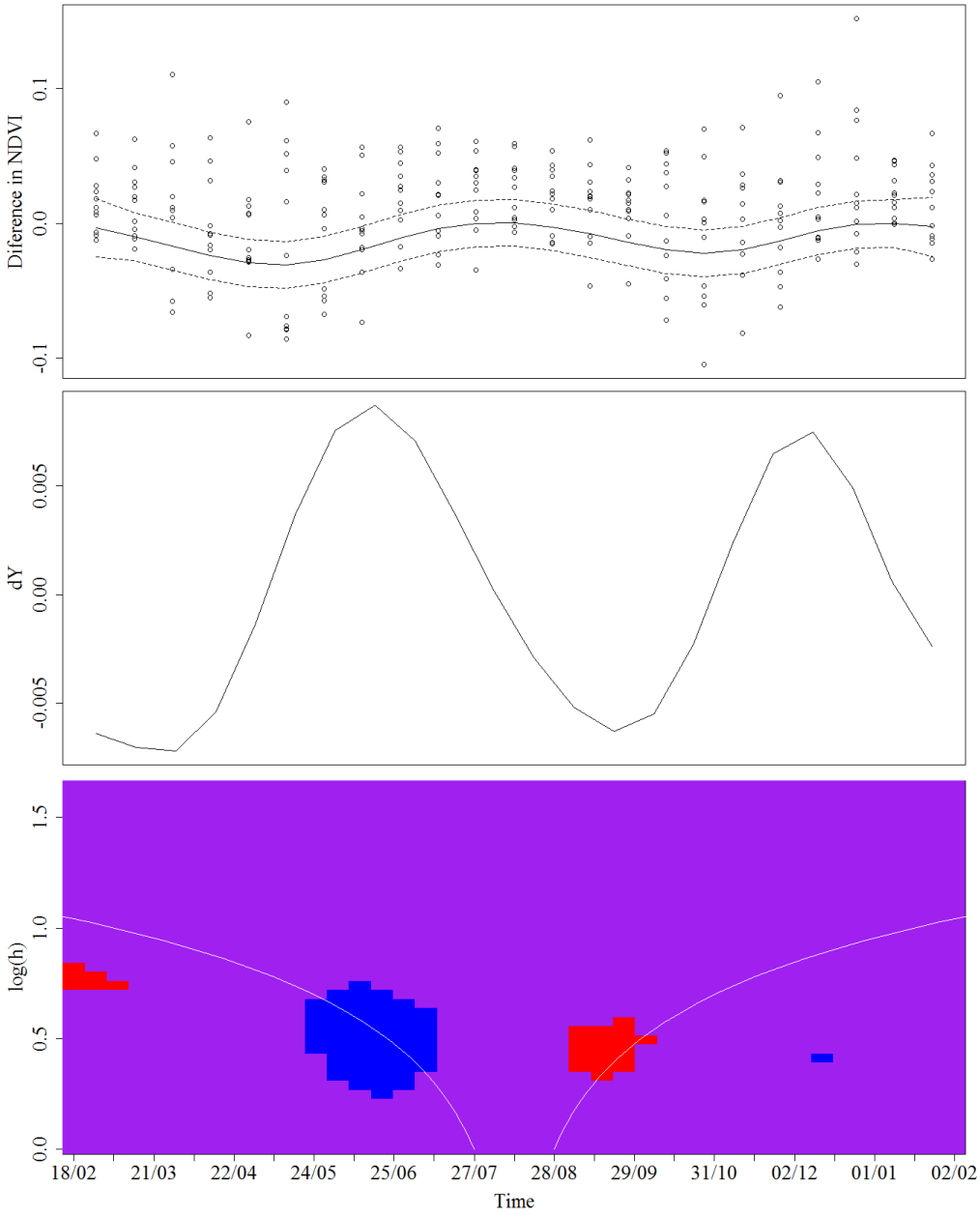


**Figure 5.10.** Shows the SiZer map of the original, unsmoothed, difference in NDVI values (camps – villages) from 2000-2005

Original Second ½

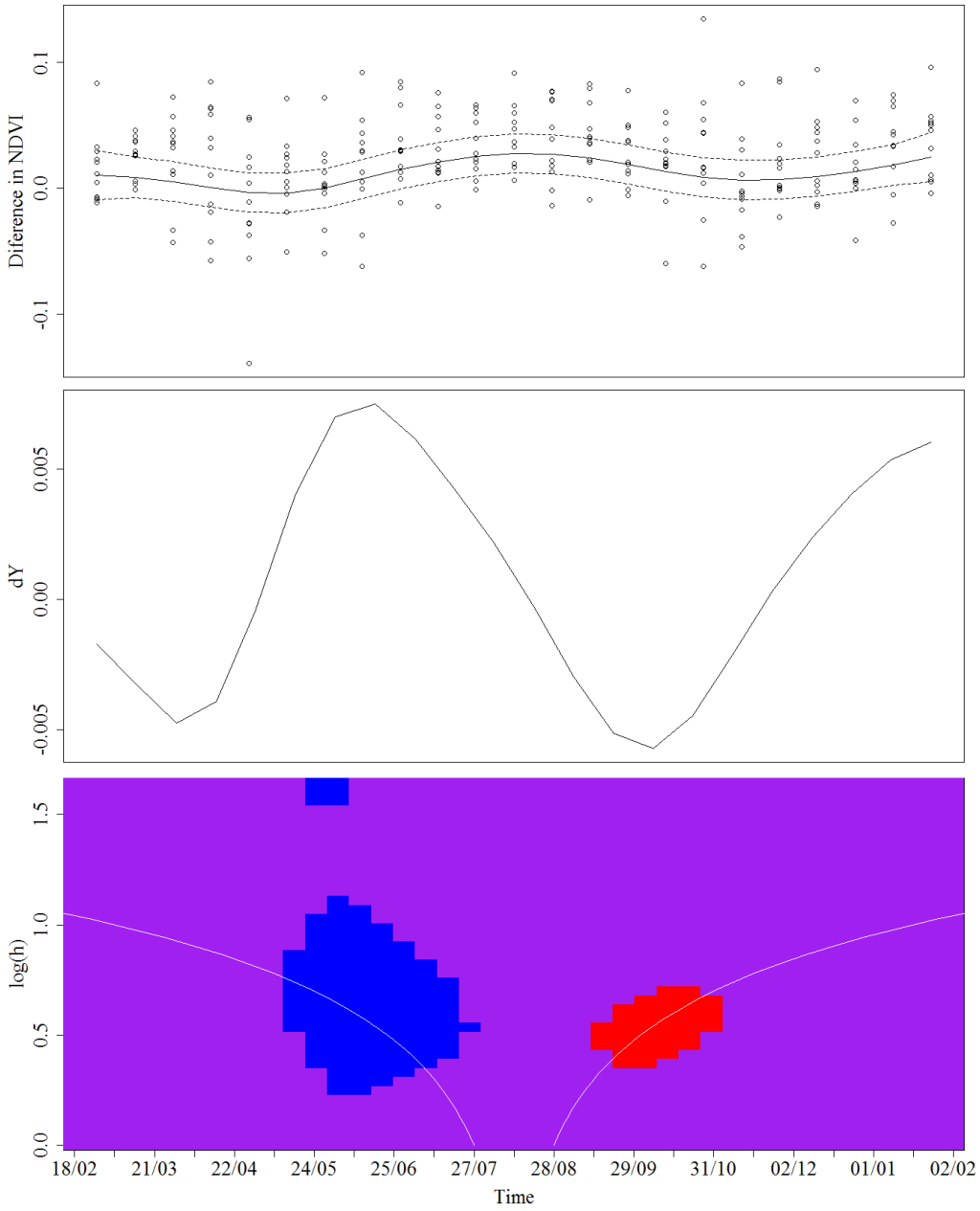


**Figure 5.11. S**hows the SiZer map of aggregated differences from 2006-2013.

Smoothed First ½

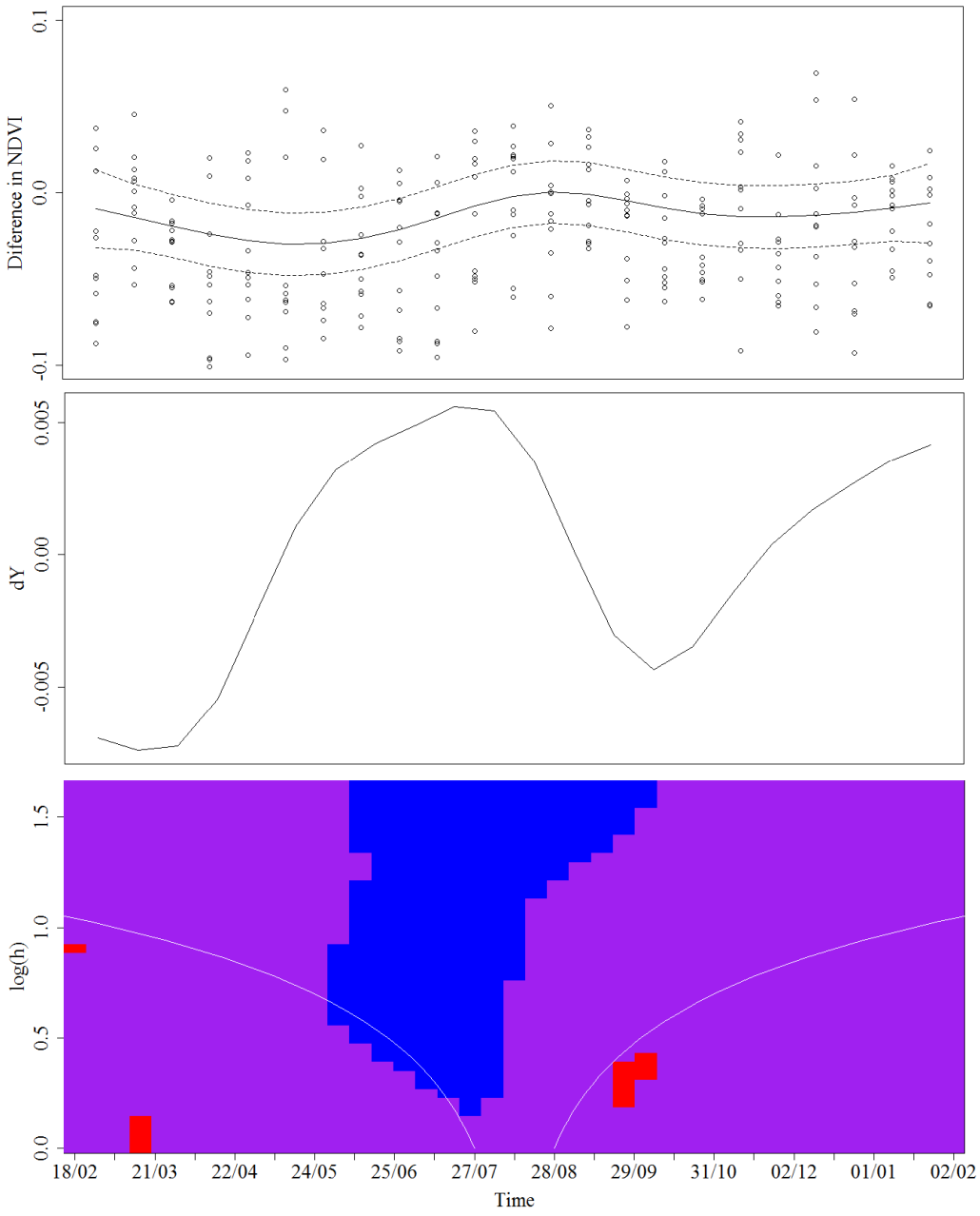


**Figure 5.12.** Shows the SiZer map of the FFT smoothed, difference in NDVI values (camps – villages) from 2000-2005.

Smoothed Second ½

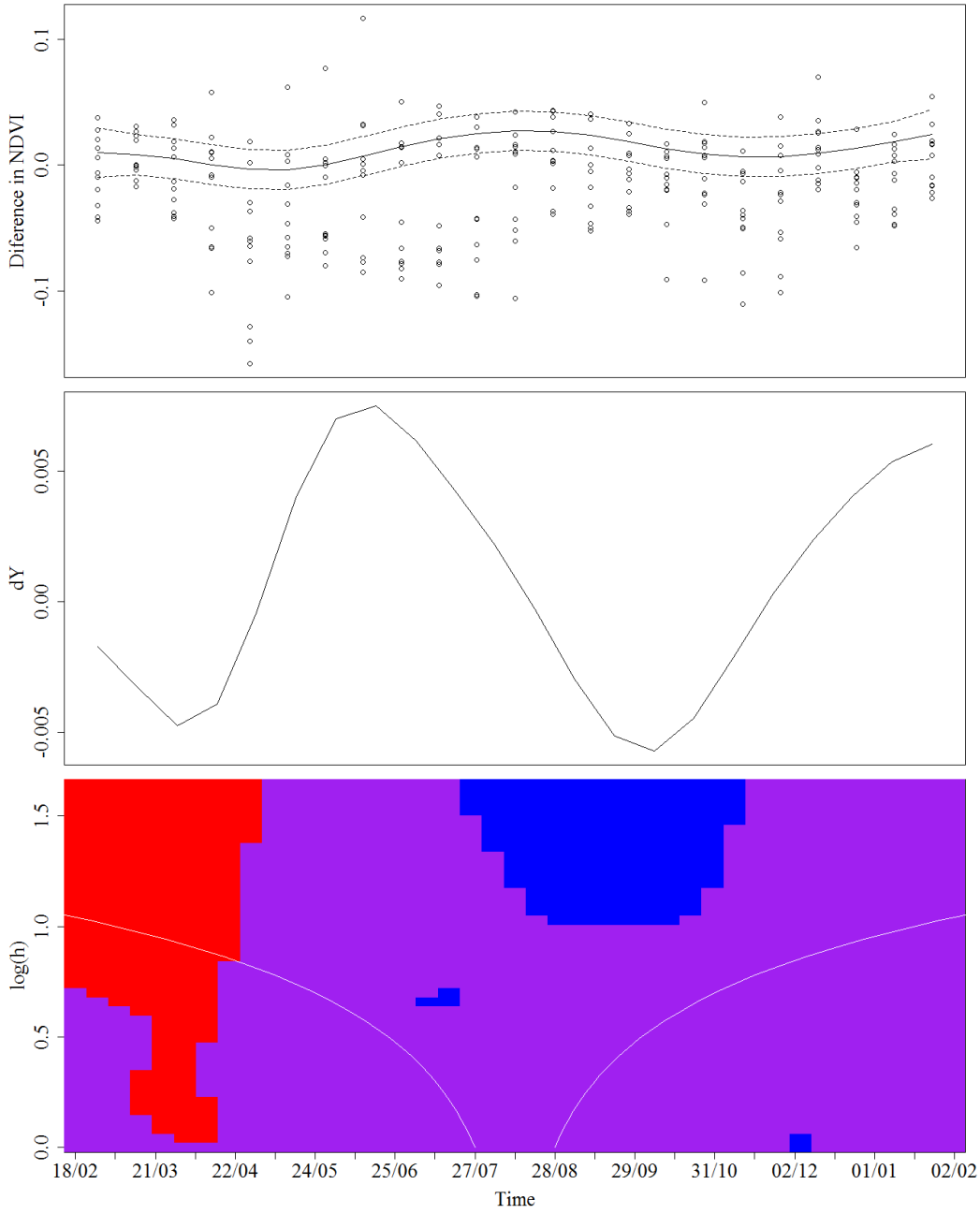


**Figure 5.13.** Shows the aggregated differences from 2006-2013.

## Conclusion

The findings of this research were consistent with my opening hypotheses. First, I was able to significantly improve the quality of medium resolution, time-series RS data using the FFT algorithm. This improvement is best demonstrated by the change in certainty illustrated in SiZer maps, seen in Figures 5.8 and 5.9. Therefore, we should accept Alternative Hypothesis 1: *H1.a: FFT-filtered, 16-day composite, 250m NDVI images will provide a cleaner time-series for further analysis.*

Second, I was able to use this smoothed time series to show a highly significant, real long-memory pattern in the data. The DD approach controlled for conflating variables by removing the effect of a 'control,' or older village, to examine only the effect of the Imidugudu program. I then split the data into two sections, 2000-2006 and 2006-2013, and showed a significant difference in the vegetation signals between the two periods. This second finding suggests we should accept Alternative Hypothesis 2: *H2.a: There is a significant long memory/low frequency trend in the difference camp and village data.* This result implies that there are long-term changes to the environment correlated with post-resettlement adaptation of refugees in Rwanda. However, this is far from a definitive result and requires a much more nuanced approach (for instance, a multiresolution strategy for identifying crops), to generate explanatory, definitive results.

My findings above are pertinent to the study of human-environment interaction in tropical-montane regions, as well as to the study of medium-resolution RS data. The outcomes of this analysis illustrate the suitability of MODIS 250m NDVI data for studying broad-scale changes in vegetation signals in tropical-montane regions. They also show the presence of a long-memory trend, possibly correlated with post-resettlement adaptation. This thesis adds to the extant

literatures on human ecology RS and documents a highly effective, frequency-domain algorithm for the improvement of 16-day composite, 250m, time-series, NDVI data. Positive results were found for both H1 and H2, motivating a great deal of potential research along these avenues.

The first goal of future research should be to gather an improved training dataset. This should be done in two parts, the formation of a ground-truth dataset and the use of higher resolution data to form a more exhaustive training dataset. Ground-truth data is expensive and time consuming, but is the only sure method to create an accurate training dataset. Furthermore, field research is needed to determine the most appropriate method for buffering village and camp extents. This ground-truth data set could be used as the final training data, or it could be used to classify higher resolution data. The refugee camps are likely distinguishable, through spectral signature, in Landsat 30m data. This method would be preferable, because it would allow for the creation on an exhaustive training dataset for the region.

The second goal of future research should be to classify MODIS pixels with ground-truth data, or by careful study of crop phenology and crop calendars for Rwanda. This analysis should use fuzzy logic to allow for mixed MODIS pixels, classified by percent of the pixel occupied by each crop. A ground-truth dataset would allow for a simple correlation between percentages of pixels, while an unsupervised classification of higher resolution data (needed to distinguish individual crops) might yield improved results by incorporating a temporal dimension. Whichever method is used, a crop-based analysis will likely show a much improved picture of the changes in vegetation signals in the refugee camps and villages of Rwanda.

A final avenue of future research, motivate by this thesis, should be to compare and improve the filtering procedure used in Chapter 5. As discussed, creation of synthetic data or further

degradation of the existing data are two methods that could much more powerfully validate the harmonic smoothing algorithm. Also, the FFT is only one frequency-domain method, but studies have shown superior results from generalized wavelet analysis (Galford et al., 2008; Ruiz et al., 2004). Wavelets better incorporate information in the time-domain in frequency-domain analysis and these methods should be explored in future work.

## Work Cited

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies*, *72*(1), 1–19. http://doi.org/10.1111/0034-6527.00321

Agarwal, C., Green, G. M., Grove, J. M., Evans, T. P., & Schweik, C. M. (2002). *A Review and Assessment of Land-use Change Models: Dynamics of Space, Time, and Human Choice*. Retrieved from http://www.treesearch.fs.fed.us/pubs/5027

Akresh, R., Walque, D., & Damien. (2008). *Armed Conflict and Schooling : Evidence from the 1994 Rwandan Genocide* (SSRN Scholarly Paper No. ID 1149109). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=1149109

Andres, L., Salas, W. A., & Skole, D. (1994). Fourier Analysis of Multi-Temporal AVHRR Data Applied to a Land Cover Classification. *International Journal of Remote Sensing*, *15*(5), 1115–1121. http://doi.org/10.1080/01431169408954145

Bernauer, T., Böhmelt, T., & Koubi, V. (2012). Environmental Changes and Violent Conflict. *Environmental Research Letters*, *7*(1), 015601. http://doi.org/10.1088/1748-9326/7/1/015601

Brandtberg, T., Warner, T. A., Landenberger, R. E., & McGraw, J. B. (2003). Detection and Analysis of Individual Leaf-off Tree Crowns in Small Footprint, High Sampling Density Lidar Data from the Eastern Deciduous Forest in North America. *Remote Sensing of Environment*, *85*(3), 290–303. http://doi.org/10.1016/S0034-4257(03)00008-7

Bremaud, P. (2002). *Mathematical Principles of Signal Processing: Fourier and Wavelet Analysis*. Springer Science & Business Media.

Chaudhuri, P., & Marron, J. S. (1999). SiZer for Exploration of Structures in Curves. *Journal of the American Statistical Association*, *94*(447), 807–823. http://doi.org/10.1080/01621459.1999.10474186

Choi, T., Qu, J. J., & Xiong, X. (2013). A Thirteen-Year Analysis of Drought in the Horn of Africa with MODIS NDVI and NWDI Measurements. In *2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics)* (pp. 302–307). http://doi.org/10.1109/Argo-Geoinformatics.2013.6621926

Diebel, J., & Norda, J. (n.d.). *Average Weather For Kigali, Rwanda* (Data Archive). Weatherspark. Retrieved from http://weatherspark.com/averages/29294/Kigali-Rwanda

Discrete Fourier Transform. (2014). In *Wolfram|Alfa*. Wolfram Alpha LLC. Retrieved from http://www.wolframalpha.com/input/?i=discrete+fourier+transform

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*(1), 185–201. http://doi.org/10.1016/S0034-4257(01)00295-4

Ford, R. E. (1990). The Dynamics of Human-Environment Interactions in the Tropical Montane Agrosystems of Rwanda: Implications for Economic Development and Environmental Stability. *Mountain Research and Development*, *10*(1), 43. http://doi.org/10.2307/3673538

Fotheringham, A. S., & Wong, D. W. S. (1991). The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environment and Planning A*, *23*(7), 1025 – 1044. http://doi.org/10.1068/a231025

Galford, G. L., Mustard, J. F., Melillo, J., Gendrin, A., Cerri, C. C., & Cerri, C. E. P. (2008). Wavelet Analysis of MODIS Time Series to Detect Expansion and Intensification of Row-Crop Agriculture in Brazil. *Remote Sensing of Environment*, *112*(2), 576–587. http://doi.org/10.1016/j.rse.2007.05.017

Greene, W. H. (2003a). Heteroskedasticity. In *Econometric Analysis* (5th ed., p. 240). Upper Saddle River, New Jersey: Pearson Education Inc.

Greene, W. H. (2003b). Time-Series Models. In *Econometric Analysis* (5th ed., p. 646). Upper Saddle River, New Jersey: Pearson Education Inc.

Hamlyn, J., & Vaughn, R. (2010a). Basics of Radiation Physics for Remote Sensing of Vegetation. In *Remote Sensing of Vegetation: Principles, Techniques, and Applications* (1st ed., pp. 7–14). Oxford, New York: Oxford University Press.

Hamlyn, J., & Vaughn, R. (2010b). Earth Observation Systems. In *Remote Sensing of Vegetation: Principles, Techniques, and Applications* (1st ed., pp. 92–127). Oxford, New York: Oxford University Press.

Havugimana, E. (2009). State Policies and Livelihoods. Rwandan Human Settlement Policy. Case Study of Ngera and Nyagahuru Villages. Retrieved from https://gupea.ub.gu.se/handle/2077/21046

History of Photography. (2014). [Encyclopedia]. Retrieved October 22, 2014, from http://www.britannica.com/EBchecked/topic/457919/history-of-photography

Justice, C. O., Townshend, J. R. G., Holben, B. N., & Tucker, C. J. (1985). Analysis of the Phenology of Global Vegetation Using Meteorological Satellite Data. *International Journal of Remote Sensing*, *6*(8), 1271–1318. http://doi.org/10.1080/01431168508948281

Kalacska, M. E., Bell, L. S., Arturo Sanchez-Azofeifa, G., & Caelli, T. (2009). The Application of Remote Sensing for Detecting Mass Graves: An Experimental Animal Case Study from Costa Rica*. *Journal of Forensic Sciences*, *54*(1), 159–166. http://doi.org/10.1111/j.1556-4029.2008.00938.x

Kaufman, Y. J., Tanré, D., Gordon, H. R., Nakajima, T., Lenoble, J., Frouin, R., … Teillet, P. M. (1997). Passive Remote Sensing of Tropospheric Aerosol and Atmospheric Correction for

the Aerosol Effect. *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16815–16830. http://doi.org/10.1029/97JD01496

Key, T., Warner, T. A., McGraw, J. B., & Fajvan, M. A. (2001). A Comparison of Multispectral and Multitemporal Information in High Spatial Resolution Imagery for Classification of Individual Tree Species in a Temperate Hardwood Forest. *Remote Sensing of Environment*, *75*(1), 100–112. http://doi.org/10.1016/S0034-4257(00)00159-0

Levin, N. (1999). *Fundametals of Remote Sensing* (Trieste 76). International Maritime Academy.

Longbotham, N., Pacifici, F., Glenn, T., Zare, A., Volpi, M., Tuia, D., Du, Q. (2012). Multi-Modal Change Detection, Application to the Detection of Flooded Areas: Outcome of the 2009 #x2013;2010 Data Fusion Contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(1), 331–342. http://doi.org/10.1109/JSTARS.2011.2179638

Lunetta, R. S., Knight, J. F., Ediriwickrema, J., Lyon, J. G., & Worthy, L. D. (2006). Land-cover Change Detection Using Multi-Temporal MODIS NDVI Data. *Remote Sensing of Environment*, *105*(2), 142–154. http://doi.org/10.1016/j.rse.2006.06.018

Madden, M., & Ross, A. (2009). Genocide and GIScience: Integrating Personal Narratives and Geographic Information Science to Study Human Rights. *The Professional Geographer*, *61*(4), 508–526. http://doi.org/10.1080/00330120903163480

McGranahan, G., Balk, D., & Anderson, B. (2007). The Rising Tide: Assessing the Risks of Climate Change and Human Settlements in low Elevation Coastal Zones. *Environment and Urbanization*, *19*(1), 17–37. http://doi.org/10.1177/0956247807076960

Mittermeier, R. A., Myers, N., Thomsen, J. B., Da Fonseca, G. A. B., & Olivieri, S. (1998). Biodiversity Hotspots and Major Tropical Wilderness Areas: Approaches to Setting

Conservation Priorities. *Conservation Biology*, *12*(3), 516–520. http://doi.org/10.1046/j.1523-1739.1998.012003516.x

Nagler, P. L., Scott, R. L., Westenburg, C., Cleverly, J. R., Glenn, E. P., & Huete, A. R. (2005). Evapotranspiration on Western U.S. Rivers Estimated Using the Enhanced Vegetation Index from MODIS and Data from Eddy Covariance and Bowen Ratio Flux Towers. *Remote Sensing of Environment*, *97*(3), 337–351. http://doi.org/10.1016/j.rse.2005.05.011

Outreach Programme on the Rwanda Genocide and the United Nations. (2014). Retrieved November 26, 2014, from http://www.un.org/en/preventgenocide/rwanda/about/bgjustice.shtml

Peña-Barragán, J. M., Ngugi, M. K., Plant, R. E., & Six, J. (2011). Object-Based Crop Identification Using Multiple Vegetation Indices, Textural Features and Crop Phenology. *Remote Sensing of Environment*, *115*(6), 1301–1316. http://doi.org/10.1016/j.rse.2011.01.009

Pender, J. L. (1999). *Rural Population Growth, Agricultural Change and Natural Resource Management in Developing Countries: A Review of Hypotheses and Some Evidence from Honduras* (EPTD discussion paper No. 48). International Food Policy Research Institute (IFPRI). Retrieved from https://ideas.repec.org/p/fpr/eptddp/48.html

Potapov, P., Stehman, S. V., Pittman, K., Turubanova, S., & Hansen, M. C. (2009). Gross Forest Cover Loss in Temperate Forests: Biome-Wide Monitoring Results Using MODIS and Landsat Data. *Journal of Applied Remote Sensing*, *3*(1), 033569–033569–23. http://doi.org/10.1117/1.3283904

Price, J. C. (2003). Comparing MODIS and ETM+ Data for Regional and Global Land Classification. *Remote Sensing of Environment*, *86*(4), 491–499. http://doi.org/10.1016/S0034-4257(03)00127-5

Roujean, J.-L., Leroy, M., & Deschamps, P.-Y. (1992). A Bidirectional Reflectance Model of the Earth's surface for the Correction of Remote Sensing Data. *Journal of Geophysical Research: Atmospheres*, *97*(D18), 20455–20468. http://doi.org/10.1029/92JD01411

Roy, D. P., Jin, Y., Lewis, P. E., & Justice, C. O. (2005). Prototyping a Global Algorithm for Systematic Fire-Affected Area Mapping Using MODIS Time Series Data. *Remote Sensing of Environment*, *97*(2), 137–162. http://doi.org/10.1016/j.rse.2005.04.007

Ruiz, L. A., Fdez-sarría, A., & Recio, J. A. (2004). Texture Feature Extraction for Classification of Remote Sensing Data Using Wavelet Decomposition: A Comparative Study. In *International Archives of Photogrammetry and Remote Sensing. Vol.XXXV*, *ISSN* (pp. 1682–1750).

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. (2005). A Crop Phenology Detection Method Using Time-Series MODIS Data. *Remote Sensing of Environment*, *96*(3–4), 366–374. http://doi.org/10.1016/j.rse.2005.03.008

Sellers, P. J., Tucker, C. J., Collatz, G. J., Los, S. O., Justice, C. O., Dazlich, D. A., & Randall, D. A. (1994). A Global 1° by 1° NDVI Data set for Climate Studies. Part 2: The Generation of Global Fields of Terrestrial Biophysical Parameters from the NDVI. *International Journal of Remote Sensing*, *15*(17), 3519–3545. http://doi.org/10.1080/01431169408954343

The World Bank. (2014). *World Development Indicators*. Retrieved from http://data.worldbank.org/

Tucker, C. J., Townshend, J. R. G., & Goff, T. E. (1985). African Land-Cover Classification Using Satellite Data. *Science*, *227*(4685), 369–375. http://doi.org/10.1126/science.227.4685.369

Turner, M. D. (2003a). Methodological Reflections on the Use of Remote Sensing and Geographic Information Science in Human Ecological Research. *Human Ecology*, *31*(2), 255–279. http://doi.org/10.1023/A:1023984813957

Turner, M. D. (2003b). Methodological Reflections on the Use of Remote Sensing and Geographic Information Science in Human Ecological Research. *Human Ecology*, *31*(2), 255–279. http://doi.org/10.1023/A:1023984813957

Van Leeuwen, M. (2001). Rwanda's Imidugudu Programme and Earlier Experiences with Villagisation and Resettlement in East Africa. *The Journal of Modern African Studies*, *39*(04), 623–644. http://doi.org/10.1017/S0022278X01003780

Van Leeuwen, W. J. D., Huete, A. R., & Laing, T. W. (1999). MODIS Vegetation Index Compositing Approach: A Prototype with AVHRR Data. *Remote Sensing of Environment*, *69*(3), 264–280. http://doi.org/10.1016/S0034-4257(99)00022-X

Wardlow, B. D., & Egbert, S. L. (2008). Large-Area Crop Mapping Using Time-Series MODIS 250 m NDVI Data: An Assessment for the U.S. Central Great Plains. *Remote Sensing of Environment*, *112*(3), 1096–1116. http://doi.org/10.1016/j.rse.2007.07.019

Wasige, J. E., Groen, T. A., Smaling, E., & Jetten, V. (2013). Monitoring Basin-Scale Land Cover Changes in Kagera Basin of Lake Victoria Using Ancillary Data and Remote Sensing. *International Journal of Applied Earth Observation and Geoinformation*, *21*, 32–42. http://doi.org/10.1016/j.jag.2012.08.005

Yin, P., Wu, M., & Liu, B. (2000). Video Transcoding by Reducing Spatial Resolution. In *2000 International Conference on Image Processing, 2000. Proceedings* (Vol. 1, pp. 972–975 vol.1). http://doi.org/10.1109/ICIP.2000.901123

# Vita

Ephraim Robert Love was born in Ypsilanti, MI, to the parents of Theresa and Jacob Love. Ephraim attended Dexter High School in Dexter, Michigan. After graduation, he traveled to Israel to study and volunteer. Upon returning to the United States, he attended the University of Michigan, where he earned the degree of Bachelor of Arts with high honors and distinction. Ephraim's undergraduate honors thesis, under the advisement of Dr. Brian Min, focused on the relationship between political institutions and public goods provision, working with remotely sensed data and GIS. This work led him to the University of Tennessee, where he studied a curriculum of remote sensing and geospatial statistics. Ephraim completed his requirements for the degree of Master of Science in May, 2015. He plans to continue developing his skills as a spatial analyst in the private sector.