

Vertical Structure, Horizontal Cover, and Temporal Change
of the North Carolina Piedmont (1985 – 2005).

by

Joseph Owen Sexton

Program in Ecology
Duke University

Date: _____

Approved:

Dean L. Urban, Supervisor

Norman L. Christensen

James S. Clark

Conghe Song

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of
of Philosophy in the Program in Ecology
in the Graduate School
of Duke University

2009

ABSTRACT

Vertical Structure, Horizontal Cover, and Temporal Change
of the North Carolina Piedmont (1985 – 2005).

by

Joseph Owen Sexton

Program in Ecology
Duke University

Date: _____

Approved:

Dean L. Urban, Supervisor

Norman L. Christensen

James S. Clark

Conghe Song

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree of
Doctor of Philosophy in the Program
in Ecology in the Graduate School
of Duke University

Copyright by
Joseph Owen Sexton
2009

Abstract

An ecosystem is a community of organisms interacting with its environment, and landscapes are spatially interactive ecosystems. Earth's burgeoning human population demands ever more from finite ecosystems; but if managed appropriately, landscapes can sustain their provision of resources and services and adapt to fulfill the changing human appetite. Sound management relies on sound information, and managing changing landscapes requires reliable spatio-temporal databases of ecologically relevant variables. Remote sensing technologies fill this niche, providing increasingly large and diverse datasets, but the algorithms to extract information from the data must be invented. I developed and compared three remotely sensed measurements of forest canopy height to one another and to *in situ* field measurements. Both the precision and the accuracy (as well as the cost) of the measurements sorted along an axis of spatial scale, with Light Detection and Ranging (lidar) proving to be the most reliable source at fine scales but prohibitively expensive over large areas, and the various radar technologies more appropriate for larger areas, especially when calibrated to the more accurate and precise lidar measurements. I also adapted traditional, single-time landcover classification algorithms to extract dense time series of maps from archival multi-spectral satellite images. The resulting map series revealed spatial heterogeneity and temporal nonlinearity in the region's change, complexities which would not be

observable either in sparser time series or by comparing maps from multiple provenances. These canopy height and landcover measurements greatly expand the potential spatio-temporal scope of landscape ecology and management, easing the current restriction to loosely connected case studies and “space-for-time substitution” toward more robust, statistical analysis based on consistent information in space and time.

Dedication

I dedicate this thesis to Trish, Tim, and Danny—whose own accomplishments rival and inspire the best of mine; to mom and dad, who support us wherever we go; and to Dean Urban, my mentor.

Contents

Abstract	iv
Dedication	vi
List of Tables	xi
List of Figures	xii
Acknowledgements	xiv
1. Ecosystem change in the Anthropocene Era	1
2. A comparison of lidar, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America	6
2.1. Introduction	6
2.2. Methods	13
2.2.1. Study Area	13
2.2.2. Data	15
2.2.2.1. Field measurement	15
2.2.2.2. Lidar	17
2.2.2.3. SRTM	18
2.2.2.4. GeoSAR	19
2.2.3. Analysis	20
2.2.3.1. Lidar scaling	20
2.2.3.2. Measurement comparison	21
2.2.3.4. SRTM calibration	24

2.3. Results	25
2.3.1. Lidar scaling.....	25
2.3.2. Measurement comparisons	28
2.3.3. SRTM calibration	31
2.4. Discussion.....	33
2.4.1. Measurement of canopy height by lidar	33
2.4.2. Measurement of canopy height by Interferometric Synthetic Aperture Radar (InSAR).....	36
2.4.4. Calibration of SRTM for canopy height measurement	40
2.4.5. Selection among methods	42
2.5. Conclusions	44
3. Landcover dynamics of the North Carolina Piedmont between 1985 and 2006 by temporal signature extension.....	45
3.1. Introduction.....	45
3.2. Methods	50
3.2.1. Study area.....	50
3.2.2. Data	52
3.2.2.1. Landsat	52
3.2.2.2. Landcover.....	55
3.2.3. Landcover classification	59
3.2.4. Accuracy assessment.....	62
3.3. Results	63

3.3.1. Landcover classification	63
3.3.2. Landcover change	65
3.4. Discussion.....	69
3.4.1. Four aspects of classification.....	69
3.4.1. Semantics.....	71
3.4.2. Causation.....	74
3.4.3. Representation.....	75
3.4.4. Uncertainty	76
3.4.2. Future research directions.....	78
3.5. Conclusions	81
3.6. Appendix: Distributional standardization of classification accuracy metrics	82
4. Concluding remarks	89
4.1. Synopsis: Landcover change of the North Carolina Piedmont (1985 – 2006)	89
4.2. Prospectus for future research.....	90
4.2.1. Methodological requirements.....	90
4.2.2. Ecological opportunities.....	91
4.2.2.1. Landcover/landuse change.....	91
4.2.2.2. Forest growth.....	92
4.2.2.3. Landcover phenology and forest succession	92
4.3. Conclusion.....	93
References	95

Biography 104

List of Tables

Table 1: Landcover classification years and input Landsat images. Reference data for model training and validation were obtained from years in bold. Intermediate (Pass 2) landcover class probabilities for terminal years (in italics) were used for temporal filtering of adjacent years, but were not converted to categorical landcover maps. 53

Table 2: Landcover proportions in the reference sample and study area (taken from NLCD 2001), and model priors derived from NLCD 2001 by splitting probabilities for class 71 equally between classes 81 and 82, and class 21 proportionally among non-urban land classes (i.e., excluding classes 11, 23, and 24). Years in bold were used for both model training and testing, whereas 1999 was reserved solely for model testing. 57

Table 3: Confusion matrix for landcover classification validated on a 4-year test sample, with producer's accuracy (PA), user's accuracy (UA), Kappa (κ), percent correctly classified (PCC) and Kappa (κ), as well as standardized percent correctly classified (PCCs) and standardized Kappa (κ_s). Standardizations were computed based on priors from NLCD 2001 (Table 2). 66

List of Figures

Figure 1: Study area arrangement and ecoregional context. Stands and inventory plots were selected from the Durham, Korstian, Blackwood, and Eno Divisions of the Duke Forest and represent forests of the Southeastern Mixed Forest Province (inset, white) (Bailey, 1995). Topography is exaggerated 5x.....	14
Figure 2: Scaling characteristics of stand-level canopy height measurements (mean, median, and mode) across a range of pixel resolutions in pine stands of different structural characteristics: narrow, linear canopy gaps between planted rows (DU-33-10); large, irregularly shaped gaps (DU-39-18); small gaps (KO-04-04); and dense, homogeneous cover (KO-06-05).....	27
Figure 3: Effects of varying aggregation of lidar returns on within-stand distributions of canopy heights measured in pine stands representing: medium-height, patchy canopy cover (DU-39-18) and tall, homogeneous canopy cover (KO-04-04).	28
Figure 4: Comparison of within-stand distributions of canopy height measurements in a representative pine stand (DU-33-10) and a representative hardwood stand (EN-19-09). Sample sizes (n) are the number of pixels—or trees, in field data—located within the 30-m radius plot.....	29
Figure 5: Comparison of field-measured and remotely sensed canopy heights among pine and hardwood plots.....	30
Figure 6: Calibration of SRTM-derived canopy height based on lidar and GeoSAR measurements in pine (n = 48) and hardwood (n = 84) plots. Calibrated curves are overlaid on reference values measured with lidar in pine plots and by field methods (Haga altimeter) in hardwood plots.....	32
Figure 7: Comparison of raw and calibrated SRTM measurements to reference height measurements in pine and hardwood plots. Reference heights in pine plots are from lidar measurements, and those in hardwood plots are from field measurements. Calibrated measurements are from a cubic fit in pine plots and a linear fit in hardwood plots.....	34
Figure 8: Schematic representation of the idealized response of X-, C-, and P-band radar interferometric heights to “true” canopy height; and the resulting canopy-sensitive ranges of GeoSAR and SRTM measurements. Axes are not drawn to scale.	39

Figure 9: Regional landcover change from three different sources: the 1992 National Landcover Database, the North Carolina Center for Geographic Information and Analysis (1996), and the 2001 National Landcover Database.....	47
Figure 10: The 1.7-Mha study area lies on the Piedmont Plateau (inset—shaded, stippled area), partially or entirely covers thirteen North Carolina counties, and is completely within WRS-2 Path 16, Row 35 (inset—shaded box). Split into separate training and test datasets, reference data were collected over 4 counties and four years.....	52
Figure 11: Dendrogram of landcover class centroids in the spectral domain, estimated by Quadratic Discriminant Analysis.	64
Figure 12: Landcover change among generalized classes within the study area from 1985 to 2006.....	67
Figure 13: Change trajectories of three selected landcover classes, with expected errors of omission (top bar) and commission (bottom bar)—see <i>Appendix for explanation</i>	69
Figure 14: Regional cover of the dominant landcover classes for three selected years, with detail of the Triangle Region (inset). (A) reservoirs built around the city of Burlington between 1987 and 1989 that accounted for the region’s greatest increase in water area over the period; (B) urban expansion in the Triangle region from the towns of Cary and Apex into the surrounding agricultural fields; (C) the Triad region, consisting of the cities of Greensboro, High Point, and Winston-Salem (off map).....	70

Acknowledgements

This research was supported by the NASA Earth Systems Science Fellowship Program, and was endorsed by the Global Land Project, a joint project of the International Geosphere-Biosphere Programme and International Human Dimensions Programme on Global Environmental Change. The work was performed in the Duke University Landscape Ecology Laboratory, under the guidance of Dean Urban, Conghe Song, Jim Clark, and Norm Christensen. It bears a strong imprint of Doug Ramsey and Dave Roberts, my official and unofficial M.S. advisors, respectively.

Landsat images were contributed by Dean Urban, Conghe Song, Pat Halpin, Pete Harrell, Jon Goodall, the Multi-Resolution Land Characteristics Consortium, the Global Land Cover Facility, the Duke University Social Science Research Institute, and the libraries of Duke University, the University of North Carolina, and North Carolina State University. Lidar data were provided by the North Carolina Floodplain Mapping Program. GeoSAR data were provided by Paul Siqueira and NASA's Jet Propulsion Laboratory. Forest inventory measurements were provided by the Office of the Duke Forest, Duke University. High-resolution reference images and supporting GIS layers were provided by the State of North Carolina.

This thesis stands on the efforts of many people. Christine Cope, Mike Donohue, Tyler Bax, and John Kerkerling, and several other students in the Master of

Environmental Management Program of the Nicholas School of the Environment at Duke University performed unpublished pilot studies. Dean Urban, Tyler Bax, Conghe Song, Paul Siqueira, Jennifer Swenson, Mike Donohue, Amanda Whitehurst, Luke Dollar, Becky Bartel, Cesar Delgado, Dalia Amor Conde, Fernando Colchero, Alex Pfaff, Jim Westervelt, and Todd BenDor co-authored manuscripts. Will Wilson, Emily Bernhardt, Dan Richter, and many others stimulated ideas with ready ecological applications. Technical assistance was generously given by Ben Best, John Fay, Pete Harrell, Gyanesh Chander, Scott Loarie, Ben Poulter, Ibrahim Alameddine, Pat Mahoney, Andy Minnis, Ben Donnelly, Gil Pontius, Forrest Hall, Song Qian, and Carl Salk. Technicians Nathan Emery, Maggie Goldman, Neoma Lavallo, Ian Varley, Phil Schutt, and Lindsey Smart processed images and other data.

1. Ecosystem change in the Anthropocene Era

Over the past 100 years, Earth's human population has increased from less than 2 billion to nearly 7 billion people (UN 1999, CIA 2008). Much of this rise is attributed to technological advances that have shifted dominance in the human-nature relationship increasingly toward human control (Crutzen 2002). Along with important social adaptations (e.g., demographic changes, migration to cities, etc.) the current "Anthropocene Era" has brought large changes in landform, biogeochemical cycles, climate, and biodiversity, as well as spatial patterns of land cover and human land use (Turner et al. 1990, Pimm et al. 1995, Schlesinger 1997, Wilkinson 2005, IPCC 2007). These past and projected future changes have led to concern over the adaptability of societies and long-term sustainability of human welfare (IPCC 2007, IMF 2008, World Bank 2008)

Understanding the complex interactions between physical, biological, and social systems is increasingly vital, not only for conserving ecosystems under growing human pressure, but also for adapting socio-economic and regulatory systems to their changing natural environment. Human-natural systems are coupled in both directions: ecosystems provide extractive natural resources (e.g., food, fiber, and fuel) and long-term services (e.g., air and water purification, climate moderation, pollination, biodiversity support, and aesthetic beauty) while humans modify the natural

environment intentionally (e.g., agricultural or urban development) and accidentally (e.g., pollution and introduction of pest species).

Studying these systems requires cooperation between the natural and social sciences. Based on the idea of *landcover*, a categorical description of landscape elements or ecosystem types (i.e., forest, field, urban, water, etc.), two current research topics are forming the interdisciplinary nexus: (1) landcover/landuse change (LCLUC) and (2) ecosystem services. Together, they describe two main feedbacks between physical-ecological and socio-economic systems: LCLUC studies focus on modeling spatially extensive changes in ecosystems (e.g., deforestation, agricultural abandonment, desertification, and urbanization) and studies of ecosystem services “close the loop” by quantifying the effects of ecological processes back onto socio-economic systems. Recognizing the importance of these feedbacks, a growing number of researchers are coordinating their efforts to better understand these coupled human-natural systems (Daily and Matson 2008).

Communication between the sciences greatly accelerates their mutual advancement. As a common database and modeling environment, the geographic information system (GIS) serves as a major platform for interdisciplinary exchange. GIS promotes cooperation by placing data from various sources—landcover, human population censuses, water quality measurements, wildlife habitat assessments, etc.—on

a common coordinate system, through which they can be related in space and time. Studying *dynamics* of coupled human-natural systems further requires that data from various sources be not just spatially coincident, but extensive in time as well. This requires consistent, long-term records.

Among the most important technological advances has been the ability to observe changes in the Earth from space. For more than 30 years, the basic measurements used to create landcover maps have been provided by sensors aboard the Landsat series of satellites. Such a long, consistent record allows scientists to study current phenomena and also to retrieve information from the past, from which it would otherwise be difficult or impossible to recover. Landsat data have supported basic research since the mission's first launch, and are increasingly used to provide governments and other parties with assessments of ecosystem health and changes due to natural and anthropogenic forcing (National Academy of Science 2005). Increasing research opportunities under the National Science Foundation Coupled Human/Natural Systems and NASA Land Cover/Land Use Change research programs reflect society's growing interest in Earth observation as a means to ensure long-term sustainability of human welfare.

This dissertation is the culmination of six years' effort in two research areas: (1) developing the spatio-temporal landcover datasets needed to study coupled human-

natural systems, and (2) analyzing the natural and anthropogenic changes that affect landscape provision of ecosystem services. Because lack of data is currently the most fundamental limitation on these studies, this thesis focuses on the first objective. Several side projects have provided proof-of-concept or pilot studies for creating and analyzing these novel datasets. These include:

- Comparing forest changes in and around Kirindy Mite National Park, Madagascar, (Whitehurst et al., 2009);
- Monitoring decadal habitat dynamics of an endangered butterfly (Bartel and Sexton, *in press*);
- Econometric analysis of the Inter-Oceanic Highway's effects on Amazonian deforestation (Delgado et al. 2009a, 2009b, *in prep*);
- Development of a GIS-based regional urban growth model (Westervelt et al., *submitted*).

The thesis presents the furthest development of two promising techniques. The second chapter is a comparison of methods for measuring forest canopy height, a fundamental characteristic of forest structure and an essential variable in forest ecology and management. The third chapter describes the results of the greatest portion of effort from the past six years—the creation of a consistent, spatio-temporal database of landcover over 23 years and 1.7 million hectares of the North Carolina Piedmont region.

Finally, the thesis closes with a synopsis of the insights gained from these efforts and a prospectus for future work.

2. A comparison of lidar, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America

2.1. Introduction

Forest canopy structure is a key determinant of forest ecosystem processes, and its measurement is essential for ecosystem monitoring, modeling, and management. As the primary attribute of vertical structure, canopy height affects plant community dynamics and composition (e.g., Welden et al. 1991, Kruger et al. 1997), boundary layer meteorology and microclimate (e.g., Raupach 2004), and wildlife habitat value (e.g., James 1971). Horizontal variations in canopy height (e.g., “edges” and “gaps”) also affect plant and animal communities and their dynamics (e.g., Andren 1988, Matlack 1994, Didham and Lawton 1999). Through allometric relationships, canopy height can be used to estimate stand age and foliage-height profile (e.g., Aber 1979), stand volume (e.g., Cadee et al. 1997), site productivity (e.g., Monserud 1984, Doolittle 1979), and biomass (Dubayah 2000).

Estimation of carbon content in forest ecosystems is of increasing importance for compliance with international environmental treaties aiming to offset greenhouse gas emissions (Balzter et. al 2007). However, considerable uncertainty surrounds regional stocks of carbon (Schimel et al. 2001), limiting the potential to effectively monitor and manage global carbon and hydrological cycles. These quantitative uncertainties could

be greatly reduced by accurate measurement of forest canopy height across broad spatial scales (Balzter et al. 2007).

Canopy height is a fundamental variable in allometric equations that estimate forest biomass and productivity (Patenaude et al. 2002, Andersen et al. 2006). Many studies have shown the power of using vegetation height from traditional and remotely sensed measurements in ecosystem models to estimate aboveground biomass and carbon stocks, from plot to landscape scales (Lefsky et al. 2002a, Drake et al. 2002, Lefsky et al. 2002b, Hurtt et al. 2004, Balzter et al. 2007). More recently, vegetation height measurements are being combined with ancillary data to develop regional estimates of aboveground biomass intended to generate a cohesive biomass and carbon database on a national scale (e.g., Kellndorfer et al. 2006).

Comparable measurements of canopy height require a consistent definition of the canopy itself, as well as what is measured as its height. Semantic distinctions become important when methods with different sensitivities are employed to measure a structure as complex as a forest canopy. A forest canopy is a complex volume, composed of the stems, branches, leaves, and other solid plant parts, as well as the gaps between them. From this, canopy height refers to the vertical distance from the ground surface to the highest tree part (i.e., non-gap) at that horizontal location. Because of the three-dimensional roughness of tree crowns and the gaps between trees (which often extend

to the ground), measurements of height will necessarily be affected by the horizontal resolution at which data are collected.

Accurate estimates of forest canopy height are therefore dependent on measurements that capture both the horizontal cover and vertical structure of aboveground vegetation (Kellndorfer et al. 2004). Traditional forest inventory methods of measuring tree height with poles or trigonometric transformations of distance and angle measurements are technically simple and inexpensive over small areas, but are difficult to apply in closed stands and impossible to implement over large areas (Andersen et al. 2006). Alternatively, airborne and satellite remotely sensed data provide an economical and efficient means of obtaining height measurements over much larger areas (Naesset 1997, Dubayah & Drake 2000, Andersen et al. 2006).

Passive optical remote sensing systems are suited to measure horizontal vegetation cover, and recent advances in active remote sensors have made it possible to acquire reliable measures of vertical forest structure as well (Waring et al. 1995, Treuhaft et al. 1996, Lefsky et al. 2002a, Treuhaft and Siqueira, 2000). Active remote sensors used for forest canopy height estimation include light detection and ranging (lidar) and interferometric synthetic-aperture radar (InSAR) (Balzter et al. 2007, Brown & Sarabandi 2003). Currently, airborne lidar sensors can estimate height with sub-meter vertical accuracy and spatial resolution (e.g., Brandtberg et al. 2003), but are economically and

computationally constrained at broad regional scales (Kelndorfer et al. 2004). In contrast, due to their capacity to systematically scan large areas and record data through cloud cover, airborne and spaceborne InSAR missions have produced an abundance of landscape- and global-scale data that have proven useful in determining canopy height in a number of studies (Balzter et al 2007, Brown & Sarabandi 2003, Dubayah et al. 2007, Hensley et al. 2001, Kelndorfer et al. 2004, Townsend 2002, Walker et al. 2007).

Lidar remote sensing uses high-frequency pulses, or “posts” of laser light to measure the distance from the sensor to target objects. Lidar sensors are most frequently carried aboard aircraft; however some satellite-based systems have recently been developed for cryosphere applications, e.g., ICESat (<http://icesat.gsfc.nasa.gov>). Calculation of height from lidar requires precise positioning of the sensor by a Differential Global Position System (DGPS) and an Inertial Measurement Unit (IMU) carried aboard the aircraft, and is based on the time elapsed from the round-trip travel of each laser pulse between the sensor and the target. Each of these lidar returns is recorded as a point in three-dimensional space; height above the ground is calculated as the vertical difference between laser returns from the target and those from the ground, or “bare earth” surface.

Lidar sensors vary according to the electromagnetic wavelength of the laser (e.g., 500 nm for bathymetric vs. 900 – 1064 nm for terrestrial applications), by power, by

duration and frequency of the pulse, and by the laser beam's cross-sectional diameter, determined by its divergence angle and flight altitude (Lefsky et al. 2002). Lidar sensors can be further divided into two types: discrete- (including multiple-) return *vs.* continuous waveform. Discrete-return lidar provides one or several height measurements (i.e., returns) per post, and is the more common of the two. These systems can be tailored to specific applications by varying aircraft speed, altitude, pulse rate, and other parameters. Alternatively, continuous-waveform lidar provides measurements of light interception continuously along the vertical axis of each post and can be used to describe the vertical distribution of biomass or foliage within vegetation canopies. Forest characterization using small-footprint, continuous-waveform lidar is an active research field, and discrete return, small footprint lidar has been successfully used for some time to measure tree height within stands (Maclean and Krabill 1986). While evolving the potential to record information from each post similar to continuous-waveform lidar (Carter et al., 2007), discrete-return lidar methods continue to be refined for many forest types as well as for detailed characterizations such as mapping of individual crowns (e.g. Roberts et al. 2005).

Interferometric SAR uses two radar receivers carried aboard air- or spacecraft and separated by a known distance, termed the "baseline". The two receivers may be two physical units or a single device passed over the target twice. These receivers record

the phase of a radar signal reflected from the ground or other targets in the “range”, or “cross-track” direction, perpendicular to the flight-line of the receivers (i.e., the “azimuth” direction). This phase measures the travel time of the radar signal and therefore the round-trip distance between the radar transmitter, target, and receivers. Through knowledge of the observing geometry, the difference between phases of the signal received at the two ends of the baseline can be translated into a topographic height via a geometric transformation (Rosen et al., 2000). In the presence of vegetation, the topography measured from interferometric SAR may be anywhere from the ground surface to the top of the canopy, depending on the signal scattering and extinction characteristics of the vegetation. These characteristics vary with radar frequency (and inversely, wavelength)—longer microwaves (30 - 100 cm and larger, commonly referred to as P-band or UHF) tend to penetrate deeper into the canopy to the ground surface, whereas shorter-wavelength signals, including C-band (~6 cm), X-band (~3 cm), and smaller, tend to reflect nearer the top of the canopy.

Methods for estimating canopy height from SAR vary. Some compare interferometric height with an independent measurement of the ground surface (Kelndorfer et al., 2006, Simard et al., 2006). Others, including polarimetric interferometric synthetic aperture radar (PolInSAR), use both interferometric height and correlation, along with multiple baselines and/or polarizations for estimating

vegetation height directly (Cloude and Papathanassiou 1998, Treuhaft and Siqueira, 2000). A third approach, applied in this study, uses the difference in sensitivity between multiple wavelengths, measuring interferometric heights at two frequencies and calculating the difference in elevation between the two measurements to estimate vegetation height (Wheeler and Hensley, 2000).

Each of these methods—lidar, radar interferometry in the various wavelengths, and field measurement—responds uniquely to the physical structure of the forest canopy. Empirical studies are necessary to understand the methods' sensitivities and comparability for various forest types, and to inform selection among the methods for specific purposes. Between February, 2000 and October, 2001, four datasets were acquired over the Duke Forest, (North Carolina, USA): (1) digital elevation data from Shuttle Radar Topography Mission (SRTM) C-band radar interferometry, (2) X- and P-band radar interferometry from an experimental flight of the airborne Geographic Synthetic Aperture Radar (GeoSAR), (3) airborne small-footprint lidar measurements acquired in leaf-off (winter) conditions, and (4) stand-inventory field measurements. Although initially for various purposes (only the GeoSAR and field measurements were originally intended for measuring canopy height), their nearly simultaneous collection presents a rare opportunity for comparing four of the dominant methods for measuring canopy height.

With this dataset, we conducted three analyses: (1) evaluation of the scaling behavior of lidar-derived canopy heights in pine forests, (2) comparison of canopy height measurements produced by the methods in pine and hardwood forests, and (3) calibration of the globally available SRTM elevation data to extract canopy height for southeastern forests. These analyses illustrate the relationships between canopy height measurements from the four methods in forests characteristic of southeastern North America, provide methods to extract canopy height measurements from available datasets, and facilitate adoption of remotely sensed methods for mapping canopy height at regional scales.

2.2. Methods

2.2.1. Study Area

The Duke Forest (36° N, 79° W) covers 7,050 ha of the Southern Appalachian Piedmont Section of the Southeastern Mixed Forest Province (Bailey, 1995) in central North Carolina, USA (Fig. 1). Forest communities, characteristic of the region, are seral mixes of planted and natural pine (predominantly loblolly pine, *Pinus taeda*) and many species of deciduous hardwoods, including oaks (*Quercus* spp.), hickories (*Carya* spp.), elms (*Ulmus* spp.), red maple (*Acer rubrum*), sweetgum (*Liquidambar styraciflua*), and tulip poplar (*Liriodendron tulipifera*). Mature forest canopies are typically closed and

two-storied with overstory pines and hardwoods reaching heights > 30 m, and shorter hardwoods in the under- and mid-story), but this varies in response to site productivity, timber management, and natural disturbances. Ground elevation ranges from 90 to 210 m, and local relief is gentle to moderate. Management of the Forest is partitioned into

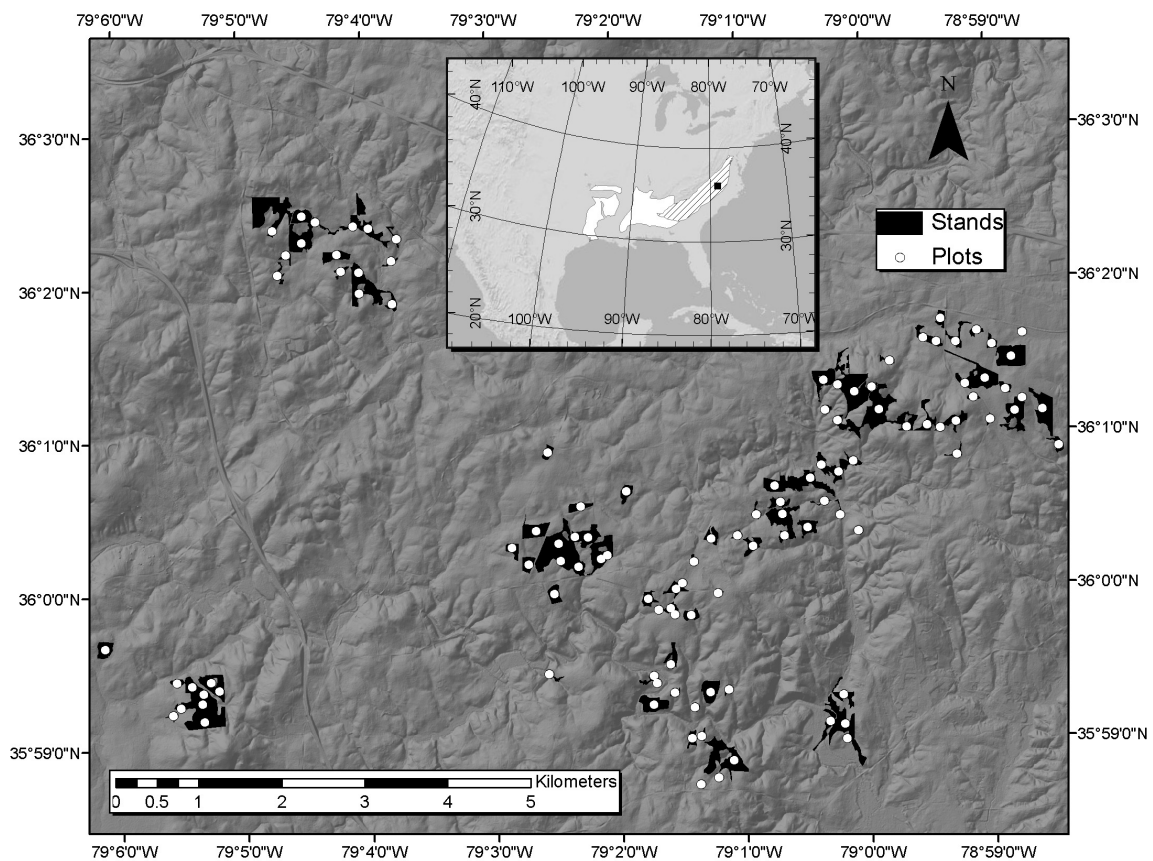


Figure 1: Study area arrangement and ecoregional context. Stands and inventory plots were selected from the Durham, Korstian, Blackwood, and Eno Divisions of the Duke Forest and represent forests of the Southeastern Mixed Forest Province (inset, white) (Bailey, 1995). Topography is exaggerated 5x.

stands of homogeneous forest composition and structure, ranging from 0.9 to 14.7 ha. Data were analyzed at two levels: (1) 112 stands covered by all four data sets (SRTM, lidar, GeoSAR, and field), including four treeless herbaceous stands digitized from ortho-imagery to represent the shortest (near-zero) canopy heights in both the pine and hardwood stands, and (2) 139 forest inventory plots, including five located in recently cleared stands (with remaining stumps, slash, and debris) to represent zero-height pine and hardwood types.

2.2.2. Data

2.2.2.1. Field measurement

A timber inventory of the Duke Forest was conducted in the summer and fall of 2000. Mensuration plot centers were spaced ~400 m apart on a regular sampling grid, and inclusion of trees within plots was determined based on a prism tally. At each plot, trees were selected with a 10 basal-area-factor (BAF) prism pivoted 1.3 m above the ground at the plot center. Border cases were decided by measuring and comparing the distance between the tree and plot center to the tree's diameter at breast height (DBH); if the distance from the plot center to the tree (in feet) was less than the DBH (in inches) x 2.75, then the tree's height was measured and recorded. Height was measured with a

Haga™ altimeter (http://www.haga-metallwaren.de/en_altimeter.htm) by an observer positioned 66 feet (~20.1 m) from the base of each tallied tree.

An index of pine dominance was used to quantify the contribution of evergreen (i.e., predominantly pine species and occasionally eastern redcedar, *Juniperus virginiana*) canopies to each plot's height measurements. The index was calculated as:

$$pineDom = \frac{\sum height_{pine}}{\sum height_{all}} \times 100, \quad (1)$$

where the numerator is the sum of heights for all pines in the tally and the denominator is the sum of heights for all trees in the tally. The index was used to categorize plots into discrete "pine" ($pineDom \geq 90$) and "hardwood" ($pineDom \leq 5$) types, which were then verified against aerial photos and the 2001 National Land Cover Database Land Cover Layer (Homer et al. 2004). Field measurements were focused on forest stands, and so did not include height measurements in short vegetation. To represent short vegetation in both pine and hardwood types, zero canopy height was assigned to five plots placed in areas of recently cleared forest, identified in 1998 aerial photos.

2.2.2.2. Lidar

In response to extensive damage from Hurricane Floyd (1999), the North Carolina Floodplain Mapping Program (NCFMP; <http://www.ncfloodmaps.com>) was established to update maps used to identify areas vulnerable to flooding. Two private engineering firms were contracted to collect lidar elevation data for the purpose of updating statewide flood insurance maps. One of the firms, 3Di (<http://www.3ditech.com/>), acquired elevation measurements with small-footprint lidar over the Cape Fear River basin in February and March 2001. Because the primary intent of the program was to map the bare-earth surface, the measurements were acquired in winter, when deciduous hardwood canopies were in dormant, “leaf-off” condition. Measurements were acquired from Datis 1, Datis 2, Optech 1225 and Optech 1210 lidar sensors flying aboard aircraft at 7100 – 8300 meters above sea level. Pulse rates were 45,000 – 50,000 pulses per second for Datis sensors and 50,000 pulses per second for Optech sensors. Average post spacing was 4 - 6 m, beam diameter was 0.5 – 1.0 m, and swath widths were ~1000 m. Vertical accuracy of lidar measurements was <25 cm RMSE.

Sensors used by NCFMP acquired up to four returns per post. Heights of first and last returns were recorded, but were not differentiated or labeled by order (Hope Morgan, NCFMP, personal communication). To estimate canopy height, we subtracted the value of a bare-earth DEM from the elevation recorded at each return and selected

the maximum height difference within pixels at a range of resolutions—from 3-m to 30-m (i.e., between the resolutions of GeoSAR and SRTM measurements, or 0.5x to 5x the average lidar posting). The bare-earth DEM used for this and the SRTM canopy measurements was derived from the NCFMP lidar measurements by converting lidar returns representing local minima to a Triangulated Irregular Network (TIN) specifying known break lines, and then interpolating the TIN to raster at 6.096-m (i.e., 20-foot) resolution.

2.2.2.3. SRTM

In February 2000, the first space-borne, fixed-baseline interferometric synthetic-aperture radar (InSAR) was carried aboard the Space Shuttle Endeavor to compile a digital topographic database of Earth's surface between 60° N and 57° S (Rabus et al. 2003). Acquisition occurred during winter in the northern hemisphere, when deciduous forest canopies were leafless. The Shuttle Radar Topography Mission (SRTM) used the SIR-C and X-SAR radar antennas to collect information in C-band (5.6cm, 5.3GHz) and X-band (3.1cm, 9.6GHz) wavelengths. The C-band data were processed at the NASA Jet Propulsion Laboratory (JPL) and released to the public at a resolution of 1-arc second (~ 30m) within the United States and 3 arc seconds (~ 90 m) elsewhere (Hensley et al. 2000). Absolute vertical accuracy for the C-band DEM was specified as < 16 m globally,

with relative accuracies of < 6 m in any 225 x 225-km area (Rabus et al 2003). Because of the short wavelength of the SRTM C-band sensor, the height response over vegetated terrain is influenced by integrated scattering from leaves, branches, and stems, measuring a phase center height that is higher than the underlying bare-Earth surface in vegetated areas.

Vegetation canopy height can be derived from the SRTM DEM by subtracting from it another DEM that represents the ground elevation (Brown & Sarabandi 2003, Kelldorfer et al. 2004, Dubayah et al. 2007). The method assumes negligible canopy penetration by the SRTM C-band radar—an assumption that is rarely met in vegetation canopies, and so calibration of the elevation differences to measured canopy heights is necessary (Kelldorfer et al. (2004). Whereas Kelldorfer et al. (2004) used the National Elevation Dataset (NED) as the bare-earth DEM, we subtracted the more precise bare-earth DEM derived from lidar from the SRTM DEM to improve comparability with our fully lidar-derived canopy height measurements.

2.2.2.4. GeoSAR

The GeoSAR instrument was initially constructed at the NASA's Jet Propulsion Laboratory (JPL), and is currently operated by Fugro, EarthData Inc. (<http://www.fugroearthdata.com>), a global geosciences mapping and remote sensing

company. GeoSAR flies aboard a Gulfstream II aircraft at a nominal altitude of 10 km and consists of two dual side-looking interferometric systems—one at P-band (86 cm wavelength) and one at X-band (3 cm wavelength)—to cover a >10-kilometer swath on either side. GeoSAR records several measurements (backscatter power, interferometric height, and interferometric correlation) for each of the two bands at 3-m resolution, with 1-m and 1-4 m vertical accuracy for X- and P-band DEMs, respectively.

In its early development stages, GeoSAR collected science data over a number of regions in the United States, including the Duke Forest and surrounding area on October 11, 2001. For this study area, we calculated canopy height as the difference between X- and P-band interferometric heights. Due to low signal-to-noise ratios in the P band in sparsely vegetated fields, missing height measurements in these areas were filled with lidar measurements for calibrating SRTM data.

2.2.3. Analysis

2.2.3.1. Lidar scaling

We selected four pine stands as case studies to assess the effect of horizontal resolution on lidar-based canopy height measurement. The cases represented stands with: (1) dense, homogeneous canopy cover; (2) small gaps; (3) narrow, linear gaps between planted rows; and (4) large, irregularly shaped gaps. Each stand boundary was internally buffered by 30 m to avoid errors due to spatial mis-registration between

datasets. Lidar heights were rasterized at resolutions equal to 0.5-, 1-, 2-, 3-, 4-, and 5-times the average posting (~6 m) and extracted from within the buffers. Pixel sizes therefore equaled the 3-m resolution of GeoSAR, the 6-m original lidar posting, and ranged up to the 30-m resolution of SRTM in increments of 6 meters. Relative frequency histograms across the range of resolution were overlaid on one another to track changes in the distribution of measured height. Measures of central tendency (mean, median, and mode) were also calculated for each stand at each resolution and plotted over resolution.

2.2.3.2. Measurement comparison

Relative frequency distributions from all measurements were compared graphically for six representative hardwood stands and five representative pine stands. Results are shown using one stand of each type. Inventory plots were then divided into pine and hardwood types ($n = 30$ and 50 , respectively) based on the field-measured pine dominance index. Means of field-measured tree heights and lidar, GeoSAR, and SRTM rasters were extracted from within 30-m-radius buffers centered on each plot. Measurements were compared via their plot means following Willmott (1982), which provides a detailed explanation of the metrics. However, we interpreted the term

“error” simply as “difference” to avoid misrepresenting any of the measurements as truth *a priori*.

The bias:

$$B = N^{-1} \sum_{i=1}^N (x_i - y_i) \quad (2)$$

quantifies the average vertical offset of one measurement (x) over another (y) across N plots. The standard deviation of bias:

$$s_B = \left[(N-1)^{-1} \sum_{i=1}^N (x_i - y_i - B)^2 \right]^{0.5} \quad (3)$$

quantifies the variability around the bias. The root-mean-squared error:

$$RMSE = \left[N^{-1} \sum_{i=1}^N (x_i - y_i)^2 \right]^{0.5} \quad (4)$$

and mean absolute error

$$MAE = N^{-1} \sum_{i=1}^N |x_i - y_i| \quad (5)$$

quantify the magnitude of difference between measurements; whereas MAE better represents the average magnitude of differences between measurements, RMSE is more sensitive to large differences and is included for comparison to previous studies.

Willmott (1982) also provides metrics describing “systematic” and “unsystematic” differences between two measurements, x and y . The systematic error:

$$MSE_s = N^{-1} \sum_{i=1}^N (\hat{x}_i - x_i)^2 \quad (6)$$

is the mean squared difference between the measurements x_i and their estimated values based on another measurement (y) that has been calibrated to x via a linear model $\hat{x} = f(y|x)$. MSE_s quantifies the error remaining between two measurements after linear calibration. The unsystematic error:

$$MSE_u = N^{-1} \sum_{i=1}^N (\hat{x}_i - y_i)^2 \quad (7)$$

is the mean “residual” error between the original measurements y_i and those that have been calibrated to x . MSE_u quantifies the variation in y removed by linear calibration to x . The two components MSE_s and MSE_u sum to MSE , with $MSE_s \approx 0$ and MSE_u approaching MSE for a well-specified, well-fit model (Willmott (1982)). We report their square roots ($RMSE_s$ and $RMSE_u$), in units of the response (m), for ease of interpretation.

2.2.3.4. SRTM calibration

Stand canopy-height averages from each method were divided into pine and hardwood classes based on the 2001 National Land Cover Database (NLCD 2001), taking the landcover value of the nearest pixel to the plot center. To represent zero-height stands of each type, the four herbaceous stands were included in both hardwood and pine classes. After removing one outlier from the pine subset, the final sample included 37 pine and 84 hardwood plots.

We calibrated the SRTM height measurements by least-squares regression, fitting functions of form:

$$H = a + b \cdot H_{SRTM} \quad (8)$$

and

$$H = a + b \cdot H_{SRTM} + c \cdot (H_{SRTM})^{1/3}, \quad (9)$$

where H is canopy height, a and b are the intercept and slope of a linear relationship between the SRTM C band (minus bare-earth DEM) measurements and H , and c fits the cube-root-transformation of SRTM-based canopy height to model the effect of canopy penetration by the C band over increasing height. Negative or zero SRTM heights were recoded to an arbitrarily small, positive value (0.0001) to accommodate the cube-root transformation, and separate regressions were fit for pine and hardwood types, with lidar used as reference in pine and field data in hardwood plots.

2.3. Results

2.3.1. Lidar scaling

Lidar-based measurement of canopy height in pine stands was sensitive to the area over which returns were aggregated (Fig. 2). At a pixel-resolution equal to half the average lidar posting (3 m), the within-stand distributions of canopy height were heavily influenced by ground returns. This effect was greatest in sparsely canopied stands and weakest in densely canopied stands. The effect diminished quickly as resolution was coarsened beyond the average posting (thereby increasing the area over

which the maximum was calculated) and was negligible in every stand at 3x the average posting (18 m). Whereas the ground effect was visible in the means, medians, and modes of the within-stand distributions, the mean was impacted most, with medians and modes stabilizing at relatively finer scales. The scale at which the measures stabilized was impacted by horizontal variations in canopy cover, with a stand exhibiting large gaps (DU-39-18) showing a greater ground effect at fine scales than the more homogeneous stands (Fig. 3). A nearly closed-canopy stand with narrow (< 5-m) linear gaps between planted rows (DU-39-18) was affected by ground returns more similarly to homogeneous, closed-canopy stands than to the stand with large, irregularly shaped gaps. Occasional tall trees—visible as secondary modes in the height distribution (Fig. 3.)—also affected stand-level estimates, by increasing the estimates at coarse resolution. At both fine and coarse resolutions, the mode was least affected by heterogeneity, followed by the median, and then the mean.

In all stands, a range of resolutions existed in which the mean, median, and mode were similar. For homogeneous stands, this range was from 2x to 4x the lidar posting (i.e., 12-m to 24-m resolution). In heterogeneous stands, the range was narrower, with canopy gaps increasing the minimum and emergent crowns decreasing the maximum of the range. Across all stands in this study, the mean, median, and mode were most similar at 3x the original posting (18 m).

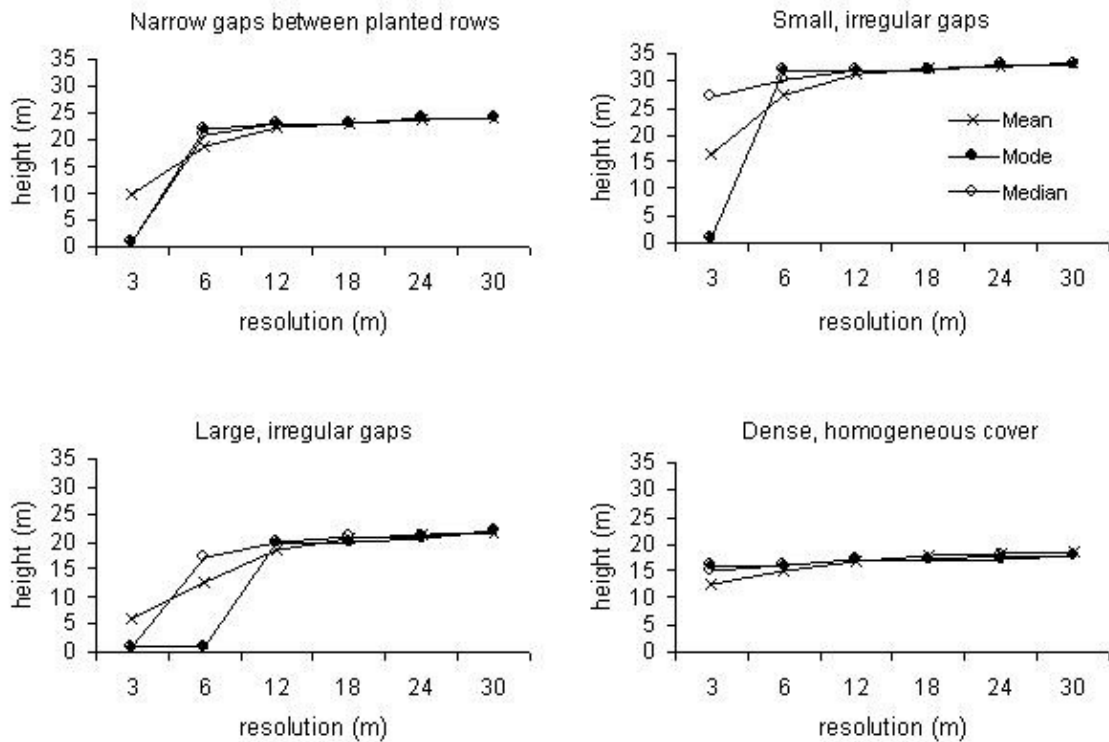


Figure 2: Scaling characteristics of stand-level canopy height measurements (mean, median, and mode) across a range of pixel resolutions in pine stands of different structural characteristics: narrow, linear canopy gaps between planted rows (DU-33-10); large, irregularly shaped gaps (DU-39-18); small gaps (KO-04-04); and dense, homogeneous cover (KO-06-05).

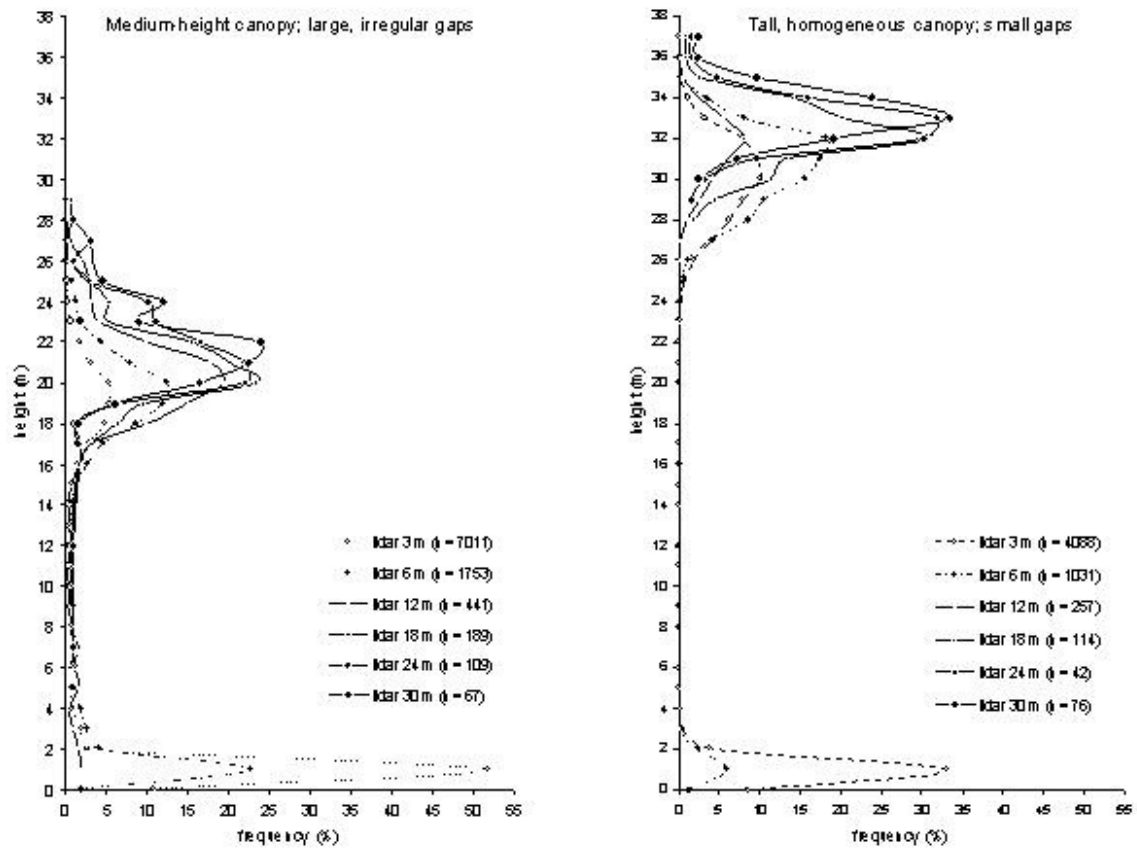


Figure 3: Effects of varying aggregation of lidar returns on within-stand distributions of canopy heights measured in pine stands representing: medium-height, patchy canopy cover (DU-39-18) and tall, homogeneous canopy cover (KO-04-04).

2.3.2. Measurement comparisons

Each of the four methods produced internally consistent, interpretable estimates of canopy height when examined graphically, but to a lesser degree for SRTM in hardwoods (Figs. 3, 4). However, direct (i.e., uncalibrated) comparison among methods was variable (Fig. 5). Both stand- and plot-level height distributions were more

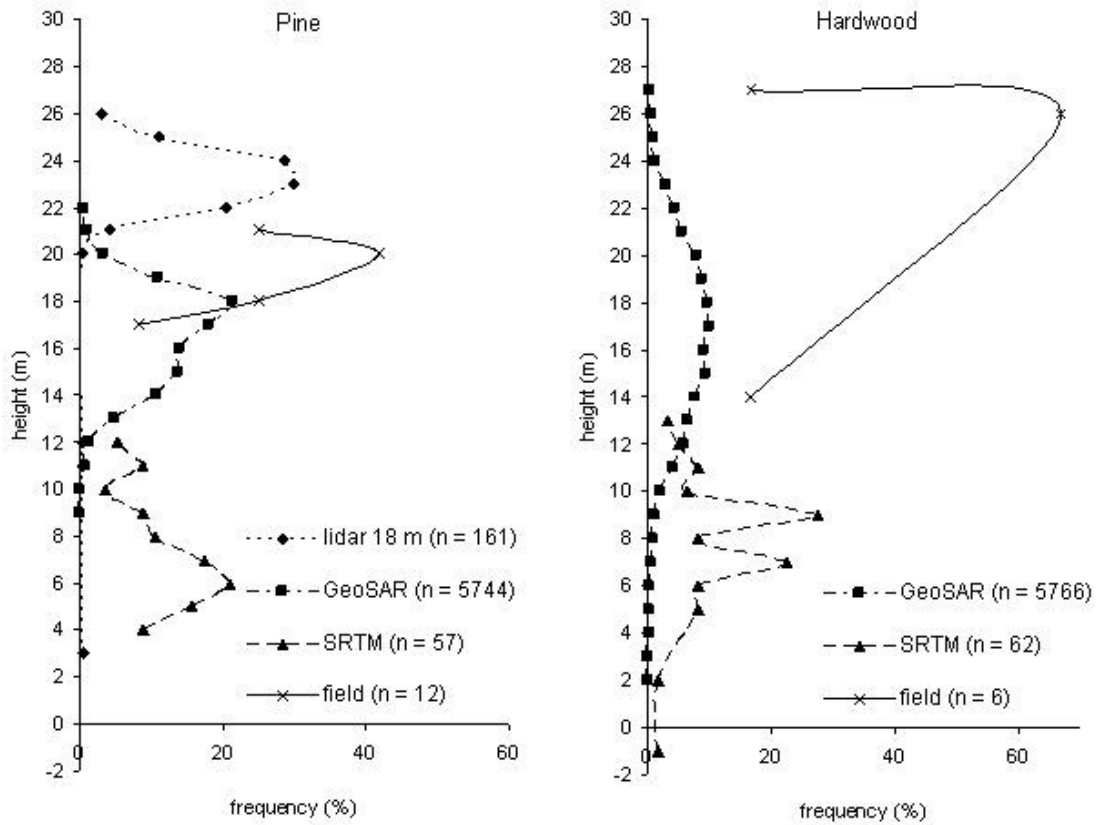


Figure 4: Comparison of within-stand distributions of canopy height measurements in a representative pine stand (DU-33-10) and a representative hardwood stand (EN-19-09). Sample sizes (n) are the number of pixels—or trees, in field data—located within the 30-m radius plot.

comparable overall in pine than in hardwood stands, with greater visual similarity between distributions (Fig. 4), approximately 1 to 2 m smaller RMSE and MAE (Table 1), and approximately 1.5 to 3 m smaller, more constant biases (Table 2) in pines than in

hardwoods. Linear calibration between measurements was also more effective in reducing differences between measurements in pines than in hardwoods (Table 3).

In both pine and hardwood forests, field and lidar methods provided the tallest canopy height measurements, with lidar measurements slightly taller than those collected in the field. In order of decreasing height, these were followed by GeoSAR and then SRTM (Table 2). Because of this sorting, GeoSAR exhibited the greatest direct (i.e.,

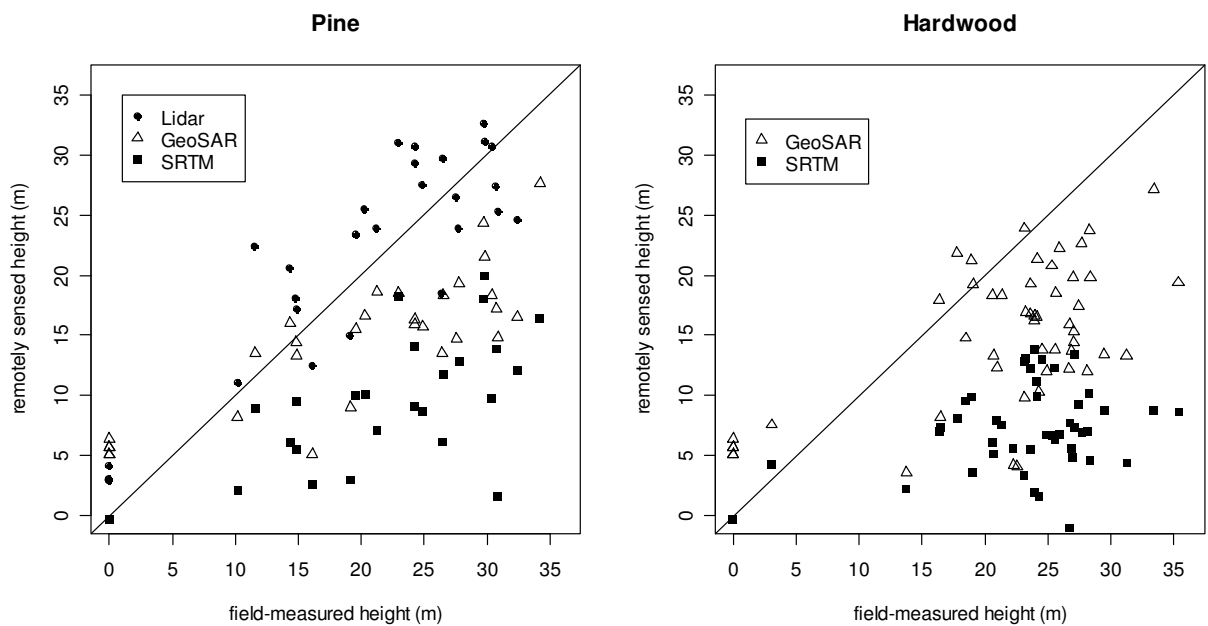


Figure 5: Comparison of field-measured and remotely sensed canopy heights among pine and hardwood plots.

uncalibrated) comparability to other measurements, followed by lidar, field, and SRTM (Table 1). Lidar measurements were most directly comparable to field data, but this could only be evaluated in pine forests. Due to its negative bias and imprecise relationships to other methods, SRTM measurements were the most weakly comparable overall.

In terms of improvements measured by R^2 , linear calibrations were most effective when based upon lidar (Table 3), with all other measurements serving similarly to one another in this regard. However, remaining “systematic” errors (RMSEs) were smallest when calibrations were based upon GeoSAR, followed by field data, lidar, and then SRTM. Likely due to shared sensitivity to within-canopy variation, linear calibration to GeoSAR provided the greatest overall reduction of RMSE in SRTM-based canopy height measurements.

2.3.3. SRTM calibration

SRTM exhibited a delayed response to canopy height in pine plots (Fig. 6), with reference (lidar) height rising more steeply than SRTM measurements in short stands where the ground surface dominated scattering of the C band. In taller pine canopies (> 10 m), where the signal’s penetration to the ground was reduced, SRTM increased in approximate 1:1 proportion to canopy height. The linear model accommodated for this

pattern with a steep slope (1.48) and a high intercept (8.93). The cube-root model followed the data more consistently over its range, with an insignificant intercept (Table 4) and a slope nearly parallel to the 1:1 line in tall (~ 20-m) pine canopies (Fig. 6).

The pattern was much less clear in hardwood canopies, with little difference in fit between the linear and quadratic models. Whereas the addition of a cube-root term decreased standard error, RMSE, and MAE by ~ 1 m in pine forests, improvements in fit from the cube-root term were slight in hardwoods (Tables 4, 5). Also in opposition to

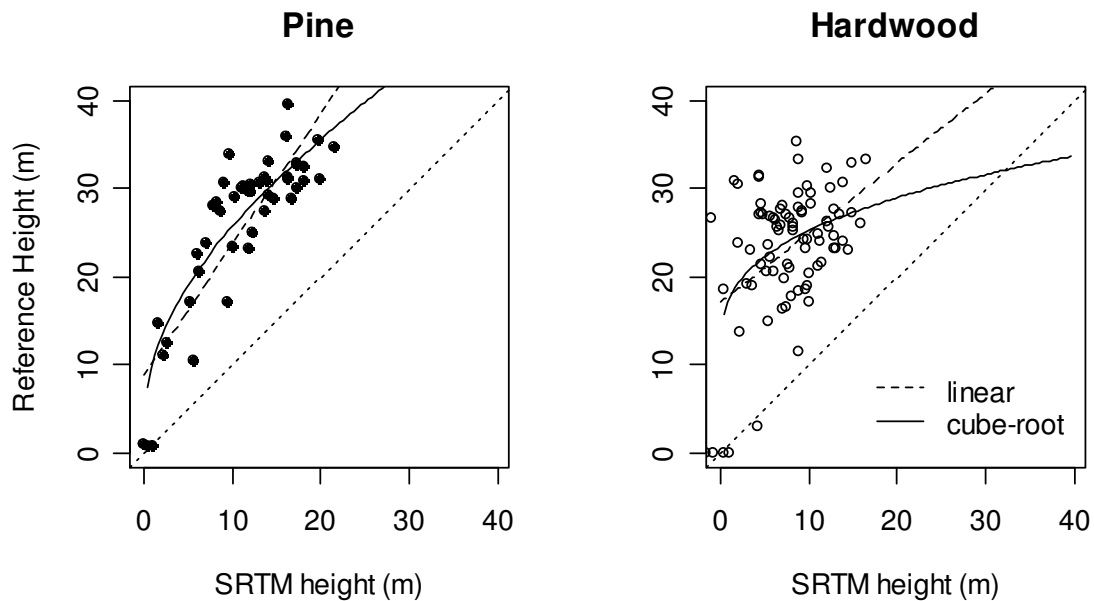


Figure 6: Calibration of SRTM-derived canopy height based on lidar and GeoSAR measurements in pine (n = 48) and hardwood (n = 84) plots. Calibrated curves are overlaid on reference values measured with lidar in pine plots and by field methods (Haga altimeter) in hardwood plots.

pine canopies, the cube-root term tended to over-correct the SRTM data in tall hardwood stands, forcing an excessively shallow slope in the tallest stands.

The greatest difference between linear and cube-root calibrations occurred in the extremes of both pine and hardwood canopy height. Whereas the linear transformations over-predicted canopy height in the shortest (~ 0 m) stands, the logarithmic corrections produced more even and accurate height estimates across the lower range of canopy heights, although the bias was not completely removed in hardwoods (Fig. 7). At the high end of the range of canopy heights, the equations diverged as well, with linear functions producing higher estimates than cube-root calibrations. The effect was extreme in hardwood stands.

2.4. Discussion

2.4.1. Measurement of canopy height by lidar

Our results corroborate studies that have confirmed the ability of discrete-return lidar to reliably estimate canopy height in coniferous forests (Naesset 1997, Dubayah & Drake 2000, Means et al 2000, Lefsky et al. 2002, Roberts 2005). Stand-level distributions of lidar height maxima provided intuitive representations of the vertical structure of pure loblolly pine stands of different sizes, including negligible mid-story cover and loss of apical dominance with increasing height (Fig 2). Canopy heights measured by lidar

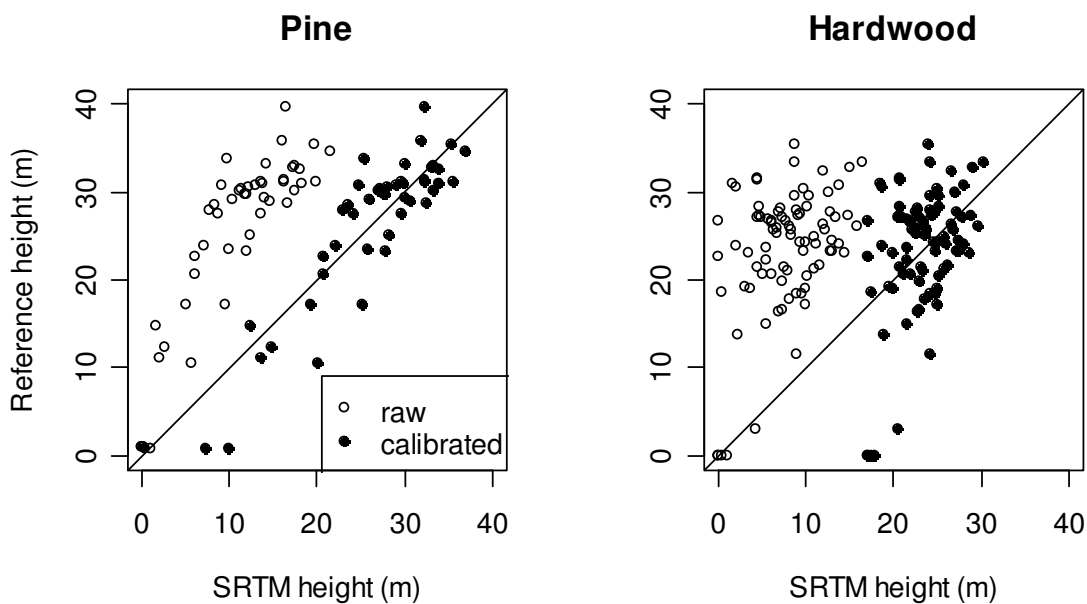


Figure 7: Comparison of raw and calibrated SRTM measurements to reference height measurements in pine and hardwood plots. Reference heights in pine plots are from lidar measurements, and those in hardwood plots are from field measurements. Calibrated measurements are from a cubic fit in pine plots and a linear fit in hardwood plots.

also had greater accuracy and fewer systematic errors with respect to field measurement in pine forests than methods based on radar interferometry. Further, judged by their increased precision relative to field data (Table 3), lidar measurements are preferable over field data for calibrating other measurements of canopy height, including those of SRTM and GeoSAR.

Canopy height measured by lidar is sensitive to the scale at which the data are collected and aggregated into raster pixels, revealing an interaction between canopy heterogeneity and data resolution. By increasing the number of returns from which per-pixel maxima are calculated, coarsening raster resolution removes the effect of gaps on canopy height measurement. This effect varies with the size and shape of canopy gaps, and so the scaling behavior of lidar measurements also has the potential to provide ancillary measurements of fine-scale (i.e., sub-pixel) canopy heterogeneity. Ongoing studies are needed to show the interaction between spatial heterogeneity of the canopy and lidar posting so that data acquisition and processing parameters can be tuned to various forest types (Dubayah and Drake 2002).

The winter acquisition of our lidar data restricted its utility in hardwood stands. In preliminary analyses, the sub-canopy and ground signals were dominant in leafless stands and were not fully removed even at 30-m resolution. High-density (~12 returns m⁻²), small-footprint lidar acquired in winter has been used to analyze vertical structure of deciduous trees at fine scales (Brandtberg et al. 2003), suggesting that regional characterizations are also possible at far coarser resolution than those studied here. However, as resolution was coarsened, hardwood pixels became increasingly contaminated by pine cover, which overshadowed the weak signal from the bare hardwood canopy. Because commercial lidar sensors typically operate at near-infrared

wavelengths, the need for data filtering could possibly be met by using the reflection intensity of each return to discriminate returns from leaves vs. those from branches etc. (Lillesand et al. 2008).

Our lidar dataset was originally developed for the purpose of mapping bare-earth elevation, but our results also encourage further study and use of this statewide dataset for mapping structural characteristics of evergreen vegetation in North Carolina. Following national guidelines (FEMA 2003), similar lidar mapping programs are being implemented in other states as well. The increasing coverage of these datasets shows great promise for mapping and studying evergreen canopy structure, biomass, growth, and habitat value at regional scales.

2.4.2. Measurement of canopy height by Interferometric Synthetic Aperture Radar (InSAR)

Scattering of microwave energy is dependent on the size and arrangement of target objects, with strong signals returned to the sensor from objects with length approximately equal to the wavelength of radiation (Lillesand et al. 2008). Objects much larger than the wavelength do not transmit, but instead tend to reflect the signal away from the sensor. Within a forest canopy, short-wavelength (e.g., X-band) signals are dominated by leaves and twigs, intermediate wavelengths (e.g., C-band) by medium-sized branches, and long wavelengths (e.g., P-band) by large-diameter trunks and the

ground surface. Due to the vertical distribution of canopy elements, a radar signal's penetration into a forest canopy is likewise proportional to its wavelength, resulting in scattering heights of X, C, and P bands sorting vertically from top to bottom in the canopy (Jensen, 2000).

This leads to a sorting of elevation values measured by interferometry based on the different wavelengths, with heights measured by the X band representing nearly the top of the canopy, those from the P band representing the bare earth, and from the C-band intermediate between these two. By pairing a long-wavelength with a short-wavelength band, canopy height is estimated by subtracting bare-earth (BE) from top-of-canopy (TOC) height measurements. In canopies shorter than the penetration depth (denoted by "*" in Fig. 8) of both the BE and TOC bands, canopy height is unreliably measured as approximately zero because the two signals are both dominated by the ground surface. In canopies with height between the penetration depths of the TOC and BE bands, subtracted canopy height is proportional to true canopy height minus the TOC penetration depth. In canopies taller than the penetration depth of both bands, height is unreliably measured as the difference between canopy-sensitive portions of each curve.

Therefore, the response of measured to true canopy height for a given pair of bands follows a three-stage curve, with the most reliable measurements obtained

between the penetration depths of the shorter (TOC) and longer (BE) bands. Although sampling was sparse below 15 m, the transition from the first to the second stage of the curve is apparent in the SRTM measurements (Fig. 5) and from the second to the third stages in the GeoSAR measurements (Figs. 5, 6). Canopy heights derived from SRTM used a relatively long-wavelength (C band) for TOC measurements and used BE elevations from lidar measurements capable of deep canopy penetration. The bias in the SRTM-derived canopy heights is therefore likely due to excessive penetration by the TOC band in short canopies, resulting in a delayed, but nearly constant negative bias in canopies taller than 10 m. Alternatively, GeoSAR, which uses a very short-wavelength TOC band, is likely more affected by incomplete penetration by its P-band BE measurements. This leads to a saturation effect with increasing canopy height, observed as a decreasing slope resulting from the measured BE surface lifting above the true ground elevation in canopies greater than approximately 20 m.

Scattering of microwave radiation in forest canopies is influenced by canopy water content, structure, and viewing geometry (Jensen 2000, Lillesand et al. 2008), so variation in any of these factors leads to imprecision in canopy height measurements. The effect of forest structure is apparent in differences between GeoSAR's precision in hardwood and pine forests. With strong apical dominance of the dominant species and

low species diversity, southeastern pine forests are structurally simpler than hardwood forests, which are richer in species composition and contain species capable of greater structural complexity than the pines. GeoSAR measured height more precisely in pine canopies than in hardwoods, suggesting that structural diversity plays a large role in measurement precision. SRTM's precision was also likely affected, but this effect was

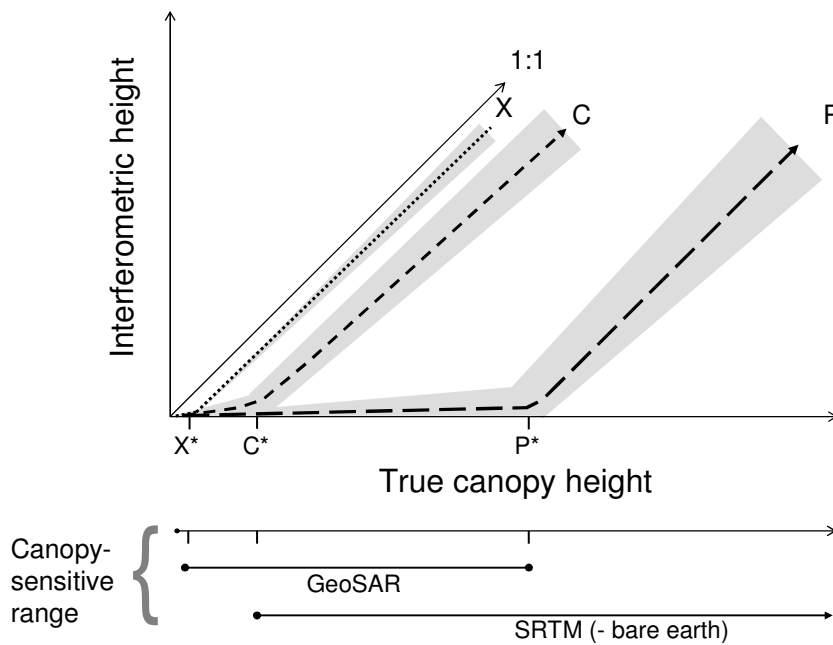


Figure 8: Schematic representation of the idealized response of X-, C-, and P-band radar interferometric heights to "true" canopy height; and the resulting canopy-sensitive ranges of GeoSAR and SRTM measurements. Axes are not drawn to scale.

overshadowed by the data's dormant-season acquisition. Whether to avoid noise from structural or other variations, frequent calibration will likely improve precision when using InSAR to measure canopy height, especially over large, heterogeneous areas.

2.4.4. Calibration of SRTM for canopy height measurement

Uncalibrated differences between the SRTM DEM and bare-earth elevation measurements are known to underestimate canopy height of various forest types due to C-band radar's penetration of the vegetation canopy (Kellndorfer et al. 2004). We also observed this bias in southeastern pine and hardwood forests, noting a substantially larger, more variable bias in dormant hardwood canopies. The increased bias in hardwood canopies shows that the effect of canopy penetration is exacerbated by the leafless condition of deciduous forests at the time of data acquisition.

Although uncalibrated SRTM-derived canopy heights underestimated true canopy height, a consistent physical relationship between SRTM measurements and canopy height allowed calibration of SRTM to more accurate measurements in pine stands. In pine forests, this relationship between canopy height (H) and the difference between the SRTM and a bare-earth DEM (H_{SRTM}) can be represented as:

$$H = 0.435 \times H_{SRTM} + 9.92 \times (H_{SRTM})^{1/3}. \quad (10)$$

Unfortunately, although errors in both pine and hardwood forests were reduced substantially by calibration (from 14.92 to 3.90 m in pine and from 16.84 to ~6.5 m in hardwood), such simple calibration does not appear nearly as effective in dormant deciduous as in fully-canopied forests. Disagreement between calibrated SRTM- and lidar-derived canopy height in this study is greater than that between SRTM and field measurement in managed pine plantations on flat topography after outlier removal (Kellndorfer et al. 2004), but is more comparable to errors in mixed coniferous forest in the Sierra Nevada Mountains in Northern California (Dubayah et al. 2007). The picture emerging from these regional studies is that both forest structural complexity and underlying bare-earth topography effect the accuracy of SRTM-based canopy height measurements (Bourgine and Baghdadi 2005, Bhang et al. 2007), but that calibration is possible in many situations.

In southeastern pine or structurally similar forests for which a reliable bare-earth DEM is available, linear or cube-root calibration of SRTM measurements substantially reduces vertical errors of SRTM-based canopy height estimates. Previous studies have recognized the potential for correcting SRTM underestimates of canopy height through simple linear regression against reference height measurements (Kellndorfer et al. 2004, Dubayah et al. 2007). Our addition of a cube-root term based on the effect of canopy

penetration is important to prevent over-predicting the height of the shortest stands. However, in regions with a mixture of extremely tall and very short stands, or where bare-earth DEMs might be influenced by the forest canopy, calibrations will need to incorporate canopy penetration artifacts from both the SRTM (as top-of-canopy) and bare-earth DEMs. Canopy-height underestimation by SRTM data is a function of vegetation structure and will likely vary if vegetation varies considerably within an area of interest (Bhang et al. 2007, Dubayah et al 2007). It would be beneficial for future efforts to focus on fitting context-dependent corrections to more accurately map vegetation height across large heterogeneous landscapes with SRTM or other C-band data.

2.4.5. Selection among methods

Consideration of sensitivity, scale, and comparability should inform selection of methods for estimating canopy height. Due to its expediency and low cost, *in situ* field measurement remains the best choice for measuring individual or small numbers of tree heights. However, field measurement requires extrapolation from small samples and is not capable of mapping canopy height over any but the smallest stands. Remote sensing provides an economically efficient means to map larger regions, with the various methods sorting along an axis of scale. Due to its fine resolution and customizable post-

spacing, lidar is best suited at scales from several to several million hectares and is useful for calibrating broader-scale InSAR measurements. With larger swaths and systematic processing, airborne InSAR sensors such as GeoSAR achieve economy of scale for larger regions—from hundreds of thousands to millions of hectares. Especially in tall pine forests, GeoSAR's precision should also support increased accuracy through calibration to lidar measurements. At the coarsest scales, an increasing number of spaceborne InSAR sensors—including the European ERS-2 and Envisat, the Japanese PALSAR, the German TerraSAR-X, the Italian COSMO SkyMed constellation, and Canadian RADARSAT series—promises the capacity to monitor forest resources at the global scale in the near future (Lillesand et al. 2008). The SRTM dataset is currently the only potential source of tree canopy height for the entire Earth. Although the SRTM DEM was collected only in 2001 and requires calibration with a separate bare-earth DEM to extract canopy height, it has the potential to provide a baseline measurement against which future measurements can be compared to measure forest growth. For this, calibrations stratified by ecoregional forest types must be conducted wherever circa-2001 canopy height datasets are available.

2.5. Conclusions

Forest canopies are complex volumes, and several methods—each with characteristic sensitivities, biases, and limitations of scale—are available to measure their height. Due to simpler structure and phenology, pine forests allow greater precision of height measurements than do hardwood forests in the southeastern United States. In southeastern pine forests, lidar measurements acquired in the dormant season have the highest precision of all methods studied, followed by those from field measurements and interferometric synthetic aperture radar (InSAR). Raster canopy-height surfaces aggregated from lidar returns are sensitive to the resolution at which they are aggregated. InSAR-derived canopy height measurements are subject to biases from factors including variable canopy penetration, but show potential for calibration based on lidar. X- and P-band interferometric radar measurements from airborne sensors (e.g., GeoSAR) provide more precise, accurate measurements than C-band data from spaceborne sensors (i.e., SIR-C/SRTM), but even the SRTM DEM can be calibrated to estimate canopy height across broad regions. Each of the remotely sensed methods studied produces reasonable and consistent depictions of canopy height that can be compared with data of similar provenance, but due to differences in underlying sensitivities between the methods, comparisons between measurements from various sources require cross-calibration and will be most useful at broad scales.

3. Landcover dynamics of the North Carolina Piedmont between 1985 and 2006 by temporal signature extension

3.1. Introduction

For more than 30 years, the Landsat satellites have imaged Earth's terrestrial and near-shore marine surfaces. These images support ecological research, which in turn provides governments and other parties with assessments of ecosystem quality, resources, services, and changes due to natural and anthropogenic forcing (National Academy of Science 2005). Such a long and consistent record allows scientists to study current phenomena, relating satellite data to "ground-truth" or other references before interpreting patterns, but also to retrieve information from the past, for which little or no reference data are available.

Landcover is a categorical description of Earth's terrestrial ecosystems. Maps of landcover are relied upon for many applications (Anderson et al. 1976), including assessments of wildlife habitat (Scott et al. 1993), watershed chemistry (Herlihy et al. 1998), and climate effects (Feddema et al. 2005). Among the most widely used, the 2001 National Land Cover Database (NLCD 2001) consists of fractional tree canopy and impervious surface cover and categorical landcover/landuse layers covering the United States. Beyond map provision *per se*, the goals of the Multi-Resolution Land Characteristics Consortium (MRLC) in creating NLCD 2001 were to foster exploration, development, application, and sharing of landcover information (Homer et al. 2004).

NLCD 2001 is an accurate and nationally consistent database for a single time, but no such data are available over multiple times.

Monitoring landcover changes over time requires consistent maps, both ontologically and statistically. That is, the maps must have identical semantics between real-world ecosystem *types* and schematic *classes*; and, unless accuracy is extremely high, they must also be based on similar correlations to the image data on which they are based. Without these two properties, it is impossible to determine whether changes observed are due to actual landscape changes or to ontological and statistical artifacts from the underlying data model.

Comparing inconsistent maps leads to unreliable conclusions. Consider the changing proportions of four landcover classes over time (Fig. 1) for a 1.7 Mha region of North Carolina (Fig 2) obtained from three publicly available landcover maps (NLCD 1992, NCCGIA 1996, and NLCD 2001). The general, 4-class classification scheme was aggregated from each of the maps' original schema to maximize comparability. In this depiction, forest cover increased slightly from 1992 to 1996, then dropped 23% of its 1996 cover (from 71 to 55% of the region's area, a change of 272,000 ha) in the next five years. Fields increased from 21 to 23% of the landscape in the first period (a large but reasonable change), but then increased 102,000 ha to 29% over the final five years. Urban area first dropped by half in the first four years, then nearly quadrupled in the next six.

Water, the most semantically and spectrally distinct landcover type, remained nearly constant over the classifications. Clearly more representative of underlying model

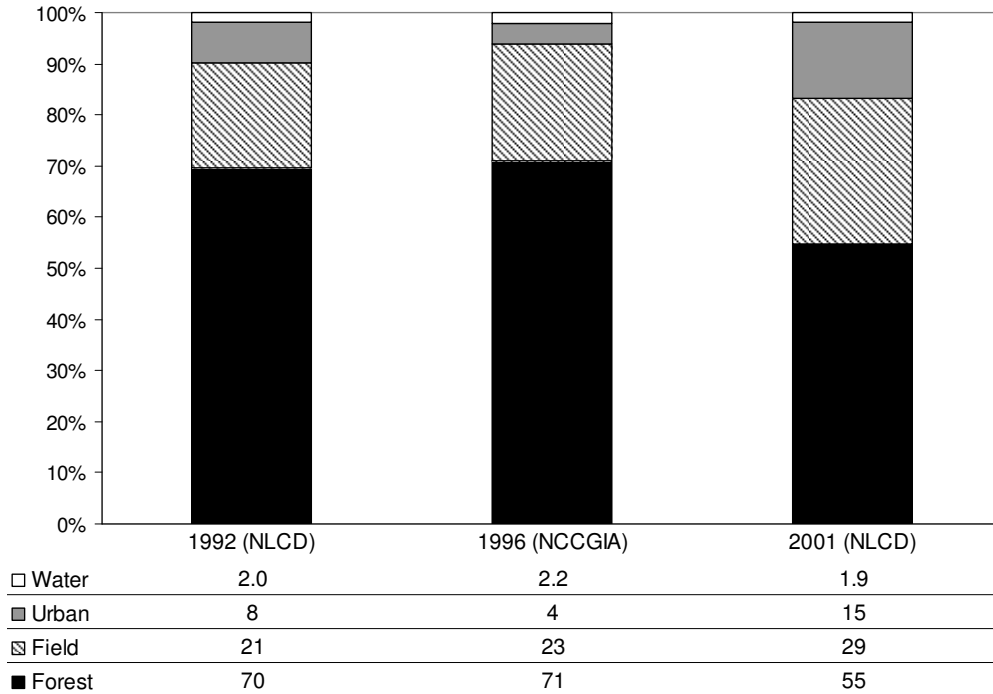


Figure 9: Regional landcover change from three different sources: the 1992 National Land Cover Database, the North Carolina Center for Geographic Information and Analysis (1996), and the 2001 National Land Cover Database.

than of changes in the real landscape, these trajectories would not be useful for further scientific analysis or policy decisions.

These ambiguities would be avoided by a single, precise model applied consistently over time. As the basis for traditional image classification and change

detection over space, this “signature extension” was an active research topic early in the history of Landsat (e.g., Minter 1978), and remains so today (Woodcock et al. 2001, Olthof 2005). The process of signature extension requires training a classification model (i.e., the spectral signature) on a small number of landcover observations matched with spectral measurements in time and space and then extrapolating the classifier on image data collected at other times and/or locations for which no such reference is available (Botkin et al. 1984, Jensen 1983, Muller 1988, Cihlar 2000).

Whether in space or in time, signature extension relies on invariance of the spectral signature. In time, this consistency can be compromised by numerous factors, including sensor degradation, atmospheric contamination, the bidirectional reflectance distribution function of the surface, and variable plant phenology (Song and Woodcock 2003). Parallax and topographic shading also present problems in areas with steep topography.

All of these effects can be accommodated, but to varying degrees. Sensor radiometry is calibrated systematically by the National Landsat Archive Production System (NLAPS) (Chander and Markham 2003), and bidirectional and atmospheric correction can be reasonably accomplished through models of varying complexity (Song et al. 2001), but phenology varies among vegetation types and environmental characteristics (Badeck et al. 2004, Schaber and Badeck 2003). Even after being

minimized by image pre-processing, phenological and other effects contribute significant noise to models of landcover signatures.

Despite these difficulties, monitoring landscape change is possible, especially in ecosystems defined by a small number of types. Pax-Lenney et al. (2001) devised an innovative generalization approach to extend spectral signatures of conifer forest across time and space based on fuzzy ARTMAP neural networks. Masek et al. (2008) used relative changes in the Disturbance Index (DI) (Healey et al. 2005) to monitor continuous, “sub-pixel” forest development and map forest disturbance and recovery across North America. Although focused solely on forest changes, these studies suggest the potential for single-model change detection in multinomial landcover systems as well.

Based on the data processing and classification scheme of the NLCD 2001 Land Cover Layer, this study demonstrates a successful adaptation of spatial signature extension methods for multi-temporal landcover classification. The study’s primary objective was to produce a time-series of comparable landcover maps over the span of the Landsat-5 mission within a single path-row scene to support ecological analysis and local land-use governance. I report the image processing, sampling, and modeling methods used to maximize the classification’s accuracy and its robustness to phenological and other image-to-image variation. I also describe modifications to

traditional validation statistics that remove distributional sampling bias, thereby allowing comparisons of accuracy across independent test samples. Through a simple landcover change analysis, the series shows nonlinear landcover trajectories that would be overshadowed by differences between multiple classifications. Based on this success, I conclude by discussing the requirements and potential avenues for future improvement to multi-temporal landcover classification.

3.2. Methods

3.2.1. Study area

Extending from New York City to central Alabama, the Piedmont Plateau is America's most populous region. On the North Carolina Piedmont, extensive pine forests established in farm fields abandoned early in the twentieth century are succeeding naturally to dominance by hardwood forests, but are also rapidly being converted to suburbs around metropolitan centers (Taverna et al.2004, McDonald and Urban 2006). Piedmont soils are spatially complex, but the region's topography is characterized by gentle relief, with broad river valleys draining extensive uplands. Phenologically, the area's agricultural fields and pastures begin "greening up" in middle- to late March, with the deciduous forests following throughout the month of April. Responding variously to drought, storms, day-length, temperature (and

ultimately harvest for herbaceous crops), senescence ranges from early September to early December.

The 1.67 million-ha (Mha) study area is defined by the intersection of the Landsat World Reference System 2 (WRS-2) Path 16/ Row 35 scene and the North Carolina counties lying completely on the Piedmont Plateau (Fig.2). From east to west, the study area spans between the Triangle region (framed by the cities of Raleigh, Durham, and Chapel Hill) and the Triad region, framed by Greensboro, High Point, and Winston-Salem. The area's combination of: (1) rapid and complex landcover changes in recent years, (2) high spatial edaphic and hydrological variability, and (3) low topographic relief make the region ideal for studying mixed natural and anthropogenic landcover dynamics.

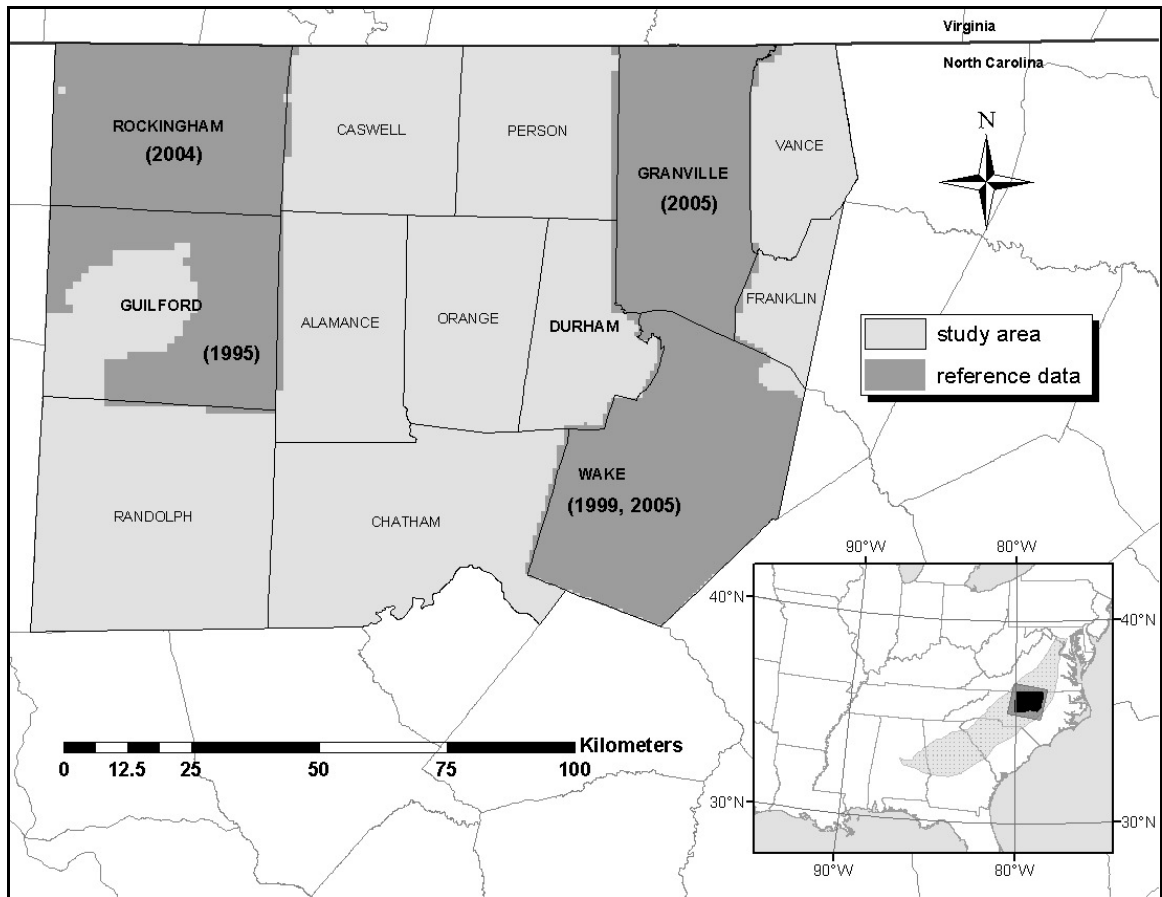


Figure 10: The 1.7-Mha study area lies on the Piedmont Plateau (inset—shaded, stippled area), partially or entirely covers thirteen North Carolina counties, and is completely within WRS-2 Path 16, Row 35 (inset—shaded box). Split into separate training and test datasets, reference data were collected over 4 counties and four years.

3.2.2. Data

3.2.2.1. Landsat

Fifty-five Landsat-5 Thematic Mapper (TM) images (Table 1) from the WRS-2 Path 16, Row 35 scene were acquired and geometrically rectified to 1-m resolution digital ortho-photographs by 1st- or 2nd-order polynomial transformation. Spatial

misregistration was less than 20 meters RMSE for most images, and every image had RMSE < 25 m. Clouds and their shadows, snow, and radiometrically “bad” pixels were identified visually and removed by on-screen digitizing. The solar-reflective (i.e., non-thermal) bands of the screened images were then converted to radiance following Chander et al. (2007) and atmospherically corrected to surface reflectance by the DOS3

Table 1: Landcover classification years and input Landsat images. Reference data for model training and validation were obtained from years in bold. Intermediate (Pass 2) landcover class probabilities for terminal years (in italics) were used for temporal filtering of adjacent years, but were not converted to categorical landcover maps.

year	Early growing season	Late growing season	Dormant season
1984	<i>June 15, 1984</i> <i>July 1, 1984</i>	<i>September 3, 1984</i> <i>September 19, 1984</i>	<i>January 9, 1985</i>
1985	May 1, 1985	September 6, 1985	January 9, 1985
1986	May 4, 1986	September 25, 1986	December 14, 1986
1987	June 8, 1987	October 14, 1987	December 14, 1986
1989	May 28, 1989	October 3, 1989	December 19, 1988
1991	May 2, 1991	October 9, 1991	January 26, 1991
1993	May 7, 1993	September 28, 1993	December 1, 1993
1994	April 24, 1994	September 15, 1994	December 12, 1993
1995	June 14, 1995	September 15, 1994	February 6, 1995
1996	July 2, 1996	September 20, 1996 October 22, 1996	December 23, 1995
1998	May 18, 1997	September 26, 1998	January 29, 1998
1999	June 9, 1999	October 15, 1999	December 2, 1999
2001	April 27, 2001	September 18, 2001 October 4, 2001	November 21, 2001
2004	May 21, 2004	September 24, 2003	January 1, 2004
2005	May 5, 2008	October 15, 2005	December 15, 2004
2006	April 25, 2006	October 15, 2005	December 5, 2005
2007	<i>May 14, 2007</i>	<i>August 18, 2007</i>	<i>December 15, 2006</i>

Dark-Object Subtraction approach (Song et al. 2001).

The surface reflectance images were masked to the study area extent and gathered into composite datasets for training, extrapolating, and validating the classification. Following Homer et al. (2004), images were divided into seasonal triplets: early growing season (“spring”) (April 24 – July 2), late growing season (“fall”) (September 3 –October 22), and dormant season (“winter”) (November 21 – February 2). Tasseled-Cap “brightness” (B) and “wetness” (W) (Kauth and Thomas 1976, Crist 1984, Crist and Cicone 1984) and the Normalized Difference Vegetation Index ($NDVI$) (Rouse et al. 1973) were computed for each image, and indices from multiple images were averaged within seasons when available.

Idiosyncrasies of crop phenology and image acquisition date frequently cause confusion of herbaceous types with urban and forest classes (Seto et al. 2002, Yuan et al 2005). To minimize errors due to phenological variation within the growing season, spring and fall brightness and $NDVI$ were combined into growing-season averages. Surface moisture is also greatly impacted by chance events, and so mean and maximum annual wetness were computed from each annual triplet to minimize variation from year to year. The final dataset thus consisted of six annual metrics: mean growing-season brightness and $NDVI$ (B_{Savg} and $NDVI_{Savg}$), winter brightness and $NDVI$ (B_W and $NDVI_W$), and annual mean and maximum wetness (W_{Yavg} and W_{Ymax}):

$$B_{Savg} = \frac{B_{spring} + B_{fall}}{2}$$

$$NDVI_{Savg} = \frac{NDVI_{spring} + NDVI_{fall}}{2}$$

$$W_{Yavg} = \frac{W_{spring} + W_{fall} + W_{winter}}{3}$$

$$W_{Ymax} = \max(W_{spring}, W_{fall}, W_{winter})$$

3.2.2.2. Landcover

The classification scheme was based on the 2001 National Land Cover Database (NLCD 2001) Land Cover Layer, with minor modifications (Table). The NLCD “grassland” class (71) was disregarded because natural grasslands are an extreme rarity on the North Carolina Piedmont. The “developed open space” class (21) was also disregarded because its identification is contextual, requiring municipal boundaries which are not available for the entire span of this study. The NLCD “shrubland” class (52) was reinterpreted as short, young forests, or “scrub”.

Table X. Landcover classification scheme, based on the 2001 National Land Cover Database Land Cover Layer (Homer et al. 2004).

Label	name	Description
11	Water	Open water (100% cover)
22	Urban, low-intensity	Mix of built and natural cover, < 50 % impervious
23	Urban, medium-intensity	Mix of built and natural cover, 50 - 75 % impervious
24	Urban, high-intensity	Mostly built cover, > 75 % impervious
31	Bare	Cleared soil or rock, with < 25 % vegetation cover
41	Forest, deciduous	Trees > 5m tall and > 50 % cover, < 25 % evergreen
42	Forest, evergreen	Trees > 5m tall and > 50 % cover, > 75 % evergreen
43	Forest, mixed	Trees > 5m tall and > 50 % cover, 25 - 75 % evergreen
52	Scrub	Mix of trees (< 5 m tall) and other natural cover, > 25 % vegetated
81	Pasture/hay	Fields with > 25 % persistent herbaceous cover, dominated by non-cereal grasses
82	Row crops	Fields with > 25 % herbaceous cover, periodically harvested
90	Wetland, woody (swamp)	Mix of water, herbaceous plants, and bare soil
95	Wetland, herbaceous (marsh)	Mix of water, trees, and bare soil

Reference landcover data were collected by a single observer via stratified random sampling and interpretation of orthorectified high-resolution digital images from Guilford County in 1995, Randolph County in 2004, and Wake and Granville Counties in 2005 (Fig. 2). The NLCD 2001 Land Cover Layer was clipped to the boundaries of each high-resolution image set and filtered to exclude spatial heterogeneity of landcover within a 50-m radius of each pixel. 150 pixels of each class (excluding classes 21 and 71) were then randomly sampled from the masked landcover layer for each reference year and converted to points labeled by NLCD 2001 landcover

Table 2: Landcover proportions in the reference sample and study area (taken from NLCD 2001), and model priors derived from NLCD 2001 by splitting probabilities for class 71 equally between classes 81 and 82, and class 21 proportionally among non-urban land classes (i.e., excluding classes 11, 23, and 24). Years in bold were used for both model training and testing, whereas 1999 was reserved solely for model testing.

class	1995		1999		2004		2005		Train (n)		Test		NLCD 2001		Model priors
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	Kha	(%)	(%)
11	115	(14.04)	109	(10.03)	112	(11.90)	130	(13.92)	144	(9.73)*	213	(17.55)*	32.00	(1.93)	1.93
21	--	--	--	--	--	--	--	--	--	--	--	--	147.27	(8.86)	0.00
22	124	(15.14)	82	(7.54)	85	(9.03)	83	(8.89)	136	(9.19)	156	(12.85)	65.31	(3.93)	4.33
23	35	(4.27)	56	(5.15)	54	(5.74)	48	(5.14)	103	(6.96)	34	(2.80)	23.80	(1.43)	1.43
24	69	(8.42)	117	(10.76)	108	(11.48)	63	(6.75)	105	(7.09)	135	(11.12)	8.53	(0.51)	0.51
31	9	(1.10)	0	(0.00)	61	(6.48)	56	(6.00)	96	(6.49)	30	(2.47)	2.98	(0.18)	0.02
41	49	(5.98)	189	(17.38)	48	(5.10)	30	(3.21)	97	(6.55)	30	(2.47)	610.98	(36.77)	40.51
42	70	(8.55)	147	(13.52)	45	(4.78)	80	(8.57)	100	(6.76)	95	(7.83)	192.83	(11.61)	12.78
43	70	(8.55)	0	(0.00)	60	(6.38)	79	(8.46)	145	(9.80)	64	(5.27)	75.11	(4.52)	4.98
52	23	(2.81)	44	(4.05)	112	(11.90)	139	(14.88)	148	(10.00)	126	(10.38)	35.86	(2.16)	2.38
71	--	--	--	--	--	--	--	--	--	--	--	--	81.24	(4.89)	0.00
81	33	(4.03)	117	(10.76)	50	(5.31)	20	(2.14)	51	(3.45)	52	(4.28)	332.22	(19.99)	24.72
82	146	(17.83)	201	(18.49)	106	(11.26)	89	(9.53)	151	(10.20)	190	(15.65)	26.87	(1.62)	4.47
90	73	(8.91)	25	(2.30)	71	(7.55)	87	(9.31)	155	(10.47)	76	(6.26)	25.84	(1.56)	1.71
95	3	(0.37)	0	(0.00)	29	(3.08)	30	(3.21)	49	(3.31)	13	(1.07)	0.69	(0.04)	0.05
total	819		1087		941		934		1480		1214		1661.52	Kha	100 %

class. These points were then overlaid on the high-resolution images and re-labeled with the image year and landcover observed in that year. NLCD 2001 labels were used as initially as a guide to help teach and maintain consistency between observers. Points were discarded if their landcover could not be determined or was mixed with other types within a 30-m radius.

The reference sample was then randomly divided into sub-samples for model training and validation. Poor image quality (e.g., in 1995) or rarity of classes led to large differences in abundance of observations among the classes, and so allocation of data between the training and validation samples was determined by its abundance in the multi-year sample (Table 2). Classes that were abundant in the sample were represented by approximately 150 observations in the training subsample, those with intermediate abundance by ~100 observations, and classes with the lowest abundance by ~50 observations. The remaining observations from each class were allocated to the test sample, to which 1087 observations made by a second observer over Wake County in 1999 were added. Thus, approximately half of the observations upon which model validation was based were spatially independent from the training data, and the remaining half was unique in terms of space, time, and observer.

3.2.3. Landcover classification

Landcover was classified by Classification and Regression Trees (CART) and Quadratic Discriminant Analysis (QDA) followed by temporal and spatial filtering, executed in three passes. CART (Breiman et al. 1984) is a nonparametric algorithm that recursively partitions the spectral domain into maximally homogeneous regions of landcover types. Most often referred to as the “Maximum Likelihood Classifier” in the remote sensing literature (Jensen 1996), QDA estimates the probability of each landcover class across the spectral domain based on the Mahalanobis distance from the class centroid. Preliminary analyses showed that CART provided superior landcover extrapolations for phenologically invariant landcover types, but that the continuous probability density functions provided by QDA were more robust to image-to-image variation and that the resulting probabilistic landcover extrapolations from QDA were more amenable to application of prior probabilities and temporal filtering. Analyses were performed in ArcGIS 9.3 (ESRI, Inc.) using the Geoprocessor object in Python 2.7 (Python Software Foundation), ERDAS IMAGINE (Leica Geosystems, L.L.C.), and the *rpart* library of the statistical software R version 2.6.1 (R Development Core Team 2007, Therneau and Atkinson 2007).

During the first pass, CART was used to determine a threshold of maximum annual wetness for discriminating water from land, based on the three-year training sample. The threshold was then applied to the series of maximum wetness images to

create a water mask for each year, assigning the probability of water in all pixels meeting the criterion to the user's accuracy (UA) calculated for the independent test sample and the probability of water for all other pixels to $1-UA$ (see *Appendix for explanation of UA and other accuracy metrics*).

In the second pass, QDA was used to estimate spectral signatures and assign probabilities of all classes, including water. Prior probabilities of the landcover classes were obtained from the NLCD 2001 Land Cover Layer (Table 2) and were applied to the QDA likelihoods following Strahler (1980) and Jensen (1996). To accommodate the class removals, the prior probability of the excluded class 21 (developed open space) was distributed equally among all remaining land classes, and the prior probability of grassland (71) was distributed equally between pasture/hay (81) and row crops (82).

After obtaining a vector of initial probabilities p'_c across the set of classes $c \in C$ from QDA and a second probability of water from CART in each pixel x and time t , a combined probability of water at each (x,t) was calculated as the maximum of the initial probabilities of water from the two classifiers at (x,t) . Two sets of classes were also regrouped based on spectral and semantic similarity by summing the probabilities of the member classes. The herbaceous classes "pasture/hay" (81) and "row crops" (82) were merged into a single herbaceous class (80), and the "mixed forest" (43), and "forested wetland" (90) were merged into the deciduous forest (41) class. The new vector of merged class probabilities p'' was then adjusted to sum to unity:

$$p_c''(x,t) = \frac{p_c'(x,t)}{\sum_{c \in C} p_c'(x,t)}. \quad (1)$$

In the third pass, filters were applied to “de-speckle” the image series in space and time. First, a temporal filter was applied to the $p_c''(x,t)$ to minimize the occurrence of transient landcover changes. The filter recalculated each class’s probability in a pixel at time t as the maximum of its probability at t vs. the square-root of the joint probability of c in the years before and after:

$$p_c'''(x,t) = \max\left(p_c''(x,t), \sqrt{p_c''(x,t_-) \times p_c''(x,t_+)}\right) \quad \text{for all } c \in C, \quad (2)$$

where t_- and t_+ are the closest years before and after the focal year t for which the necessary p_c'' were available. The new probabilities of each pixel were then adjusted to sum to unity via equation (1) and the pixel was assigned the label of the landcover with the maximum adjusted probability. The spatial “Clump” and “Eliminate” operations were then performed in ERDAS IMAGINE on the categorical landcover maps to further reduce speckle and spurious landcover changes due to spatial misregistration. The clump function identifies each pixel with a group of equally valued neighbors, and the eliminate function recodes the landcover of pixels in groups whose area is below a given

threshold. I clumped the data using an 8-neighbor rule and eliminated clumps with area < 1 ha.

3.2.4. Accuracy assessment

Accuracy assessment was based on confusion matrices calculated for the test sample, using producer's accuracy (PA), user's accuracy (UA), percent correctly classified (PCC), and the Kappa statistic (κ) as validation metrics. Class proportions in the sample were neither uniform nor representative of the study area at any time, so it was necessary to adjust PCC and Kappa estimates to neutralize sampling bias (Congalton 1991, Byrt et al. 1993, Appendix A). A vector of class weights:

$$w_i = \frac{p_i^*}{p_i}, \quad (3)$$

in which p_i is the proportion of class i in the reference sample and p_i^* is the probability of i from a distribution chosen *a priori* as a standard, was applied to calculations of PCC and Kappa. As in model training, p_i^* was derived from NLCD 2001 (Table 2). The standardized percent correctly classified (PCC_s) and Kappa (κ_s) metrics were thus calculated from the joint class frequencies (f_{ij}) of the confusion matrix as:

$$PCC_s = \frac{1}{n} \sum_{i=j} w_i f_{ij} \quad (4)$$

and

$$\kappa_s = \frac{\sum_{i=j} w_i f_{ij} - \sum_i w_i f_i f_j}{n^2 - \sum_i w_i f_i f_j}. \quad (5)$$

3.3. Results

3.3.1. Landcover classification

A loose hierarchy of similarities existed between the classes (Fig. 3). Water was spectrally distinct from all of the land classes. Forests formed a cohesive group, but the various gradations of deciduous forest—i.e., deciduous forest (41) and mixed forest (43), and (predominantly deciduous) swamps (90)—were more similar to scrub (52) and herbaceous marshes (95) than they were to evergreen pine forests (42). This broad forest group was distinct from the urban and field classes, within which high-density urban (24) was spectrally unique. The remaining urban, field, and bare classes were spectrally similar to one another, with medium-density urban (23) and bare ground (31) exhibiting the greatest similarity and the field classes—pasture/hay (81) and row crops (82)—also similar to one another.

The optimal threshold of maximum annual wetness estimated by CART for discriminating water from land was -0.01588. Above this threshold, the probability of water (i.e., user's accuracy, *UA*) was 0.97. Among the individual years, user's accuracies for water ranged from 83.7 (2004) to 0.98 (1995). Most water misclassifications by CART were made in forested and herbaceous wetlands, but a small minority were also committed in urban types.

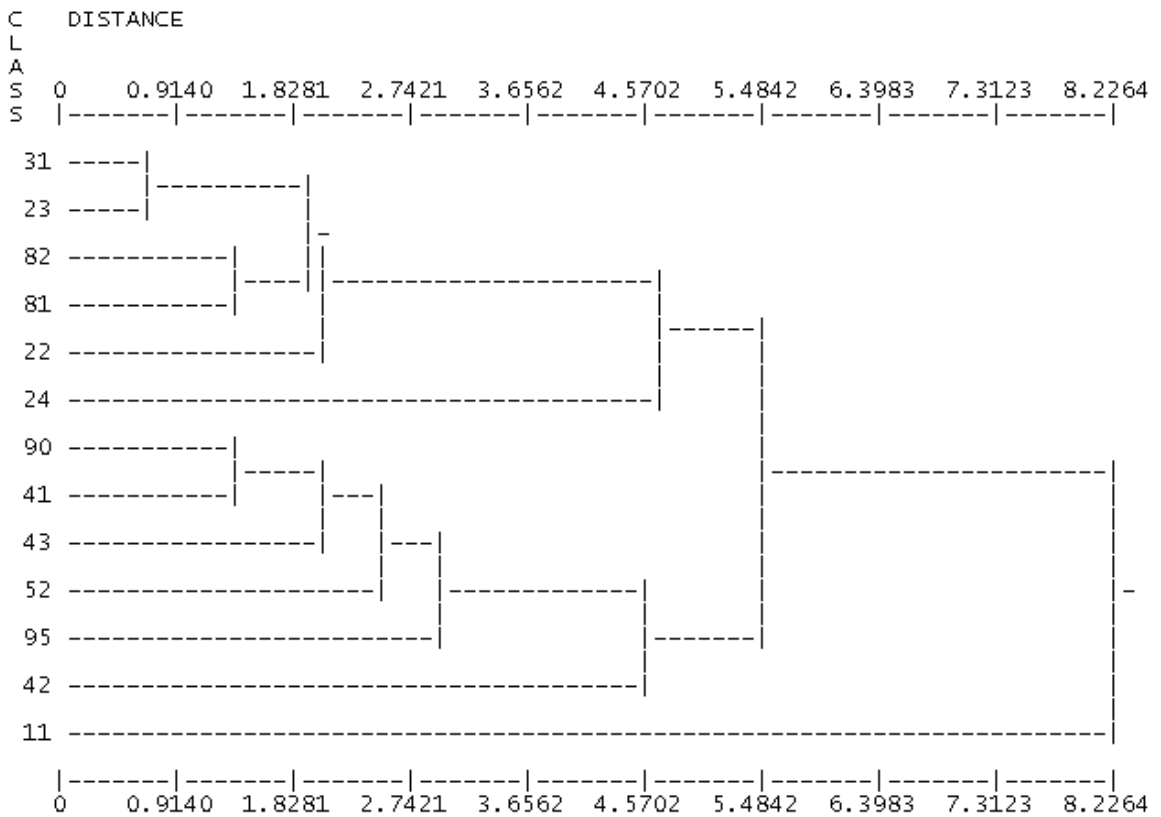


Figure 11: Dendrogram of landcover class centroids in the spectral domain, estimated by Quadratic Discriminant Analysis.

The test sample was classified with $PCC = 75\%$ and $\kappa = 0.70$. After adjusting for the distribution of classes, the standardized (map) accuracy rose to $PCC_s = 88\%$ and $\kappa_s = 0.84$. Water was accurately classified without bias, with producer's accuracy (PA) and UA both equaling 98%. Within urban types, classification was biased toward lower densities, with high-density urban (24) more often labeled as medium-density urban (23) and medium-density urban more often labeled as low-density (22). Low-density urban was under-predicted, most frequently in favor of herbaceous fields (80). Fields themselves were over-predicted at the expense of all types except water, but were especially greedy of the low-density urban type. Both evergreen and deciduous forests were classified with very high accuracy. The rare types—bare (31), scrub (52), and marsh (95)—were classified poorly.

3.3.2. Landcover change

Landcover change over the study area was slight overall, but showed long-term trends, temporal nonlinearity, and strong spatial asymmetries. The four general types—forest, field, urban, and water—gained or lost between 5 and 20% of their 1985 respective areas, with greater proportional changes in the rarer classes (Fig. 12).

Proportional to the entire study area, forest cover increased from ~50 to ~53 % over the period, expanding mostly at the expense of herbaceous fields, which decreased

Table 3: Confusion matrix for landcover classification validated on a 4-year test sample, with producer’s accuracy (PA), user’s accuracy (UA), Kappa (κ), percent correctly classified (PCC) and Kappa (κ), as well as standardized percent correctly classified (PCCs) and standardized Kappa (κ_s). Standardizations were computed based on priors from NLCD 2001 (Table 2).

Class		Prediction										PA (%)
		11	22	23	24	31	41	42	52	80	95	
Reference	11	317	1	0	0	0	2	0	0	2	0	98.4
	22	0	97	10	0	0	20	11	0	100	0	40.8
	23	0	51	31	2	0	0	2	0	4	0	34.4
	24	0	7	110	121	0	0	0	0	14	0	48.0
	31	1	0	3	5	0	0	0	0	21	0	0
	41	0	3	2	0	0	351	16	2	9	1	91.4
	42	0	1	0	0	0	20	209	0	12	0	86.4
	52	1	3	0	0	0	48	7	46	65	0	27.1
	80	0	1	1	1	0	3	0	0	554	0	98.9
	95	3	0	0	0	0	6	0	0	3	1	7.7
UA (%)		98.4	59.2	19.7	94.8	--	78.0	85.3	95.8	70.7	50.0	n = 2301
		PCC = 75.1		PCC _s = 88.1		$\kappa = 0.70$		$\kappa_s = 0.85$				

from 42-46 % in the late 1980s to 38-39 % around 2005. Urban area increased from 3-5% in the beginning of the period to 6-7% at the end of the period. The coverage of water increased from 1.9 % of the landscape in 1985 to 2.1 % in 2006, with the greatest gain associated with the filling of reservoirs around the city of Burlington between 1987 and 1989.

High-frequency temporal dynamics were blurred by errors in individual years. Three years—1991, 1993, and 2001—especially exhibited transient landcover changes that were more likely due to phenological or other sources of noise than to true changes. However, the misclassifications were not consistent across the outlier years. Urban area

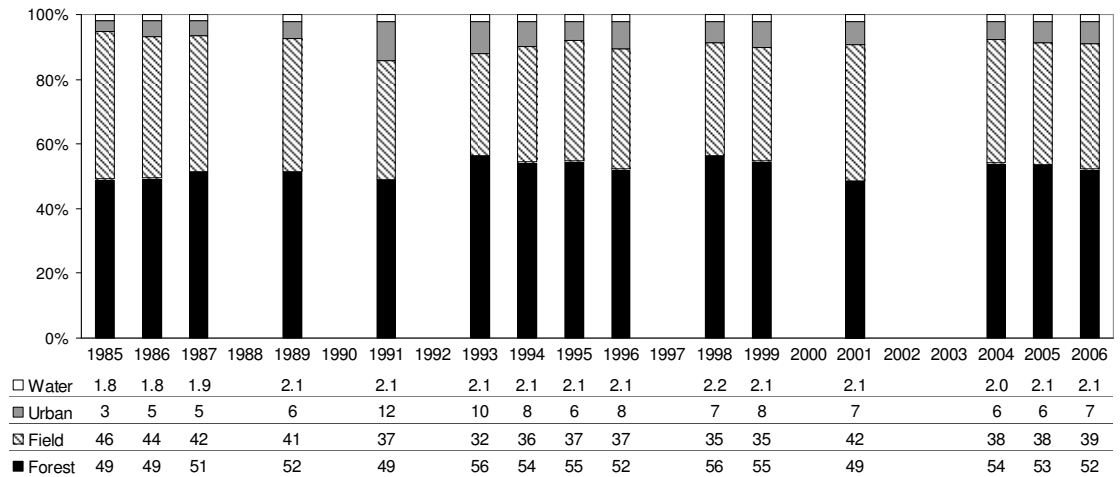


Figure 12: Landcover change among generalized classes within the study area from 1985 to 2006.

increased erratically in both 1991 and 1993—in 1991 at the expense of both forests and fields but in 1993 at the expense of fields alone. In 2001, fields increased erratically at the expense of forest area, with urban area mostly unaffected.

Ignoring these outlier years, the region’s landcover changes exhibited slight nonlinearity over time. The area of agricultural fields dropped from its maximum near 45 % in the 1980s to a minimum around 36 % in the 1990s, then rose again slightly to ~38 % at the end of the period. Forest area changed complementarily, increasing from a minimum in the 1980s to a maximum of ~55 % in the 1990s and dropping again slightly to ~53 % in 2005. Acceleration of urbanization was more difficult to discern due to imprecision in the low-density urban class, but it appears that urban growth was most rapid in the 1980s and 1990s and slowed thereafter.

The effect of classification error is made apparent by comparison of three classes with different error structures (Fig. 13). The imposition of a single set of reference priors over time resulted in constant predictions of the upper margin of error for each class, but lower margins varied in response to varying posterior class probabilities. Thus, the lower prediction interval for each class expanded in proportion to its predicted cover.

Strong asymmetries between the upper and lower bounds of class 80 resulted from biased classification of fields (high *PA*, low *UA*), with over-prediction leading to negatively biased prediction intervals. Due to this over-prediction, the observed changes in field cover were smaller than the class's margin of classification error. The cover of low-density suburbs at any time is also surrounded by broad intervals relative to the magnitude of predicted cover, but roughly equal errors of omission and commission led to more equal distribution of uncertainty in this class. The high precision and minimal bias associated with the classification of evergreen forests resulted in narrow, symmetric prediction intervals of this class over time.

Landcover changes were disproportionately spread across the study area, with greatest change near cities (Fig 14). Development was most pronounced in the Triangle region, where the urban cover of Raleigh and Durham and smaller towns of Cary and Apex expanded predominantly into the surrounding agricultural fields. Urban area also

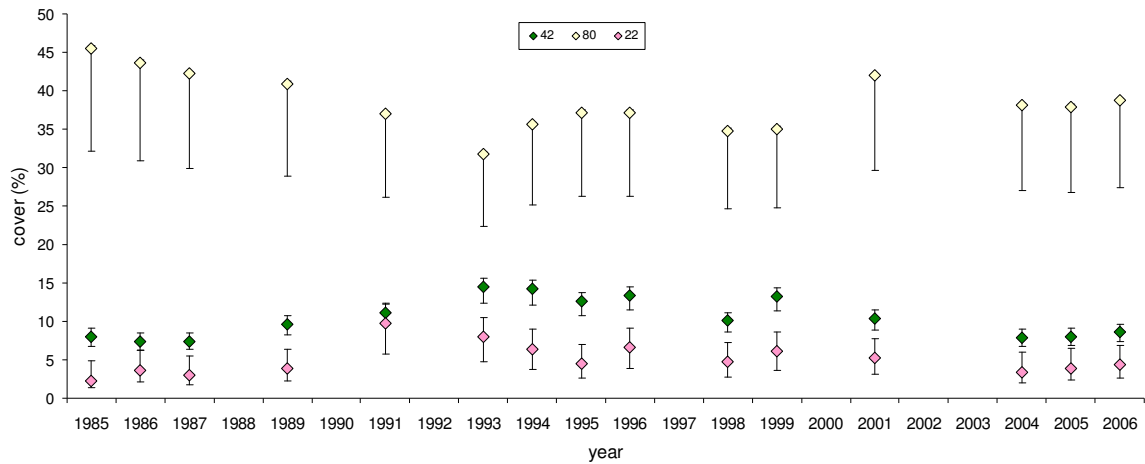


Figure 13: Change trajectories of three selected landcover classes, with expected errors of omission (top bar) and commission (bottom bar)—see Appendix for explanation.

expanded significantly in the Triad, but in comparison to the larger agglomerations, the relatively isolated city of Burlington expanded only slightly. Overall, the region exhibited a broad initial conversion of agricultural fields to forests from 1985 to 1995, followed by partial recovery of fields over the next decade, but these rural changes were diffuse, with no clear spatial pattern.

3.4. Discussion

3.4.1. Four aspects of classification

Multi-temporal landcover classification adds another dimension to purely spatial approaches, but the added complexity does not fundamentally change the traditional problem. Past efforts at temporal signature extension have been hindered not so much by inherent difficulty of the problem as by insufficient data availability—both of

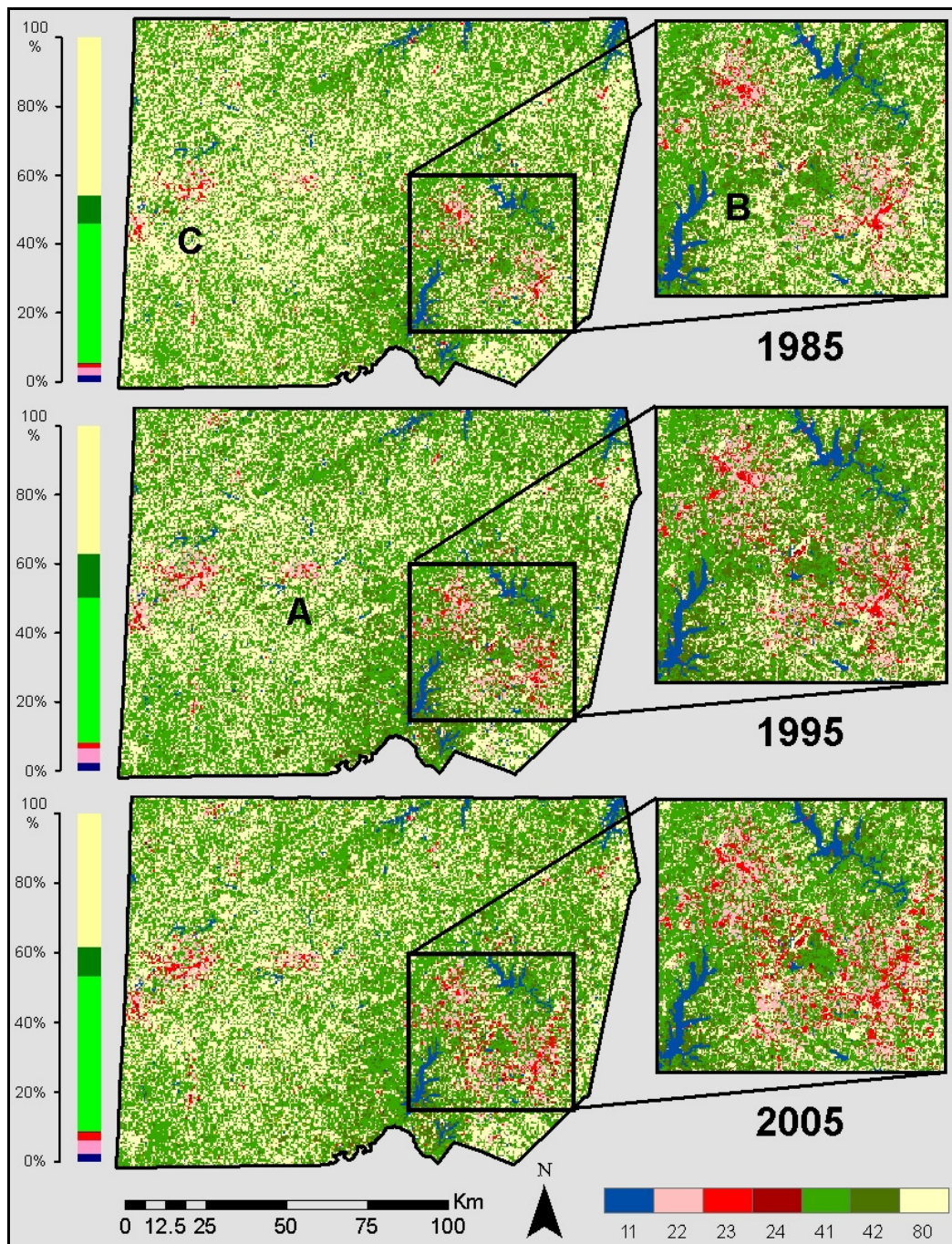


Figure 14: Regional cover of the dominant landcover classes for three selected years, with detail of the Triangle Region (inset). (A) reservoirs built around the city of Burlington between 1987 and 1989 that accounted for the region's greatest increase in water area over the period; (B) urban expansion in the Triangle region from the towns of Cary and Apex into the surrounding agricultural fields; (C) the Triad region, consisting of the cities of Greensboro, High Point, and Winston-Salem (off map).

high-resolution images used for reference and of Landsat data, which were prohibitively expensive to purchase in large quantities until their acquisition became free in early 2009. These limitations led to greater effort spent toward improving accuracies of single-date classifications than toward increasing inter-annual robustness.

Variation in spectral signatures is compounded when models are extrapolated over space and time, but the reliability of multi-temporal landcover classifications can be maximized with attention to both the semantic and statistical aspects of the classification. In particular, four key aspects must be addressed:

- (1) semantic relationships between the model and reality,
- (2) causal relationships between landcover and the data used to predict it,
- (3) representation of variability over space and time, and
- (4) uncertainty accumulated in the modeling process.

Proper attention to these aspects facilitates adaptation of spatial signature extension for spatio-temporal application.

3.4.1. Semantics

The ontological stage of landcover classification is unavoidably subjective. Defining the set of classes to be assumed *a priori* by all subsequent analyses requires a

great deal of inductive reasoning and foresight, as well as compromise between the classes desired and those that are discernible from the data. This effort is well spent however, as loose definitions can lead to spurious conclusions and bad decisions (Comber et al. 2005).

The information potential of remotely sensed imagery is partitioned among its spatial, temporal, and spectral (or radiometric) domains (Jensen 1996). Likewise, information in a remotely sensed geospatial dataset is partitioned between its spatial, temporal, and thematic domains. Given a finite limit of available information, this leads to a tradeoff of scales. That is, if one wishes to expand the representation of one domain (by either increasing resolution or widening extent) without collecting more data, he or she is forced to reduce information in one or both of the others. Two examples illustrate this point: (1) refining the classification scheme of a landcover map (e.g., from Anderson Level I to Anderson Level II) can be achieved by applying convolution filters, which incorporate information from multiple observations in space and/or time; and (2) basing a classification on annual metrics (e.g., mean NDVI) can improve discrimination between forests and fields but might necessitate lumping evergreen and deciduous forests into a more general “forest” class. Although particular to each dataset and study system, in general this balance explains the smaller landcover legends and lower accuracy standards employed in change studies compared to those of single-date classifications (Rogan et al. 2003).

The most effective solution to this tradeoff is to collect more data, augmenting the database by denser, more extensive, or altogether different measurements (e.g., lidar, radar, hyperspectral). However, this solution is only applicable for post-classification change detection, and therein only where and when the ancillary data are available. Although remotely sensed measurements are diversifying to meet an expanding range of needs, today's measurements are not available for the past. Extrapolating remotely sensed signatures over time requires limiting the database to measurements that are both available and maintain consistent signatures over time. Thus constrained, the analyst must rely on robust features of the spectral, spatial, and temporal signatures of the landcover classes.

In this study, confusion was great within three sets of NLCD 2001 landcover types. Confusion among the herbaceous field classes was due to two reasons: the variable practice of planting “cover” crops (i.e., alfalfa, clover, etc.) in harvested row-crop fields over the winter and to poor visual discrimination between mowed hay and harvested row-crop fields in the aerial photographs used as reference images. Confusion within the forest classes was due to the arbitrary partitioning of this continuous gradient into three classes—deciduous/evergreen/mixed forest. Confusion among high-, medium-, and low-intensity urban was also due to arbitrary partitioning, as well as to the inherently mixed nature of low-intensity urban class. Errors can thus arise from intra-annual mixing of classes (as in planting harvested crops in hay fields for a portion

of the year), imprecise identification of classes in reference images, and arbitrary binning of continuous gradients into discrete partitions, the last of which can be either between gradients of similar or very different types—e.g., deciduous and evergreen forest (i.e., mixed forest) *vs.* deciduous forests and impervious surfaces (i.e., low-intensity urban). To the degree possible, classes should be defined such that they are pure at the temporal and spatial grain of analysis, clearly identifiable in reference datasets, and based on discrete types in reality.

3.4.2. Causation

Environmental variables are those that have an ecological effect on landcover change, such as topographic gradients on vegetation change or infrastructural boundaries on landuse change. These “ancillary” variables have long been used to increase classification accuracy (Jensen 1996) and have also recently been used for similar purposes in change detection studies (Rogan et al 2003). Despite their benefits to accuracy, incorporating the variables that cause landcover changes into the classification logic invalidates their use in subsequent analyses.

This is especially true for variables measured at a single time but whose dynamics are poorly represented by a single date. Municipal boundaries for a single time might well improve classification accuracy, even in years far from the date of the boundary’s relevance. But how can a later analyst, studying the drivers of urbanization,

distinguish the real effect of the municipality from the artifact imposed by the classification algorithm? Classifications intended for subsequent analysis should be based on spectral reflectance alone to avoid circularity of logic; and when increases in accuracy are so great as to make their use worthwhile, their inclusion should be reported so that subsequent researchers can avoid spurious or circular interpretations.

3.4.3. Representation

Multispectral remote sensing frequently relies on indirect correlations between electromagnetic reflectance and the occurrence of landcover types—e.g., we do not remotely sense “urban” land *per se*, but rather that which is visibly bright, un-vegetated, possibly shadowy, and more or less constant in these characteristics throughout the year. What is required by the classification is not that modeled relationships between the classes and the measurements are true in an absolute sense, but that they remain robust at least over the domain of study. In this regard, extrapolation of models over time is no different than extrapolation in space – if the relationship remains true across the defined interval of space and time, the classification will be reliable within that domain.

In traditional (i.e., purely spatial) classification, it is common to cover as much of the study area as possible to represent the variation in spectral signatures of the classes. In spatio-temporal applications, it is likewise necessary to distribute the sample over widely over the spatio-temporal domain. The latter is potentially more difficult because

of the varying availability and quality of reference data over time, but the growing archive of high-resolution aerial imagery promises to ease this limitation.

High-resolution images are often collected at the county level, with individual county collections occurring in different years. It is therefore possible in many cases to collect several years' worth of training data within a study area simply by distributing the reference sample over several counties, making sure that the years of high-resolution images match those of satellite data. When too few years' high-resolution data are available to adequately capture the variability within the study domain, the relative stability of landcover over time can be used to augment the reference sample in time. Landcover identified in one year's high resolution image can be matched to satellite data from nearby years—possibly after removing pixels that have undergone landcover change by change-detection analysis, thus multiplying the size of otherwise small reference samples to adequate representation.

3.4.4. Uncertainty

Despite the best efforts of experienced analysts, errors are an inextricable component of variation in any landcover dataset. Through adoption of metrics such as percent correctly classified, user's and producer's accuracy, and the Kappa statistic, accuracy assessment has become common practice in the production of landcover maps (Stehman 1997, Foody 2002). Rigorous appraisal, reporting, and analysis of errors over

time have undoubtedly led the way to incremental improvements in all aspects of modeling.

Users of landcover maps require estimates of uncertainty that are relevant to their particular samples, but assessments based on validation samples are rarely pertinent to the various samples drawn by users (Appendix A). This is due in large part simply to distributional differences between the validation sample and the samples drawn for later analysis. However, these biases can be neutralized by incorporating informative class probability distributions into the familiar metrics of accuracy. Thus weighted, estimates of uncertainty can be extrapolated from validation samples to any other sample of the landscape, providing more relevant accuracy estimates for different study areas.

Landcover maps are analyzed as data—if not as truth itself—by a wide variety of sciences. Just as comparison of means requires variances, sound analysis by a growing user community requires knowledge of the uncertainty in the data they are using. Quantification and propagation of uncertainty have long been important foci of applied remote sensing and geographic information systems (Hunsaker et al. 2001), and multi-temporal landcover classifications add another dimension of complexity to this problem. Continued research is needed to develop methods for estimating uncertainty in time-serial landcover maps; in the meantime, earnest appraisal and reporting of uncertainty can minimize its downstream impact, and possibly its magnitude as well.

3.4.2. Future research directions

Although the problems are fundamentally equivalent, spatio-temporal landcover classification also has advantages and disadvantages that purely spatial classifications do not. With continued effort and attention to the four areas described above, further improvements to multi-temporal landcover classifications are likely. Based on this study, I suggest a few specific areas where gains may be close at hand.

Our approach modified the NCLD 2001 landcover classification scheme *post hoc*, after estimation of class spectral signatures, to increase robustness in the face of image differences over time. Classification precision could be improved by defining a set of classes whose spectral characteristics are robust despite phenological, atmospheric, and other differences between years. Similarly to standard exploratory analyses for purely spatial mapping efforts, cluster analysis—the common precursor to unsupervised classification—of multi-temporal image datasets could help to identify a robust set of classes *a priori*. As in this study, archival high-resolution images over time will be critical to identifying the landcover types of the spectral clusters. Given our difficulty in separating the NLCD urban herbaceous classes, special attention should be paid to these types. Growing interest in impervious surfaces (Schueler 1994) and the carbon sequestration potential of agricultural fields (Henebry 2009) make this an especially important priority.

Given the set of classes, our algorithm fits a single classification model to a multi-temporal reference dataset and applies that classifier across all extrapolation years, regardless of the similarity between the reference and extrapolation datasets. Accuracy in individual years might be improved by fitting several models—one for each reference year, and selecting the best model for each extrapolation year. As an added alternative to applying the single best model for each year, the predictions from several models could be combined by weighted averaging, with weights determined by spectral similarity between the extrapolation year and each model's reference year. Model selection could be based entirely upon the images themselves, by comparing distributional characteristics of the reference and extrapolation image datasets.

We extrapolated our classifier to all years from 1984 – 2007 based on the distribution of landcover classes from 2001. Improvements might be gained by recursively shifting this constraint from a single prior to varying posterior distributions over multiple model iterations, using the 2001 distribution to initialize the recursion and then iteratively adapting the priors toward greater influence from the years themselves. Excessive reliance on priors from individual years might destabilize results in years with poor fit of the spectral signature, possibly leading to runaway landcover distributions over the iteration. To solve this problem, various techniques for smoothing the prior distributions over temporal windows might be applied.

Post-classification filters—including “clump-and-eliminate” (Leica Geosystems, Inc. 2006) and “boundary-clean” (ESRI 2008) operations—have been applied in post-classification steps for many years. Relying on assumptions of spatial homogeneity, these operations compromise the spatial resolution of the map (i.e., the “minimum mapping unit”) in favor of increased overall accuracy. In comparison, the constancy of landcover over time is often far greater than its homogeneity in space, especially in highly fragmented landscapes. Landcover change is a relatively rare phenomenon, and even rarer are “flip-flop” changes back and forth between classes over very short times. By comparing each location’s landcover over small temporal windows, this stability of landcover over time can be used to filter out erroneous changes and increase the accuracy of each landcover map in a time-series. Although this approach has been known for three decades (e.g., the “cascade classifier, Swain 1978), it been applied only rarely. Our filter is just one of many possibilities.

Further, pre-classification temporal filters, including weekly, daily, and annual integrations, have been found here and elsewhere to minimize noise. The sample available within each window depends on the number of cloud-free observations for each location. Landsat-5 and -7, each of which orbit on a 16-day cycle, provide far fewer usable images than do daily-repeat sensors such as MODIS. Multi-day MODIS composites rely on cloud-removal algorithms (King et al. 1997), and so improvements in cloud-detection from the Landsat bands would also likely increase the utility of partially

cloudy Landsat images for multi-date composites and thereby increase the number of observations available for temporal composites.

3.5. Conclusions

Spatial extrapolation of landcover classification models is common practice, and the recent public release of Landsat images has removed a major data limitation on extrapolation in time as well. Using a 55-image series of Landsat images from a single path-row scene between 1984 and 2007, I successfully adapted traditional (i.e., purely spatial) landcover classification techniques for temporal extension. Modifications to the classification scheme, sampling design, variable selection, and accuracy assessment resulted in a time-series of 16 landcover maps over the period with a spatial extent of 1.7×10^6 ha and minimum mapping unit of 1 ha. Accuracy varied among classes, but was comparable overall to the NLCD Land Cover Layer for the region. The resulting landcover map series reveals spatial heterogeneity and temporal nonlinearity in the region's change, complexities which would not be observable either in sparser time series or by comparing maps from varying provenances. Increases in data availability will continue to facilitate such studies, and attention to semantics, causation, representation, and uncertainty in the classification process will maximize reliability of maps series for a receptive user community of scientists and managers.

3.6. Appendix: Distributional standardization of classification accuracy metrics.

Accuracy assessment for landcover classifications is based on the $i \times j$ confusion matrix, wherein the columns (j) tally reference, or “truth” class assignments and the rows (i) tally “model” assignments resulting from applying the classifier to a dataset of predictor variables (Congalton et al. 1983). The confusion matrix can be expressed either in terms of frequencies f_{ij} or in terms of proportions p_{ij} , by dividing each frequency by the total sample size, n :

$$p_{ij} = \frac{f_{ij}}{n} \quad (1)$$

Because the reference and model frequencies (and therefore proportions) are both based on the same sample, $n = n_i = n_j$.

The vector of proportional column sums:

$$p_j = \sum_{i=1 \dots k} p_{ij} = \frac{1}{n} \sum_{i=1 \dots k} f_{ij} \quad (2)$$

gives the reference probability distribution of the classes in the sample, and the vector of row sums:

$$p_i = \sum_{j=1..k} p_{ij} = \frac{1}{n} \sum_{i=1..k} f_{ij} \quad (3)$$

gives the model probability distribution of classes in the sample.

As an estimate the overall accuracy of the model, the Percent Correctly Classified metric, *PCC* is calculated as the sum of the diagonal ($i=j$) elements of the proportional confusion matrix, rescaled to percentages:

$$PCC = 100 \times \sum_{i=j} p_{ij} = \frac{100}{n} \sum_{i=j} f_{ij} . \quad (4)$$

Conversely estimating the accuracy of the model specifically with respect to the reference class assignments, “producer’s” accuracy for each class j is calculated as the proportion of reference assignments to class j also assigned to j by the model:

$$PA_j = \frac{p_{i=j}}{\sum_i p_{ij}} . \quad (5)$$

This metric is so named because it informs the map producer of the accuracy of his or her classifier based on information to which a user of the map does not have access—i.e., the reference, or truth” data. The inverse of producer’s accuracy ($1-PA$) is interpreted as

errors of *omission*. Conversely, “user’s” accuracy quantifies model accuracy relative to the predictions themselves, and is calculated as the proportion of model assignments to class i in agreement with the reference assignments:

$$UA_i = \frac{P_{i=j}}{\sum_j P_{ij}}. \quad (6)$$

The inverse of user’s accuracy ($1-UA$) is interpreted as errors of *commission*.

The Kappa coefficient (Cohen 1968) is the proportion of agreement between a reference and model classification, adjusted for chance:

$$\kappa = \frac{p_o - p_c}{1 - p_c}, \quad (7)$$

where p_o is the observed proportion of agreement between model and reference and p_c is the proportion of agreement expected from a random assignment of classes within the sample. Originally developed to measure agreement between observers in psychometric studies (Cohen 1968), Kappa has been widely adopted by the remote sensing community for accuracy assessment in landcover classifications (Foody 2002, Liu et al 2007).

The observed proportion of agreement p_o is equivalent to PCC on the unit scale:

$$p_o = \sum_{i=j} p_{ij}, \quad (8)$$

and the chance proportion of agreement p_c is calculated as the sum of the row \times column (i.e., marginal) products of the confusion matrix:

$$p_c = \sum_{i=j} p_i \times p_j \quad (9)$$

The Kappa coefficient is therefore calculated as:

$$\kappa = \frac{\sum_{i=j} p_{ij} - \sum_{i=j} p_i p_j}{1 - \sum_{i=j} p_i p_j} = \frac{n \sum_{i=j} f_{ij} - \sum_{i=j} f_i f_j}{n^2 - \sum_{i=j} f_i f_j}. \quad (10)$$

Assuming that changes to the sample distribution affect model errors proportionally (i.e., that observations added or removed bear the same covariance structure as the original sample, or that changes to the sample impart no *relational* bias), PA and UA are unaffected by the class frequency distribution of the sample. But because PCC and Kappa summarize accuracy across the distribution of classes, these metrics are affected by the relative class frequencies within the sample. This is known as the

prevalence problem (Byrt et al. 1993)—leading to different accuracy estimates from various samplings of the population, prevalence bias greatly diminishes relevance of accuracy estimates based to populations of which the sample is not representative.

Several modifications have been proposed to adjust Kappa (Cohen 1968, Foody 1992, Ma and Redmond 1995). Specifically, weighted Kappa κ_w (Cohen 1968) can be defined to remove the effect of sampling bias by standardizing the metric to a reference distribution. This is accomplished by defining the weights used to compute κ_w as vector w of class weights, with each element w_j equal to the ratio of the population proportion to the sample (reference) proportion for each class:

$$w_j = \frac{N_j}{N} \div \frac{n_j}{n} = \frac{n \times N_j}{N \times n_j}, \quad (12)$$

i.e.,

$$w_j = \frac{P_j}{p_j}. \quad (13)$$

The standardized Kappa coefficient is thus calculated as:

$$\kappa_s = \frac{\sum_{i=j} w_j p_{ij} - \sum_i w_j p_i p_j}{1 - \sum_i w_j p_i p_j} = \frac{\sum_{i=j} w_j f_{ij} - \sum_i w_j f_i f_j}{n^2 - \sum_i w_j f_i f_j} \quad (14)$$

Similarly, the vector w can be applied to the diagonal of the confusion matrix to remove sampling bias from PCC :

$$PCC_s = \sum_{i=j} w_i p_{ij} = \frac{1}{n} \sum_{i=j} w_i f_{ij} \cdot \quad (15)$$

This distribution-based weighting scheme has the effect of standardizing PCC and the Kappa coefficient computed from any sample to some *a priori* distribution of classes, thus allowing consistent comparisons regardless of sampling idiosyncrasies. Further, any distribution may be chosen as the standard, including uniform (i.e., $n_i = n_j$ for all i, j) or some accepted empirical distribution such as the U.S. National Land Cover Database 2001 (Homer et al. 2004).

UA and PA can also be weighted based on a class frequency distribution to estimate errors for a class over a predicted landcover map:

$$PA_{map}(i) = p_i \times PA \quad (16)$$

$$UA_{map}(i) = p_i \times UA, \quad (17)$$

where p_i is the probability or proportion of class i over the area of interest. However, the two metrics must use different distributions as their standards. As a metric of the errors of *omission*, $1-PA_{map}$ must be based on some reference distribution that is accepted as truth—identical to the adjustment of Kappa and PCC above. But as a metric of the map's errors of *commission*, $1-UA_{map}$ is based on the distribution of the predictions themselves.

Based on these two metrics, the range of frequencies (or proportions) can be predicted for a landcover map. The expected upper bound for the frequency of class i on the map, f_i^+ , is:

$$f_i^+ = f_i + (p_i^* \times (1 - PA_i)), \quad (18)$$

where p_i^* is the proportion of i in the reference distribution. The expected lower bound for the frequency of i on the map is:

$$f_i^- = f_i + (\hat{p}_i \times (1 - UA_i)). \quad (19)$$

where \hat{p}_i is the posterior probability of i —i.e., its proportion on the predicted map.

4. Concluding remarks

The overarching goal of this research has been to quantify landscape changes as affected by natural and anthropogenic forces. The effort has been predominantly methodological, developing the techniques necessary to extract ecologically pertinent information from remotely sensed datasets. As the techniques are refined, the data's transparency will continue to increase, focusing insights further and deeper into the spatio-temporal processes governing coupled human-natural systems. The data's richness also necessitates development of new statistical methods to analyze ecosystem dynamics in space and time. Meanwhile, the first version of the database is providing an unprecedented view of changes in a dynamic landscape over two decades.

4.1. Synopsis: Landcover change of the North Carolina Piedmont (1985 – 2006)

The landcover map series shows a changing pattern of urban, forest, and herbaceous cover on the North Carolina Piedmont from 1985 to 2006. These changes were complex in time and space. Temporal trajectories revealed transient nonlinearities superimposed on long-term trends, and spatial patterns showed centers of rapid local development in a context of broader regional stability.

Region-wide, net changes were slight relative to the overall coverage of the classes. The greatest proportional change occurred in urban area, which—although rare overall—nearly doubled over the period. Fields and forests—the region's two dominant

types—exchanged area diffusely over the entire study area, with forest clearing in the latter half of the period nearly compensating earlier agricultural reforestation.

Change was unevenly distributed across the region. As suburban growth agglomerated neighboring cities, the zones around and between existing urban centers developed more rapidly than outlying rural areas. Although the North Carolina Piedmont has no natural lakes, its coverage of water increased due to the filling of large reservoirs around growing metropolitan centers.

4.2. Prospectus for future research

The observed landcover changes reveal spatial and temporal complexity sufficient to inspire extensive further analysis. However, full use of spatio-temporal datasets such as this one will require methodological adaptations, both in their production and analysis. The disciplines of remote sensing and landscape ecology will likely continue to co-evolve in this relationship.

4.2.1. Methodological requirements

Analysis of uncertainty will be a critical area of future research, especially given the increases in error incurred with model extrapolation over time. Hall et al. (1991) and Pontius and Li (2008) provide methods for estimating true change rates from measured transition and confusion matrices. These methods can also be simplified to estimate true

landcover proportions at any single time, and the combined toolbox of techniques will greatly increase robustness in future analyses of landcover dynamics.

Currently, landscape ecology relies heavily on sparse samplings of landcover over time, with the preponderance of analyses based on data from a single year. Despite wide disciplinary recognition of the importance of temporal processes, this reliance is an accommodation forced by the rarity of true spatio-temporal datasets. Consistent spatio-temporal databases loosen this constraint, releasing landscape ecology from its reliance on selective case studies and tenuous assumptions of equivalence between space and time and facilitating wider adoption of data-driven, statistical regression-based approaches.

4.2.2. Ecological opportunities

4.2.2.1. Landcover/landuse change

The primary use of landcover data will continue to be for landcover/landuse change (LCLUC) analysis, and denser temporal sampling of changing landscapes will increase our ability to observe the complexities of coupled human-natural systems in finer detail. Dense time-series of landcover will increase the sophistication of LCLUC models, allowing incorporation of acceleration and higher-order nonlinearities.

Interleaving landcover with other spatio-temporal measurements such as transportation

networks, conservation reserves, and various socio-economic indices will promote analyses of the feedbacks operating in LCLUC systems over time.

4.2.2.2. Forest growth

Remote sensing of vertical forest structure is still in its infancy, and data collection parameters have yet to be sufficiently standardized for robust comparisons over time. However, single-time measurements of structure overlaid on maps of forest establishment can provide chronosequences of forest growth. Overlaid on hydrological, edaphic, and solar-radiative datasets, these chronosequences will be valuable for modeling forest growth over environmental gradients. As a geospatial analog of traditional site-index curves, the resulting models of early height growth can be extrapolated based on extensive geospatial datasets to predict the potential growth and yield of loblolly pine forests over large areas.

4.2.2.3. Landcover phenology and forest succession

Just as single-date landcover maps are used to stratify and extract other spatial information, dynamic landcover maps can be used as multi-temporal filters for spatio-temporally coincident datasets. Various indices of time-varying ecosystem properties (e.g., NDVI, NDWI, surface temperature) can be extracted to document development at intra- and inter-annual time scales to analyze ecosystem changes at seasonal and decadal

scales. The North Carolina Piedmont specifically has experienced punctuated disturbances due to drought (2002, 2007), a hurricane (1996), and an ice storm (2002), as well as potentially other, more chronic influences from climate change, forest pests, landuse legacies, urban heat islands, and forest edge effects. Stratified by landcover type, the record of remotely sensed measurements provides a view of phenological and successional responses to static environmental gradients and variable disturbances.

4.3. Conclusion

Spatio-temporal datasets provide a unique and irreplaceable view of landscape dynamics. In addition to increasing temporal resolution of a complex phenomenon, multi-temporal landcover maps can also be used for a much wider array of ecological analyses beyond those of uni- or bi-temporal maps. Overlaying maps identifying dates of forest establishment with remotely sensed measurements of forest structure will provide chronosequences of forest growth. Landcover maps also serve as filters, providing criteria through which metrics of ecosystem properties can be extracted to analyze ecosystem dynamics at various temporal scales. At 20-year and million-hectare scales, it is impossible to study ecological phenomena without explicitly addressing human systems. With sustained progress in this direction, remote sensing and landscape ecology will continue to explore the dynamics of increasingly coupled human-natural systems.

References

- Aber, J. D. 1979. Foliage-height profiles and succession in northern hardwood forests. *Ecology* 60:18-23.
- Andersen, H. E., R. E. Reutebuch and R. J. McGaughey. 2006. A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. *Canadian Journal of Remote Sensing* 32:355-366.
- Anderson, J.R., E.E> Hardy, J.T. Roach, and R.E. Witmer. 1976. A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964, 28 pp.
- Andren, H. and P. Angelstam. Elevated predation rates as an edge effect in habitat islands: experimental evidence. *Ecology* 69:544-547.
- Bailey, R. G. Descriptions of the Ecoregions of the United States (2nd ed.). Miscellaneous Publication No. 1391, Map scale 1:7,500,000. U.S. Dept. of Agriculture, Forest Service. 108 pp.
- Balzter, H., C. S. Rowland and P. Saich. 2007. Forest canopy height and carbon estimation at Monks Wood National Nature Reserve, UK, using dual-wavelength SAR interferometry. *Remote Sensing of Environment* 108:224-239.
- Bartel, R.A and J.O. Sexton. in press. Monitoring habitat dynamics for rare and endangered species using satellite images and niche-based models. *Ecography*.
- Bhang, K. J., F. W. Schwartz, A. Braun. 2007. Verification of the vertical error in C-band SRTM DEM using ICESat and Landsat-7, Otter Tail County, MN. *IEEE Transactions on Geoscience and Remote Sensing* 45(1): 36-44.
- Bourgine B., and Baghdadi N., 2005. Assessment of C-band SRTM DEM in a dense equatorial forest zone. *Comptes Rendus Geoscience* 337: 1225-1234.
- Brandtberg, T., T. A. Warner, R. E. Landenberger, J. B. McGraw. 2003. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar from the eastern deciduous forest in North America. *Remote Sensing of Environment* 85:290-303.
- Brown, G. C., K. Sarabandi. 2003. Estimation of Red Pine tree height using Shuttle Radar Topography Mission and Ancillary Data. *Geoscience and Remote Sensing Symposium* 2850-2852.

- Byrt, T., J. Bishop, J.B. Carlin. 1993. Bias, prevalence, and Kappa. *Journal of Clinical Epidemiology* 46(5): 423-429.
- Cadee, G. C., S. E. Walker, K. W. Flessa, E. Naasset. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment* 61:246-253.
- Carter, W.E., R.L. Shrestha, K.C. Slatton. 2007. Geodetic laser scanning. *Physics Today*. December: 41-46.
- The Central Intelligence Agency (CIA) of the United States of America. 2008. The World Factbook. <https://www.cia.gov/library/publications/the-world-factbook/print/xx.html> [accessed 2008.05.07]
- Chander, G., B. Markham, and J.A. Barsi. 2007. Revised Landsat-5 Thematic Mapper radiometric calibration. *IEEE Geoscience and Remote Sensing Letters* 4(3): 490 – 494.
- Cloude, S. and K.P. Papathanassiou. 1998. Polarimetric SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing* 36:1551-1565.
- Comber, A.P. Fisher, R. Wadsworth. 2005. What is landcover? *Environment and Planning B* 32: 199-209.
- Crist, E. C. 1985. A TM Tasseled Cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment* 17: 301-306.
- Crist, E.C. and R.C. Cicone. 1984. Application of the Tasseled Cap concept to simulated Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing* 50: 343 – 352.
- Crutzen, P.J. 2002. The Anthropocene. *Journal de Physique IV* 10:1-5. DOI 10.1051/jp4:20020447
- Daily, G. and P. Matson. 2008. Ecosystem services: from theory to implementation. *Proceedings of the National Academy of Sciences* 105: 9455-9456.
- Delgado, C., D. Amor, J.O. Sexton, A. Pfaff. 2008a. Do protected areas constrain road deforestation? *Proceedings of the Amazon in Perspective International Scientific Conference, Integrated Science for a Sustainable Future/ LBA (Large Scale Biosphere-Atmosphere Project in the Amazon) / GEOMA (Amazonian Environmental Modeling Network) / PPBio (Biodiversity Research Program)*. November 17 - 20, 2008 Manaus – Brazil.

- Delgado, C., D. Amor, J.O. Sexton, F. Colchero., A. Pfaff. 2008b. Short- and long-term road impacts on frontier deforestation. Proceedings of the Amazon in Perspective International Scientific Conference, Integrated Science for a Sustainable Future/ LBA (Large Scale Biosphere-Atmosphere Project in the Amazon) / GEOMA (Amazonian Environmental Modeling Network) / PPBio (Biodiversity Research Program). November 17 - 20, 2008 Manaus – Brazil.
- Didham, R. K. and J. H. Lawton. 1999. Edge structure determines the magnitude of changes in microclimate and vegetation structure in tropical forest fragments. *Biotropica* 31:17-30.
- Doolittle, W. T. 1979. Site index comparisons for several forest species in the Southern Appalachians. *Soil Science Society of America Journal* 22:455-458.
- Drake, J. B., R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton et al. 2002. Estimation of tropical forest structural characteristics using large footprint lidar. *Remote Sensing of Environment* 79: 305-319.
- Dubayah, R. O. and J. B. Drake 2000. Lidar remote sensing for forestry. *Journal of Forestry* 98:44-46.
- Federal Emergency Management Agency. 2003. Guidelines and Specifications for Flood Hazard Mapping Partners. Appendix 4b: Airborne Light Detection And Ranging Systems. Accessed Oct 2008: <http://www.fema.gov/library/viewRecord.do?id=2206>
- Foody, G.M. 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80: 185-201.
- ESRI, Inc. 2008. ArcGIS Desktop 9.3, Service Pack 1.
- Hall, F.G., D.B. Botkin, D.E. Strelbel, K.D. Woods, and S.J. Goetz. 1991. Large-scale patterns of forest succession as determined by remote sensing. *Ecology* 72: 628-640.
- Hensley, S., P., Rosen and E. Gurrola. 2000. Topographic map generation for the Shuttle Radar Topography Mission C-band SCANSAR interferometry. *Proceedings of SPIE* 4125:179-189.
- Henebry, G.M. 2009. Carbon in idle croplands. *Nature* 457: 1089-1090.
- Herlihy, A.T., J.L. Stoddard, C.B. Johnson. 1998. The relationship between stream chemistry and watershed land cover data in the Mid-Atlantic region, U.S. *Water, Air, and Soil Pollution* 105: 377-386.

- Homer, C., C. Huang, L. Yang, B. Wylie, and M Coan. 2004. Development of a 2001 National Land-cover database for the United States. *Photogrammetric Engineering and Remote Sensing* 70(7): 829-840.
- Hunsaker, C.T., Goodchild, M.F., Friedl, M.A, and Case, T.J. (Eds.). 2001. *Spatial uncertainty in ecology: implications for remote sensing and GIS applications*. Springer, New York.
- Hurttt, G. C., R. Dubayah, J. Drake, P. R. Moorcroft, S. W. Pacala and J. B. Blair, et al. 2004. Beyond potential vegetation: Combining lidar data and a height-structured model for carbon studies. *Ecological Applications* 14:873-883.
- Hyde, P., R. Nelson, D. Kimes and E. Levine. 2007. Exploring LiDAR-RaDAR synergy-predicting aboveground biomass in a southwestern ponderosa pine forest using LiDAR, SAR and InSAR. *Remote Sensing of Environment* 106:28-38.
- Intergovernmental Panel on Climate Change (IPCC). 2007. *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II, and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland. 104 pp.
- The International Monetary Fund (IMF). 2008. *Climate Change and the Global Economy. Chapter 4 in The World Economic Outlook, April 2008: Housing and the Business Cycle*. International Monetary Fund, Washington, DC. 303 pp.
- James, F. C. 1971. Ordinations of habitat relationships among breeding birds. *The Wilson Bulletin* 83:215-236.
- Jensen, J.R. 1983. Biophysical remote sensing. *Annals of the Association of American Geographers*, 73(1), 111-132.
- Jensen, J.R. 1996. *Introductory digital image processing: a remote sensing perspective*. Prentice Hall, Inc. New Jersey, USA.
- Jensen, J. R. 2000. *Remote Sensing of the Environment: an Earth Resource Perspective*. Prentice Hall, New Jersey. 544 pp.
- Kauth, R.J. and G.S. Thomas. 1976. The Tasseled Cap—A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*. West Lafayette, IN: Laboratory for Applications of Remote Sensing, pp. 41 – 51.

- Kellndorfer, J., W. Walker, L. Pierce, C. Dobson, J. A. Fites, C. Hunsaker, J. Vona and M. Cutter. 2004. Vegetation height estimation from Shuttle Radar Topography Mission and National Elevation Datasets. *Remote Sensing of Environment* 93:339-358.
- Kellndorfer, J., W. Walker, E. LaPoint, M. Hoppus and J. Westfall. 2006. Modeling Height, Biomass, and Carbon in U.S. Forest from FIA, SRTM, and Ancillary National Scale Data Sets. *IEEE International Conference on Geoscience and Remote Sensing Symposium* 3591-3594.
- Kenyi, L., R. Dubayah, M. Hofton, J.B. Blair and M. Schardt. 2007. Comparison of SRTM-NED Data to LIDAR Derived Canopy Metrics. *Geoscience and Remote Sensing Symposium* 2825:2829.
- M. D. King, S.-C. Tsay, S. E. Platnick, M. Wang, and K. N. Liou, "Cloud retrieval algorithms for MODIS: Optical thickness, effective particle radius and thermodynamic phase," *MODIS Algorithm Theoretical Basis Document: NASA*, 1997.
- Kruger., L. M., J. J. Midgely, and R. M. Cowling. 1997. Resprouters vs reseeders in South African forest trees; a model based on forest canopy height. *Functional Ecology* 11:101-105.
- Lefsky, M., W. B. Cohen, G. G. Parker and D. J. Harding. 2002a. Lidar Remote Sensing for Ecosystem Studies. *BioScience* 52:19-30.
- Lefsky, M., W. B. Cohen, D. J. Harding, G. G. Parker, S. A. Acker and S. T. Gower. 2002b. Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology & Biogeography* 11:393-399.
- Leica Geosystems Geospatial Imaging, L.L.C. 2006. ERDAS IMAGINE version 9.1.
- Lillesand, T. M., R. W. Kiefer, and J. W. Chipman. 2008. *Remote Sensing and Image Interpretation*. John Wiley & Sons, Inc. U.S.A. 756 pp.
- Lim, K., P. Treitz, K. Baldwin, I. Morrison and J. Green. 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forest. *Canadian Journal of Remote Sensing* 29:658-678.
- Maclean GA, Krabill WB. 1986. Gross-merchantable timber volume estimation using an airborne LIDAR system. *Canadian Journal of Remote Sensing* 12:7-18.

- Mansfield, C., S.K. Pattanayak, W. McDow, R.I. McDonald, and P.N. Halpin. 2005. Shades of green: measuring the value of urban forests in the housing market. *Journal of Forest Economics* 11(3):177-199.
- Matlack, G. R. 1994. Vegetation dynamics of the forest edge—trends in space and successional time. *The Journal of Ecology* 82:113-123.
- McDonald, R.I., and D.L. Urban. 2006. Spatially varying rules of landscape change: lessons from a case study. *Landscape and Urban Planning B* 74:7-20.
- Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning LiDAR. *Photogrammetric Engineering & Remote Sensing* 66:1367–1371.
- Minter, T.C. 1978. Methods of extending crop signatures from one area to another, *Proceedings, the LACIE symposium, a technical description of the large area crop inventory experiment (LACIE), October 23–26, 1978, Houston, TX.*
- Monserud, R. A. Height growth and site index curves for inland Douglas-fir based on stem analysis data and forest habitat type. *Forest Science* 30:943-965.
- Muller, J.-P.A.L. 1988. Key issues in image understanding in remote sensing. *Philosophical Transactions of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 324(1579), 381-395.
- Naesset, E. 1996. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry & Remote Sensing* 52:49-56.
- National Academy of Science. 2005. *Earth Science Applications from Space: Urgent needs and opportunities to serve the nation. Committee on Earth Science and Applications from Space: A Community Assessment and Strategy for the Future.* Space Studies Board, Division on Engineering and Physical Sciences, National Research Council of the National Academies. The National Academies Press, Washington, D.C. 48 pp.
- Patenaude, G., R. Milne and T.P. Dawson. 2005. Synthesis of remote sensing approaches for forest carbon estimation: reporting to the Kyoto Protocol. *Environmental Science & Policy* 8:161-178.
- Pimm, S.L., G.J. Russell, J.L. Gittleman, T.M. Brooks. 1995. The future of biodiversity. *Science* 21(5222): 347-350.

- Pontius, R.G. and XiaoXiao Li. 2008. Estimating the land transition matrix from erroneous maps. *Proceedings of Studying, Modeling, and Sense Making of Planet Earth*. Mytline, Greece. 8 p.
- Python Software Foundation. 2001-2008. Python version 2.7.
- R Development Core Team 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>
- Rabus, B., M. Eineder, A. Roth, and R. Bamler. 2003. The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing* 57:241-262.
- Raupach, M. R. 2004. Simplified expressions for vegetation roughness length and zero-plane displacement as functions of canopy height and area. *Boundary-Layer Meteorology* 71:211-216.
- Roberts, S.D., Dean, T.J., Evans, D.L., McCombs, J.W., Harrington, R.L., Glass, P.A. 2005. Estimating individual tree leaf area in loblolly pine plantations using LiDAR-derived measurements of height and crown dimensions. *Forest Ecology and Management* 213:54-70.
- Rosen, P.A., S. Hensley, I.R. Joughin, F.K. Li, S.N. Madsen, E. Rodriguez and R. Goldstein. 2002. Synthetic Aperture Radar Interferometry. *Proceedings of the IEEE* 88:333-382.
- Rouse, J.W., Jr., R.H. Haas, J.A. Schell, and D.W. Deering. 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. *Prog. Rep. RSC 1978-1*, Remote Sensing Center, Texas A&M Univ., College Station. 93 pp. (NTIS No. E73-106393).
- Schimel, D. S., J. I. House, K. A. Hibbard, P. Bousquet, P. Ciais, P. Peylin et al. 2001. Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature* 414: 169-172.
- Schlesinger, W.H. 1997. *Biogeochemistry: An Analysis of Global Change*. 2nd ed. Academic Press, San Diego. 558 p.
- Schueler, T.R. 1994. The importance of imperviousness. *Watershed Protection Techniques* 1: 100-111.

- Seto, K.C., Woodcock, C.E., Song, C., Huang, X., Lu, J., Kaufmann, R.K. 2002. Monitoring land-use change in the Pearl River Delta using Landsat TM. *International Journal of Remote Sensing*, 23(10): 1985-2004.
- Simard, M., K. Zhang, V.H. Rivera-Mnroy, M. Ross., P. Ruiz, E. Castañeda-Moya, E. Twilley, E. Rodriguez. 2006. Mapping height and biomass of mangrove forests in the Everglades National Park with SRTM elevation data. *Photogrammetric Engineering and Remote Sensing*. 72:299-311.
- Song, C. and Woodcock, C.E. 2003. [Monitoring forest succession with multitemporal Landsat images: factors of uncertainty](#). *IEEE Transactions on Geoscience and Remote Sensing*, 41(11):2557-2567.
- Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., & Macomber, S.A. 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment*, 75, 230-244.
- Stehman, S.V. 1997. Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment* 62: 77-89.
- Strahler, A.H. 1980. The use of prior probabilities in Maximum Likelihood Classification of remotely sensed data. *Remote Sensing of Environment* 10: 135-163.
- Therneau, T.M. and B. Atkinson. 2007. rpart: Recursive Partitioning. R package version 3.1-3.8. S-PLUS 6.x original at <http://mayoresearch.mayo.edu/mayo/research/biostat/splusfunctions.cfm>
- Townsend, P. A. 2002. Estimating forest structure in wetlands using multitemporal SAR. *Remote Sensing of Environment* 79:288-304.
- Treuhaft, R.N., S.N. Madsen, M. Moghaddam, and J.vanZyl. 1996. Vegetation characteristics and underlying topography from interferometric radar. *Radio Science* 31:1449-1485.
- Treuhaft, R.N. and P.R. Siqueira. 2000. The vertical structure of vegetated land surfaces from interferometric and polarimetric radar. *Radio Science* 25:141-177.
- Turner, B.L. II, W.C. Clark, R.W. Kates, J.F. Richards, J.T. Mathews, W.B. Meyer (eds.). 1990. *The Earth as Transformed by Human Action: Global and Regional Changes in the Biosphere Over the Past 300 Years*. Cambridge Univ. Press, Cambridge.

- The United Nations (UN).1999. The world at six billion.
<http://www.un.org/esa/population/publications/sixbillion/sixbillion.htm> [accessed 2008.05.07]
- Urban, D.L., R.V. O'Neill, H.H. Shugart, Jr. 1987. Landscape ecology. *BioScience* 37(2): 119-127.
- Walker, W., J.M., Kellndorfer and L. Pierce. 2007. Quality assessment of SRTM C- and X-band interferometric data: Implications for retrieval of vegetation canopy height. *Remote Sensing of Environment* 106:428-448.
- Waring, R. H., J. Way, E. R. Hunt, Jr., L. Morrissey, K. J. Ranson, J. F. Weishampel, R. Oren, S. E. Franklin. 1995. Imagine radar for ecosystem studies. *BioScience* 45:715-723.
- Welden, C. W., S. W. Hewett, S. P. Hubbell, and R. B. Foster. 1991. Sapling survival, growth, and recruitment: relationship to canopy height in a neotropical forest. *Ecology* 72:35-50.
- Wheeler, K. and S. Hensley. 2000. The GeoSAR Airborne mapping system. *IEEE International Radar Conference Proceedings* 831-835.
- Wilkinson, B.H. 2005. Humans as geologic agents: a deep-time perspective. *Geology* 33: 1361-164.
- Westervelt, J., T. BenDor, J.O. Sexton. submitted manuscript. A technique for rapidly forecasting regional urban growth. *Environment and Planning B*.
- Whitehurst, A.W., J.O. Sexton, L. Dollar. 2009. Landcover change in western Madagascar's dry deciduous forests: a comparison of changes in and around Kirindy Mite National Park. *Oryx* 43: 275-283.
- Willmott, C. J. 1982. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society* 63:1309-1313.
- The World Bank. 2008. Global Monitoring Report 2008. MDGs and the Environment. Agenda for Inclusive and Sustainable Development. Washington, DC. 290 pp.
- Yuan, F., K.E. Sawaya. B.C. Loeffelholz, M.E. Bauer. 2005. Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal remote sensing. *Remote Sensing of Environment* 98(2-3): 317-328.

Biography

I was born at 8:16 PM on August 12, 1976. I was raised in the neighborhood of Pleasant Ridge in Cincinnati, Ohio. During my morning walks to Nativity elementary school, I occasionally had to cover my nose and mouth with the sleeve of my sweater, taking short, shallow breaths because the acrid air burned my lungs. It was the Reagan Era. Nearly every weekend until high school and using trails only when the blackberry bushes were impenetrable, my dad, my siblings, and I explored the wildernesses of Winton Woods or French Park.

Not much happened in high school.

During my Bachelor's degree in the Department of Wildlife Ecology and Conservation at the University of Florida, I was taught natural resource management in the classroom and studied animal effects on ecosystem processes as a work-study technician for doctors Lyn Branch and Ken Clark over the nights and weekends. Motivating a group of some twenty or so project volunteers with food, beer, and access to a locked preserve, I camped under every full moon of 1997. Once I described my research to my grandparents; when I had finished explaining how pocket gophers affect the decomposition of oak leaves and pine needles, my grandmother paused... and asked emphatically, "...who...CARES?!?". That stuck with me. My first GIS project was the fire management database for the Ordway Preserve, which I handed to the departmental

chair at my graduation barbecue. I was unaware that there was an undergraduate thesis program.

I took four years to earn my Master's degree from the College of Natural Resources at Utah State University. It's not easy to go from the flatwoods and sandhills of north-central Florida to Utah's aspen-dotted Tavaputs Plateau and come home for Christmas making any sense to grandma. I scrambled to find a theory that fit all the cases, and thanks to discussions with people like Paul Box, Dave Roberts, and Dave Stoner, I wrote my own.

In my fourth year in Utah, I applied for a postdoctoral position in Dean Urban's lab. It was mostly to get a chance to talk shop with *the* Dean Urban, but I was a cocky sonofabitch as well—if I didn't get the job, I wanted to know who beat me so I could win next time. Dean and I hit it off, the first candidate for the job went elsewhere, and I was next in line. Colleagues and friends at USU told me this was my shot at the big leagues. With only an M.S. to my name, Dean offered to turn the postdoc job into a doctoral assistantship. That's how I got into Duke.

While here, I have struggled to find my place among geniuses of all kinds—ecological and many others as well. In the past six years, I've visited Nepal, bought a house, gotten married, gotten a NASA fellowship, gotten divorced, sold the house, celebrated my 30th birthday in Machu Picchu, established myself as a landscape ecologist/remote sener/GIS expert, learned to dance salsa, been given an award for

excellence in mentoring, and collaborated on research projects in Madagascar, North Carolina, and South America. This dissertation is the product of these times and efforts. The work is not nearly done, but it's my best shot at something useful, innovative, and beautiful. I hope it gives back as much as I put in—and got out.