

5-1-2011

FOREST CARBON MAPPING AND SPATIAL UNCERTAINTY ANALYSIS: COMBINING NATIONAL FOREST INVENTORY DATA AND LANDSAT TM IMAGES

Andrew Lawrence Fleming

Southern Illinois University Carbondale, afleming04@alumni.unity.edu

Follow this and additional works at: <http://opensiuc.lib.siu.edu/theses>

Recommended Citation

Fleming, Andrew Lawrence, "FOREST CARBON MAPPING AND SPATIAL UNCERTAINTY ANALYSIS: COMBINING NATIONAL FOREST INVENTORY DATA AND LANDSAT TM IMAGES" (2011). *Theses*. Paper 584.

This Open Access Thesis is brought to you for free and open access by the Theses and Dissertations at OpenSIUC. It has been accepted for inclusion in Theses by an authorized administrator of OpenSIUC. For more information, please contact opensiuc@lib.siu.edu.

FOREST CARBON MAPPING AND SPATIAL UNCERTAINTY ANALYSIS:
COMBINING NATIONAL FOREST INVENTORY DATA AND LANDSAT
TM IMAGES

by

Andrew Lawrence Fleming

B.S., Unity College, 2008

A Thesis

Submitted in Partial Fulfillment of the Requirements for the
Master of Science Degree

Department of Geography and Environmental Resources
In the Graduate School
Southern Illinois University Carbondale
May 2011

THESIS Approval

FOREST CARBON MAPPING AND SPATIAL UNCERTAINTY ANALYSIS:
COMBINING NATIONAL FOREST INVENTORY DATA AND LANDSAT TM
IMAGES

by

Andrew Lawrence Fleming

A Thesis Submitted in Partial
Fulfillment of the Requirements
for the Degree of
Master of Science
in the field of Geography and Environmental Resources

Approved by:

Dr. Guangxing Wang, Chair

Dr. Justin Schoof

Dr. Matthew Therrell

Graduate School
Southern Illinois University Carbondale
March 25, 2011

AN ABSTRACT OF THE THESIS OF

Andrew Fleming, for the Master of Science degree in Geography and Environmental Resources, presented on March 25, 2011, at Southern Illinois University Carbondale.

TITLE: FOREST CARBON MAPPING AND SPATIAL UNCERTAINTY ANALYSIS: COMBINING NATIONAL FOREST INVENTORY DATA AND LANDSAT TM IMAGES

MAJOR PROFESSOR: Dr. Guangxing Wang

Being able to accurately map forest carbon is a critical step in the global carbon cycle modeling and management process. This project is aimed at enhancing the current methodologies used for forest carbon mapping, and applying a method to account for any errors produced. By doing so, more accurate decisions can be made based on the knowledge gained from forest carbon maps; such as policy decisions on how to manage forests, or how to mitigate climate change. The use of remotely sensed images, in combination with Forest Inventory and Analysis (FIA) data, is one such way of doing this. This study compared three different methods; including linear regression, cosimulation, and up-scaled cosimulation to interpolate forest carbon based on a defined relationship between sample plots of national FIA data and satellite images. An uncertainty analysis was completed in an effort to quantify, and separate the different sources of error produced within a cosimulation mapping effort.

The results indicated that the band ratio of $TM4 / TM5 + TM4 / TM7$ had the highest correlation coefficient, around 0.56, with the FIA forest carbon values. At a resolution of $90\text{ m} \times 90\text{ m}$, co-simulation predicted carbon values from about 14 Mg/ha, to 135 Mg/ha. The regression model, at the same resolution, estimated carbon values from about -17 Mg/ha, to 2,400 Mg/ha. Up-scaled cosimulation at a resolution of $990\text{ m} \times 990\text{ m}$, predicted carbon values of ranging from 16 Mg/ha, to 133 Mg/ha. The

uncertainty analysis was unable to produce any statistically significant results, with all R^2 values below 0.1. These results showed that using a linear regression produced some impossible estimates, while cosimulation led to more realistic values. However, no conclusion can be made when comparing the methods based on the map validation techniques used. Although limited validation of the results was conducted, using both the FIA data and some independent sampling data; further work that focuses on validation is recommended.

ACKNOWLEDGMENTS

I would like to thank my adviser, Dr. Guangxing Wang for his assistance throughout this project. He provided an immense amount of help, and it is greatly appreciated. I would also like to thank Dr. Justin Schoof for his assistance with MatLab, and Dr. Matthew Therrell and Dr. Schoof for their comments and input on this thesis. Many thanks go out to Dr. Mark H. Hansen for his assistance with the FIA based biomass calculations. Additionally I would like to thank Ronald E. McRoberts for his assistance and support throughout this project; as well as for the unaltered FIA coordinates. Lastly I would like to thank Southern Illinois University Carbondale for the project funding.

TABLE OF CONTENTS

<u>CHAPTER</u>	<u>PAGE</u>
ABSTRACT	i
ACKNOWLEDGMENTS	iii
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER 1	1
INTRODUCTION.....	1
Background.....	1
Study Objectives.....	4
Justification.....	4
Research Questions and Hypotheses	6
CHAPTER 2	8
LITERATURE REVIEW	8
Forest Carbon Mapping Methods	8
Use of Remotely Sensed Images	12
Interpolation Techniques	13
Uncertainty Analysis	16
CHAPTER 3	18

STUDY AREA AND DATASETS	18
Study Area	18
FIA Sampling and Measuring Procedures.....	19
Additional Field Plots	21
Landsat TM Images	22
CHAPTER 4	23
METHODS.....	23
Calculation of Tree Carbon	23
Image Processing and Enhancement	28
Regression Model.....	34
Sequential Gaussian Cosimulation	34
Up-Scaled Sequential Gaussian Cosimulation	36
Accuracy Assessment	37
Uncertainty Analysis	38
CHAPTER 5	42
RESULTS.....	42
Correlation Between Images and Plot Forest Carbon	42
Regression Models	44
Cosimulation.....	48

Up-Scaled Cosimulation.....	52
Annual Carbon Sequestration.....	53
Accuracy Assessment.....	55
Uncertainty Analysis.....	58
CHAPTER 6.....	62
DISCUSSION AND CONCLUSION.....	62
Interpolation Methodologies.....	62
Uncertainty Analysis.....	64
Research Questions and Hypotheses Revisited.....	65
Conclusion and Recommendations.....	70
REFERENCES.....	75
APPENDIX	
APPENDIX A.....	82
VITA.....	84

LIST OF TABLES

<u>TABLE</u>	<u>PAGE</u>
Table 1.....	24
Table 2.....	25
Table 3.....	26
Table 4.....	32
Table 5.....	68

LIST OF FIGURES

<u>FIGURE</u>	<u>PAGE</u>
Figure 1.....	19
Figure 2.....	27
Figure 3.....	29
Figure 4.....	31
Figure 5.....	39
Figure 6.....	40
Figure 7.....	43
Figure 8.....	44
Figure 9.....	45
Figure 10.....	46
Figure 11.....	47
Figure 12.....	47
Figure 13.....	48
Figure 14.....	49
Figure 15.....	50
Figure 16.....	51
Figure 17.....	51
Figure 18.....	54
Figure 19.....	54
Figure 20.....	55

Figure 21.....	56
Figure 22.....	57
Figure 23.....	59
Figure 24.....	60
Figure 25.....	60
Figure 26.....	61

CHAPTER 1

INTRODUCTION

BACKGROUND

The anthropogenic effects on global climate change are of key interest to the scientific community, and the general public right now. As such much of the global community is looking for ways to help control the anthropogenic consequences of global climate change from further worsening. Forest carbon sinks are one such way to do this. It is well known that CO₂ is one of the major greenhouses gasses, and is contributing to global climate change (Nowak and Crane 2002). Through the process of photosynthesis, forests and other vegetation pull carbon out of the atmosphere and sequester it within their woody biomass (Neilson et al. 2007; Nowak and Crane 2002). This carbon path does not stop when carbon is sequestered within a plant. Plants also act as a conduit by which carbon is incorporated into the soil. As leaf litter and other vegetation falls to the forest floor, it can become incorporated into the soil or inserted via rhizodeposition. The decomposition of soil organic matter releases carbon back to the atmosphere, as do fires (Jandl et al. 2007; Tomlinson 2005). As such, soil too is a viable carbon sink, but adds additional complications. Generally, it is the estimation of a tree's biomass that researchers use to generate forest carbon estimates (Hall et al. 2006). Forests can either be forest carbon sinks, or sources. A growing forest is a carbon sink, whereas a dying and/or deforested forest is a carbon source (Neilson et al. 2007; Tan et al. 2007).

Ideally, the delineation of forest carbon sinks will be an all-encompassing process, that not only involves an estimation of the amount of carbon contained within

the trees of the forest, but also the amount of carbon contained below ground. While this study will not be able to quantify all the carbon pools including: soil, belowground root carbon, ground surface carbon, and aboveground tree carbon, it does create estimates and predicted values for the carbon contained within trees (both above and belowground). Nonetheless, the different ways in which carbon can be sequestered is a process that needs to be understood as thoroughly as possible if accurate estimates of the world's carbon are to be obtained.

Ocean carbon and soil carbon also have an important place in the world of carbon sequestration. Carbon is sequestered within oceans via the burial of marine sediments and is directly related to plankton productivity (Mollenhauer et al. 2004; Ruddiman 2008). It has been observed that the chlorophyll-like pigments produced where plankton productivity occurs is correlated with the amount of carbon being stored in the oceans (Mollenhauer et al. 2004). On the other hand, Feng et al. (2006) has found that the reflectance of soil can be correlated specifically to the amount of soil organic carbon (SOC) within the soil. By use of sediment cores procured via grid sampling, they determined that an accurate recreation of SOC was possible when using a satellite image with these cores. But, this methodology is not easily reproduced in forested areas, since the sediment cores that were linked to imagery were not from forests. It should also be pointed out that ocean and soil carbon sinks are not only very difficult to estimate, but even more difficult to manage. Although forest carbon is just one part of the total global carbon sink, it poses the greatest potential in being accurately mapped and managed for the purposes of carbon sequestration.

It is estimated that temperate and boreal regions of the earth hold 1-2 gigatons of carbon (Bousquet et al. 2001), which is approximately equal 15-30% of the annual anthropogenic carbon emissions (Dong et al. 2003). Forests, in general, hold 60-85% of the aboveground terrestrial carbon (Dong et al. 2003; Tan et al. 2007), and temperate and boreal forests cover 1.4-1.5 billion hectares (ha) (Dong et al. 2003). Of that 1.5 billion ha, 60% percent of the active carbon sink is found in North America (Woodbury et al. 2007). As one can see, forests have the ability to sequester a large amount of carbon (800 gigatons within the forest and soil (Sohngen and Mendelsohn 2003)), and to pull even more from the atmosphere (87 gigatons) (Sohngen and Mendelsohn 2003). Therefore, there is potential for humans to manage forests in a way to help mitigate the effects of global climate change. To do this however, it is necessary to understand the dynamics and patterns of forest carbon sequestration (Blackard et al. 2008; Chen et al. 2003; Dong et al. 2003; Hall et al. 2006; Wang et al. 2009).

The potential environmental consequences from global climate change have attracted the attention of scientists and policy makers from around the world. The scientific community is facing a great challenge on how to reduce carbon sources without a significant loss in gross domestic productivity. A potential resource that couple be deployed to reduce the human carbon footprint is to develop a forest carbon market. Obviously, this requires accurate assessment maps of global forest carbon.

STUDY OBJECTIVES

Based on the large proportion of land that is covered by forests, and the ability of these forests to hold vast amounts of carbon; it is quite apparent that more work is needed to better understand the processes of carbon sequestration in a global context. More importantly, there is a need to be able to locate potential carbon sinks. This pilot study attempts to do this by using existing Forest Inventory and Analysis (FIA) plot data in tandem with Landsat Thematic Mapper (TM) images to generate forest carbon maps in the United States for the years 2004 and 2005. Three different methods will be used to generate these forest carbon maps, and will be compared to see which one is the most accurate. The methodologies are: a regression model, sequential Gaussian co-simulation, and up-scaled sequential Gaussian co-simulation. Lastly, this study will attempt to conduct an uncertainty analysis using a polynomial regression, so that the various sources of uncertainties can be separated and analyzed. Within the uncertainty analysis, measurement error will be calculated from additional field data collected specifically for that purpose.

JUSTIFICATION

The purpose of this pilot study is to try to fill in some of the gaps that are associated with forest carbon mapping. For instance, the quantification of the spatial patterns and distribution of forest carbon is one such major gap (Wang et al. 2009). Although there is a significant amount of literature pertaining to different methods used to map forest carbon; no real consensus has been reached in the terms of a valid and

reliable method. Scaling up point data, or data from smaller map units to larger map units or regions, is another major problem associated with carbon mapping (Chen et al. 2007; Wang et al. 2009). Data that is gathered at large scales is often broad, and misses the spatial variability of the landscape. Also budget, or cost, limitations usually limit the collection of field observations at larger scales (Wang et al. 2009). On the other hand, data that are collected at a fine scale will only work on small or regional levels unless it is scaled up (or aggregated) to cover larger areas, in which case this brings about additional uncertainties and errors. Consequently, there is a need to find a way to generate accurate carbon maps, and scale up the spatial data. This study plans to answer this challenge with the comparison of the different map generation methodologies and an uncertainty analysis.

The uncertainties associated with forest carbon mapping often have to do with the techniques used to evaluate and interpolate the data, and the quality of the data used (Wang et al. 2009). These have obvious effects on the outcomes of the forest carbon map estimates that are generated. Additionally, there is also a disconnect between the physiological processes of forests, and the models that are being used to generate the forest carbon maps (Wang et al. 2009). Some studies try to handle this problem by attempting to duplicate every environmental factor and variable they can locate, and quantify it. While this is a good idea in theory, it too has its fair share of problems (which will be discussed later in this thesis). This proposed study plans to overcome some of these issues. The dissection of the different errors and uncertainties that are introduced and produced throughout the study, should help in the creation of more accurate methods in the long run.

Additionally, there is also a basic need to generate carbon maps for the United States. There have been a few studies that have attempted this (e.g. Birdsey 1992; Birdsey and Heath 1995; Brown et al. 1999; Meng et al. 2009; Woodbury et al. 2007, Zheng et al. 2008), but none so far has filled the gaps mentioned above. So, in addition to scaling up maps, this study will also be satisfying the basic knowledge gaps of forest carbon sink mapping within the United States while using adequate interpolation techniques, which propagate uncertainties when maps are up-scaled. Lastly this study will include methods that can be used for conducting uncertainty budgets. This need is well justified by the large proportion of the global temperate forests in the U.S. (Woodbury et al. 2007).

RESEARCH QUESTIONS AND HYPOTHESES

This study aims to fill some of the gaps associated with mapping forest carbon sinks, and will attempt to answer the following questions:

- 1) How should the US FIA plot data be combined with Landsat TM images when the sample plots do not match the image pixels in size?
- 2) Are the co-simulation based interpolation methods more accurate than traditional regression modeling to generate forest carbon maps, when combining the FIA data and TM images?

- 3) Will the measurement errors of forest variables contribute more to the uncertainties of forest carbon estimates, than the variation of forest carbon and spectral variables?

In researching these questions, this study poses these hypotheses:

- 1) Landsat TM images used in tandem with FIA data will not provide linkage between forest carbon from sampled locations, to unobserved locations.
- 2) Co-simulation and the regression based modeling will produce maps with equal errors while attempting to reproduce the spatial variability of forest carbon.
- 3) Uncertainties from the error associated with FIA produced forest carbon will not contribute more to the uncertainties of forest carbon estimates, than Landsat TM images.
- 4) Measurement errors will not make a significant contribution to the total uncertainty within the study.

CHAPTER 2

LITERATURE REVIEW

FOREST CARBON MAPPING METHODS

A large amount of literature exists on forest carbon that needs to be taken into consideration when developing a methodology that proposes to solve many of the problems that are associated with this type of multi-faceted project. This literature review attempts to demonstrate the need for this pilot study by pointing out several key aspects that are important for the creation of this study. The identification of what has been done in the past is the first step, followed by identifying the problems and gaps that are associated with the past studies. This literature review covers the topics of; general forest carbon mapping methodology, the gaps that are associated with forest carbon mapping, and the different types of interpolation algorithms that are available. Lastly, it discusses the uncertainty analysis.

There are a number of ways to generate forest carbon sink maps, including the conversion of forest inventory plot data straight into carbon estimates (Fang et al. 2001; Nowak and Crane 2002; Peltoniemi et al. 2006); the use of forest growth models to generate carbon estimates (Neilson et al. 2007; Ryan 1991); the use of process models (Chen et al. 2007; Chen et al. 2000a; Chen et al. 2003; Chen et al. 2000b; Running et al. 1989; Running 1994); and the use of remote sensing imagery coupled with forest inventory plot data (Dong et al. 2003; Gibbs et al. 2007; Myneni et al. 2001; Tan et al. 2007; Wang et al. 2009; Zheng et al. 2007). All of these methods have their own negative and positive aspects.

The conversion of forest inventory plot data directly into carbon estimates is a method that provides excellent local spatial resolution (Fang et al. 2001), and quite often excellent temporal resolution (Fang et al. 2001; Peltoniemi et al. 2006). This method is relatively simple and only requires forest inventory data. However, there are several flaws with this method. For instance, field data needs to be reevaluated year to year to determine if any changes have occurred (Peltoniemi et al. 2006). There are also many unknown uncertainties associated with the conversion factors from volumes to biomass, and further from biomass to carbon (Wang et al. 2009). Additionally, this methodology is only spatially explicit for the areas in which forest variables are observed (Wang et al. 2009). All of the natural spatial variability that occurs in forests, particularly in different terrains and regions, cannot be accurately accounted for with just ground data for one location, or region.

Forest growth models are a similar technique to direct conversion, except that modeling of tree growth processes is attempted. In Ryan's (1991) study, the main focus was on the rates of photosynthesis and respiration of the forested area being studied. He could then use these parameters to generate estimates of how much carbon would be flowing in and out of a forest ecosystem. Neilson et al. (2007) used a similar methodological paradigm, in that he too focused on how forests grow. In this case, however, they used their high caliber inventory data and plugged it into various simulators to produce their estimates. The carbon estimates created by these studies give excellent temporal resolution (Neilson et al. 2007), but their focus on particular areas to get the detailed data that was required, limits these studies spatially (Wang et al. 2009). Neilson et al. (2007) said that their studies results could be only regionally accurate based

on the tree species composition of that area, compared to others. Furthermore, there are also uncertainties created by these models and simulations that are difficult to identify (Wang et al. 2009).

The use of process models is another method that can be used to generate carbon estimates. These models often incorporate hundreds of parameters (Wang et al. 2009), and attempt to account for all the regional characteristics that make up an ecosystem. This method has a strong following because of its paradigm, which in theory can create a very accurate carbon estimate. Unfortunately, there are too many uncertainties associated with the different parameters used in these models, and these will greatly reduce the accuracy of the results (Wang et al. 2009). For instance, in Chen et al. (2003), they state that forests grow more slowly in colder climates, or in Running (1989) where he discusses the need to get snow melt coefficients. These examples are just two of the parameters that need to be taken into consideration when using a process model, and exemplifies the difficulty in applying such a model to a continental or even global scale. There are just too many factors that must be considered, and too much accumulating uncertainty. To use this method to generate carbon estimates for the entire globe on any kind of annual timeframe would require the input and measurement of so many different variables that it would be unusable, both financially and temporally.

One other methodology that is commonly used is the combination of forest inventory plot data with remotely sensed images. This method is promising because it has the potential to be both spatially and temporally explicit (Wang et al. 2009). The use of forest inventory data, along with satellite image data greatly increases the utility of the dataset (Zheng et al. 2007). While many studies do not get the anticipated accuracy (Tan

et al. 2007; Zheng et al. 2007), there is a lot of potential for this method in general, which is why this pilot study will be utilizing it. Ideally, with the use of a proper interpolation technique, an accurate forest carbon sink map will be produced. Also, this methodology has great replication potential, depending on the source of data. Forest inventory programs have existed in many countries for several decades. The use of free data means that it is inexpensive, and the cost-effectiveness of the method, along with the simplicity of the data (Wang et al. 2009), will enable more studies to be conducted using this method.

In general, there are still many gaps and uncertainties associated with forest carbon mapping. Many studies are trying to replicate the complex physiological processes of plants, and are just too complicated for practical use (e.g. Chen et al. 2003; Running et al. 1989; Running 1994; Ryan 1991). Other studies have had sample sizes that are too small to be statistically valid (e.g. Zheng et al. 2007), or satellite images that had spatial resolution that was too coarse (e.g. Dong et al. 2003; Tan et al. 2007) or poor quality. There have been many different types of statistical methodologies used to create the carbon estimates. The nearest neighbor technique (Kimes et al. 1999) is one such example, as well as various types of regressions (Brown et al. 1999; Dong et al. 2003; Woodbury et al. 2007 and others). However, these methods lack the ability to propagate uncertainties of a forest landscape over multiple scales (Wang et al. 2009). Many studies have also ignored an uncertainty analysis (e.g. Chen et al. 2003; Chen et al. 2007; Woodbury et al. 2007; among others). This is real need for the proper examination of the forest carbon map results, and a need to be able to increase the performance of the models.

Therefore, there is a true need to fill these gaps. Statistically sound forest inventory data with direct field measurements (Fang et al. 2001) needs to be used, and has been shown to correlate well with Landsat images (Zheng et al. 2007). The use of a methodology to accurately scale up image and field data is vital for the generation of forest carbon maps. The identification of uncertainties in a quantifiable manner to reveal their relative contributions is also very important for delivering forest carbon products (Wang et al. 2009). To this point, it can be said that there simply is a basic need to improve forest carbon sink mapping methods (Wang et al. 2009).

USE OF REMOTELY SENSED IMAGES

The use of remotely sensed images, particularly satellite images is well justified. These images provide land cover types and biophysical parameters (Chen et al. 2003) on a reliable temporal and spatial scale (Myneni et al. 2001; Running et al. 1989; Tan et al. 2007). Biophysical parameters such as Leaf Area Index (LAI) and Normalized Difference Vegetation Index (NDVI) are commonly used as surrogates to correlate the satellite imagery to the forest inventory measured biomass (Chen et al. 2003; Dong et al. 2003; Myneni et al. 2001; Tan et al. 2007; Zheng et al. 2007). NDVI is the measure of the contrast between red and near infrared solar reflection, and as such is a measure of greenness. Many studies have shown that there is a strong correlation between NDVI and biomass (Dong et al. 2003; Tan et al. 2007). NDVI does have a tendency to saturate in dense canopies, such as the canopies of tropical and old growth forests (Dong et al. 2003; Myneni et al. 2001; Zheng et al. 2007). In addition to this tendency, reflectance curves of

leaves of different species and age vary, and so there is added uncertainty from the foliage optical properties. However this added uncertainty is debatable depending on the variance thresholds being used (Knyazikhin et al. 1999).

These flaws imply the need for other methods, for example, identification and use of forest types, to get strong correlations between the different parameters (Zheng et al. 2007). Using remote sensing images, with some assistance from biophysical parameters can determine the forest cover types (Brown et al. 1999; Gibbs et al. 2007; Knyazikhin et al. 1999; Myneni et al. 2001; Neilson et al. 2007; Tan et al. 2007; Zheng et al. 2007). Land cover categories are important to help distinguish between the different carbon amounts that are present in different forest types. The most common categories are: broadleaf forests, needle leaf forests, mixed forests, and woody savannas (Dong et al. 2003; Myneni et al. 2001) with some basic alterations depending on the study area (Tan et al. 2007). Because satellites cannot directly measure forest carbon, the combination of ground data (In Gibbs et al. 2007) and the information acquired from satellite images is an important part of increasing the accuracy of forest carbon estimates.

INTERPOLATION TECHNIQUES

There are many interpolation mapping algorithms that are used to generate the forest carbon maps; however, none so far have solved many of the gaps associated with this practice. In the mapping algorithm world, there is a need for a methodology that creates maps with unbiased population and local estimates (Wang et al. 2002). In other words, there is a need for a method which will reproduce the natural variety that is found

in vegetation cover (Meng et al. 2009). A way to measure the uncertainty produced from the mapping is also needed (Wang et al. 2002; 2009). There have been studies that have attempted to do so, but rarely do they attempt to estimate relevant forest variables, such as carbon content (Meng et al. 2009).

Essentially, interpolation algorithms that use two sources of data: primary data and secondary; data provide the potential to increase the accuracy of forest carbon estimates. Previously there were many problems associated with the integration of two sources of data (Mowrer 1997), but in Recent years these problems have been dealt with through the use of new techniques (Berterretche et al. 2005; Meng et al. 2009). Using remotely sensed images (serving as a secondary variable), it is possible to create a link between observed locations, and un-observed locations when a high correlation is found between the two sources of data, and this increases the accuracy of the predicted values (Gertner et al. 2002). The regression, co-simulation, and up-scaled co-simulation are three interpolation techniques that use primary and secondary data.

Regressions are a common aspatial technique used to predict spatial forest variables (Meng et al. 2009). Regressions have been used in the past to model relationships between spectral indices and primary variables (Berterretche et al. 2005), while maintaining spatial heterogeneity within the estimated carbon maps (Gertner et al. 2002). In fact, a study by Zirlwagen et al. (2007) stated that no validation techniques are needed for a multiple regression if the coefficient of determination is close to 100%, and if all of the requirements are given (Zirlwagen et al. 2007). Zirlwagen et al. (2007) even preferred the regression to other geostatistical techniques. However, there are a number of problems with the regression. It only works well if the relationship between

the variables is highly significant (Wang et al. 2005). Moreover, the maps generated by regression do not preserve the covariance structure of the variables. That is to say, the regression is unable to replicate or track the natural variability that is found within a landscape, and how this transcends when paired with a secondary variable. Lastly, no uncertainties are generated for the mapped values when using the regression (Gertner et al. 2002).

On the other hand, co-simulation and up-scaled co-simulation (to be discussed later) produce unbiased results with minimal error variances and corresponding variance maps; and so are excellent ways to fill the gaps mentioned above (Wang et al. 2002). However, these two techniques are not widely used. Therefore there is a need to explore their utility as interpolation techniques for forest carbon mapping, and to further investigate the advantages and disadvantages. For instance, co-simulation is based on a semivariogram and Gaussian distribution (Berterretche et al. 2005), which reproduces the spatial variability of the primary variable, but does not necessarily produce the best individual predictions (Wang et al. 2002). Most importantly, the co-simulation techniques provide spatially explicit predicted values, and spatial uncertainty measures of the predicted values (Wang et al. 2009).

There are many problems associated with the up-scaling of forest carbon to large scales as well. The problems begin with the nonlinearity of the variables that each pixel contain (Wang et al. 2009), so that when values are aggregated to a coarser scale, variability found at the smaller scale is often lost (Chen et al. 2007). A methodology is needed to tackle this portion of the problem. There are several methods that are used to do this, such as the K-nearest neighbor method (Kimes et al. 1999; Tomppo et al. 2008),

block averaging, and block co-simulation (Wang et al. 2009). K-nearest neighbor and block averaging lack the ability to derive and propagate the uncertainties of the estimates from finer spatial resolutions to coarser ones, whilst block co-simulation is capable of such tasks.

UNCERTAINTY ANALYSIS

Just as there is a need to create more accurate estimates of forest carbon, there is also a need to conduct uncertainty and error budgets for the resulting maps. Forest carbon sinks can be presumed to be highly uncertain (Peltoniemi et al. 2006), and the need for a measure of reliability has been recognized (Sales et al. 2007; Wang et al. 2009). Uncertainty and error budgets can be used to calculate the relative variance contributions from different sources of error in a way similar to that of an ANOVA table (Gertner et al. 2002). By providing uncertainty budgets, studies can better inform their audiences (Heath and Smith 2000) by improving ecological interpretation, helping assess errors in a spatial context, and decreasing the losses and risks in policy and management decision making (Berterretche et al. 2005).

Several methods have been used to assess uncertainty, such as the Monte Carlo methods (Heath and Smith 2000; Mowrer 1997), Fourier Amplitude Sensitivity Test (FAST), Taylor series, Sobol's method (Wang et al. 2005), cross validation (Berterretche et al. 2005; Meng et al. 2009; Zirlwagen et al. 2007), Bayesian statistics (Deckmyn et al. 2009; Tang and Zhuang 2009) and the polynomial regression (Gertner et al. 2002). However, many of these are not actually designed for an uncertainty analysis, and few

have the ability to calculate relative uncertainty (Wang et al. 2005). Of these methods, the polynomial regression has the ability to partition the error components down to their sources (Gertner et al. 2002; Wang et al. 2005). Additionally, uncertainties can be tracked as they propagate through the model (Gertner et al. 2002; Heath and Smith 2000). In this way, an error budget can be produced to show the relative errors of the individual sources, and future studies can focus their efforts on the portions of the methodology that need the most improvement (Heath and Smith 2000; Wang et al. 2009).

CHAPTER 3

STUDY AREA AND DATASETS

STUDY AREA

This study took place in Shawnee National Forest (SNF; defined by U.S. Department of Agriculture, Forest Service (d) 2009; Figure 1). This area is located in the Southern Illinois - Illinois Ozark or Shawnee Hills section of Illinois, and resides atop unglaciated soil resulting in greater topographical variation than Illinois is traditionally known for (USDA Forest Service (c) 2009). This area ranges in elevation from 89 meters (m) to 312 m, and has slopes ranging from 0 to 46 degrees (Illinois Natural Resources Geospatial Data Clearinghouse). The average winter temperature for the area is 2.2 degrees Celsius, and the average summer temperature is 24.4 degrees Celsius. Average rainfall for the winter is 290 millimeter (mm), and average summer precipitation is 310 mm (winter defined as December through February, and summer as June through August; Illinois State Water Survey). Shawnee National Forest consists predominantly of oak-hickory stands, and cover about 107,000 ha located across 11 counties (USDA Forest Service (c) 2009), including Alexander, Gallatin, Hardin, Jackson, Johnson, Massac, Pope, Pulaski, Saline, Union, and Williamson.

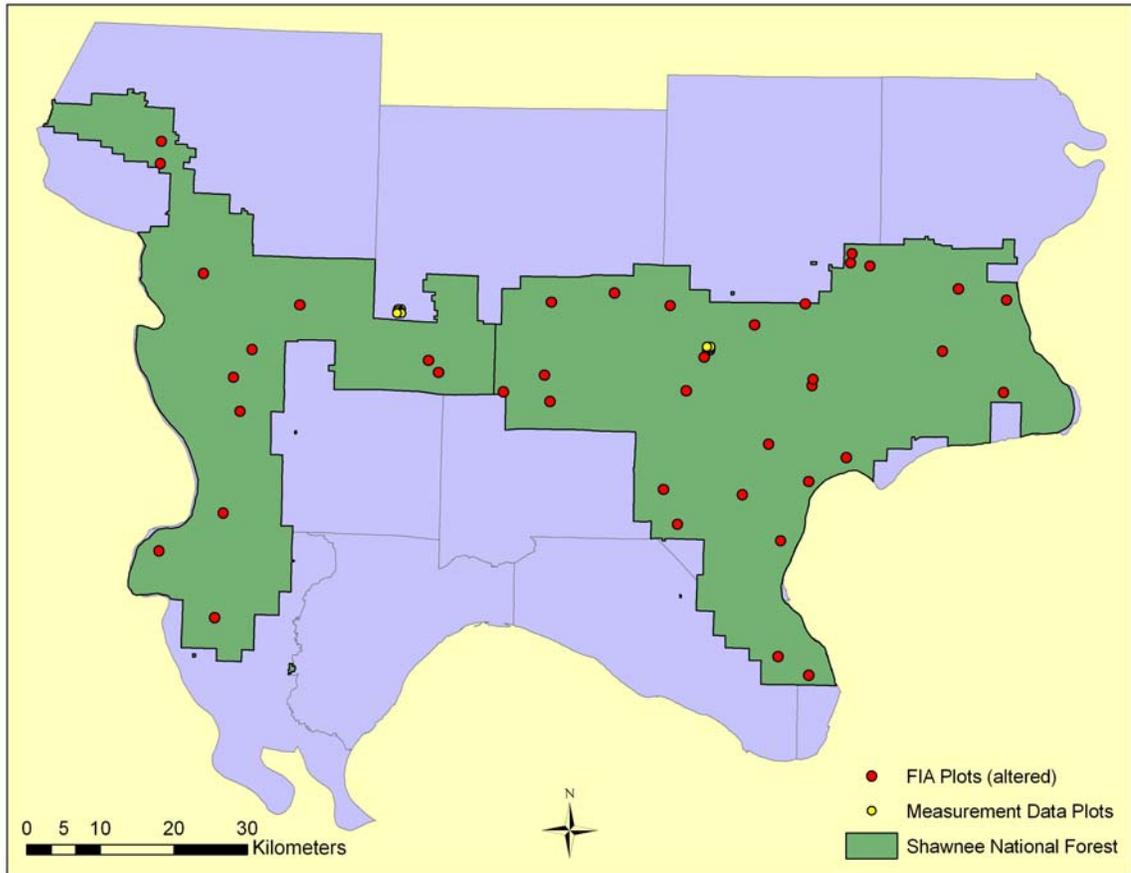


Figure 1: The study area is in Shawnee National Forest, located in southern Illinois.

FIA SAMPLING AND MEASURING PROCEDURS

Within this area, there are on average 40 National FIA sample plots that are measured each year and can be used to map forest carbon sinks. The FIA datasets were provided by the U.S. Department of Agriculture, Forest Service Datamart website ((a) 2009). The sampling design of FIA was described in detail by Bechtold and Patterson (2005). The FIA subdivides the total land area of the U.S. into mutually exclusive populations defined by county boundaries. A double sampling stratification system, that is, photo points or satellite pixels as the first stratum with field permanent plots as the

second stratum, is used. The sampling framework is based on a triangular grid, but the cells surrounding each point on a triangular grid form a hexagonal shape. Thus, the sampling framework is a network of hexagons, and each hexagon represents an area of 2,403 ha. The centers of the hexagons in a given panel form a triangular pattern of equidistant points for a five-panel rotation designed to accommodate the measurement frequency of 20 percent per year. The re-visiting time is five years. Therefore, FIA is a systematic national forest inventory sampling program with periodic measurements of forest variables and provides the potential to monitoring forest carbon dynamics.

Each FIA plots consists of four subplots with a tri-areal design. That is, one subplot is central with three surrounding subplots located 36.6 m from center to center, and at azimuths of 0, 120, and 240 degrees from the central plot, respectively. Each subplot has a 7.3 m radius, and the four subplots total 0.07 ha in size. Within each subplot, trees with a diameter at breast height (DBH) greater than 12.7 centimeters (cm) are measured. Tree species, state of health (living or dead), distance from centroid of plot, and azimuth from centroid of the plot is also measured. Tree height is measured on some trees, depending on how many trees are located within each plot. In each of the subplots there is also a microplot offset from the center of each subplot by 3.7 m , at an azimuth of 90 degrees. The 4 microplots together cover a total of 0.005 ha. In each microplot, trees with a DBH of 2.54 – 12.6 cm- are measured to the same degree as in the subplots.

The FIA plot data was downloaded from the national FIA database. Forty-one plots were obtained for the year 2004, and 38 plots for the year 2005. Each dataset contained tree heights and diameters, as well as pre-calculated volumes, biomass values,

and aboveground and belowground carbon content for the corresponding locations for the years 2004 and 2005. Normally, the released FIA plot locations are only approximate, because the Forest Service is required by law not to release the exact locations of the plots to maintain the privacy of plots on private land, and the integrity of the plots themselves (McRoberts et al. 2005). However, for the purposes of this study, we acquired the unaltered GPS coordinates for the FIA plot locations.

ADDITIONAL FIELD PLOTS

In order to quantify measurement errors and their contributions of the total uncertainty for the entire model, additional field plot data were collected in 2009. The additional field plots were designed and measured according to FIA sampling procedures (Bechtold and Patterson 2005). Two 1 km × 1 km blocks were selected at two different locations in this study area. Within each of these blocks are 9 plots laid out in grid formation with 250 m intervals, and each plot consists of 4 subplots laid out in a tri-areal design as in FIA program. Within each subplot, the forest variables were measured according to FIA sampling procedures. Within the two blocks, a total of 18 plots were measured and could be used to conduct primary validation of forest carbon estimates. It was anticipated that more additional field plots would be measured for the validation; however this was limited because of cost. In addition, within one block the 9th plot that consisted of 4 subplots was measured 9 times so that measurement errors of tree diameter and height could be calculated.

LANDSAT TM IMAGES

Landsat Thematic Mapper (TM) images of Southern Illinois were downloaded for the years 2004 and 2005 (U.S. Department of the Interior, U.S. Geological Survey). The study area is covered by two scenes of Landsat TM satellite images, so at least two images for each year are collected. The images were taken between the months of July and October, with a focus on the month of September. Cloud cover is not more than 10% in any image. For each year, two images, an east and a west, were selected with the least amount of cloud cover, and the so that the date taken was as close to one another as possible. The Landsat TM images used consist of six bands at spatial resolution of 30 m, including band 1: 0.450-0.515 μm , band 2: 0.525-0.605 μm , band 3: 0.630-0.690 μm , band 4: 0.76-0.90 μm , band 5: 1.55-1.75 μm , and band 7: 2.09-2.35 μm , and one band at spatial resolution of 120m (band 6): 10.4-12.5 μm .

CHAPTER 4

METHODS

This study aims to reduce many of the gaps and uncertainties that are associated with mapping forest carbon. To do this, existing FIA data will be used together with Landsat TM images to create forest carbon sink maps. The Landsat TM images provide the basis for the interpolation of forest carbon in sampled locations, to the locations that have not been sampled. To create the maps, three different interpolation methods will be used and compared to one another to try and discern the best methodology. Also, an uncertainty analysis will be performed using a polynomial regression to link the input uncertainty with the output uncertainty, and in doing so be able to calculate the relative contributions of the errors involved.

CALCULATION OF TREE CARBON

The acquired FIA plots had pre-calculated carbon values, but since this was not the case for the additional field plots; a recreation of the carbon calculation process was required. To do this, FIA volume and biomass equations from the FIA Database: Database Description and Users manual Version 4.0 for Phase 2 (Table 1; US Department of Agriculture, Forest Service (b) 2009), were used to calculate the total aboveground volumes for the trees based on the DBH and height. The tree biomass equations were based on species group specific parameters (Table 2; Jenkins et al. 2004), which make for more accurate biomass calculations than using tree type (coniferous or deciduous). The carbon content was calculated from the biomass using the widely used

coefficient of 0.5 (Blackard et al. 2008; Dong et al. 2003; Gibbs et al. 2007; Peltoniemi et al. 2006; Tan et al. 2007), which the FIA also uses.

Table 1: FIA biomass equations.

Tree Part	Equation
Bole Biomass	= (VOLCFSND * (BARK_VOL_PCT/100.0) * (BARK_SPGR_GREENVOL_DRYWT * 62.4)) + (VOLCFSND * (WOOD_SPGR_GREENVOL_DRYWT * 62.4))
Crown Biomass	= (total_agb - (total_agb * (stem_ratio + bark_ratio + foliage_ratio)) - stump_biomass) * AdjFac
Stump Biomass	= stump_biomass * Adjustment Factor
Root Biomass	= total_agb * root_ratio
Sapling Biomass	= (total_agb - (total_agb * foliage_ratio)) * JENKINS_SAPLING_ADJUSTMENT
Adjustment Factor	= Bole Biomass / (total_agb * (stem_ratio + bark_ratio))
Total Tree Biomass	= Bole Biomass + Top Biomass + Stump Biomass + Root Biomass
	agb = aboveground biomass
	For further explanation of equations, see citation US Department of Agriculture, Forest Service (b) 2009

Table 2: Species group specific parameters used for tree biomass calculations.

Jenkins SPGRCD	Parameters								
	total b1	total b2	stem wood b1	stem wood b2	bark b1	bark b2	foliage b1	foliage b2	
CL	-2.0336	2.2592	-0.3737	-1.8055	-2.098	-1.1432	-2.9584	4.4766	
PI	-2.5356	2.4349	-0.3737	-1.8055	-2.098	-1.1432	-2.9584	4.4766	
MH	-2.48	2.4835	-0.3065	-5.424	-2.0129	-1.6805	-4.0813	5.8816	
MO	-2.0127	2.4342	-0.3065	-5.424	-2.0129	-1.6805	-4.0813	5.8816	
Jenkins SPGRCD	Parameters (cont)								
	root b1	root b2	sap adj	stump DO	stump b1	stump b2	Wood SpGr	Bark Sp Gr	Bark %
CL	-1.5619	0.6614	0.63086	0.12667	0.914	0.11975	0.4	0.4	14.5
PI	-1.5619	0.6614	0.45789	0.08091	0.90698	0.08469	0.43	0.4	16.15
MH	-1.6911	0.816	0.79248	0.12766	0.91979	0.12152	0.5	0.51	14.35
MO	-1.6911	0.816	0.76989	0.12798	0.92267	0.12506	0.6	0.59	18.06

However, because species specific tree information was not collected, a reconstructed species group composition profile for the additional field data had to be created (Figure 2). This was done using the FIA provided species specific data (Jenkins et al. 2004, US Department of Agriculture, Forest Service (a) 2009, US Department of Agriculture, Forest Service (b) 2009). Each species from the SNF FIA data was grouped into the standard FIA species groups, from which the group percentage of the total was calculated, to see what groups were dominant (Table 3). The two most prevalent species groups for both deciduous and coniferous trees were chosen. Mixed Hardwoods (MH) and Hard Maple/Oak/Hickory/Beech (MO) were given equal representation among the deciduous trees; while Pine (PI) was given 66% of the coniferous representation, leaving Cedar/Larch (CL) with 33%. Biomass was then calculated for each species group for every measurement value using the appropriate parameters (Table 2).

Table 3: Percent species group representation from FIA SNF data, grouped by FIA approved species groups.

Jenkins SPGRP Name	Jenkins SP %
Mixed Hardwoods	43.61%
Hard Maple/Oak/Hickory/Beech	38.66%
Soft Maple/Birch	6.72%
Pine	6.53%
Cedar/Larch	2.54%
Aspen/Alder/Cottonwood-willow	1.92%

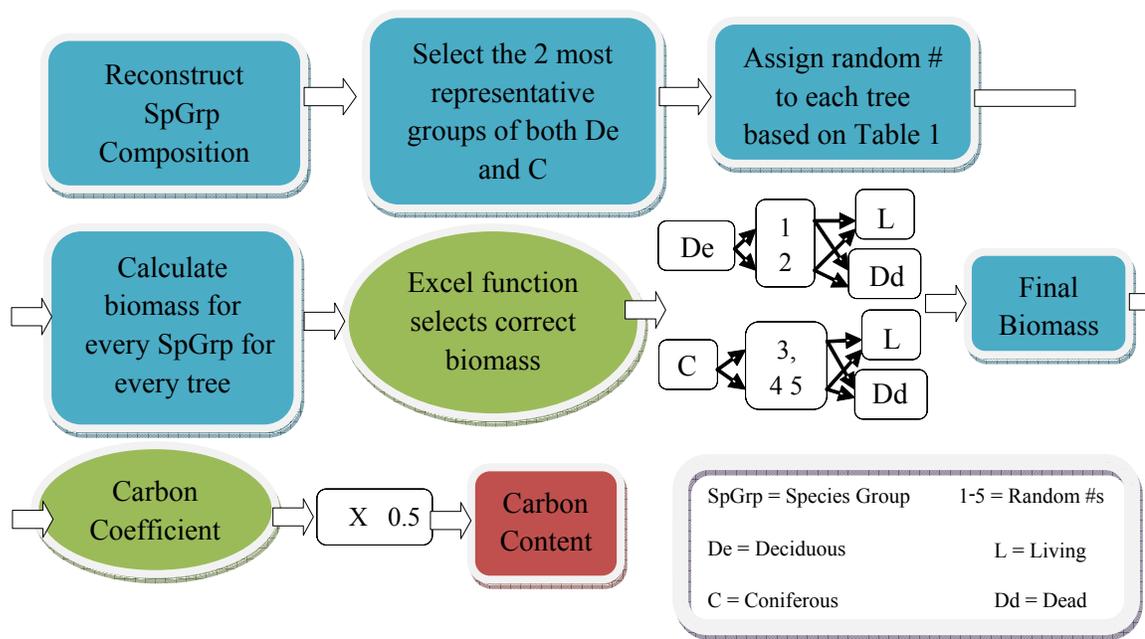


Figure 2: Flow chart representation of how carbon content was calculated for the measurement data.

Once total tree biomass was obtained for every tree and tree group, random numbers were assigned to each tree corresponding to the likelihood of that particular tree group being selected (Table 3). Deciduous trees received a random value of 1, or 2. A “1” was used to represent MH, while a “2” was used to represent MO. Coniferous trees received random values between 3 and 5. Three and “4” were used to represent PI, and “5” was used to represent CL. An Excel “IF” function was then written that would select which ever biomass total was specified depending on if the tree was deciduous or

coniferous, which random number was assigned to that tree, and whether or not the tree was recorded as living or dead. At this point the carbon content could then be calculated by multiplying the final biomass by the standard coefficient of 0.5 (Blackard et al. 2008; Dong et al. 2003; Gibbs et al. 2007; Peltoniemi et al. 2006; Tan et al. 2007). This data was then used for the measurement error tabulation, and map accuracy assessment.

IMAGE PROCESSING AND ENHANCEMENT

The Landsat TM image processing started with selecting and merging two images covering the study area for each year (Figure 3). Each east and west image, for each year was cropped from the original Landsat TM image so that only the greater Shawnee National Forest area remained. Geometric correction was done for each image, using ArcMap and topographic images obtained from Illinois Natural Resources Geospatial Data Clearinghouse (University of Illinois at Urbana-Champaign) as the reference image. The topographic maps were at a 1:100,000 scale, in Digital Raster Graphic format, and obtained from the US Geological Survey. The maps were projected in the WGS 1984 UTM Zone 16N coordinate system. The resultant geometrically corrected images had a root mean square error of 15 meters, or half a pixel size; which is an acceptable amount of error.

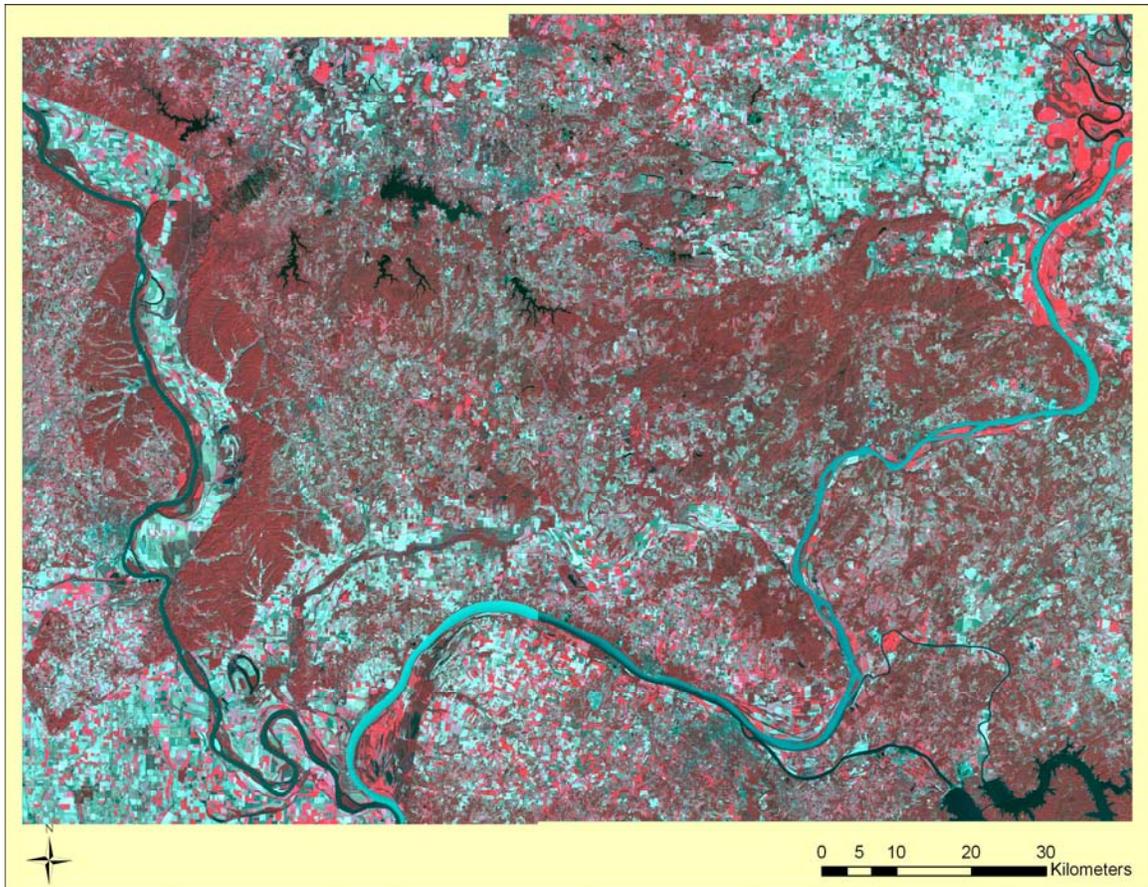


Figure 3: False color composite Landsat TM image for the year 2004. Forests and other vegetation appears as dark red, agricultural land is bright red and light blue; and water is dark blue in most cases.

Radiometric correction (haze reduction and empirical line calibration) was performed on each image, band by band. Band 6, the far-infrared band or thermal band, and was discarded. The bands were then put back together using the Layer Stack function from ERDAS Imagine (the same function that separated them). ERDAS Imagine was used to complete Haze Reduction, which uses a Tasseled Cap transformation and an absolute atmospheric correction method to remove the haze. The

layer stack function was again used to separate all the bands from each image, and Empirical Line Calibration was also conducted band by band. Empirical line calibration converted the original digital numbers of the image to reflectance values, based on the selected known white and dark bodies on each image. This too was done using ERDAS Imagine on each band of each image.

The bands were then put back together, to form one image for east, and one image for west, and were then merged together using mosaic function in ERDAS Imagine. The combined images were then subsetted (clipped) again so that all images had the same dimensions (Figure 4). The bands from each image were once again removed, and ArcMap was used to resample the image, which changed the pixel size of each image from $30\text{ m} \times 30\text{ m}$ to $90\text{ m} \times 90\text{ m}$ using a window average function. The resulting images, with a pixel size of $90\text{ m} \times 90\text{ m}$ most closely matched the size of one FIA plot; and thus it was expected that the effects of scale would be reduced (Wang et al. 2001).

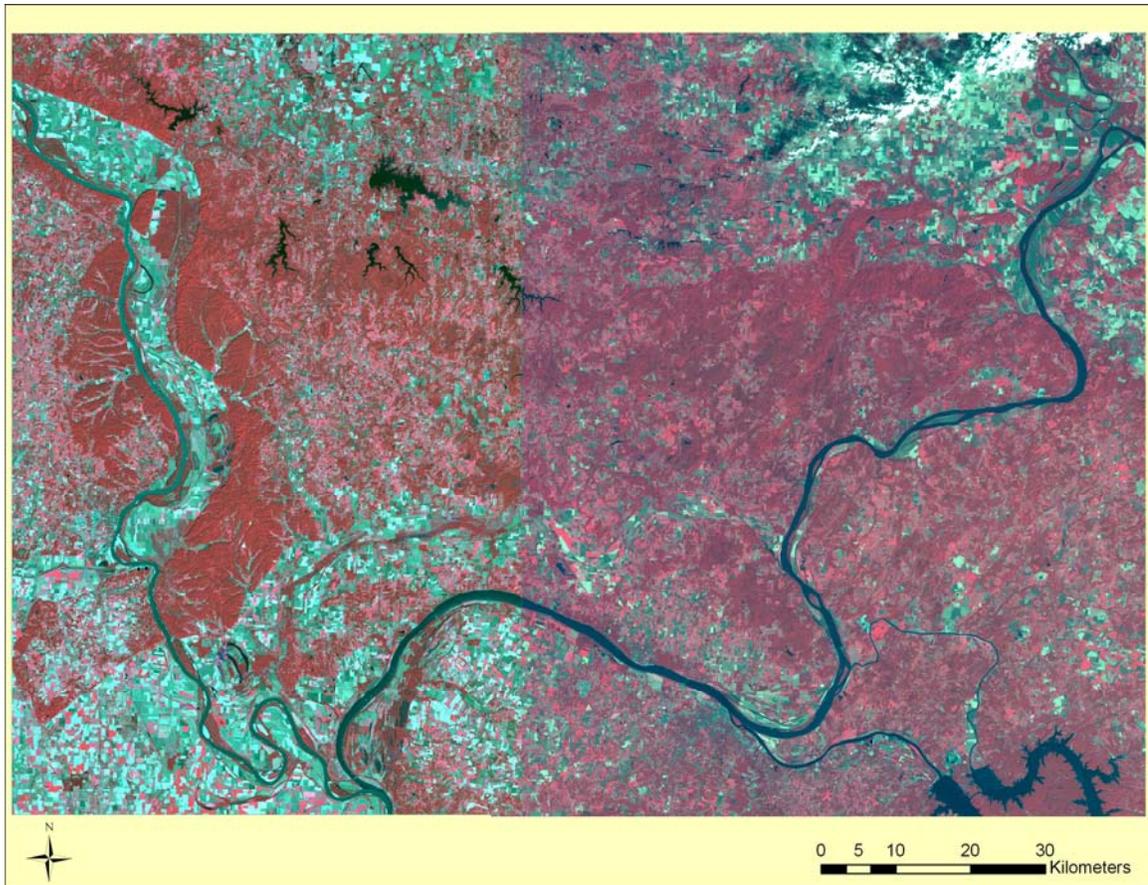


Figure 4: 2005 Landsat TM image post haze reduction, empirical line calibration, and clipped to uniform size. Forests and other vegetation appear as dark red, agricultural land is bright red and light blue; and water is dark blue in most cases.

The Landsat TM images were enhanced by creating band ratios from the different spectral bands. Various band ratios such as the Normalized Difference Vegetation Index (band 4 - band 3 / band 4 + band 3), Simple Vegetation Index (band 4 / band 3), and others were calculated (Table 4). In addition, principle components analysis (PCA) was performed on the images, and the same ratios were calculated using these images. PCA was used on the images to try and reduce the number of bands that would be needed, and

isolate the most important variables within each image. PCA transforms an original image into a smaller dataset of uncorrelated variables which represent the majority of the information from the original images. The first component accounts for the maximum proportion of the variance of the original image, with the second orthogonal component accounting for the next largest amount of variance, and so on (Jensen 2005). Thus PCA would ideally be able to isolate the most important spectral information from the original images, and these too could be used for band ratios. The information with the most variance, representing the largest amount of information could increase the correlation found between the FIA data and image, because more of the variance recorded by the satellite of the ground, could be attributed to the variance on the ground, as measured by the FIA. The images with the highest correlation to the sample plot carbon content were used for interpolation.

Table 4: Coefficients of correlation between band ratio combinations and FIA carbon values at FIA locations.

Landsat TM Band	Correlation	
	2004	2005
tm1	-0.342	-0.514
tm 2	-0.214	-0.265
tm 3	-0.254	-0.379
tm 4	0.223	0.135
tm 5	-0.374	-0.363
tm 7	-0.444	-0.485
(4-3)/(4+3) (NDVI)	0.363	0.541
4/3 (SR)	0.346	0.55
2+3+5/7	0.476	0.336
4/3+5 (SLAVI)	0.504	0.566
(1+2+3)/(4+5+7)	-0.031	-0.408

Table 4 (cont): Coefficients of correlation between band ratio combinations and FIA
carbon values at FIA locations.

Landsat TM Band (cont)	Correlation (cont)	
	2004	2005
1*5/7	0.163	-0.127
4/7	0.553	0.559
4/5	0.55	0.555
4/7*5	0.403	0.217
2/4	-0.362	-0.507
4/5+4/7	0.562	0.568
4-5/4+5	0.545	0.549
7/4	-0.519	-0.521
5/4	-0.54	-0.529
4/2	0.347	0.482
5/4*2	-0.44	-0.444
2/4+4/7	0.553	0.541
4/5+2/4	0.553	0.536
4/7+2/4	0.553	0.541
(4/5)/(4/7)	-0.422	-0.352
2/7*4	0.534	0.21
4/5*7	-0.211	-0.152
7/2*5	-0.453	-0.463
4*5	-0.161	-0.173
(2+4+5)/7	0.534	0.496
(4-7)/(4+7)	0.531	0.542
pca tm 1	-0.355	-0.32
pca tm 2	-0.271	0.306
pca tm 3	-0.466	-0.234
pca tm 4	-0.093	0.244
pca tm 5	-0.277	-0.314
pca tm 7	0.261	-0.341
pca (4-3)/(4+3)	0.177	-0.018
pca 3/4	0.179	-0.219
pca (2+3+5)/7	0.355	0.377
pca 1/2	0.496	-0.503
pca 2/3	-0.109	-0.049
pca 2/4	-0.154	0.189
pca (1+7)/(1+5)	0.07	0.196
pca 1/7+1/5	0.078	0.163

REGRESSION MODEL

Three different methods were used to generate forest carbon maps from the FIA data and Landsat TM images. The first method was the linear regression model. This method took the pixel values that have co-locations with the FIA plot tree carbon values, and generated a linear model that was used to assign carbon values to un-observed pixels. The equation $y = 19,841.39x - 21,107.50$ was defined for the year 2004, and the equation $y = 19,147.19x - 25,948.08$ for 2005. These equations use the image ratio pixel values as the independent variable, and the FIA carbon values as the dependent variable. This model has reversed the previously defined relationship between FIA carbon values and band ratio pixel values, so that the image values can then be used to estimate carbon for areas where no FIA data is available. The model will therefore be estimating carbon Megagrams (Mg)/ha values based on what the image value is, at the location the estimate is being made. This modeling was done using MatLab, and produced the first version of the forest carbon maps.

SEQUENTIAL GAUSSIAN COSIMULATION

Gaussian cosimulation is a geostatistical method that was used to produce a forest carbon map. This methodology assigns each pixel a predicted value of forest carbon (Mg/ha) based on the spatial autocorrelation of forest carbon values from both the known FIA carbon values and selected satellite image. The spatial autocorrelation was calculated using semivariograms (variograms) by determining the variance of forest carbon between any two sample locations separated by a vector called lag h (distance).

These values were then graphed with the distance between the locations as the x-axis and the variance as the y-axis: this is how the autocorrelation was calculated. The nugget, range, and partial sill of the variogram model were obtained by fitting a spheric variance model to the sample data in the graph. The nugget value indicates the within plot variation of forest carbon, while the partial sill indicates the spatial variation of forest carbon across the study area. The range suggests a maximum distance of spatial autocorrelation, and beyond this distance the data values are considered to be independent from each other. The range can be used to determine a neighborhood within which sample data can be used for predicting forest carbon for each of the unknown locations.

The forest carbon values for unknown locations were calculated by weighting each known sample plot within a given neighborhood, based on the spatial autocorrelation of forest carbon. The weighted average from the obtained values was then calculated. Within this neighborhood the predicted values, if any, were also taken into account. In addition, the image value of the co-located pixel was weighted. The obtained weights were then used to calculate a variance for the predicted value of this unknown location. A Gaussian distribution graph was created based on these predicted values and variances. A random value was then derived from the Gaussian distribution of the individual pixel, based on the probability of the Gaussian distribution. A 'random path' function was applied, to move the process on to the next pixel, until the entire image was processed.

This entire process was repeated 400 (L) times with different random paths. The number of simulations (L) was determined from the relationship between the variance of

the block estimates, and the number of simulations. Usually, when more simulations are used, less variance is associated with the average of the predicted values from all of the maps. The number of simulations chosen should be the proper balance between a smaller degree of variance, and shorter program run time (Wang et al. 2009). In this case, it was 400 simulations. To generate the final map from the hundreds of simulated maps, the average of each pixel was calculated, leading to an E-type, or final predicted value map of forest carbon.

Concurrently, the variance of the E-type predicted values for each pixel was calculated. This variance reveals the uncertainty of the predicted values, not only from the configuration of the sample data in relation to the location to be estimated, but also from the local variation of the data themselves based on the variograms. This is a measure of the true uncertainty for the predicted value at that specific location. In this study, it was assumed that forest carbon had a normal distribution. However, the presence of a normal distribution within the data was not verified. A normal score transformation was performed on the FIA plot and image data to ensure it existed, which helps to ensure autocorrelation is found within the variogram. Autocorrelation is often not found, if sample sizes are small (as in the case of the FIA data). Ensuring that the data is normally distributed allows for autocorrelation to be found, if it is present.

UP-SCALED SEQUENTIAL GAUSSIAN COSIMULATION

Up-scaling, or aggregation, means that spatial data is scaled up, or aggregated from a finer spatial resolutions to coarser ones. This process is needed when national

forest carbon maps have to be generated with large pixel sizes, (e.g. $1 \text{ km} \times 1 \text{ km}$) using the FIA sample plot data collected at smaller measurement units (e.g. $90 \text{ m} \times 90 \text{ m}$). In this study, the up-scaled forest carbon map was created using ‘block co-simulation’. This method is fundamentally the same as the regular co-simulation method described above. However, the neighborhood defined not only incorporates the value of the FIA data, but also the values of the image kernel. The image kernel defines the size of the up-scaling that will occur (so a 3×3 kernel will up-scale an image pixel size three times). The neighborhood with the FIA values, and kernel image values, is then used to produce the variogram and Gaussian distribution, just like in sequential Gaussian cosimulation. The processes from here on out is the same, except that the final predicted value is assigned to the up-scaled block size, instead of the center kernel pixel. In this study, this block co-simulation was done with a kernel of 11×11 pixels, that is, a window of $990 \text{ m} \times 990 \text{ m}$. The kernel’s size depends on how large a pixel is required for the final up-scaled map result.

ACCURACY ASSESSMENT

The next step was to conduct an accuracy assessment and uncertainty analysis. Validation was used to determine the estimation error of the different methods, using two sources of data. One validation was done by comparing the FIA values used to create the maps, with their collocated pixel values. A known location was presumed to be unknown and estimated, or predicted, using one of the previously mentioned methods (Sales et al. 2007). The known values and predicted values were then compared. This

was done for all of the FIA plot locations using each of the three methods. The second validation was done in a similar manner, by comparing the additional field plot values of forest carbon to their corresponding pixel estimates or predicted values. Because the additional field plot values were collected in 2009, 4 to 5 years of carbon accumulation had to be subtracted from the values so that they could accurately be compared to the carbon estimation maps. The figure of 1,188 kg/ha/yr of carbon accumulation was used, which is an average amount of carbon accumulation per year in Illinois forests (Birdsey 1992). The accuracies of the maps from the three methods were then compared.

UNCERTAINTY ANALYSIS

A polynomial regression was developed to create a relationship between the uncertainties from the input datasets to the uncertainties of the output estimates. Three different primary sources of error were taken into account: measurement error, the variation of forest carbon observations within the FIA data, and the variance from the Landsat TM band ratio images. The different sources of error were quantified by using a Fortran program that calculated the variance between pixel values or sample plots. The image variance of each pixel for the map at pixel size of 90 m was calculated using the pixel values at the resolution of 30m × 30m within a window of 3 × 3 pixels. The input forest carbon variance of each pixel was derived using 3 to 5 sample plots depending on how many sample plots were used for predicting this location. This program could be run directly on the band ratio images, and the FIA plot data. However, additional steps had to be taken to produce the raster variance image output for the measurement error.

The measurement error data were collected from one small plot, so it could not be used directly for the entire study area. Instead, a relationship was defined between the average tree diameter, and the variance from the carbon content of those trees (Figure 5), with values beyond one standard error removed. The defined relationship could then be used to extrapolate the measurement errors to FIA plot data locations. At this point, the variance between the measurement error values at the FIA plot locations could be calculated using the same Fortran program used on the other two sources of error.

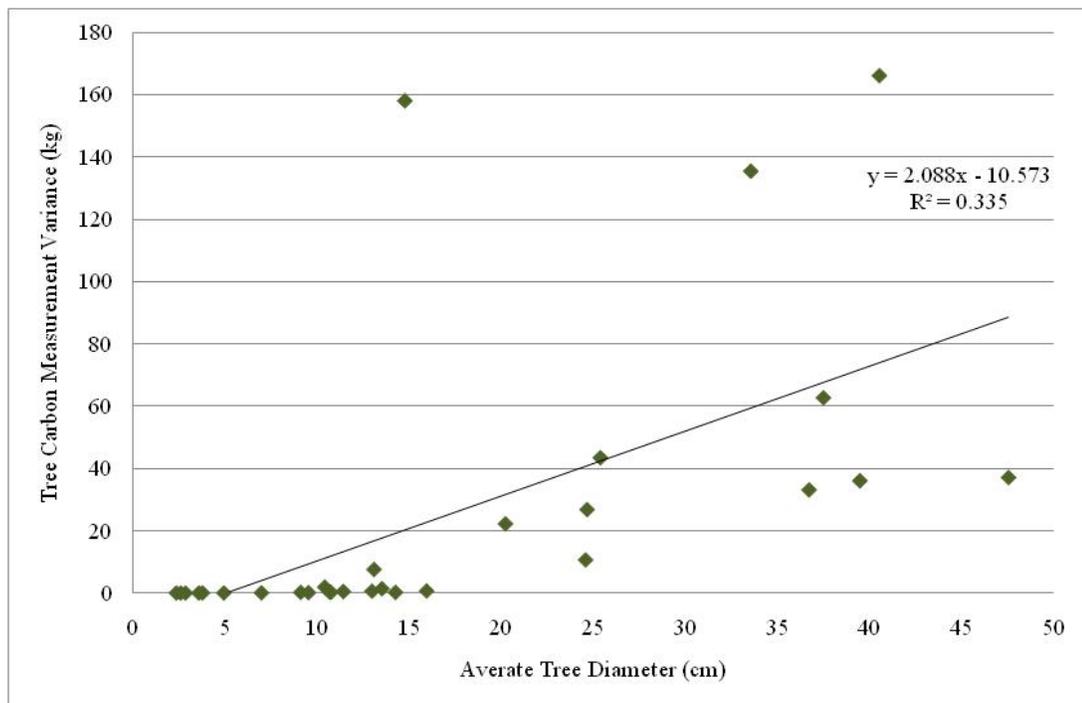


Figure 5: Relationship between the average tree diameter and tree carbon variance, which was used to extrapolate the localized data to the FIA data.

A random grid consisting of 379 points was then created to cover all of southern Illinois. For each of the random points, the three errors mentioned above were extracted

(Figure 6). The same random points were also used to extract the variances of predicted forest carbon values from the co-simulation variance maps. The different sources of input error were then used as independent variables to create a polynomial regression with the output variance as a dependent variable. In this way, the defined relationship was used to calculate the relative contributions of the input uncertainties to the output uncertainties.

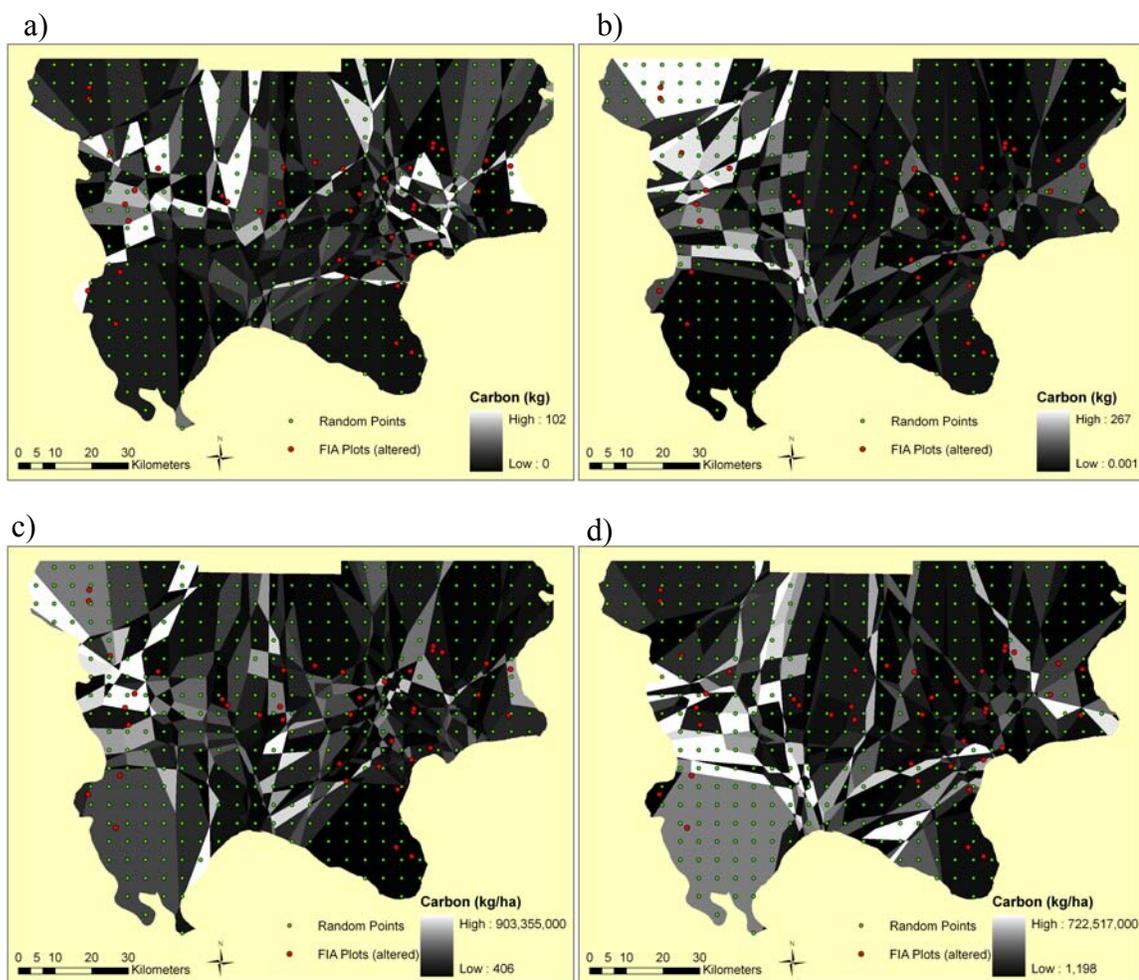


Figure 6: Figure and caption continue on the next page.

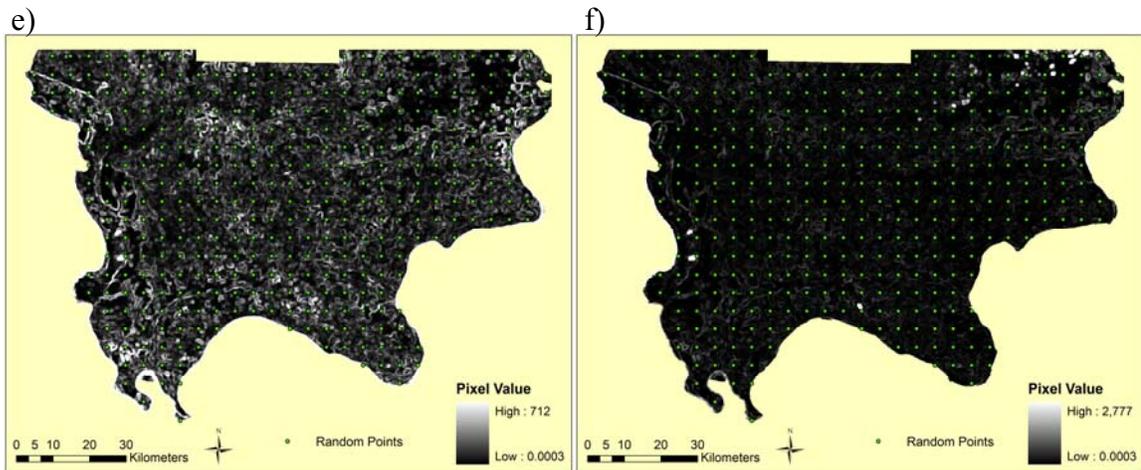


Figure 6 (cont): Random grid, and images from which the uncertainty analysis data were collected from. Image (a) displays the 2004 measurement variance, (b) the 2005 measurement variance, (c) the 2004 FIA plot variance, (d) the 2005 FIA plot variance, (e) the 2004 TM image band ratio variance, and (f) the 2005 TM image band ratio variance.

The TM image band ratio variance is displayed using histogram equalization, which emphasizes the variance so that it can be seen.

The above methodology was used to produce three different forest carbon maps, and two forest carbon uncertainty maps per year; as well as an uncertainty analysis. These were then compared to calculate the relative differences between the techniques. The uncertainty analysis, or budget, with the help of the technique comparison can then allow for an improvement in the forest carbon mapping methodologies, providing reinforcement to the use of co-simulation and up-scaled co-simulation as an accurate way to produce carbon maps by combining forest inventory data and remotely sensed images.

CHAPTER 5

RESULTS

CORRELATION BETWEEN IMAGES AND PLOT FOREST CARBON

Based on Table 4, from all the original and transformed band ratio images, the band ratio: (band 4 / band 5 + band 4 / band 7) had the highest correlation coefficient with plot forest carbon for both years 2004 and 2005. The correlation coefficient for the year 2004 was 0.562, which is statistically significantly different from zero using the equation $r_\alpha = \sqrt{\frac{t_\alpha^2}{(n-2+t_\alpha^2)}}$ at $\alpha = 5\%$, and $n = 41$. The correlation coefficient for the year 2005 was 0.568, which is also statistically significant at the 5% significant level using the same equation, but when $n = 38$. The images produced from the band ratios are shown in Figures 7 and 8. Significant correlations were found because the bands used are in the infrared spectrum, which is a spectrum that often correlates well with vegetation (Jensen 2005). Vegetation, particularly tree leaves, reflect significantly more infrared light than visible light (Jensen 2005), so by using bands with these properties; more ground characteristics of vegetation could be extrapolated out of the pixel values. It is likely that this ratio was able to discern forest density and age from the reflectance pixel values which, which could be attributed to why a significant correlation was found.

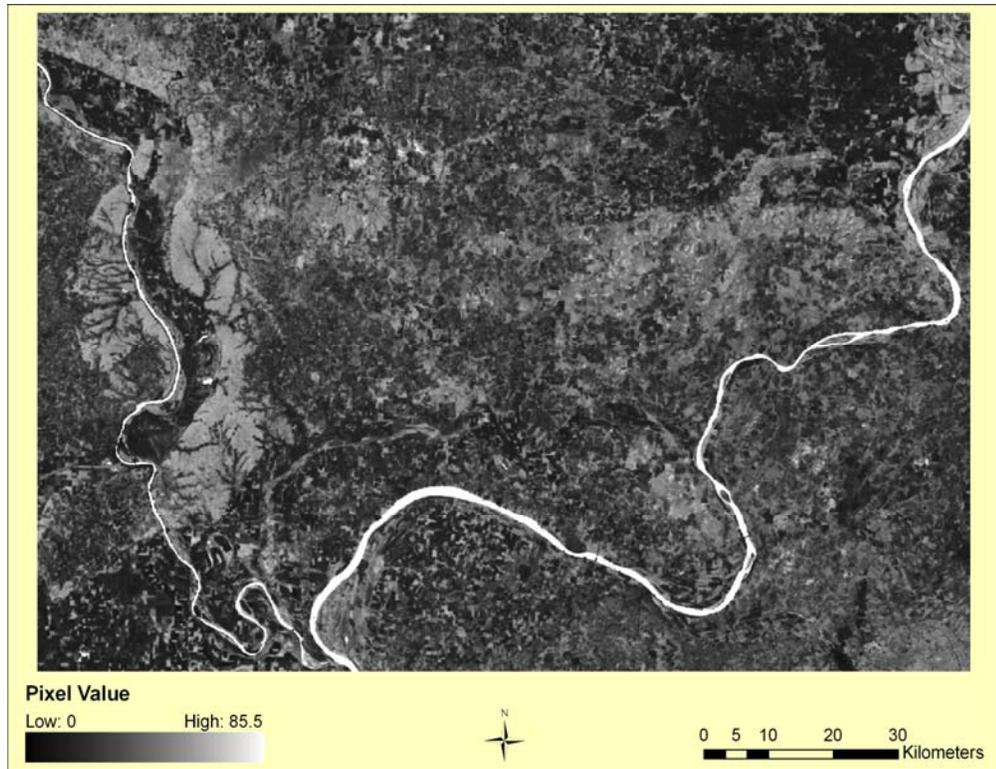


Figure 7: 2004 band ratio image displaying band combination (band 4 / band 5 + band 4 / band 7).

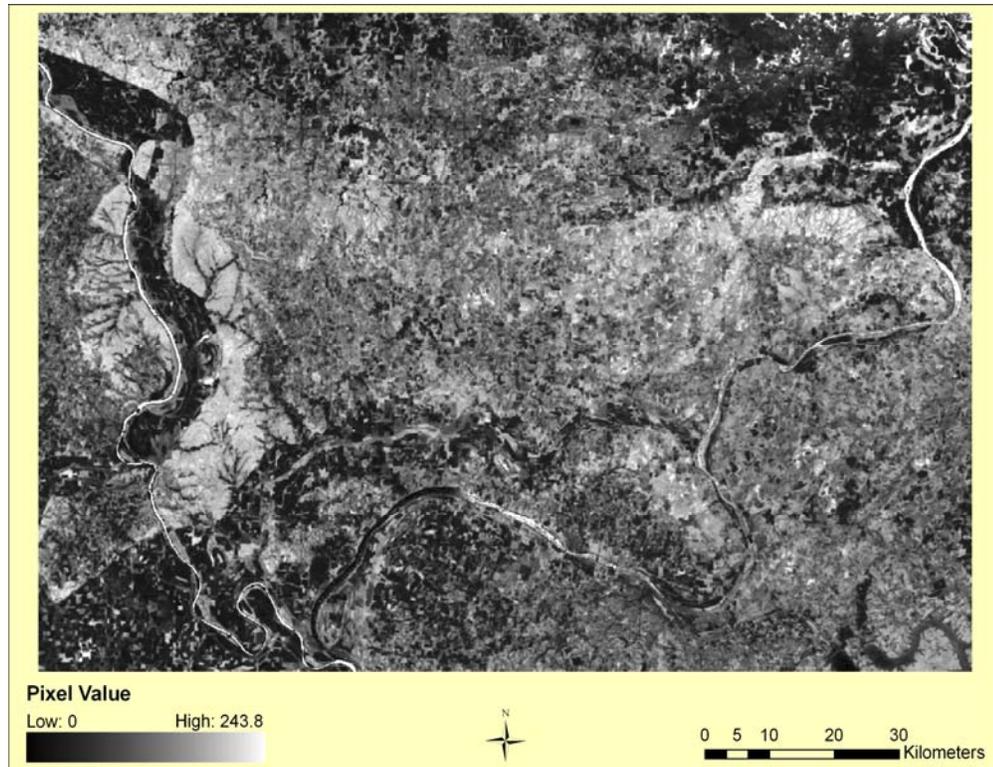


Figure 8: 2005 band ratio image displaying band combination (band 4 / band 5 + band 4 / band 7).

REGRESSION MODELS

The initial step in the usage of the regression model was the calculation of the relationship that existed between the FIA plot forest carbon values and the collocated image pixel values (Figures 9 and 10). This relationship, as defined in the previous results section, is significantly different from zero. This means that relationship is significant enough to proceed with the regression modeling. This will be done however, knowing that this relationship does not explain about 70% of the variance.

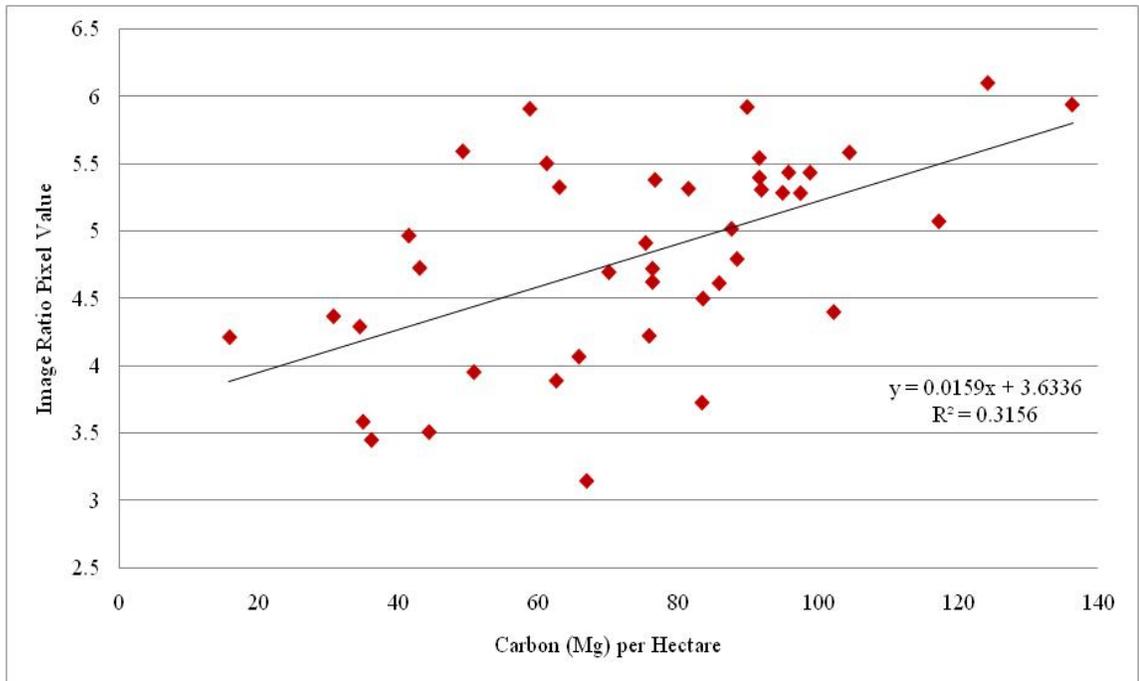


Figure 9: The relationship between forest carbon (kg/ha) and the image ratio, which represents the regression model used to create a forest carbon map for the year 2004.

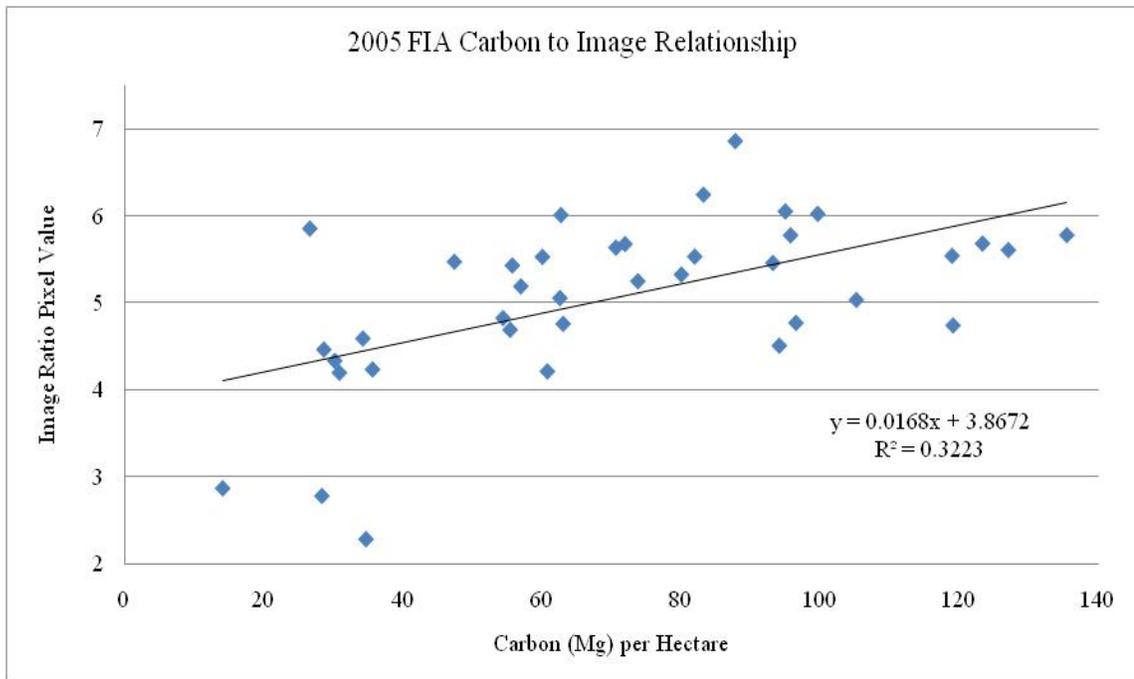


Figure 10: The relationship between carbon (kg/ha) and the image ratio, which represents the regression model used to create a forest carbon map for the year 2005.

The regression outputs for both years (Figure 11) depicted a visually agreeable representation of the Southern Illinois forests. High levels of forest carbon were estimated for areas where there are dense forests, and low levels of carbon were estimated for areas that are dominated by crops. However, some of the actual values estimated by the regression model do not appear as reliable. The highest estimated carbon was almost 1,700 Mg/ha for the year 2004, and over 2,400 Mg/ha for the year 2005. Additionally, this method's estimated low values were negative; with -9 Mg/ha for 2004, and -17 Mg/ha for 2005. Obviously these extreme values are not physically possible. In the 2004 result, these extreme values do not seem actually represent the average values corresponding to forest landcover or agricultural landcover, but to

extraneous places such as a small lake or edge of a river. However the 2005 map does have negative values associated with a fair amount of agricultural land. Ad hoc averages of the high carbon estimated areas (forests) ranged between 60 to 100 Mg/ha, and around 2.4 Mg/ha for low carbon estimates (agricultural fields).

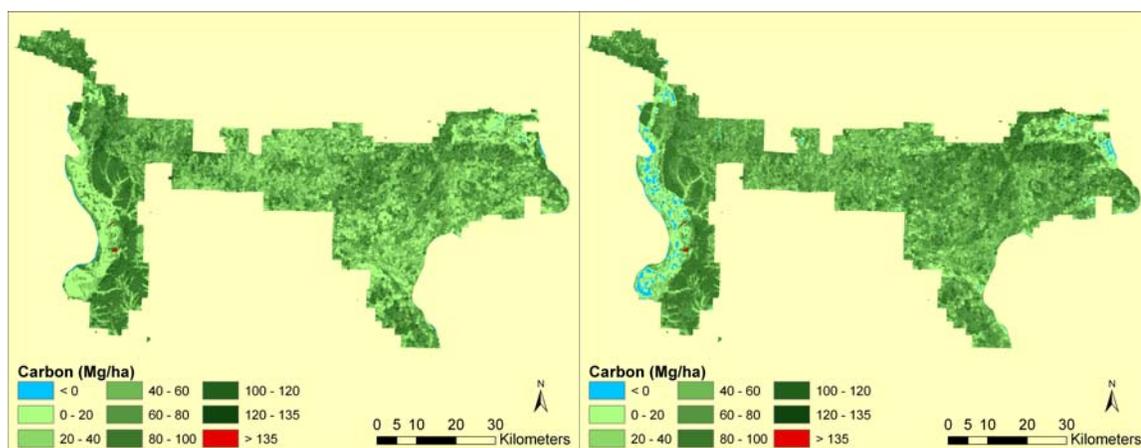


Figure 11: 2004 and 2005 regression based spatial distribution of forest carbon.

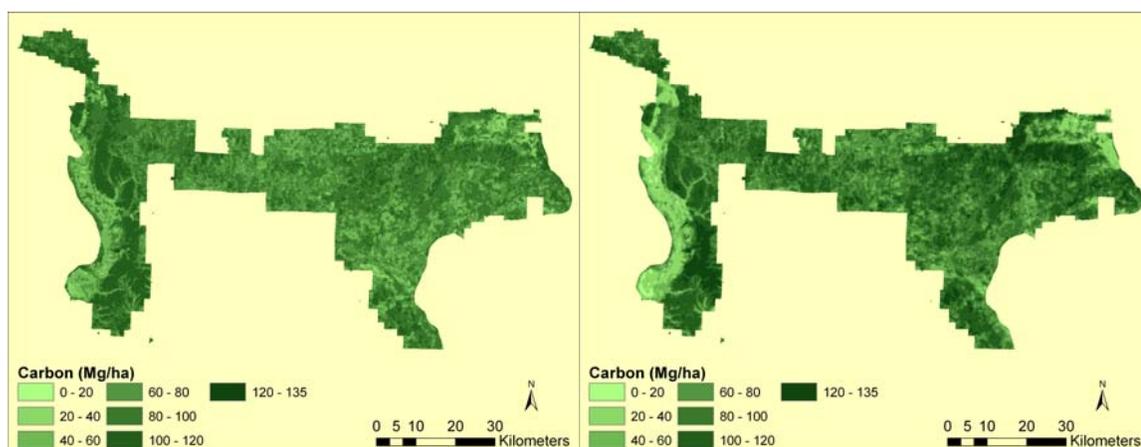


Figure 12: 2004 and 2005 cosimulation based spatial distribution of forest carbon.

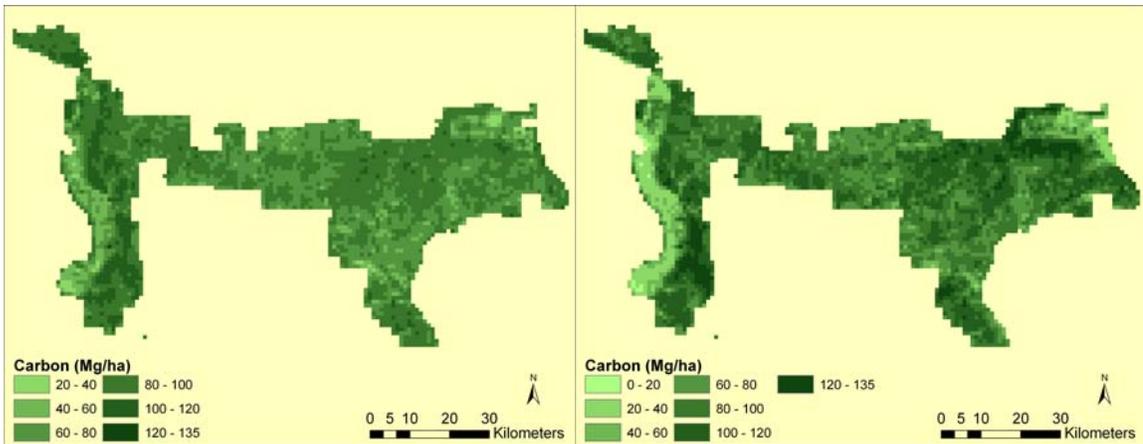


Figure 13: 2004 and 2005 up-scaled cosimulation based spatial distribution of forest carbon.

COSIMULATION

An important component of the cosimulation of forest carbon was the creation of a variogram for each image; from which the forest carbon maps were partially generated. Figure 14 depicts the omnidirectional, standardized variogram for the year 2004. A large nugget (0.85) was observed, indicating that there is a high degree of variance within the sample plots, regardless of distance between the plots. This indication of high ‘background’ variance was supported by the presence of both a short partial sill (0.15), and short range (13,200 m): indicating that only a small amount of autocorrelation exists between the FIA plot points. This large nugget was partly caused by large distances between the FIA sample plots, and partly by the FIA plot configuration. Because the FIA sampling procedure is based around the measurement of trees within subplots and plots that are equal distances apart; it cuts back on the amount of autocorrelation that can be

found. That is to say, with so few differences in the distances between plots, a standard autocorrelation trend is difficult to measure. This is not made easier by the fact that the FIA plots are spaced widely apart, which gives the models a kind of “all or nothing” sort of data. Still, the year 2005 variogram (Figure 15) showed a smaller nugget (0.36), and a larger partial sill (0.64), which indicated that greater autocorrelation within a range of 14,400 m was observed.

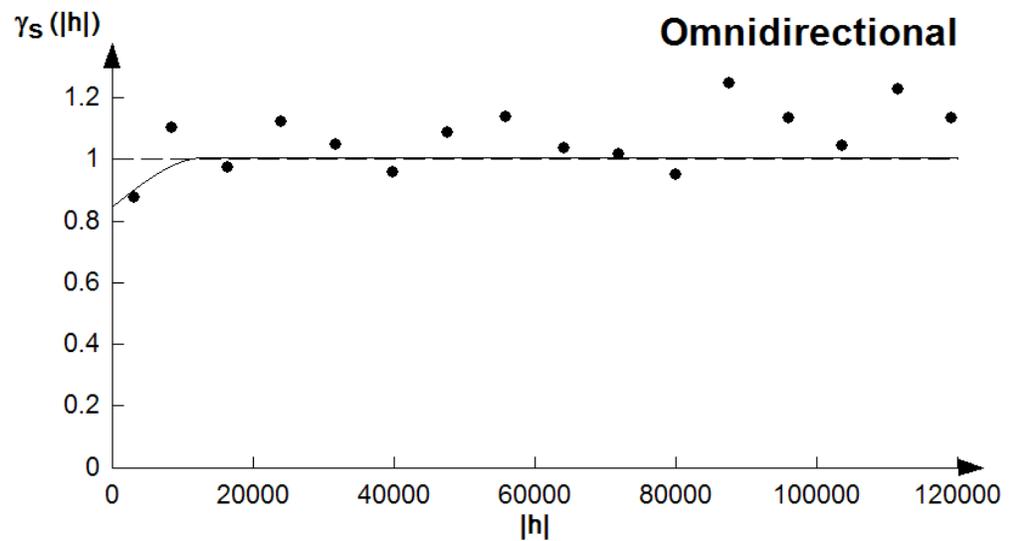


Figure 14: Standardized variogram representing the autocorrelation present between forest carbon plots in the year 2004. “ h ” represents distance in meters.

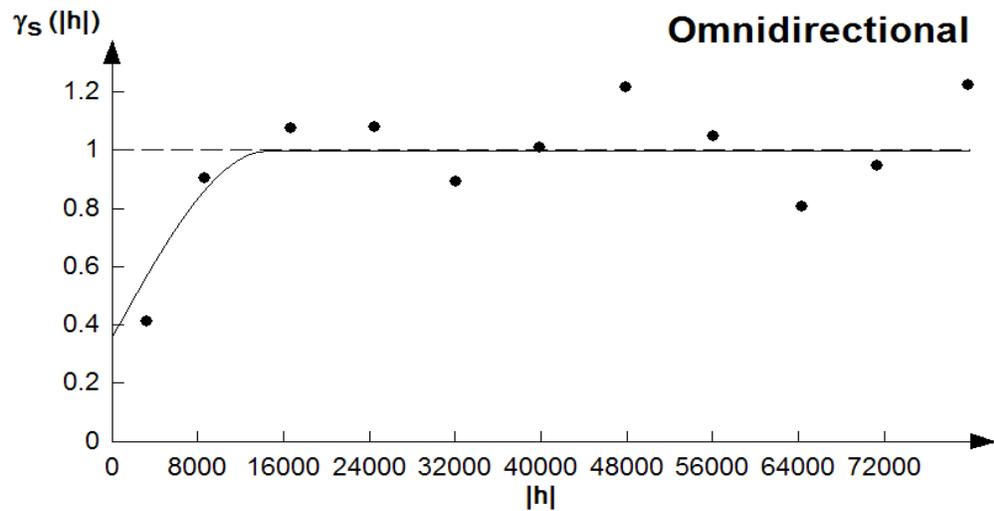


Figure 15: Standardized variogram representing the autocorrelation present between forest carbon plots in the year 2005. “ h ” represents distance in meters.

The cosimulation outputs (Figure 12) also produced carbon representations that mimicked the spatial variability of Shawnee National Forest. Here however, as compared to the regression outputs, the predicted minimum and maximum carbon values are within reason. The ad hoc average of forest carbon from the cosimulation outputs ranged from 100 to 112 Mg/ha for the heavily forested areas, and were around 17 to 22 Mg/ha for the dense agriculture areas.

Uncertainty maps were produced for each cosimulation predicted value map. These maps (Figures 16) show the standard deviation (Mg/ha) of each predicted value. Low standard deviations occur in areas with extremely low or high predicted values (e.g., dense forests tracks or highly agricultural areas). More moderate to high standard deviations are seen in areas with average predicted values. This larger uncertainty could

be because of the mixed vegetation cover that dominates southern Illinois, and the lack FIA data to extend the image to ground correlation, to this type of ground cover. The map for 2005 depicts low variances associated with low and high predicted values more strongly than does the 2004 map, which likely corresponds to more autocorrelation being found within the 2005 variogram.

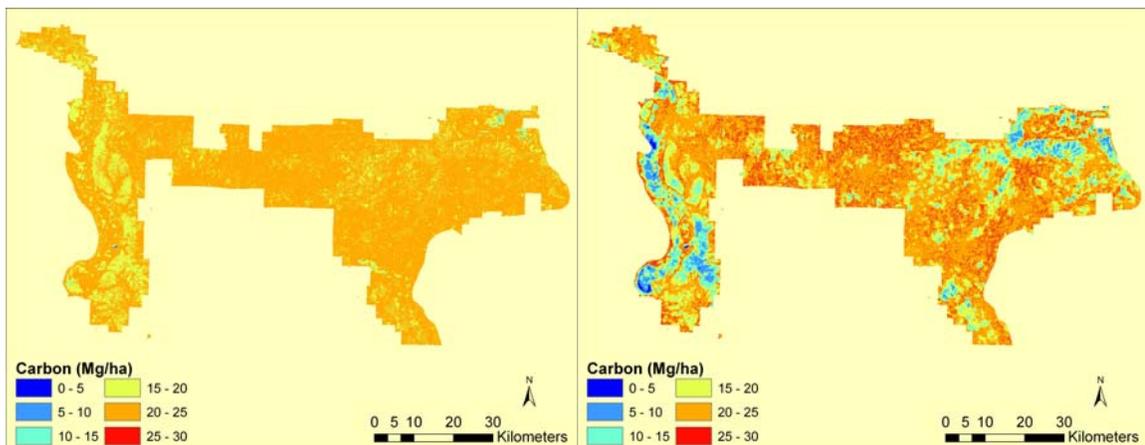


Figure 16: 2004 and 2005 cosimulation based spatial distribution of forest carbon predicted value standard deviations.

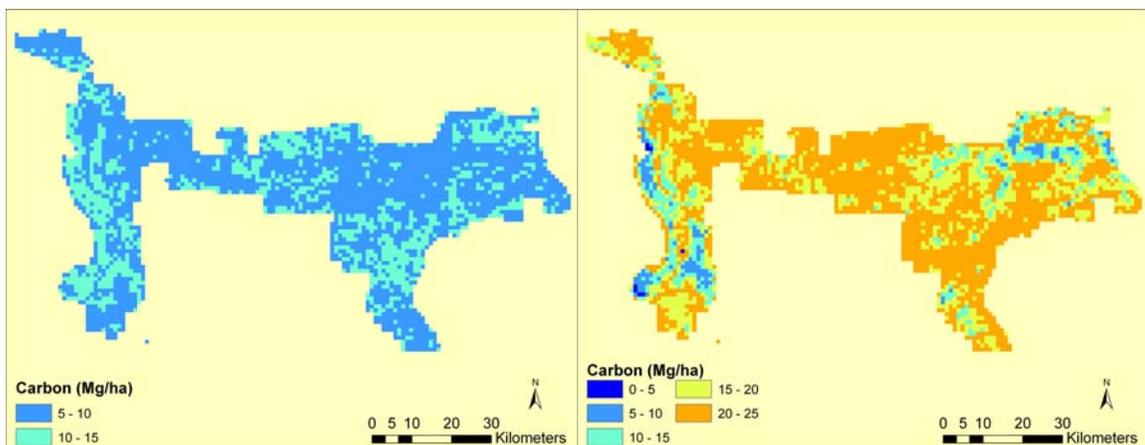


Figure 17: 2004 and 2005 up-scaled cosimulation based spatial distribution of forest carbon predicted value standard deviations.

UP-SCALED COSIMULATION

The up-scaled cosimulation maps (Figure 13) produced results very similar to cosimulation without up-scaling. This does not come as a real surprise since the methods are so similar. These maps, up-scaled to nearly 1 km resolution (990 m) are still able to reproduce the natural spatial variability of forest carbon, just as the maps without up-scaling did. The resulting forest carbon predicted values are very similar to that of the non up-scaled maps.

The up-scaled cosimulation standard deviation output maps (Figure 17) also resemble the non up-scaled cosimulation outputs. However, it is worth noting that the 2004 standard deviations have noticeably lower variability, with respect to the high values recorded in the 2005 map (for example, 14 Mg/ha in 2004, compared to 26 Mg/ha in 2005). Additionally, low variability is observed not only in areas where extreme values are predicted, but also across the entire Shawnee National Forest study area. The 2005 standard deviation map has similar spatial patterns to 2004, but the spatial variability is more extreme.

These maps of predicted forest carbon values provide insight into the spatial distribution of forest carbon, which can enhance our ability to make rational decisions for forest management. If the error produced by these maps is brought to an insignificant level, they will provide the ability to track the changes in forest carbon sequestration over time; which will greatly enhance our understanding. In addition, the output variance maps of predicted carbon values can be used for conducting uncertainty analysis to further this understanding.

ANNUAL CARBON SEQUESTRATION

To assess any potential gain in carbon sequestration, the forest carbon maps were subtracted from one another (2005-2004). This process reveals how much carbon is lost or gained in one year. All three interpolation techniques produced annual carbon accumulation maps that were comprised of losses and gains in forest carbon (Figure 18-20). All three maps show the western edge of SNF to have lost carbon, and the eastern portion of SNF to have gained carbon. The areas where forests dominate, seem to have generally had an increase in carbon, while the areas dominated by agriculture seem to have lost carbon. However, these results are only displayed for demonstrational purposes, since the error associated with the estimated and predicted values of the forest carbon maps could very well account for the changes measured here. To be able to accurately measure the changes in the amount of carbon stored in sinks, more sample plots are needed to produce accurate maps.

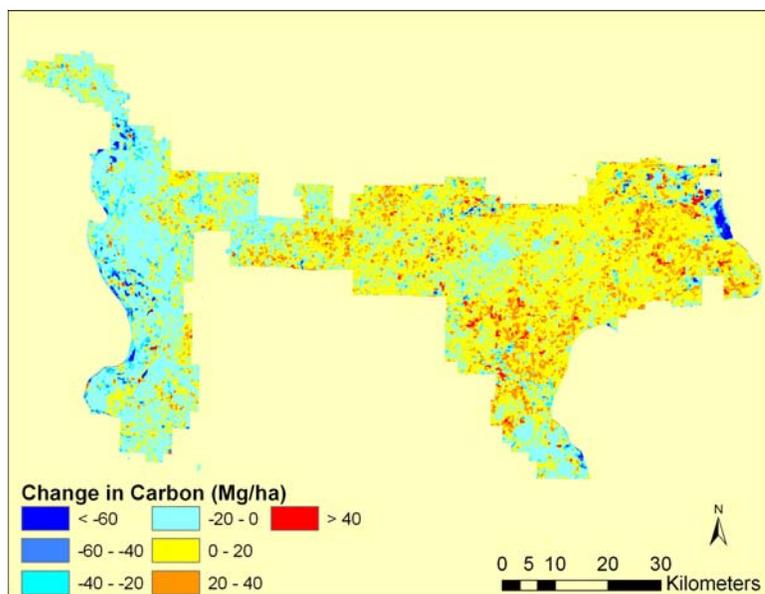


Figure 18: Changes of forest carbon between the years 2005 and 2004 based on the regression interpolation technique.

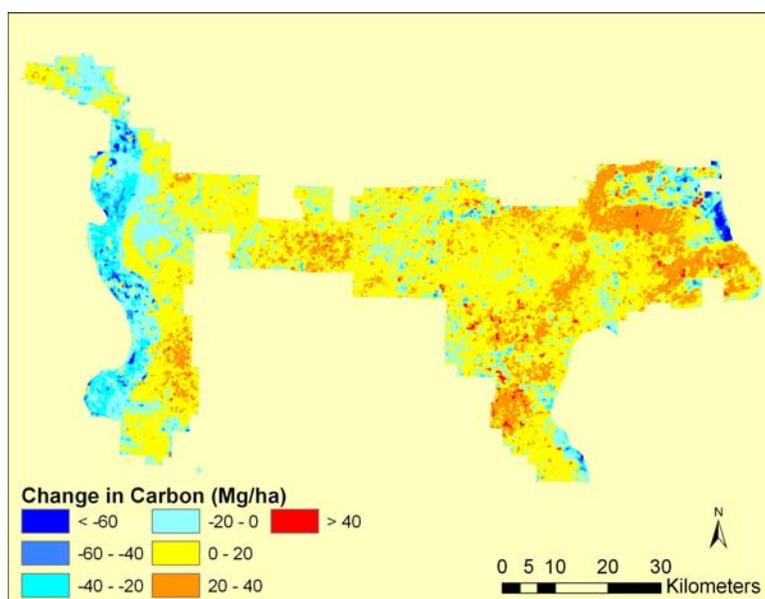


Figure 19: Changes of forest carbon between the years 2005 and 2004 based on the cosimulation interpolation technique.

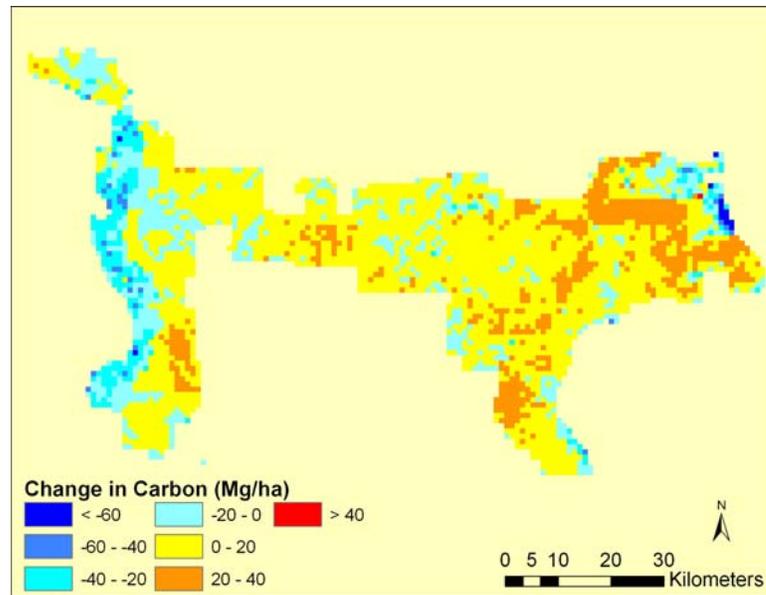


Figure 20: Changes of forest carbon between the years 2005 and 2004 based on the up-scaled cosimulation interpolation technique.

ACCURACY ASSESSMENT

Two kinds of validation were calculated on the maps produced by this study. Validation was performed using the FIA plots used for the interpolation (Figure 21), and from the additional field plots collected by students at Southern Illinois University (Figure 22). The validation using the FIA plots and validation using the additional field data yielded similar, yet contrasting results. Validation using the FIA plots estimated an average regression bias of 1.8% for the two years the maps were produced. Bias is defined as the percent difference between the actual values of the FIA plots, and the estimated values. A small bias indicates a small error associated with the created estimates. Thus the regression model produced accurate maps according to the bias

produced from the validation using FIA data. The cosimulation and up-scaled cosimulation produced overall averaged biases of 28.6% and 28.7%. The overall averaged standard errors for the regression model, cosimulation, and up-scaled cosimulation were 5.6, 4.3, and 8.3 Mg/ha, respectively.

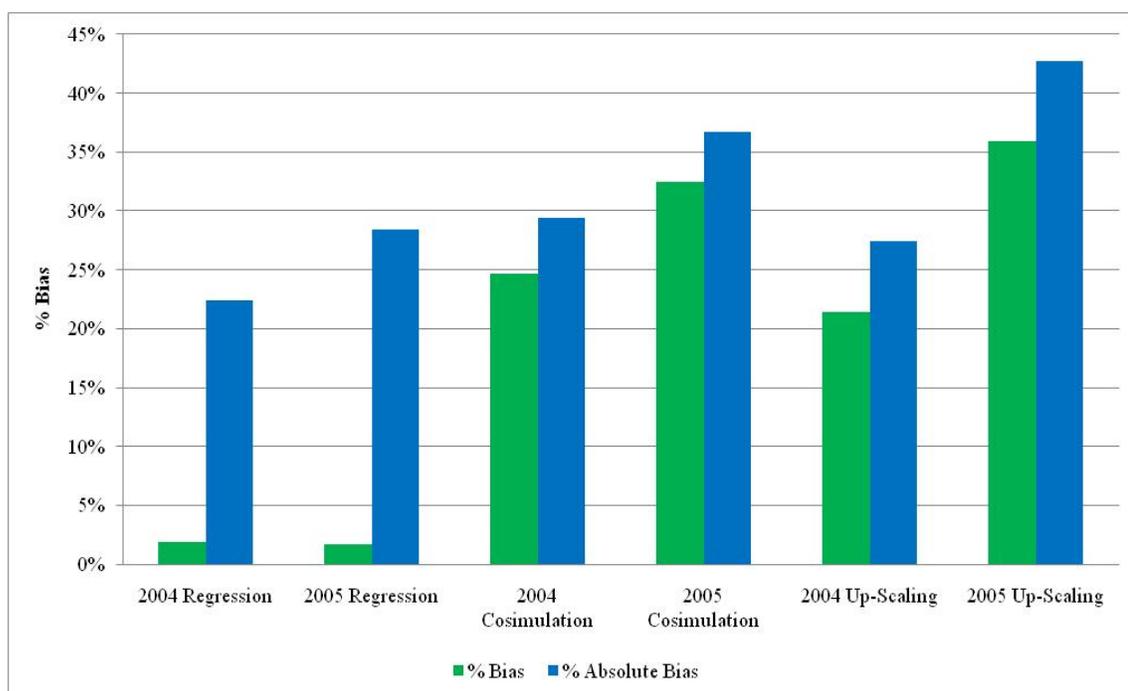


Figure 21: Validation by use of the FIA plot data.

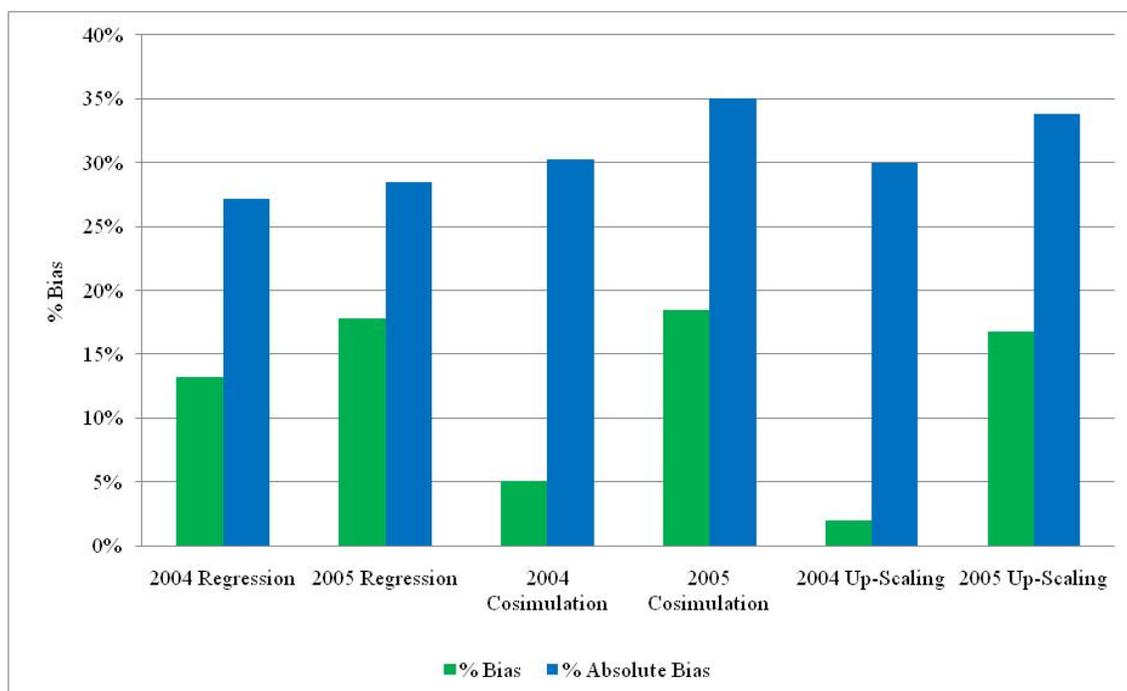


Figure 22: Validation by use of the additional field plot data.

The validation results using the additional field plot data reveal an overall average bias of 15.5% for the regression, 11.8% for the cosimulation, and 9.4% for the up-scaled cosimulation. The up-scaled cosimulation, although limited in collocated pixels to compare in the validation; still allows a limited accuracy assessment to be made. It is also worth noting that the cosimulation techniques are the most accurate based on the validation using the additional plot data. The standard error for the regression model was 9.9 Mg/ha, 6.6 Mg/ha for cosimulation, and 4.3 Mg/ha for up-scaled cosimulation.

Finally, an absolute bias was calculated to assess each of the interpolation techniques. The absolute bias is defined as the percent absolute difference between the measurements and the predicted values in the maps. In this way, absolute bias does not

allow the negative and positive differences between the ground truth and predicted values to cancel one another out. When comparing the absolute biases, both validations using FIA data and the additional field plot data yield similar results. Validation using FIA data's absolute bias of the regression model, cosimulation, and up-scaled cosimulation was 25.4%, 33.1%, and 35.1%, respectively. Similarly, the validation using the additional field plot data resulted in absolute bias of 27.8%, 32.7%, and 31.9% for the regression model, cosimulation, and up-scaled cosimulation, respectively.

UNCERTAINTY ANALYSIS

A correlation coefficient of 0.579 between the average tree diameter and tree carbon variance in Figure 5 was found to be statistically significant at the 95% confidence level with $n = 29$. Thus, this model was used to extrapolate the measurement error to the FIA plot locations.

The uncertainties from the input measurement variance, FIA plot variance, and TM band ratio image variance were used as independent variables, and the variance of the predicted forest carbon values produced via cosimulation were used as the dependent variable. A four order polynomial regression was attempted. The resulting regression model had an R^2 value of 0.0042 for year 2004 (Figure 23) and 0.0036 for year 2005 (Figure 24). In an attempt to improve the results, the points that fell outside of two standard errors, and then one standard error, were removed. This increased the 2004 model's R^2 value to 0.0109 with two standard errors and 0.0329 with one standard error

(Figure 25) for year 2004. Similar improvements were observed for year 2005 (Figure 26). However, none of the results were significant at a risk level of $p = 0.05$.

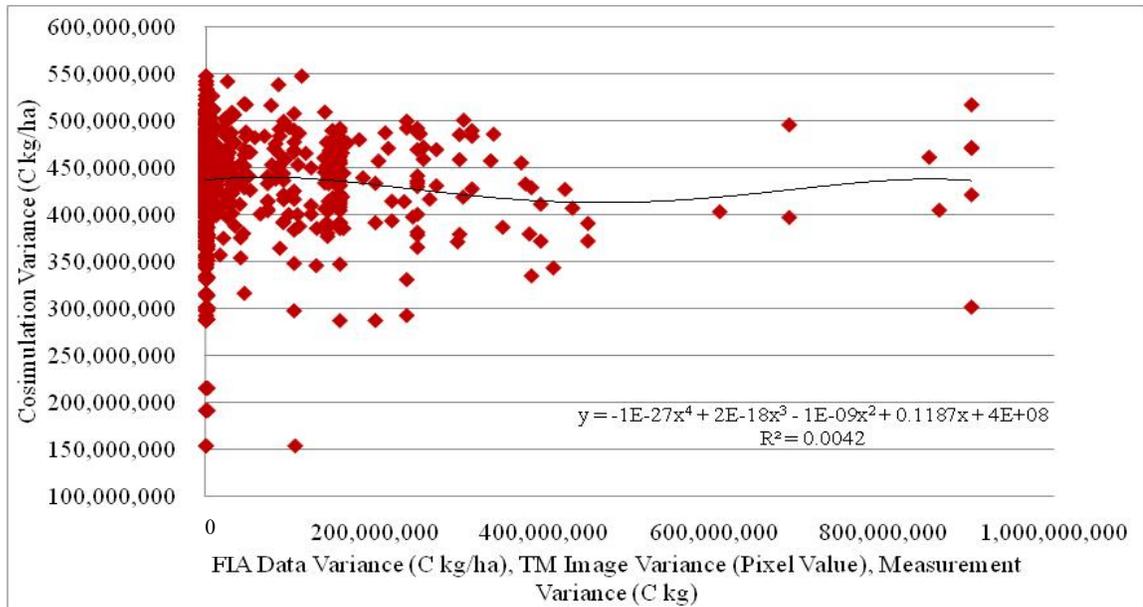


Figure 23: A 4th order polynomial regression of the input variances with the output variances for 2004 uncertainty analysis.

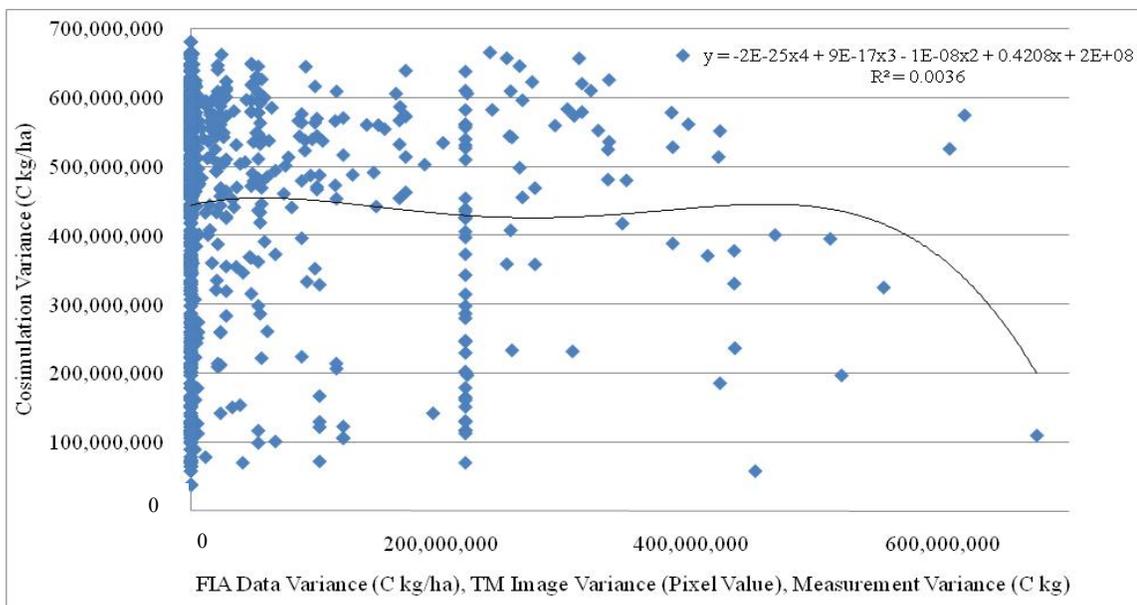


Figure 24: A 4th order polynomial regression of the input variances with the output variances for 2005 uncertainty analysis.

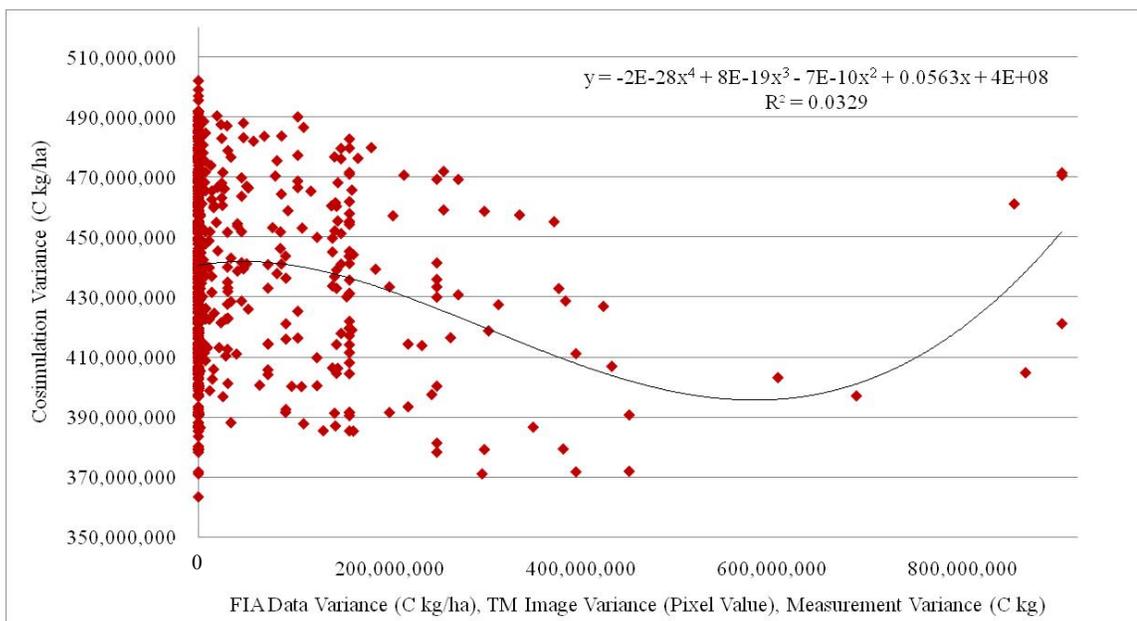


Figure 25: A 4th order polynomial regression of the input variances with the output variances for 2004 uncertainty analysis when the values outside of one standard error were removed.

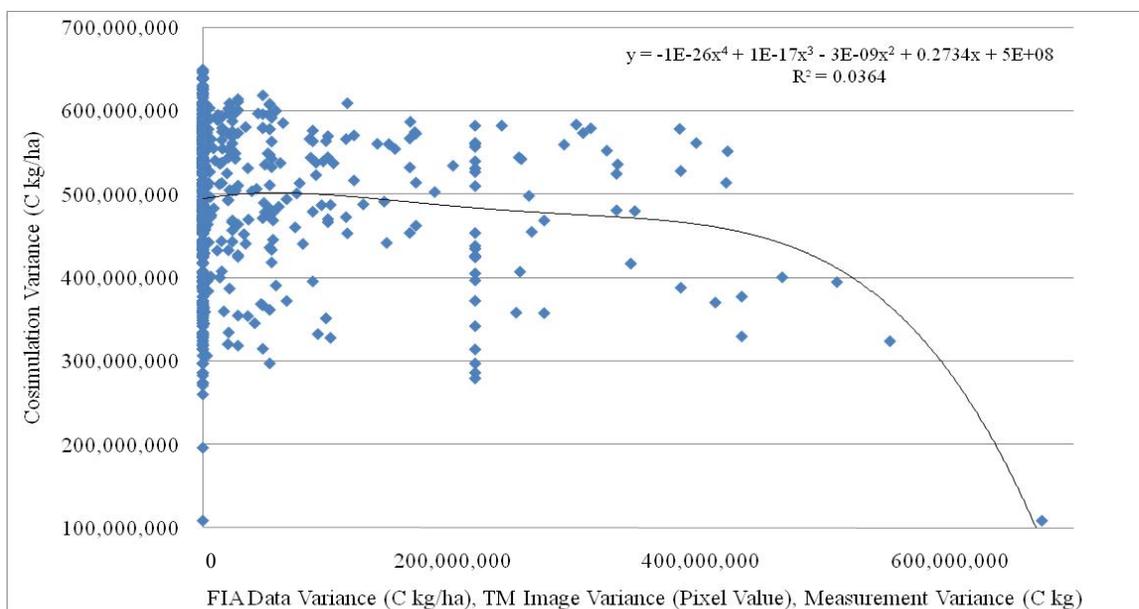


Figure 26: A 4th order polynomial regression of the input variances with the output variances for 2005 uncertainty analysis when the values outside of one standard error were removed.

CHAPTER 6

DISCUSSION AND CONCLUSION

INTERPOLATION METHODOLOGIES

The various interpolation methodologies all produced reasonable spatial patterns of forest carbon stocks, and the results are similar to those of other forest carbon studies (e.g., Nowak and Crane 2002; Wang et al 2009; Zheng et al. 2007, Zheng et al. 2008). The validation techniques reveal estimated and predicted values that are similar to the ground observations. This indicates that the methods provide the potential to generate forest carbon maps by combination of FIA plot and Landsat TM image data. However, the methods differ in their interpolation abilities.

The regression methodology, while up to 98% accurate according to the validation using the FIA plot data, does have the large fault of creating extremely large and negative values. Additionally, the regression estimates have a bias of 13% and 18% for the years 2004, and 2005, respectively, according to the validation using the additional plot data. This is arguably more realistic in terms of the validation, because these additional plots were not used to develop the regression model, as was the case with the FIA plots. On the other hand, the additional plots were located in two local areas, where as the FIA plots were spread throughout the entire study area. There is also the potential that the lack of species specific data for the additional plots, and the resultant tree species group approximation that was calculated; might also have led to the differences in validation techniques. Not only does the lack of species data make it difficult to be confident in the accuracy assessment, but also the fact that four to five

years of approximate annual carbon accumulation had to be subtracted from the additional plot data as well due to their time differences from FIA plots.

As a result, the two conflicting validation techniques are inconclusive and no significant results can be determined from them. However, the map outputs do provide another method of drawing a conclusion. The regression methodology, whilst potentially accurate in meaningful areas; does have the significant drawback of estimating large and negative values. Despite the fact that these erroneous values are not representative to the method as a whole; they do impact the reliability and there is nothing that can be done to prevent such estimates from being created. They are simply a byproduct of using a regression model.

On the other hand, the 2004 cosimulation and up-scaling cosimulation outputs produced a bias below 5% which is significant at the 95% confidence interval. The low validation bias is contradicted by the higher FIA cross validation results. The interpretation of these results is dependent on the relative weight given to the various validation techniques, which unfortunately are inconclusive. In either case the cosimulation outputs have the distinct advantage of creating carbon maps where all the predicted values are reasonable. This gives cosimulation the potential to be a more functional methodology, but further validation of these methods is needed to clarify this.

The real significance of the up-scaled cosimulation maps is the demonstration of their feasibility. Producing an up-scaled map at this regional scale is not, strictly speaking, necessary because the processing time for an image of greater resolution is still within normal capabilities. The reason this study produced the up-scaled cosimulation

maps, was to demonstrate the viability of using this methodology at larger geographic scales. The objective of attempting the up-scaling methodology in this study is to demonstrate that it can produce reasonable values of forest carbon, which indicates that the up-scaling would be appropriate to use at the national or even global scales.

UNCERTAINTY ANALYSIS

The uncertainty analysis did not produce the expected results; with R^2 values below 0.1, there was no significance to any kind of errors that can be quantified from these equations. This was mainly due to the large spread of the input and output variances. By further investigation, it was found that these variances had three different categories. The average variance of 2004 TM image was 1.15, the average variance associated with the measurement errors was 35.66 kg, and the average variance from the FIA plots data was about 112,000 Mg/ha. These large differences led to the large spread of the uncertainties and made it impossible to establish a relationship between the input uncertainties and output uncertainties. This result might be caused by using the smaller and different units of the variables. Potential improvement may be achieved by standardization of the input variances.

The manner in which the uncertainty analysis data values were collected may also have contributed to the failure to reject the hypothesis that a polynomial regression could be used to account for the variance. Many of the values collected were sampled from outside of SNF. Both the FIA plot uncertainty and measurement error uncertainty were calculated from within SNF, which means that all variance values used in the uncertainty

analysis outside of SNF will have much more error associated with the accuracy of these values.

Additionally, the regression model that was used to extrapolate the measurement variance from the additional plot data locations, to the FIA plot locations may not have been ideal. The defined relationship was significantly different from zero, but it only accounted for about 30% of the variance found with measurement error. A much stronger relationship is needed to extrapolate data, because these values were calculated for FIA plot locations which unfortunately were not all that numerous to begin with. The small FIA sample size therefore not only influenced the accuracy of the carbon maps, but also the accuracy of the uncertainty analysis. The calculation of forest carbon mapping input variances is a complex process, and is something that future studies should investigate further.

RESEARCH QUESTIONS AND HYPOTHESES REVISED

This study aimed at closing some of the gaps associated with forest carbon mapping, and correspondingly posed some key questions that it attempted to answer. The first question to be answered, dealt with how to match Landsat TM images with the FIA sample plot data in terms of different spatial resolution. In this study, the Landsat TM images were re-sampled using a window average method to increase their pixel size from $30\text{ m} \times 30\text{ m}$, to $90\text{ m} \times 90\text{ m}$ which approximated the area covered by one FIA sample plot. The use of re-sampling allowed changes to be made to the spatial resolution of images, so that the sample plots and image pixels matched each other and represented

approximately the same area. This reduced the error associated with the predicted values made using the plot data and TM images. This was one way in which this study managed the effects of scale. The use of the up-scaling cosimulation technique in this study also provided the potential to mitigate the effects of scale because it allowed the forest carbon maps to be produced at any spatial resolution desired. Thus, this method is very flexible because it is possible to combine FIA sample plots and remotely sensed images at any spatial resolutions for generation of forest carbon maps.

In response to question 2, the regression and cosimulation techniques both proved to be about equal in their abilities to generate accurate forest carbon maps; when judged using the results from both validation using FIA data, and validation using the additional field plots. The validation using the additional field plot data had a greater potential to provide unbiased results of how accurate the outputs were. However, because the additional field plots used in this study were spatially limited, the validation results did not provide the evidence needed for a conclusion to be made, and therefore a failure to reject hypothesis two. More measurement error or validation based field plots are needed, and at the same time the plots should be more widely spread throughout the study area. Despite the lack of conclusion based on the validation techniques; it has to be pointed out that the regression methodology falls short of the cosimulation techniques because of its inability to produce maps with reasonable low and high values. These extreme low and high values from the regression method are not possible, which means the forest carbon maps created are useless without image post-processing. In addition, the co-simulation methods produced, not only more reasonable predicted values of forest

carbon, but also produced predicted carbon variances, that could provide useful information for forest carbon management relevant decision making.

The third and fourth questions pertained to the contribution of the different sources of error. This study attempted to develop the polynomial regression that linked the input variances with the uncertainties of the estimates. Unfortunately, the obtained polynomial regression model was not statistically significant and thus no partitioning of the different sources of error could be completed. This therefore means that no conclusions can be made about measurement error contribution from the uncertainty analysis. However, it is observable that the overall and maximum standard deviation contribution to measurement error is 1.64% and 13.68%, respectively (Table 5). This error contribution was calculated by dividing the standard deviation of the measured diameters, by the diameter. This creates a percent value that should indicate how much of an effect measurement error would have on that particular measurement. An average measurement error percentage of 1.64 is very small, thus it is unlikely that the measurement errors would contribute much to the total uncertainties involved.

Table 5: Tree diameter measurement error contribution as a percentage of itself.

Subplot	Tree #	Diameter StDev	Diameter Variance	Average Diameter (cm)	stdev/avg diameter
1	1	0.046	0.002	14.257	0.32%
1	2	0.076	0.006	10.663	0.72%
1	3	0.258	0.066	13.093	1.97%
1	4	0.367	0.135	34.08	1.08%
1	5	2.178	4.742	15.912	13.68%
1	6	0.129	0.017	24.574	0.53%
1	7	0.248	0.061	25.376	0.98%
1	8	0.066	0.004	9.532	0.69%
1	9	1.06	1.124	14.766	7.18%
1	10	0.117	0.014	47.531	0.25%
2	1	0.299	0.09	33.539	0.89%
2	2	0.058	0.003	15.951	0.37%
2	3	0.073	0.005	12.98	0.56%
2	4	0.125	0.016	3.784	3.30%
2	5	0.032	0.001	2.6	1.22%
2	6	0.076	0.006	2.366	3.23%
2	7	0.073	0.005	6.985	1.05%
3	1	0.309	0.096	40.521	0.76%
3	2	0.3	0.09	20.223	1.48%
3	3	0.132	0.017	13.525	0.97%
3	4	0.159	0.025	36.705	0.43%
3	5	0.059	0.003	10.752	0.55%
3	6	0.174	0.03	10.419	1.67%
3	7	0.073	0.005	3.586	2.03%
3	8	0.034	0.001	2.861	1.17%
4	1	0.15	0.023	39.46	0.38%
4	2	0.204	0.042	24.655	0.83%
4	3	0.419	0.176	51.63	0.81%
4	4	0.212	0.045	37.483	0.57%
4	5	0.097	0.009	11.427	0.85%
4	6	0.085	0.007	9.111	0.94%
4	7	0.053	0.003	4.944	1.07%
				Average:	1.64%

Landsat TM images, when used in tandem with FIA data; did provide adequate linkage of forest carbon from the sampled locations to unobserved locations (Hypothesis

1). Statistically significant correlations were found between the band ratio (band 4 /

band 5 + band 4 / band 7) and the FIA data, which indicated that this hypothesis was rejected. However, a greater correlation is achievable. There were a relatively small number of FIA plots used in this study, and the FIA plot data used had a relatively small range of forest carbon values; this compared to the number of image pixels and the range of pixel values. The reason was mainly because the forest carbon values were associated with forest types, this study area was relatively small and flat, and thus the forest types were very simple. These characteristics limited the ability of the obtained regression model, and made it easy to create an output with extremely low and high pixel values. The ways to improve this correlation would begin with using a larger sample size of FIA plots. The use of more plots will increase the representation of the spatial variability within a forest area such as Shawnee. Additionally it will give the results more statistical confidence and increase the spatial autocorrelation.

Therefore, adding more forest types would be quite useful. This would force the regression trend line to have a wide range by adding forest carbon values from the forests that have very low and high productivity, especially those that are located in a complicated topography. This would create more collocated pixel locations over the study area and increase the confidence of the models. Overcoming the shortcomings that exist in this current study would without doubt increase the user's confidence with these methods.

CONCLUSION AND RECOMMENDATIONS

In conclusion, this study found that both cosimulation based interpolation techniques were more accurate in creating forest carbon maps than the regression method, based solely on the map outputs. However, all three different methods still posed significant errors when examining the validation results. These methods could be improved with the use of more FIA sample plots, as well as more diversity present within the these data. The results could also be interpreted with more confidence if more validation data were available. The uncertainty analysis proved to be invalid as a result of the large variation within the data used. More work is needed to investigate the impacts of the various input errors on the uncertainties of the map outputs.

This study did effectively demonstrate how forest carbon sink maps could be produced using free FIA plot data and Landsat TM images, and with some ways of tracking the propagation of uncertainty associated with the creation of the maps; although the ability to model the propagation of uncertainty eluded us. Additionally, the creation of the maps did provide a starting point for determining the need and scope of forest management plans to increase the amount of forest carbon being stored within a region.

Future studies may want to investigate the use of landcover classification on the images being used, with the above methods. If landcover classification is used, it will add some sources of uncertainty; these being related to the accuracy of the classification, but should reduce the errors involved in the estimation or prediction of carbon values. However, by doing landcover classification; forest carbon, and even soil carbon, could be more accurately interpolated. Landcover classification would allow for more accurate

carbon estimates by classifying an image into various landcover types, such as bare ground, prairie, forest, and even forests types. The pixels associated with these classifications could then be extracted so that a relationship between image pixel values, and ground truth values could be calculated exclusively based on that landcover type. This method would allow for more specific correlations to be calculated per landcover type, and allow for more landcover types to be used. As Feng et al. (2006) mentioned, SOC has been found to correlate with satellite imagery. Therefore, this method could not only map forest carbon, but potentially soil and other ground sources of carbon; and all at the same time, visible on the same map. It is therefore recommended that landcover classification be used in conjunction with an interpolation technique such as cosimulation, and this could greatly enhance the accuracy of carbon sink mapping.

Basic management strategies could be based on an economic study of forest carbon sequestration according to Sohngen and Mendelsohn (2003), in which the most effective global management strategy was determined to be land use change. Results varied by region, but by reducing the amount of deforestation taking place, particularly in low and bottomland forests; significant amounts of additional carbon could be sequestered. Increasing the time between crop rotations, or harvesting, is the next most successful economically produced method of increasing carbon sequestration (Sohngen and Mendelsohn 2003). Although the processes of carbon sequestration is by no means cheap; both timber and carbon markets could be sustainable when managing with carbon sequestration in mind (Sohngen and Mendelsohn 2003; Jandl et al. 2007).

Slowing deforestation will allow more trees to remain within a forest as a carbon stock, and allow for additional growth (Sohngen and Mendelsohn 2003). Not only will

this aid in forest carbon sequestration, but also in soil carbon sequestration. Soil carbon is primarily lost through disturbances, such as tillage (Farage et al. 2007). When deforestation occurs, not only is the carbon stored within the trees lost, but when the ground is tilled for crops, additional carbon is lost (Farage et al. 2007). Tilling the soil is not the only way in which soil carbon is lost either. The act of harvesting trees also disturbs the soil enough that significant amounts of carbon can be lost here as well. Thus it is important, if carbon sequestration is to be maximized, both above and belowground; to harvest only when necessary, and reduce the disturbances to the soil as much as possible (Jandl et al. 2007).

Similarly to slowing deforestation, afforestation has obvious benefits in regards to forest carbon sequestration. Places where trees were not present before are often excellent places for afforestation to occur, but not all places are ideal. For example, pasture land typically has a significant amount of carbon stored within the soil and roots. The process of planting more trees may actually decrease net carbon sequestration because of the disturbance to the soil. Former cropland on the other hand, is an excellent candidate (Jandl et al. 2007).

Also, by increasing the length of time between harvesting rotations, a tree has more time to grow and sequester more carbon. Typically, a tree is harvested at peak economic maturity, which happens to often coincide with a period of substantial growth. By allowing for more time to elapse between harvests, the forest will have sequestered more carbon and have produced larger trees. Essentially, over time the timber industry will be able to produce more wood from longer rotational periods, and simultaneously be able to sequester more carbon (Sohngen and Mendelsohn 2003). The effects of fire

however, can counteract some of the advantages associated with increased stand rotation. Because fires release 0.5 gigatons of carbon, or 4% of the annual US emissions a year (Daigneault et al. 2010); it poses a significant threat to a forest carbon sequestration and wood production management strategy. In essence, longer stand rotations will increase the likelihood that a fire will occur, and thus shorter stand rotations are ideal in areas where fire is more prevalent (Daigneault et al. 2010). It is therefore important to take into consideration the climate in which carbon sequestration is being attempted, for areas with a higher fire risk will be less ideal than those with a lower risk (Daigneault et al. 2010).

Although fire does play a risk in some areas, there are additional reasons that longer rotations between harvests can help sequester more carbon. Not only does a longer rotational period increase the amount a tree can grow above ground, but some evidence indicates that more mature trees will sequester more carbon within their roots. This is also one of the primary ways in which carbon is sequestered within the soil. Longer rotational periods also allow for fewer soil disturbances, which allows for more carbon to remain within the soil. Lastly, by increasing stand rotation, more time is given to carbon on the forest floor to be converted into a more stable mineral carbon that can remain within the soil, and act as a more long term sink (Jandl et al. 2007). Therefore, there are many advantages associated with longer stand rotation periods.

Forest carbon sequestration based management is a complex mix of balancing economic profit produced by the land, and the amount of carbon that can be sequestered. However, since current climate models are suggesting that mitigation is truly necessary (Karl and Trenberth 2003), the cost of sequestration should become less of an issue

(Sohngen and Mendelsohn 2003). The maps produced in this study provide a carbon sink base, from which management plans can be constructed.

The delineation of Illinois, the United States, and even the world forest carbon sinks is a topic of great importance. The described methodology can be used to assist with the important policy decisions that need to be made in regards to global climate change; as well as providing another resource for managing global climate change as defined by the Kyoto protocol (Dong et al. 2003; Gibbs et al. 2007; Peltoniemi et al. 2006; Wang et al. 2009). This study provides a methodology that utilizes free data to be processed in a simple format (Wang et al. 2009), and creates spatially explicit estimates of forest carbon amounts, as well as uncertainty values (in the case of the cosimulation maps). This methodology is easy replicated, which should allow for different regions of the world to determine the extent of their forest carbon resource. This study contributes to the scientific knowledge of forest carbon allocation, and to global political needs of the country and world. Lastly, it provides a starting point from which future work can expand and improve the methodologies described within this study.

REFERENCES

- Bechtold, W. A., and P. L. Patterson. 2005. The enhanced forest inventory and analysis program - national sampling design and estimation procedures. In: General Technical Report SRS-80, U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC, 85 p.
- Berterretche, Mercedes, Andrew T. Hudak, Warren B. Cohen, Thomas K. Maier-sperger, Stith T. Gower, and Jennifer Dungan. 2005. Comparison of regression and geostatistical methods for mapping Leaf Area Index (LAI) with Landsat ETM+ data over a boreal forest. *Remote Sensing of Environment* 96:49-61.
- Birdsey, R. A. 1992. Carbon Storage and Accumulation in United States Forest Ecosystems. USDA Forest Service, Washington Office, GTR-WO-59, Washington, D.C.
- Birdsey, R. A., and L. S. Heath. 1995. Carbon changes in US forests. In: Joyce, L.A. (Ed.), *Climate Change and the Productivity of America's Forests*. USDA, Forest Service, Ft. Collins, CO, Rocky Mountain Forest and Range Exp. Station General Technical Report, pp. 56–70.
- Blackard, J. A., M. V. Finco, E. H. Helmer, G. R. Holden, M. L. Hoppus, D. M. Jacobs, A. J. Lister, G. G. Moisen, M. D. Nelson, R. Riemann, B. Ruefenacht, D. Salajanu, D. L. Weyermann, K. C. Winterberger, T. J. Brandeis, R. L. Czaplewski, R. E. McRoberts, P. L. Patterson, and R. P. Tymcio. 2008. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment* 112:1658-1677.
- Bousquet, Philippe, Philippe Peylin, Philippe Ciais, Corinne Le Quere, Pierre Friedlingstein, and Pieter P. Tans. 2000. Regional changes in carbon dioxide fluxes of land and oceans since 1980. *Science* 290:1342-1346.
- Brown, Sandra L., Paul Schroeder, and Jeffrey S. Kern. 1999. Spatial distribution of biomass in forests of the eastern USA. *Forest Ecology and Management* 123:81-90.
- Bytnerowicz, A., B. Godzik, W. Fraczek, K. Grodzińska, M. Krywult, O. Badea, P. Barančok, O. Blum, M. Černý, S. Godzik, B. Mankoska, W. Manning, P. Moravčík, R. Musselman, J. Oszlanyi, D. Postelnicu, J. Szdźuj, M. Varšavova, and M. Zota. 2002. Distribution of ozone and other air pollutants in forests of the Carpathian Mountains in central Europe. *Environmental Pollution* 116:3-25.
- Chen, Baozhang, Jing M. Chen, Gang Mo, Chiu-Wai Yuen, Hank Margolis, Kaz Higuchi, and Douglas Chan. 2007. Modeling and scaling coupled energy, water, and carbon fluxes based on remote sensing: An application to Canada's landmass. *Journal of Hydrometeorology* 8:123-143.

- Chen, Feng, David E. Kissel, Larry T. West, W. Adkins, Doug Rickman, and J. C. Luvall. 2006. Feature selection and similarity analysis of crop fields for mapping organic carbon concentrations in soil. *Computers and Electronics in Agriculture* 54: 8-21.
- Chen, Jing, Wengjun Chen, Jane Liu, and Josef Cihlar. 2000a. Annual carbon balance of Canada's forests during 1895-1996. *Global Biogeochemical Cycles* 14 (3): 839-849.
- Chen, Jing M., Weimin Ju, Josef Cihlar, David Price, Jane Lie, Wengjun Chen, Jianjun Pan, Andy Black, and Alan Barr. 2003. Spatial distribution of carbon sources and sinks in Canada's forests. *Tellus* 55B:622-641.
- Chen, Wengjun, Jing Chen, Jane Liu, and Josef Cihlar. 2000b. Approaches for reducing uncertainties in regional forest carbon balance. *Global Biogeochemical Cycles* 14 (3): 827-838.
- Daigneault, Adam J., Mario J. Miranda, and Brent Sohngen. 2010. Optimal forest management with carbon sequestration credits and endogenous fire risk. *Land Economics* 86(1): 155-172.
- Deckmyn, Gaby, Bostjan Mali, Hojka Kraigher, Niko Torelli, Maarten Op de Beeck, and Reinhard Ceulemans. 2009. Using the process-based stand model ANAFORE including Bayesian optimization to predict wood quality and quantity and their uncertainty in Slovenian Beech. *Silva Fennica* 43 (3): 523-534.
- Dong, Jiarui, Robert K. Kaufmann, Ranga B. Myneni, Compton J. Tucker, Pekka E. Kauppi, Jari Liski, Wolfgang Buermann, V. Alexeyev, and Malcolm K. Hughes. 2003. Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. *Remote Sensing of Environment* 84:393-410.
- Fang, Jingyun, Anping Chen, Changhui Peng, Shuqing Zhao, and Longjun Ci. 2001. Changes in forest biomass carbon storage in China between 1949 and 1998. *Science* 292:2320-2322.
- Farage, P. K., J. Ardo, L. Olsson, E. A. Rienzi, A. S. Ball, and J. N. Pretty. 2007. The potential for soil carbon sequestration in three tropical dryland farming systems of Africa and Latin America: A modeling approach. *Soil & Tillage Research* 94: 457-472.
- Gertner, George, Guangxing Wang, Shoufan Fang, and Alan B. Anderson. 2002. Mapping and uncertainty of predictions based on multiple primary variables from joint co-simulation with Landsat TM image and polynomial regression. *Remote Sensing of Environment* 83:498-510.
- Gibbs, Holly K., Sandra Brown, John O. Niles, and Jonathan A. Foley. 2007. Monitoring and estimating tropical forest carbon stocks: Making REDD a reality. *Environmental Research Letters* 2 (4) 045023.

- Hahn, Jerold T. 1984. Tree volume and biomass equations for the Lake States. Res. Pap. NC-250. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 10p.
- Hahn, Jerold T. and Mark H. Hansen. 1991. Cubic and board foot volume models for the central states. *Northern Journal of Applied Forestry* 8:47-57.
- Hall, R. J., R. S. Skakun, E. J. Arsenault, and B. S. Case. 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *Forest Ecology and Management* 225:358-390.
- Heath, Linda S., and James E. Smith. 2000. An assessment of uncertainty in forest carbon budget projections. *Environmental Science & Policy* 3:73-82.
- Illinois Natural Resources Geospatial Data Clearinghouse. Surficial Geology and Elevation Data. University of Illinois at Urbana-Champaign. <http://www.isgs.uiuc.edu/nsdihome/webdocs/st-geolq.html#shaderel1> (last accessed December 6, 2009).
- Illinois State Water Survey. Illinois State Climatologist Data: Results – Anna. University of Illinois at Urbana-Champaign. <http://www.isws.illinois.edu/data/climatedb/data.asp> (last accessed December 6, 2009).
- Jandl, Robert, Marcus Lindner, Lars Vesterdal, Bram Bauwens, Rainer Baritz, Frank Hagedorn, Dale W. Johnson, Kari Minkinen and Kenneth A. Byrne. 2007. How strongly can forest management influence soil carbon sequestration? *Geoderma* 137: 253-268.
- Jenkins, Jennifer C., David C. Chojnacky, Linda S. Heath, and Richard A. Birdsey. 2004. Comprehensive database of diameter-based biomass regressions for North American tree species. General Technical Report NE-319. U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 48p.
- Jensen, John R. 2005. “Principle Components Analysis.” In *Introductory Digital Image Processing: A Remote Sensing Perspective 3rd ed*, Keith C. Clarke, 296-301. Pearson Prentice Hall, Upper Saddle River, NJ.
- Karl, Thomas R., and Kevin E. Trenberth. 2003. Modern global climate change. *Science* 302: 1719-1723.
- Kimes, D. S., R. F. Nelson, W. A. Salas and D. L. Skole. 1999. Mapping secondary tropical forest and forest age from SPOT HRV data. *International Journal of Remote Sensing* 20 (17): 3625-3640.
- Knyazikhin Y., J. Glassy, J. L. Privette, Y. Tian, A. Lotsch, Y. Zhang, Y. Wang, J. T. Morisette, P. Votava, R.B. Myneni, R. R. Nemani, S. W. Running. 1999. MODIS

Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15): Algorithm theoretical basis document Version 4.0. <http://eosps0.gsfc.nasa.gov/atbd/modistables.html>.

- McRoberts, Ronald E., Geoffrey R. Holden, Mark D. Nelson, Greg C. Liknes, Warren K. Moser, Andrew J. Lister, Susan L. King, Elizabeth B. LaPoint, John W. Coulston, W. Brad Smith, and Gregory A. Reams. 2005. Estimating and circumventing the effects of perturbing and swapping inventory plot locations. *Journal of Forestry* 103:275-279.
- Meng, Qingmin, Chris Ciezewski, and Marguerite Madden. 2009. Large area forest inventory using Landsat ETM+: A geostatistical approach. *ISPRS Journal of Photogrammetry and Remote Sensing* 64:27-36.
- Mollenhauer, Gesine, Ralph R. Schneider, Tim Jennerjahn, Peter J. Muller, and Gerold Wefer. 2004. Organic carbon accumulation in the South Atlantic Ocean: its modern, mid-Holocene and last glacial distribution. *Global and Planetary Change* 40: 249-266.
- Mowrer, H. Todd. 1997. Propagating uncertainty through spatial estimation processes for old-growth subalpine forests using sequential Gaussian simulation in GIS. *Ecological Modeling* 98:73-86.
- Myneni, R. B., J. Dong, C. J. Tucker, R. K. Kaufmann, P. E. Kauppi, J. Liski, L. Zhou, V. Alexeyev, and M. K. Hughes. 2001. A large carbon sink in the woody biomass of northern forests. *Proceedings of the National Academy of Sciences* 98 (26): 14784-14789.
- Neilson, E. T., D. A. MacLean, F. -R. Meng, and P. A. Arp. 2007. Spatial distribution of carbon in natural and managed stands in an industrial forest in New Brunswick, Canada. *Forest Ecology and Management* 253:148-160.
- Nowak, David J. and Daniel E. Crane. 2002. Carbon storage and sequestration by urban trees in the USA. *Environmental Pollution* 116:381-389.
- Peltoniemi, Mikko, Taru Palosuo, Suvi Monni, and Raisa Makipaa. 2006. Factors affecting the uncertainty of sinks and stocks of carbon in Finnish forests soils and vegetation. *Forest Ecology and Management* 232:75-85.
- Ruddiman, William F. 2008 "Climatic Data." In *Earth's Climate: Past and Future 2nd ed*, 26-29. W. H. Freeman and Company, New York, NY.
- Running, Steven W. 1994. Testing FOREST-BGC ecosystem process simulations across a climate gradient in Oregon. *Ecological Applications* 4 (2): 238-247.
- Running, Steven W., Ramakrishna R. Nemani, David L. Peterson, Larry E. Band, Donald F. Potts, Lars L. Pierce, and Michael A. Spanner. 1989. Mapping regional forest

- evapotranspiration and photosynthesis by coupling satellite data with ecosystem simulation. *Ecology* 70 (4): 1090-1101.
- Ryan, Michael G. 1991. A simple method for estimating gross carbon budgets for vegetation in forest ecosystems. *Tree Physiology* 9:255-266.
- Sales, Marcio H., Carlos M. Souza, Jr, Phaedon C. Kyriakidis, Dar A. Roberts, and Edson Vidal. 2007. Improving spatial distribution estimation of forest biomass with geostatistics: A case study for Rondonia, Brazil. *Ecological Modeling* 205:221-230.
- Smith W. Brad. 1985. Factors and equations to estimate forest biomass in the North Central Region. Res. Pap. NC-268. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 6p.
- Sohngen, Brent and Robert Mendelsohn. 2003. An optimal control model of forest carbon sequestration. *American Journal of Agricultural Economics* 85(2): 448-457.
- Tan, Kun, Shilong Piao, Changhui Peng, and Jingyun Fang. 2007. Satellite-based estimation of biomass carbon stocks for northeast China's forests between 1982 and 1999. *Forest Ecology and Management* 240:114-121.
- Tang, Jinyun and Qianlai Zhuang. 2009. A global sensitivity analysis and Bayesian inference framework for improving the parameter estimation and prediction of a process-based terrestrial ecosystem model. *Journal of Geophysical Research* 114: D15303, doi:10.1029/2009JD011724.
- Tomlinson, R. W. 2005. Soil carbon stocks and changes in the Republic of Ireland. *Journal of Environmental Management* 76: 77-93.
- Tomppo, Erkki, Hakan Olsson, Goran Stahl, Mats Nilsson, Olle Hagner, and Matti Katila. 2008. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment* 112:1982-1999.
- U.S. Department of Agriculture, Forest Service (a). Forest Inventory and Analysis Program. FIA DataMart version 4.0. <http://199.128.173.17/fiadb4-downloads/datamart.html> (last accessed October 14, 2009).
- U.S. Department of Agriculture, Forest Service (b). Forest Inventory and Analysis Program. 2009. The forest inventory and analysis database: Database description and users manual version 4.0 for Phase 2, Draft Revision 1. Prepared by the Forest Inventory and Analysis National Office. Arlington, VA. 344p.
- U.S. Department of Agriculture, Forest Service (c). Shawnee National Forest: Forest facts. http://www.fs.fed.us/r9/forests/shawnee/about/forest_facts/ (last accessed November, 10, 2009).

- U.S. Department of Agriculture, Forest Service (d). Shawnee National Forest shapefile. Washington, D.C. (received on November 20, 2009).
- U.S. Department of the Interior, U.S. Geological Survey. Earth resources observation and science center: USGS global visualization viewer. <http://glovis.usgs.gov/> (last accessed December 10, 2009).
- Wang, Guangxing, George Gertner, Xiangyn Xiao, Steven Went, and Alan B. Anderson. 2001. Appropriate plot size and spatial resolution for mapping multiple vegetation types. *Photogrammetric Engineering and Remote Sensing* 67 (5): 575-584.
- Wang, G. S. Went, G. Z. Gertner, and A. Anderson. 2002. Improvement in mapping vegetation cover factor for the universal soil loss equation by geostatistical methods with Landsat Thematic Mapper images. *International Journal of Remote Sensing* 23 (18): 3649-3667.
- Wang, Guangxing, George Z. Gertner, Shoufan Fang, and Alan B. Anderson. 2005. A methodology for spatial uncertainty analysis of remote sensing and GIS products. *Photogrammetric Engineering and Remote Sensing* 71 (12): 1423-1432.
- Wang, Guangxing, Tonny Oyana, Maozhen Zhang , Samuel Adu-Prah, Siqi Zeng, Hui Lin, Jiyun Se. 2009. Mapping and spatial uncertainty analysis of forest vegetation carbon by combining national forest inventory data and satellite images. *Forest Ecology and Management* 258 (7): 1275-1283.
- Wang, Guangxing, Maozhen Zhang, George Z. Gertner, Tonny Oyana, Ronald E. McRoberts and Hongli Ge. 2011. Uncertainties of mapping aboveground forest carbon due to plot locations using national forest inventory plot and remotely sensed data. *Scandinavian Journal of Forest Research* iFirst article: 1-14.
- Woodbury, Peter B., James E. Smith, and Linda S. Heath. 2007. Carbon sequestration in the U.S. forest sector from 1990 to 2010. *Forest Ecology and Management* 241:14-27.
- Zheng, Daolan, Linda S. Heath, and Mark J. Ducey. 2008. Spatial distribution of forest aboveground biomass estimated from remote sensing and forest inventory data in New England, USA. *Journal of Applied Remote Sensing* 2: 021502.
- Zheng, G., J. M. Chen, Q. J. Tian, W. M. Ju, and X. Q. Xia. 2007. Combining remote sensing imagery and forest age inventory for biomass mapping. *Journal of Environmental Management* 85:616-623.
- Zirlewagen, Dietmar, Gerhard Raben, and Markus Weise. 2007. Zoning of forest health conditions based on a set of soil, topographic and vegetation parameters. *Forest Ecology and Management* 248:43-55.

APPENDIX

APPENDIX A

Measurement error calculation data. Live/Dead and Deciduous/Coniferous categories removed because all trees in this plot were live and deciduous.

Diameter cm	Rand #	Live, Dead L=1, D=3	8 Mix Hard	9 Maple Oak	1 Ced Lar	4 Pine	Final Total Biomass Excluding Dead Trees	Carbon kg
39.44	1	1	1697.27	2259.53	1163.2	1342.89	1951.64	442.63
24.61	2	1	525.9	716.59	400.65	425.75	782.53	177.47
51.34	2	1	3267.89	4294.24	2110.94	2552.65	4491.96	1018.76
37.34	1	1	1481.53	1977.64	1027.89	1175.32	1711.06	388.06
11.33	1	1	76.67	108.54	69.51	64.45	100.15	22.71
9.14	2	1	44.9	64.24	42.72	38.14	60.82	13.79
4.93	2	1	9.72	14.34	10.62	8.51	13.55	3.07
39.5	1	1	1704.08	2268.41	1167.45	1348.18	1959.23	444.34
24.8	2	1	536.09	730.2	407.71	433.84	796.98	180.75
51.03	2	1	3217.81	4229.73	2081.49	2514.29	4425.86	1003.77
37.66	1	1	1513.09	2018.94	1047.79	1199.87	1746.32	396.06
11.46	1	1	78.83	111.54	71.28	66.23	102.82	23.32
9.17	2	1	45.29	64.79	43.06	38.47	61.34	13.91
4.9	2	1	9.57	14.12	10.47	8.38	13.33	3.02
39.53	1	1	1707.49	2272.87	1169.57	1350.83	1963.02	445.21
24.8	2	1	536.09	730.2	407.71	433.84	796.98	180.75
51.38	2	1	3272.93	4300.73	2113.89	2556.51	4498.6	1020.26
37.31	1	1	1478.39	1973.54	1025.91	1172.88	1707.56	387.27
11.36	1	1	77.21	109.29	69.95	64.9	100.81	22.86
9.26	2	1	46.47	66.45	44.08	39.45	62.9	14.27
4.93	2	1	9.72	14.34	10.62	8.51	13.55	3.07

Diameter cm	Rand #	Live, Dead L=1, D=3	8 Mix Hard	9 Maple Oak	1 Ced Lar	4 Pine	Final Total Biomass Excluding Dead Trees	Carbon kg
39.15	1	1	1666.81	2219.78	1144.2	1319.27	1917.73	434.93
24.29	2	1	509.16	694.23	389.03	412.46	758.78	172.09
51.41	2	1	3277.96	4307.22	2116.85	2560.36	4505.24	1021.77
9.07	2	1	44.13	63.16	42.05	37.5	59.79	13.56
4.93	2	1	9.72	14.34	10.62	8.51	13.55	3.07
39.6	1	1	1714.33	2281.79	1173.83	1356.13	1970.63	446.93
24.51	2	1	520.84	709.84	397.14	421.73	775.36	175.85
51.34	2	1	3267.89	4294.24	2110.94	2552.65	4491.96	1018.76
37.27	1	1	1475.26	1969.45	1023.93	1170.45	1704.06	386.47
11.33	1	1	76.67	108.54	69.51	64.45	100.15	22.71
9.14	2	1	44.9	64.24	42.72	38.14	60.82	13.79
4.9	2	1	9.57	14.12	10.47	8.38	13.33	3.02
39.5	1	1	1704.08	2268.41	1167.45	1348.18	1959.23	444.34
24.92	2	1	542.96	739.36	412.45	439.28	806.7	182.96
51.95	2	1	3364.33	4418.41	2167.53	2626.48	4619.12	1047.6
37.88	1	1	1535.43	2048.14	1061.85	1217.23	1771.25	401.71
11.55	1	1	80.47	113.81	72.63	67.58	104.85	23.78
8.98	2	1	42.98	61.55	41.06	36.55	58.27	13.22
5.06	2	1	10.36	15.26	11.25	9.06	14.42	3.27
39.28	1	1	1680.31	2237.39	1152.62	1329.74	1932.76	438.34
24.54	2	1	522.53	712.08	398.31	423.07	777.74	176.39
51.92	2	1	3359.21	4411.83	2164.53	2622.57	4612.37	1046.07
37.37	1	1	1484.67	1981.75	1029.87	1177.76	1714.57	388.86
11.59	1	1	81.02	114.58	73.08	68.04	105.53	23.93

Diameter cm	Rand #	Live, Dead L=1, D=3	8 Mix Hard	9 Maple Oak	1 Ced Lar	4 Pine	Final Total Biomass Excluding Dead Trees	Carbon kg
9.01	2	1	43.36	62.09	41.39	36.86	58.78	13.33
4.9	2	1	9.57	14.12	10.47	8.38	13.33	3.02
39.53	1	1	1707.49	2272.87	1169.57	1350.83	1963.02	445.21
37.62	1	1	1509.92	2014.78	1045.79	1197.4	1742.77	395.25
11.46	1	1	78.83	111.54	71.28	66.23	102.82	23.32
9.14	2	1	44.9	64.24	42.72	38.14	60.82	13.79
4.93	2	1	9.72	14.34	10.62	8.51	13.55	3.07
39.6	1	1	1714.33	2281.79	1173.83	1356.13	1970.63	446.93
24.57	2	1	524.21	714.33	399.48	424.41	780.13	176.93
52.27	2	1	3415.76	4484.61	2197.65	2665.84	4686.88	1062.97
37.59	1	1	1506.75	2010.64	1043.79	1194.93	1739.23	394.45
11.43	1	1	78.29	110.78	70.84	65.78	102.15	23.17
9.1	2	1	44.51	63.7	42.38	37.82	60.31	13.68
5	2	1	10.04	14.8	10.93	8.78	13.98	3.17

VITA

Graduate School

Southern Illinois University

Andrew Lawrence Fleming

Date of Birth: April 11, 1986

12304 Beech Forest Rd. Box 2, Laurel, Maryland, 20708

afleming04@alumni.unity.edu

Unity College

Bachelors of Science, Wildlife Biology, May 2008

Thesis Title:

Forest Carbon Mapping and Spatial Uncertainty Analysis: Combining National Forest Inventory Data and Landsat TM Images.

Major Professor: Dr. Guangxing Wang