ABSTRACT

ANDERSON, ERIC SCOTT. Spatial Prediction of Forest Soil Carbon: Spatial Modeling and Geostatistical Approaches (under the direction of James. A. Thompson)

Understanding the carbon cycle is one of the most difficult challenges facing scientists studying the global environment. Efforts to balance the global C budget have focused attention on terrestrial carbon storage in temperate ecosystems. Historically, most estimates of soil organic C (SOC) are based on means extrapolated from broad categories of soils and vegetation on a regional scale. Forest ecosystems of North America are of particular interest because of their ability to provide long-term C storage in both the forest vegetation and soils. Understanding spatial patterns in forest SOC may result in future development of techniques for conserving and enhancing terrestrial C pools. A series of studies were undertaken to explore a number of current issues that contribute to our inability to model SOC on a regional or landscape scale. Investigation into the spatial distribution of SOC occurred on a 32,500 ha forest ecosystem located entirely within the bounds of Hofmann Forest. Hofmann Forest is located in Jones and Onslow Counties of eastern North Carolina, USA. The objectives of the research were to compile and compare a remotely-sensed high-resolution digital elevation model (DEM) to other commonly available DEM sources, (ii) to utilize landscape attributes and selected soil properties to develop and validate an explicit, quantitative, and spatially realistic model of SOC for a 32,500 ha forest ecosystem; (iii) to determine if the spatial scale of environmental variables

affects model predictions; and (iv) to quantify SOC on an areal basis using the newly parameterized spatial models. The first issue at hand was to derive a highly precise, highly accurate DEM. A newly emerging technology, light detecting and ranging (LiDAR), was selected as a source of highly precise and accurate source of elevation data. Attempts to produce landscape scaled DEM lead to a series of issues that inhibited their production. To resolve these issues a series of geostatistical approaches were developed to reduce the LiDAR data sets while maintaining their precision and accuracy. A study was conducted to evaluate the effects of inverse distance weighted (IDW) and ordinary kriging (OK) linear interpolators on datasets of various levels of data reduction. A series of 10 forested 1000-ha LIDAR tiles on the Lower Coastal Plain of eastern North Carolina was used. Results indicated, for LiDAR interpolation on low relief landscapes, that IDW was favored over OK methods. Further testing indicated that a substantial data set reduction was possible while still maintaining a high level of accuracy. An additional study was implemented to evaluate the effects of data density on production of DEM of various resolutions using a series of 61 LIDAR tiles (1000 ha) covering the spatial extent of the Hofmann Forest in the Lower Coastal Plain of eastern North Carolina. The study area was of relatively little relief where we anticipated limited impacts from data reduction. The LIDAR data set was reduced to 50%, 25%, 10%, 5%, and 1% of the original density. We created 5-m, 10-m, and 30-m DEM with 0.1 m vertical precision for each density level and used paired t-tests to determine if the true mean of their

differences were equal to zero. Differences indicated that for the 30-m DEM, LIDAR data sets could be reduced to 10% of their original data density without statistically altering the resultant DEM. The LIDAR data could only be reduced to 25% of the original data set before statistically altering the 10-m DEM only to 50% for the 5-m DEM.

To address the issue of spatial stores of SOC a study was conducted to estimate SOC with in the Hofmann Forest. This study attempted to develop explicit spatial models to predict landscape-scale SOC storage. Soil-landscape modeling is an approach to analyzing soil variability in response to exogenous environmental variables related to topographic and hydrologic parameters. A series of spatial models were developed to predict SOC storage. The models included a geostatistical approach, OK, and two landscape models including a model that used land use and landscape attributes (LandTopo), and a second model that used landscape attributes only (Topo). The results of this study found that LandTopo was able to explain 18% of the variability in SOC while Topo was only able to explain 10% of the variability. Both models used landscape attributes of different scales, yet were relatively unsuccessful in their ability to predict SOC. Total SOC stores for the Hofmann Forest were calculated, with all three models predicted ~40 Gt C stored in the top 1m of Hofmann Forest.

SPATIAL PREDICTION OF FOREST SOIL CARBON: SPATIAL MODELING AND GEOSTATISTICAL APPROACHES

by

Eric S. Anderson

**A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy**

Soil Science

July 2004

_____   _____

Co-Chair of Advisory Committee   Co-Chair of Advisory Committee

_____   _____

_____

BIOGRAPHY

Eric Anderson was born and raised in the small East Texas community of Hallsville. He is the son of Paul and Linda. He has a brother, Paul, and a sister, Amanda. He completed a Bachelor of Science in Environmental Science (1998) and Master of Science in Forestry (2000) from Stephen F. Austin State University in Nacogdoches, TX.

He is married to Charlotte Young Anderson of Hallsville, TX, and is confident that without her unconditional love and support he would have never completed this degree. I love you Charlotte.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 1 AN INTRODUCTION TO SOIL CARBON, LIGHT DETECTING AND RANGING (LIDAR) AND SPATIAL MODELLING AND GEOSTATISTICS

RELEVANCE

Terrestrial Carbon Stocks

Understanding the carbon cycle is one of the most difficult challenges facing scientists studying the global environment. Our limited knowledge of the global carbon cycle is illustrated by our inability to balance the present-day global carbon (C) budget (Dixon *et al.*, 1994; Tans *et al.*, 1990). Efforts to balance the global C budget have focused attention on terrestrial carbon storage in temperate ecosystems (Murray *et al.*, 2000; Fan *et al.*, 1998; Dixon *et al.*, 1994; Tans *et al.*, 1990). Soils constitute the major terrestrial C reservoir, 1400-1600 Pg (10^{15} g) C globally (Falloon 1998; Sundquist, 1993), approximately three to five times the amount of C contained in terrestrial biomass (Brady and Weil, 2000; Houghton and Woodwell, 1989). Roughly 78-85 Pg of SOC resides in the contiguous United States (U.S.)(Kern, 1990) with an estimated 18-19 Pg in forest soils (Turner *et al.*, 1995).

Historically, most estimates of soil organic C (SOC) are based on means extrapolated from broad categories of soils and vegetation on a regional scale (Kern, 1990; Post *et al.* 1982). However, over the past 20 years a number of estimates of regional and global C stocks have been made on extrapolations from more localized empirical data (Fan *et al.*, 1998; Eswaran *et al.*, 1993; Post *et al.*, 1982; Schlesinger, 1977). Better analysis and forecast of spatial patterns of soil properties such as SOC is important for sustainable land management (Florinsky *et al.*, 2002) and for potentially offsetting global C emissions.

Forest ecosystems of North America are of particular interest because of their ability to provide long-term C storage in both the forest vegetation and soils (Johnsen *et al.*, 2001). North American forests ecosystems sequester 3 to 4 Mg ha^{-1} yr^{-1} (Fan *et al.*, 1998) with biomass C accumulation accounting for only 1.4 Mg ha^{-1} yr^{-1} (Dixon *et al.*, 1994), suggesting C accretion in forest soils (Johnson *et al.*, 2002). Hudson *et al.* (1994) found that approximately 40% of the total global carbon (C) inventory resides in forest ecosystems, with approximately 60% of forest ecosystem C residing in soil organic matter (Birdsey *et al.*, 1993; Musselman and Fox, 1991).

Spatial Patterns of Soil Carbon

Relationships between SOC, topography and land management are important in helping to formulate and evaluate global and regional process models as well as the effects of future climate and land use changes (Post *et al.*, 1996). Furthermore, understanding spatial patterns in forest SOC may result in future development of techniques for conserving and enhancing terrestrial C pools. The importance of conservation and enhancement of SOC is owed to its positive influence on forest growth and long-term sustainability of soils. Knowledge of spatial patterns of total, labile, and recalcitrant SOC pools will aid in the forest management decision process. Additionally, development of mesoscale (1:50,000 to 1:100,000) spatial models of SOC may provide more precise measurements of regional carbon inventories.

Spatial models derived from geographic information systems (GIS) may potentially improve spatial predictions of environmental variables. Digital elevation models (DEM) provide the basis for spatial representation of environmental variables, while global positioning systems (GPS) enable accurate registration of sample locations within the DEM based environmental coverages. Our dependence on DEM for spatial modeling has contributed to a growing need for digital elevation data that provides a better representation of the earth's surface than those derived from small-scale photogrammetric sources.

Prediction of SOC and other soil properties is dependent upon the selection of pedologically important proxy landscape attributes and soil properties for use in explicit spatial models (Gessler *et al.*, 2000). A study were conducted on the Lower Coastal Plain (LCP) of eastern North Carolina to develop spatial models that can be utilized for prediction of total, labile, and recalcitrant pools of SOC by using selected soil physico-chemical properties and pedologically important topographic variables. Spatial models of SOC were developed and tested around two hypothesis: (i) that spatial patterns of SOC on a watershed scale are predictable by models based on pedological relationships displayed by topographic variation; and (ii) spatial carbon storage patterns of forest soils are affected by the methods of forest management.

OBJECTIVES

The specific objectives of this study are:

1. to compile and compare a remotely-sensed high-resolution DEM to other commonly available DEM sources

2. to utilize landscape attributes and selected soil properties to develop and validate an explicit, quantitative, and spatially realistic model of SOC for a 32,500 ha forest ecosystem

3. to determine if the spatial scale of environmental variables affects model predictions

4. to quantify SOC on an areal basis using the newly parameterized spatial models

BACKGROUND

Spatial Modeling

Although we are beginning to understand patterns of SOC storage at the site or hillslope scale (Gessler *et al.* 2000), we need better methods to scale our findings to larger landscapes. Mental models developed by soil scientist based on landscape attributes, vegetation, hydrology, and other environmental variables have long been integrated in soil science for soil mapping purposes. However, these methods result in qualitative models that produce broad schemes that attempt to encompass the soil continuum and seek to provide simplistic classification regimes (Cook *et al.*, 1996). With the emergence of quantitative pedologic measurement and modeling techniques, or pedometrics, in the 1960's (Webster, 1994) soil scientists have sought a more quantitative approach to modeling the spatial distribution of soil properties. Pedometrics

resulted in the development of statistical based approaches that incorporate surrogate environmental and edaphic explanatory variables and provide estimations of selected soil properties (McBratney *et al.*, 2000). Several approaches have historically been applied to quantitatively predict soil properties on various scales.

Spatial models have been developed that integrate standard measures of variables within discrete land units. This modeling approach has been referred to as "measure and multiply" (Schimel and Porter, 1995). Essentially, this modeling technique measures a soil property and multiplies by the area of a functionally classified land unit. Spatial anisotropy in soil properties within discrete map units hinders the ability of these models to properly represent fine-scale variability within a landscape. Despite their shortcomings, numerous studies have utilized these models to quantitatively represent soil properties. Models developed by Post (1982) used world life zones to provide discrete finite land units to quantify SOC. Similarly, Xu and Prisely (2000) used soil mapping units for areal representations of SOC and further simplified estimates on a countywide basis. Batjes (2000) utilized small- scale (1:5,000,000) maps of soil zones to predict SOC stocks for South America. "Measure and multiply" models provide a coarse estimation of selected soil properties but often lack the ability to predict soil properties on scales of 1:50,000 to 1:100,000 needed for intensive land management (McKenzie and Ryan, 1999).

Conversely, a spatial extrapolation approach referred to as "paint by numbers" (Schimel and Potter, 1995) integrates a series of independent soil and environmental variable classes with a known relationship to the dependent soil variable. Thus, discrete classes with defined combinations of explanatory variables are formed for model parameterization. An output value is determined through a deterministic function typically based on empirical data. Outputs are associated with each combinatory class and all classes are summed to represent an ecosystem response. Johnson and Kern (2003) and Falloon *et al.* (1998) utilized a similar modeling approach to predict SOC stores on an areal basis. In both cases, combinations of land units were integrated to produce prediction systems within a landscape.

Similarly, an approach to spatial extrapolation that has found recent use in soil science and geomorphology is soil-landscape modeling. Soil-landscape modeling is an approach to analyzing soil variability in response to exogenous environmental variables related to topographic and hydrologic parameters (McSweeney *et al.*, 1994; Paustian *et al.*, 1997; Thompson *et al.*, 2001). McSweeney *et al.* (1994) incorporated three stages for soil-landscape modeling: (i) physiographic representation through DEM and terrain attributes; (ii) georeferenced training data with information about soil properties; and (iii) development and validation of explicit quantitative models. The approach provided a hierarchal regime of dissimilarly scaled variables for soil-landscape modeling. This is an important concept of landscape models because it allows

for multi-resolution modeling of the soil properties. Furthermore, soil-landscape models provide quantification of soil properties through proxy variables and are not intended to provide a process-level understanding of individual soil properties.

Soil-landscape models utilize discrete land units of similar vegetation, soils or ecological zones to guide representative sampling strategies to help integrate process dynamics on the landscape. Most models are founded in Jenny's (1941) "Factors of Soil Formation" equation:

$$S = f\,(cl,\ o,\ r,\ p,\ t)$$

where S is a selected soil property as a function of climate (*cl*), organisms (*o*), relief (*r*), parent material (*p*), and time (*t*). Soil-landscape models normally assume that at the county (10 km^2) or regional (10 km^3) scale, (Ryan *et al.*, 2000) variability within *cl, p* and *t* are controlled across the study site (Paustian *et al.*, 1997). To this end, sampling strategies are designed to control for those factors that vary across the area of interest. Thus, the driver of soil-landscape modeling is that variation in topography and vegetation provide responsive proxy variables that can be utilized for prediction of soil properties.

Topographic-based spatial models derived using GIS have been utilized for spatial predictions of soil properties (Gessler, 2000; Ryan *et al.* 2000; Moore *et al.*, 1993), including forest SOC pools (McKenzie and Austin, 1993). A number of soil-landscape modeling techniques that use readily available geomorphic and pedologic based environmental explanatory variables to quantitatively predict

spatial patterns of soil properties have been developed (McSweeney *et al.*, 1994; Odeh *et al.* 1994, Gessler *et al.*, 1995; McKenzie and Ryan, 1999; McKenzie *et al.*, 2000; Florinsky *et al.*, 2002). Studies using soil-landscape spatial models have been able to predict and quantify specific soil properties such as A-horizon depth (Moore *et al.*, 1993; Bell *et al.*, 1994; Gessler *et al.*, 1995; McKenzie and Ryan, 1999; Gessler *et al.*, 2000), organic matter content (Moore *et al.*, 1993), and total SOC (Arrouays *et al.*, 1998; McKenzie and Ryan, 1999; Gessler *et al.*, 2000; Ryan *et al.*, 2000). These studies were able to explain 40 to 85% of the variability in the predicted soil properties.

Digital Elevation Models

Scaling is a serious issue in the study of C cycling in terrestrial ecosystems (Schimel and Potter, 1995). One underlying problem is that the factors that control soil variability occur across a range of scales. Proximal factors (i.e. pH, soil moisture) that influence SOC contribute to the variability on a much smaller scale. Distal factors (*viz.* Jenny's (1941) factors) vary over a much larger scale. Integration of multiple attributes at varying scales imposes a problem for developing models that predictably explain SOC variability across a landscape. Moreover, there is an inherent discord between the scale at which soil property dynamics occur (microscale), the scale at which they are measured (mesoscale) and the scale at which they are modeled (macroscale). With each increase in coarseness of scale, the variability, and therefore the uncertainty of the prediction increases (Kern, 1994; Gessler *et al.*, 2000; Ryan *et al.*, 2000).

An issue in soil-landscape modeling is that competing models often use DEM from different sources with various resolutions and precisions. Resolution is defined as the dimensions of raster grid cells. Precision refers to the number of significant digits used to report a measurement. Comparison between the models is seldom straightforward and may be confounded by DEM differences rather than by pedological differences. Landscape attribute prediction exhibits a direct dependency on the qualities of the DEM used for surface representation and attribute derivation (Jenson, 1991; Thieken *et al.*, 1999). Thieken *et al.* (1999) and Thompson *et al.* (2001) indicated that DEM resolution contributes to differences in the distribution and representation of landscape attributes. Gessler *et al.* (2000) found little difference among landscape models based on a landscape attribute derived from a series of DEM with 2- to 10-m resolutions. Likewise, Chaplot *et al.* (2000) found that 10- to 30-m DEM generally provided an unbiased prediction of landscape terrain but prediction was influenced as DEM resolution increased to 50-m. Florinsky and Kuryakova (2000) emphasized that DEM resolution was highly dependent on the scale of the process modeled, concluding that high resolution (between 2.25 and 3.25 m) DEM were important for modeling processes at the microscale. In general, these studies indicated that attribute value ranges increased and predictive capabilities decreased as DEM grid size increased. The amount of relief on a landscape contributes to the effects of DEM resolution on terrain attributes, with low relief landscapes being

less sensitive to resolution impacts. Consequently, large study areas that may

incorporate a larger range of relief may require higher resolution DEM.

Thompson *et al.* (2001) found statistical differences in landscape attributes

(specific catchment area, compound topographic index) when comparing DEM of

1 m and 0.1 m vertical precisions. Paired data revealed that lower precision (1

m) DEM had higher slope gradients and lower values for specific catchment area

and compound topographic index (Thompson *et al.*, 2001). Gyasi-Agyei *et al.*

(1995) concluded that changes in vertical precision, particularly in low relief

landscapes, affected individual cell values for terrain attributes such as slope,

specific catchment area and compound topographic index, but did not affect the

cumulative distribution of theses attributes. When precision was held constant,

Thompson *et al.* (2001) indicated that there was a dependency on horizontal

resolution, with low resolution DEM creating smoother transitions between

adjacent cells than did high resolution DEM. However, Sasowsky *et al.* (1992)

found that low vertical precision often results in a "stair stepped" appearance.

Similarly, Thompson *et al.* (2001) found that decreased vertical precision created

a greater segregation of slope values that included a large number of zero slope

areas and steeply sloped areas. A possible solution to this may be found in

calculating attributes over greater distances rather than by using only adjacent

cells (Sasowsky *et al.*, 1992).

The horizontal and vertical qualities of a DEM are directly linked to the

source of data used for its production. Traditionally, DEM have been derived

from photogrammetric techniques (includes contour mapping) and ground elevation surveys, with more recent application of remotely sensed elevation data acquired through interferometric synthetic aperture radar (IFSAR), light detecting and ranging (LiDAR) or similar technologies. The most widely used DEM within the United States have historically been the 30-m DEM (level 1) produced by the U.S. Geological Survey (USGS). DEM were produced using stereocorrolation techniques that estimate elevations from relief displacement from areas within the stereomodel. A lattice of elevation points within a pair of stereomodels was developed and resampled to a digital 30-m grid to create the DEM. A more recent approach has been the use of radar- or laser-based remotely sensed elevation data for model derivation. Application of remotely sensed data requires digital spot elevation data interpolation into raster-based DEM.

Landscape modeling studies have made use of a large variety of sources of elevation data for terrain representation. Photogrammetrically derived DEM at multiple horizontal resolutions and vertical precisions have been used to provide landscape attribute characterization, comparison, and environmental modeling (Gessler, 2000; Arrouays *et al.*, 1998; McKenzie and Ryan, 1999; Ryan *et al.*, 2000; Thompson *et al.*, 2001). Moore *et al.* (1993), Gessler *et al.* (2000), Thompson *et al.* (2001) and Florinsky *et al.* (2002) utilized DEM derived from intensive land-based elevation surveys for DEM comparisons, landscape attribute evaluations, and environmental modeling. Additionally, Thieken *et al.* (1999), White and Wang (2002) and, Woolard and Colby (2002) used DEM

12

produced from remotely sensed elevation data in similar capacities. Continued

exploration of available elevation data sources of various precisions and

resolutions will enhance our ability to model the soil landscape continuum.

Carbon Pools

Relating SOC, topography, and land management is an important step in

formulating landscape models to aid in understanding the effects of future climate

and land use changes (Post *et al.*, 1996). Sequestration of C into soil systems

has been conceptualized as a possible mechanism for offsetting C containing

atmospheric gases. Southern managed pine forests play an important role in

sequestering and maintaining SOC. Johnsen *et al.* (2001) proposed that future

carbon accretion in managed forests is dependent on three factors: (i) land-use

changes and management regimes that increase biomass carbon; (ii) remaining

recalcitrant SOC following timber harvest; and (iii) long-term carbon storage in

forest products. Understanding the spatial distribution of terrestrial C pools is

paramount in initializing a C sequestration program. Furthermore, knowledge of

spatial patterns of total, labile, and recalcitrant SOC pools will aid in the forest

management decision process.

The SOC component within mineral soil, defined as that which passes

through a 2 mm sieve or that which is associated with mineral material (Johnsen

et al., 2001), is the keystone of perennial C storage. However, SOC accretions

in mineral soil are often slow in the southern forests because of high

decomposition rates of new carbon added to the soil system (Richter *et al.*,

1999). Thus, knowledge of the landscape or regional scale quantities of discrete C pools will allow for prescription of long-term land management regimes designed to enhance our ability to sequester atmospheric C. Several studies have indicated that a reasonable estimate for total SOC in southeastern forest ecosystems should range from 6-20 kg C m^{-2} for mineral soils and around 80 kg C m^{-2} for organic soils to a depth of 1 m (Johnson and Kern, 2003; Birdsey and Lewis, 2003; Garten *et al.*, 1999).

Location of SOC within the soil profile also plays a critical role in determining turnover rates of specific fractions. Garten *et al.* (1999) found that the upper 30 cm accounted for 70-90% of all C in forest soils. Additionally, they found that within the upper 5 cm, approximately 31% of the total SOC was in the labile fraction, but fell to 20% in the 5-30 cm depth.

Catenary landforms may provide a basis for predicting SOC on a landscape scale. Moore *et al.* (1993) rationalized that in many landscapes catenary soil development occurred in response to surface and subsurface water movement across the landscape. Bell *et al.* (2000) found significant relationships between SOC and landscape position and characterization including slope gradients, elevation above peatlands, and distance from peatlands. Moore *et al.* (1993) and Arrouays *et al.* (1998) found a spatial dependence of organic matter accretion associated with landscape morphology and position. Gessler *et al.* (2000) found landform and landscape position to be significant in SOC prediction, but was openly curious about the efficiency of similar models applied to low-relief

14

landscapes. However, it is thought that pedologic processes in the LCP of eastern NC are responsive to changes in hydrology across broad interstream divides (Daniels *et al.*, 1999). These broad low-relief interfluves are composed of mineral soils that often encircle Histisols containing large stores of SOC. These large organic flats are a result of shallow groundwater and are generally located in areas furthest from major drainages. Thus it is anticipated that landscape position could contribute to SOC stores, and that these stores could be predicted based on landscape attributes.

Land Use and Forest Management

Humans have the potential to alter the magnitude and direction of forest SOC stores through forest management activities and land-use change (Brown *et al.* 2002). The U.S. is composed of nearly 204 million ha of forest ecosystems (Birdsey and Lewis, 2003). The southern U.S. contains roughly 81 million ha of land classified as timberland (Southern Forest Resource Assessment, 2002), including an approximate 15 million ha of managed plantation forests (Birdsey and Lewis, 2003). Management activities in production forestry plantations generally practiced in the southeastern U.S. include mechanical site preparation, fertilization and competition control (Harding and Jokela, 1994). These management regimes govern our ability to sequester SOC and sustain current SOC stores on lands actively managed for timber resources through manipulation of edaphic and vegetative components.

Some uncertainty remains as to the effects of forest management activities on the accretion, allocation, and dynamics of total forest C (Harding and Jokela, 1994). Even less is known about SOC dynamics in forest soils because many of the controlling variables lack extensive research (Sanchez and Eaton, 2001). Several studies have addressed the rates of change in SOC under specific conditions at selected sites, however the quantity of SOC on regional scales remains unknown (Johnson *et al.*, 2002). Understanding the effects of land use and forest management allow for appropriate stratified sampling regimes for integration of controlling factors across regional-scaled landscapes and subsequent improved spatial modeling of soil resources.

Land use and land change play a crucial role in determining morphology, physiology, and quantities of SOC. It is well understood that conversion of non-forested land to forests (afforestation) causes increased C sequestration, and conversely that conversion of forests to non-forested land causes decreased SOC storage. Afforestation of abandoned agricultural lands has been shown to increase SOC at a rate of 0.03 to 0.89 Mg C ha^{-1} yr^{-1} (Huntington, 1995; Richter *et al.*, 1995; Van Lear *et al.*, 1995). Hu *et al.* (1997) found that total SOC was 2 to 25 times greater for forest soils as compared to adjacent agricultural lands. This large range was likely due to the inherent variability of forest SOC pools, yet Trettin *et al.* (1999) found that SOC pools in undisturbed forests tend to remain relatively stable over long periods. In a review of published and unpublished

studies, Paul *et al.* (2002) found that SOC data were highly variable, particularly in reestablished stands less than 10 yr of age.

Mechanical and chemical site preparation and mid-rotation amendments of intensively managed pine plantations result in transitory perturbations in SOC pools (Carter *et al.*, 2002; Shan *et al.*, 2001; Johnson, 1992). Mechanical site preparations such as bedding, disking, and sub-soiling are performed to ameliorate soil physical properties. Laiho *et al.* (2003) found that intensive site preparation such as bedding, fertilization, and herbicide increased SOC stores. Additionally, a literature review by Paul *et al.* (2002) found that there were no significant effects of disturbance levels on SOC and that the lack of litter input following the site preparation application was potentially the cause for some of the decreases detailed in the reviewed articles. However, Burger and Pritchett (1984) found nearly 68% reduction in SOC following disking and bedding. Existing SOC levels prior to mechanical site preparations would likely influence the total change in SOC levels.

Contrasting results were seen in studies that focused on the effects of fire and herbicide on SOC stores. These treatments primarily release pine plantations from competing vegetation, improve equipment mobility within sites, or achieve tertiary management goals. After a review of some 15 studies, Johnson (1992) concluded that low intensity prescribed fire resulted in minor changes in SOC while more intense burns resulted in large SOC losses. A more recent review indicated that prescribed fire resulted in SOC loss in the A horizons

while wildfires positively influence SOC levels (Johnson and Curtis, 2001). Herbicide treatments have generally proved deleterious to SOC, resulting in as much as 20% decrease by herbicide treatment alone (Burgess *et al.*, 1995). Likewise, Shan *et al.* (2001) found that herbicide control of competing vegetation in slash pine (*Pinus elliottii*) reduced SOC. They concluded it was a result of decreased root mass and litter inputs in conjunction with elevated microbial decomposition rates.

Fertilizer applications are often used in multiple applications to improve the nutrient status of production pine plantation sites (Harding and Jokela, 1994). A study conducted by Carter *et al.* (2002) found no effects on SOC in the upper 60 cm from nitrogen (N) fertilization 16 months after application on a loblolly pine (*Pinus taeda*) stand. Similarly, Harding and Jokela (1994) found no effects on SOC levels to a depth of 91 cm 25 yrs after an application of superphosphate. However, these results were in contrast to the general trend of increasing SOC with fertilization as reported in a meta-analysis by Johnson (1992). Consequently, it would appear that there are conflicting reports as to the overall effect of fertilization on SOC levels. Fertilization results in improved biomass production and increased microbial decomposition of new organic inputs, thus regulating the net change in SOC.

Current assessment models assume a 20% decline in SOC following clear-cut harvesting and intensive site preparation (Birdsey, 1996). However, a recent review by Johnson and Curtis (2001) concluded that forest harvesting had

18

little impact on SOC in the A horizon, yet indicated that sawlog harvesting had a slight increase and whole-tree harvesting had a slight decease in SOC levels. Johnson *et al.* (2002) found no lasting effect of sawlog or whole-tree harvesting on SOC in slash pine flatwoods sites in Florida or in loblolly pine sites in South Carolina. However, harvesting resulted in an initial increase in SOC during the first 2-5 yrs (Laiho *et al.*, 2003; Johnson *et al.*, 2002) after harvest that subsequently declined to levels slightly elevated from pre-harvest stores by year 4 (Johnson *et al.* 2002). Knoepp and Swank (1997) found similar results on white pine (*Pinus strobus*) in western North Carolina, with an initial increase over a 3 yr period following harvest that eventually reverted to pre-harvest levels. Van Lear *et al.* (2000) suggested that increased SOC levels following harvest were due to decomposing root systems. A study conducted by Carter *et al.* (2002) found that harvesting had no effects on SOC compared to pre-harvest levels and meta-analysis (Johnson, 1992) of numerous studies indicated that harvesting introduced a ± 10% change in SOC pools. In general, small short-term changes in SOC could be expected but over the long run SOC will remain relatively stable following harvests of timber.

Environmental dynamics coupled with numerous permutations of possible management regimes makes prediction of SOC a difficult task. Often site manipulations are made in conjunction with one another and on a sliding temporal scale. Inherent and management induced variability of SOC in forest soils emphasizes the need for wariness in modeling efforts yet also suggests the

need for improved spatial modeling techniques. However, it appears that in general forest management has only modest effects on long term SOC. In addition, its important to add that spatial modeling within this study concerns residual SOC pools and will not be focused on obtain or predicting carbon accretion rates.

APPROACH

Study Site

Investigation into the spatial distribution of SOC will occur on a 32,500 ha forest ecosystem located entirely within the bounds of Hofmann Forest. Hofmann Forest is located in Jones and Onslow Counties of eastern North Carolina, USA (Figure 1) and lies on the LCP Wicomico and Talbot morphostratigraphic units of the mid-Atlantic seaboard. Located in the temperate climate zone, the study area is characterized by warm summers and mild winters with a mean summer temperature of 25°C and a mean winter temperature of 7°C. Mean annual precipitation is 1400 mm with a large portion of the rainfall received in late summer. Elevations range from 12 to 20 m above mean sea level (Daniels *et al.*, 1977), with the landscape characterized by broad, flat interfluves. In some areas, relief may be as low as 1.5 m elevation difference in 3 to 4 km (Daniels *et al.*, 1999).

Soils

Soils of the Hofmann Forest were derived from surficial marine sediments of the Wicomico and Talbot morphostratigraphic units, alluvial deposits, and

organic deposits on low-relief interfluves. The soils are predominately poorly to somewhat-poorly drained Saprists, Aquults and Aquepts (Figure 1). Organic soils dominate Hofmann Forest, representing nearly 24,000 ha (Daniels *et al.*, 1977). The Hofmann Forest landscape indicates poorly drained organic soils occurring on broad low-relief interfluves and better-drained soils occurring in close proximity to drainages. Mineral soils fringe the broad interstream divides and are typified by deep water tables with light surface and subsurface horizons. Daniels *et al.* (1977) provides a detailed description of the stratigraphy, geomorphology and pedological units with the Hofmann Forest.

Hydrology

The unique hydrological conditions within Hofmann Forest appear to be beneficial to the accumulation of C within interfluve areas. Preliminary examination of soil survey information indicates a topographic influence on spatial patterns of SOC, with C accretions occurring in interfluves furthest from major drainages. Thick organic layers occur within the Forest interior and grade to mineral soils with minor SOC components nearest to streams. Daniels *et al.* (1977) indicated an influence of distances to nearest major drainage on organic material accumulation within Hofmann Forest. The wide spacing and low slope between natural drains inhibit lateral water movement, thus creating large partially saturated regions within the centers of the forest. These areas, commonly referred to as pocosins, are often characterized by substantial accumulation of organic materials (Figure 2). Gilliam (1991) found that pocosins

are generally infertile due to deficient base cations, low effective cation exchange

capacity and reducing conditions below the water table. Without intensive

management, pocosins do not normally support productive loblolly pine stands.

Natural pocosins occupy large extents of the landscape but their role in the

regional and global carbon and water cycles is undetermined (Brinson, 1991).

Vegetation

The Hofmann Forest contains a large range of vegetation, both as a result

of natural regeneration and intensive forest management. Natural palustrine

wetland plant communities dominate the pocosin area with the presence of pond

pine (*Pinus serotina*), redbay (*Persea borbonia*), loblolly bay (*Gordonia

lasianthus*), sweetbay (*Magnolia virginiana*), bamboo (*Smilax laurifolia*) and

gallberry (*Ilex glabra*), and many others. A major portion of the land surrounding

the pocosin includes managed pineland dominated by loblolly, slash and longleaf

pine (*Pinus palustris*) plantations. Other vegetation land use/land classification in

Hofmann Forest includes bottomland hardwood, hardwood flats, headwater

swamps, non-managed pine flats, swamp forests, agricultural fields and pine

savannas (unpublished data).

Forest Management

Moderate to intensively managed pinelands occupy nearly 15,000 ha of

Hofmann Forest. Drainage networks have been installed throughout most of the

managed pine plantation lands with minor drainage networks on approximately

100 m spacing between adjacent drains. Management goals mandate

establishing plantation to meet target survival rates of roughly 205 trees ha^{-1}.

Bedding is ubiquitous and is generally performed at 3-4 m spacings. Often sites are windrowed, however this practice has been reduced somewhat because of suspected negative impacts on long term site sustainability. Due to a relative universal phosphorus (P) deficiency in the soils of Hofmann Forest, sites normally receive at least 35 kg ha^{-1}of P as diammonium phosphate (DAP) incorporated into the planting bed. Fire is common as a pre-planting silvicultural prescription and on 3-5 yr intervals as an understory vegetation control mechanism.

Most stands are thinned every 5 to 10 yrs with fertilization with N and P often occurring following mid-rotation thinnings. Pine plantations in Hofmann Forest range in age from 0 to 67 yrs, fairly distributed across the age groups of 0-5, 5-10, 10-15, 15-20, 20-25, 25-30, 30-35, 35-40 yrs. However, the 20-25 yrs age group was by far the largest, representing over 3000 ha (20% of total plantations).   Harvesting is performed on a commercial contract basis. Harvest regimes are site specific and may include whole-tree and sawlog-only harvests. Digital information regarding tree volume removals and other details of the harvest are acquired and incorporated into a GIS database for Hofmann Forest. Other information including soils, vegetation, land use/land classification, and age class are digitally georeferenced and integrated into the database.

Analysis of LiDAR Data

The need for high-resolution, high accuracy elevation data for purposes of landscape-scale modeling has resulted in the application of various technologies for digital elevation acquisition. Light Detecting and Ranging (LiDAR) technology is a source of high-precision, high-accuracy data and may be used to produce high-resolution, high-accuracy DEM. The LiDAR system uses thousands of laser pulses (4,000-50,000 returns sec^{-1}) directed from an airborne transmitter to accurately measure distances to ground features. LiDAR returns are generated and recorded as spot elevation data of known accuracy within a given land class (forest criteria: RMSE = 20 cm; agriculture criteria: RMSE = 15 cm). Compilation of the spot elevation data with subsequent data interpolation could provide a high-resolution, high-accuracy DEM for landscape-scale spatial modeling. Actively acquired LiDAR spot elevation data of known accuracy were utilized for derivation of a high resolution DEM.

A series of 10 replicate 1000-ha LiDAR data tiles were acquired from the North Carolina Flood Mapping Program. The tiles will represent LiDAR acquired digital spot elevation data for forested sites in the Lower Coastal Plain of eastern North Carolina. Data densities are expected to range from roughly 100 to 300 spot elevations ha^{-1}. While an abundance of data has traditionally been positively viewed, it has come to our attention that typical LiDAR datasets impose serious computing constraints and are often problematic. The initial step were to perform a sequential reduction in dataset size through random selection of a

24

predetermined percentage of the original LiDAR dataset (Figure 3). The total LiDAR dataset were initially reduced by 50% to better accommodate computing requirements. One half were used as a training dataset and a source of subsequent datasets while the remaining half of the dataset were used as a validation dataset for a prediction error analysis. The sequential reduction were performed in this manner so as to not include any validation data points in the training datasets.

An exploratory analysis of the spatial structure of the LiDAR dataset were performed. The global exploratory analysis were performed on the 50% training dataset produced from the original LiDAR dataset. The initial step in the exploratory analysis were to determine an appropriate lag interval for producing the semivariograms. The novariogram procedure were employed using the spatial module of SAS (SAS Institute, 2000) to evaluate and determine a lag distance that appears to be most beneficial for further analysis of the spatial structure of each dataset. A restricted maximum likelihood (REML) analysis were used to parameterize kriging variograms to provide the "best" possible fit of the theoretical semivariogram with the empirical data. Omni-directional semivariograms were produced for each tile to analyze global trends in the data. Four-directional semivariograms were produced to evaluate global anisotropic tendencies within each LiDAR dataset. Spherical, log, or other mathematical transformations were employed to correct anisotropy within the datasets. A subsequent evaluation were performed at small lags to adjust the scale of the

25

evaluation so as to focus on variance at smaller distances (particularly within 30-50 m).

There are two main groupings of linear interpolation: deterministic and geostatistical. Deterministic interpolation techniques like inverse distance weighted create surfaces from measured points based on the extent of similarity. Geostatistical interpolation methodologies such as those imposed by kriging utilize the statistical properties of the measured points. Two common interpolation procedures were evaluated across the density range of training datasets. Inverse distance weighted squared (IDW^2) and ordinary kriging (OK) were utilized to produce interpolation models. IDW^2 assumes that each measured point has a local influence that diminishes with distance. It weights the points closer to the prediction location greater than those farther away. Kriging forms weights from surrounding measured values to predict values at unmeasured locations. As with IDW^2, the closest measured values usually have the most influence. Kriging weights come from a theoretical semivariogram fitted to empirical data that was developed by looking at the spatial structure of the data in question (Kitanidis, 1997).

Inverse distance weighted squared takes the form $w(d) = 1 / d^2$ where $w(d)$ is the weight at distance d that is determined by the inverse of the distance at a user defined power (in this case power = 2)(Wilson and Gallant, 2000). Ordinary kriging takes the form $Z(\mathbf{s}) = \mu(\mathbf{s}) + \varepsilon(\mathbf{s})$ where $Z(\mathbf{s})$ is the variable of

interest, decomposed into a deterministic trend μ(**s**) and a random,

autocorrelated error in the form of ε(**s**) (Kitanidis, 1997).

These methods were evaluated by cross-validation and through prediction

and standardized error analysis of the validation dataset. Parameters gained

from global exploratory analysis were used to parameterize both the IDW^2 and

OK models. Model evaluation were performed in ArcInfo/ArcGIS (ESRI, 2002)

with data analysis, model parameterization and prediction and standardized error

analysis performed in the geostatistical module. Interpolation models were used

to predict elevations of the independent validation datasets. Prediction and

standardized errors for the validation dataset were based on the irregularly

spaced LiDAR spot elevation locations. The same validation dataset were used

at each training density within a given LiDAR tile to ensure comparable results

across the range of LiDAR densities.

Selection of the most efficient and statistically valid LiDAR interpolation

technique were used to produce a DEM for Hofmann Forest. It is assumed that

given that Hofmann Forest is predominantly composed of forested land and

located within the same physiographic region as the analyzed LiDAR tiles, similar

techniques may be used to produce a valid DEM.

Terrain Modeling

The topographic features of the landscape display pedologic variation

within the Hofmann Forest. Terrain modeling and landscape characterization to

elucidate edaphic relationships is key to understanding biogeochemical cycling

27

because they often indicate spatial distributions of soil processes (Gessler *et al.*, 2000). Spatial models derived from GIS may potentially improve spatial predictions of environmental variables. These models have myriad uses that are applicable to scientific, economic, and political disciplines.

Most spatial models depend on DEM to provide the basis for spatial representation of environmental and landscape variables. The initial step in terrain modeling is developing or acquiring a DEM that provides suitable spatial representation of landscape morphology. Photogrammetrically derived 30-m horizontal resolution, 1-m vertical precision seamless National Elevation Dataset (NED) DEM for Hofmann Forest were acquired. Though NED-DEM are commonly used for landscape-scaled surface representations, many discrepancies are often apparent in the broad low-relief interfluves of the LCP.

A comparison of NED-DEM and LiDAR-DEM produced from selected interpolation procedure with 17 North Carolina Geodetic Survey (NCGS) control points located within Hofmann Forest were performed. This will provide an evaluation of each DEM source to accepted "true" ground elevations within Hofmann Forest. Also, a direct comparison were made between the NED- and LiDAR-DEM. This comparison will focus on differences generated as a result of source information and not an evaluation of precisions. The two sources of DEM were compared for differences in elevation and other terrain attributes between co-registered grid cells. Differences were evaluated for spatial autocorrelation to measure strengths in tendencies for differences from nearby regions to be more

(or less) alike than differences from regions further apart. Moran's I and Geary's c statistics for continuous data were utilized for autocorrelation analysis of difference grids produced from competing DEM.

It is anticipated that the low relief landscape of the Lower Coastal Plain of eastern North Carolina were best represented with a coarse horizontal resolution, high vertical precision DEM. A statistical evaluation of the two competing DEM and an assessment of the required precision were performed. Based on the information gleaned from this analysis, a elevation data source and interpolation technique were selected for producing the final DEM for Hofmann Forest.

Terrain Attributes

Terrain attributes were calculated from the selected DEM that may be used in the prediction of soil properties were compiled. Terrain attributes are designed to provide quantitative parameters indicative of landform shape, connectivity, and adjacency that control external landscape geomorphology and represent hydrologic tendencies (Gessler *et al.*, 2000).   Topographic descriptors correlated to soil properties generally provide some indication of overland or subsurface lateral flow of water across the landscape.

Digital modeling uses a host of terrain attributes for describing landscape morphology and often point toward landscape functionality and control biogeochemical cycling mechanisms. Terrain attributes can be divided into primary and secondary attributes. Primary attributes such as elevation and slope gradient are derived directly from DEM, whereas secondary attributes are

derived from combinations of primary attributes (Moore *et al.*, 1991). Secondary

attributes serve as surrogates for complex hydrological, geomorphological and

pedological processes (Moore *et al.*, 1991). Studies by Bell *et al.* (1995) and

Moore *et al.* (1993) explored the rationale and importance of elevation data and

derived terrain attributes for prediction of soil properties. A list of terrain

attributes that may provide relational patterns to soil properties were compiled,

yet does not represent an exhaustive list of all possible terrain attributes

desirable for this study (Table 1). These terrain attributes were included in a

statistical evaluation and model parameterization of the spatial patterns of SOC.

Sample Stratification Regime

A stratified random sampling scheme for collection of SOC samples were

used to minimize variability in physical and chemical differences between

samples, while maintaining adequate coverage of the watershed. Variation may

be reduced when sampling on a stratification scheme based on *a priori*

understanding of factors that potentially influence the prediction variable (Ryan

et al., 2000; Turner and Lambert, 2000; Zinke and Stangenberger, 2000). The

stratified sampling regime were based on three criteria: (i) plantation pine versus

natural pocosin plant communities, (ii) age groupings within pine plantations, and

(iii) distance from major streams.

Hofmann Forest were divided into two major vegetation classes

(plantation and pocosin) for initial stratification. In the LCP, vegetation and land

use are often indicative of underlying soil conditions. These two vegetation

groups were selected based on their influence on C storage and their dominant areal extent within Hofmann Forest. Pocosin areas are anticipated to represent areas of greatest SOC accumulation. Plantation areas are suspected to occupy drier, less organic soils or have less SOC as a result of anthropogenic activities.

Plantation and pocosin vegetation classes were subsequently stratified into 4 groups representing distances of 0-2000m, 2000-3000m, 3000-4000m and 4000[+]m to major natural drainages. It is suspected that distance to natural drainage is extremely influential in the genesis of soils within Hofmann Forest. Previous research in Hofmann Forest by Daniels *et al.* (1977) indicated that morphological differences occurred on the broad flat interfluves as result of distance from major streams and drains.

A further stratification of pine plantations were based on plantation age. Numerous research studies indicate the highly variable nature of SOC pools within different aged pine stands (Huntington, 1995; Richter *et al.*, 1995; Van Lear *et al.*, 1995; Trettin *et al.*, 1999; Paul *et al.*, 2002). Age groupings will include 0-5, 5-15, 15-25, 25-35 and 35[+] yr. This strategy should help ensure proper stratification based on stand establishment (0-5yr), initial stand closure, thinning, and fertilization (5-15yr), late rotation thins or other silvicultural activities (15-25yr), stands scheduled for harvest (25-35yr) and mature older plantations (35[+]yr).

Soil Sampling and SOC Analysis

Soil samples were collected at 190 georeferenced locations throughout the study site for chemical analysis and determination of soil bulk density. Four sub-samples were taken at roughly 7.5 m at approximate cardinal directions from each prescribed geo-referenced sampling location. Sub-samples on bedded plantation sites were oriented to provide 2 inter-bed and 2 intra-bed samples. Soils were collected in 3 cm butyrate plastic liners with a stainless steel soil recovery probe with slide hammer attachment (JMC Soils-ESP soil sampler). Measurements from small diameter coring systems have similar variability as large diameter (15 cm) coring systems (Ruark and Zarnoch, 1992) and provide an amicable means of capturing intact soil cores on the remote sites within Hofmann Forest.

Intact volumetric soil samples were collected to a depth of approximately 1 m from the surface, including the organic horizon. Samples were stored indoors at 16 to 20°C in the plastic butyrate liners during the sample collection period. Subsequently, soil samples were extruded and divided into 20 cm depth increments. Each sample were dried in a forced-air oven at approximately 40°C for 72h. Bulk density were measured on oven-dried samples by standard technique using the cylindrical volume and soil mass within a given depth increment. Moisture corrected mass of 5 g sub-samples were calculated and entered into the overall bulk density calculation. Bulk density will represent an average of the 4 sub-samples or the total number of uninterrupted cores from a

given location. Soil samples will then be aggregated with the other sub-samples from each sampling location.   Cores from interrupted samples were composited with others from the given sampling location, but these composites will only be used for chemical analysis of total SOC. Total SOC of whole soil samples were determined by dry combustion in a Perkin-Elmer Series II 2400 CNH analyzer (Nelson and Sommers, 1996; Bremner, 1996). All soil samples were ground to pass a 0.250 mm (no. 60) sieve and analyzed for total SOC.

Statistical Analysis

The development of useful statistical models for soil distribution is dependent on the explanatory variables (environmental variables) being more easily observed or quantified than the soil properties of interest. The task is to determine the minimal quantity of the physical resource needed to produce the spatial model in compliance within a success criterion threshold or error estimates. Likewise, we continue to desire new methods to evaluate and validate our predictions. Model evaluation helps to determine under what circumstances and with what reliability a model will function. Landscape attributes coordinated with geo-referenced soil samples were integrated with measured SOC to produce spatial models. Spatial models were interfaced into GIS and used to spatially quantify SOC.

Seventy-five percent of the 190 geo-referenced sampling locations were used as model training data with the remaining 25% of the sites serving as a validation dataset for a prediction error analysis. A range of statistical

methodologies were employed on the training data to develop models for spatial

prediction of the SOC. Methods such as backward and forward stepwise multiple

regression analysis (McKenzie and Austin, 1993; Gessler *et al.*, 1995; Gessler *et al.*, 2000), tree-based methods (McKenzie and Ryan, 1999), maximum and

minimum R^2 analysis (SAS Institute, 2000) and geostatistical techniques were

used for development of model parameters and spatial predictions. Primary and

secondary landscape attributes were used as explanatory variables in the

prediction models.

Stepwise multiple regression methods relate target variable (SOC and N)

to explanatory prediction variables. A predefined significance level were

assigned and an *F*-test statistic were used to decide whether an explanatory

variable were added to the regression model (Neter *et al.*, 1989; SAS Institute,

2000). Forward stepwise regression successively adds variables to the

prediction model that exceed the established partial *F*-test statistic threshold, yet

only adds variables that improve the overall predictive capabilities of the model

(Neter *et al.*, 1989; SAS Institute, 2000). Backward stepwise regression begins

with all potential variables and successively removes variables from the model.

Variables are removed from the model when they are less than the minimum

required *F* value (Neter *et al.*, 1989; SAS Institute, 2000). Maximum and

minimum R^2 evaluations work in similar fashion but rather than looking at *F*

values, it focuses on the change in model R^2 as a result of adding variables to

the model (SAS Institute, 2000). The difference between the stepwise regression

34

methods and the maximum/minimum methods is that all possible variable

replacements are evaluated before any replacement is implemented (SAS

Institute, 2000).

Regression trees are regarded as a variant of decision trees, designed to

approximate values instead of being used for classification tasks. The final

results of using tree methods for regression can be summarized in a series of

logical if-then conditions at tree nodes (Breiman, 1984). Therefore, there is no

implicit assumption that the underlying relationships between the predictor

variables and the dependent variable are linear, follow some specific non-linear

function (Breiman, 1984). They are particularly useful in that they work well

when regression variables are a mixture of categorical and continuous variables.

They may allow for regression models using categorical separation of the

landscape (e.g. vegetation classes) in conjunction with continuous data from

calculated landscape attributes (e.g. elevation, profile curvature).

Landscape attributes for the prediction model were developed from digital

elevation models (DEM) of Hofmann Forest using standard GIS methodologies

as described by Moore *et al.* (1991), Moore *et al.* (1993), Ryan *et al.* (2000),

Wilson and Gallant (2000). Primary and secondary landscape attributes will

include, but are not limited to those attributes described in the previous sections.

Regression models were implemented into the GIS and displayed on a raster

basis. Map algebra were employed to provide quantification with raster units.

Predictive maps and models were evaluated against the remaining 25% subset established for model testing. Both validation and cross-validation techniques were used to provide error estimates of the SOC multiple regression model. Mean predicative error and RMSE were calculated for the testing dataset. The end result were a predictive map of areal quantities of SOC with known error estimates and predictive capabilities.

LITERATURE CITED

Amelung, W., K. Kaiser, G. Kammerer, and G. Sauer. 2002. Organic carbon at soil particle surfaces – evidence from x-ray photoelectron spectroscopy and surface abrasion. Soil Sci. Soc. Am. J. 66:1526-1530.

Arrouays, D., and P. Pelissier. 1993. Modeling carbon storage profiles in temperate forest humic loamy soils of France. Soil Sci. 157(3):185-192.

Arrouays, D., J. Daroussin, J. Luc Licin and P. Hassika. 1998. Improving topsoil carbon storage prediction using a digital elevation model in temperate forest soils of France. Soil Sci. 163:103-108.

Bajtes, N.H. 2000. Effects of mapped variation in soil conditions on estimates of soil carbon and nitrogen stocks for South America. Geoderma 97:135-144.

Barrios, E., R.J. Buresh, and J.I. Sprent. 1996. Organic matter in soil particle size and density fractions from maize and legume cropping systems. Soil Biol. Biochem. 28(2):185-193.

Beare, M.H., P.F. Hendrix, and D.C. Coleman. 1994. Water-stable aggregates and organic matter fractions in conventional- and no-tillage soils. Soil Sci. Soc. Am. J. 58:777-786.

Beare, M.H., M.L. Cabrera, P.F. Hendrix, and D.C. Coleman. 1994. Aggregate-protected and unprotected organic matter pools in conventional- and no-tillage soils. Soil Sci. Soc. Am. J. 58:787-795.

Bell, J.C., D.F. Grigal, and P. Bates. 2000. A soil-terrain model for estimating spatial patterns of soil organic carbon. P. 295-310. *In* Wilson and Gallant (eds.) Terrain analysis: principles and applications. John Wiley and Sons, New York, NY.

Birdsey, R.A. and G.M. Lewis. 2003. Current and historical trends in use, management, and disturbance of U.S. forestlands. p. 15-33. *In* Kimble *et al.* (eds.) The potential of U.S. forest soils to sequester carbon and mitigate the greenhouse effect. CRC Press, Boca Raton, FL.

Birdsey, R.A., A.J. Plantinga, and L.S. Heath. 1993. Past and prospective carbon storage in United States forest. For. Ecol. Manage. 58:33-40.

Boone, R.D. 1994. Light-fraction soil organic matter: origin and contribution to net mineralization. Soil Biol. Biochem. 26(11):1459-1468.

Brady, N.C. and R.R. Weil. 2000. Elements of the nature and properties of soils. Prentice Hall, Upper Saddle River, NJ.

Bremner, J.M. 1996. Nitrogen-Total. p. 1048. *In* Page *et al.* 1996. Methods of soil analysis. Part 3: Chemical Methods. ASA, CSSA, and SSSA, Madison, WI.

Breiman, L., J. Friedman, R. Olshen, and C.J. Stone. 1984. Classification and

    Regression Trees. Chapman and Hall, New York.

Brinson, M.M. 1991. Landscape properties of pocosins and associated wetlands.

    Wetlands 11:4441-4465.

Brown, S., I.R. Swingland, R. Hanbury-Tenison, G. T. Prance, and N. Myers.

    2002. Changes in the use and management of forests for abating carbon

    emissions: issues and challenges under the Kyoto Protocol. Phil. Trans. R.

    Soc. Lond. 360:1593-1605.

Burger, J.A. and W.L. Pritchett. 1984. Effects of clearfelling and site preparation

    on nitrogen mineralization in a southern pine stand. Soil Sci. Soc. Am. J.

    48:1432-147.

Burgess, D., J.A. Baldock, S. Wetzell, D.G. Brand. Scarification, fertilization and

    herbicide treatment effects on planted conifers and soil fertility. Plant Soil

    169:513-522.

Cambardella, C.A. and E.T. Elliot. 1992. Particulate soil organic-matter changes

    across a grassland cultivation sequence. Soil Sci. Soc. Am. J. 56:777-783.

Cannell, M.G.R. 1999. Growing trees to sequester carbon in the UK: answers to

    some common questions. Forestry 72(3):237-247.

Chapela, I.H., L.J. Osher, T.R. Horton, and M.R. Henn. 2001. Ectomycorrhizal

    fungi introduced with exotic pine plantations induce soil carbon depletion. Soil

    Biol. Biochem. 33:1733-1740.

Chaplot, V., C. Walter, and P. Curmi. 2000. Improving soil hydromorphy prediction according to DEM resolution and available pedogical data. Geoderma 97:405-422.

Christensen B.T. 1992. Physical fractionation of soil and organic matter in primary particle size and density separates. Advances in Soil Science 20:1-90.

Clapp, C.E., M.H.B. Hayes, and R.S. Swift. 1993. Isolation, fractionation, functionalities, and concepts of structures of soil organic macromolecules. *In* Beck *et al.* (eds.) Organic substances in soil and water: natural constituents and their influences on contaminant behavior. Royal Soc. Chem. Special Pub. 135. Thomas Graham House, Cambridge.

Cook, S.E., R.J. Corner, G. Grealish, P.E. Gessler, and C.J. Chartes. 1996. A rule-based system to map soil properties. Soil Sci. Soc. Am. J. 60:1893-1900.

Cromack, K., R.E. Miller, O.T. Helgerson, R.B. Smith, and H.W. Anderson. 1999. Soil carbon and nutrients in a coastal Douglas-fir plantation with red alder. Soil Sci. Soc. Am. J. 63:232-239.

Daniels, R.B., E.E. Gamble, W.H. Wheeler, and C.S. Holzhey. 1977. The stratigraphy and geomorphology of the Hofmann Forest Pocosin, North Carolina. Soil Sci. Soc. Am. J. 41:1175-1180.

Daniels, R.B., S.W. Buol, H.J. Kleiss, and C.A. Ditzler. 1999. Soil systems of North Carolina. Tech. Bull. 314. North Carolina State Univ., Raleigh.

Dixon, R.K., S. Brown, R.A. Houghton, A.M. Solomon, M.C. Trexler and J. Wisniewski. 1994. Carbon pools and flux of global forest ecosystems. Science 263:185-190.

Eswaran, H., E. Van Den Berg, and P. Reich. 1993. Organic carbon in soils of the world. Soil Sci. Soc. Am. J. 57:192-194.

Falloon, P.D., P. Smith, J. U. Smith, J. Szabo, K. Coleman, and S. Marshall. 1998. Regional estimates of carbon sequestration potential: linking the Rothamsted Carbon Model to GIS databases. Biol. Fert. Soils. 27:236-241.

Fan, S., M. Gloor, J. Mahlman, S. Pacala, J. Sarmiento, T. Takahashi, and P. Tans. 1998. A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models. Science 282:442-446.

Florinsky, I.V., R.G. Eilers, G.R. Manning, and L.G. Fuller. 2002. Prediction of soil properties by digital terrain modeling. Environ. Model. Soft. 17:295-311.

Florinsky, I.V. and G.A. Kuryakova. 2000. Determination of grid size for digital terrain modeling in landscape investigations – exemplified by soil moisture

Garten, C.T., W.M. Post, P.J. Hanson, and L.W. Cooper. 1999. Forest soil carbon inventories and dynamics along an elevation gradient in the southern Appalachian Mountains. Biogeochem. 45:115-145.

Gaudinski, J.L., S.E. Trumbore, E.A Davidson, and S. Zheng. 2000. Soil carbon cycling in a temperate forest: radiocarbon-based estimates of residence times, sequestration rates and portioning. Biogeochem. 51:33-69.

Gessler, P.E., O.A. Chadwick, F. Chamran, L. Althouse, and K. Holmes. 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Sci. Soc. Am. J. 64:2046-2056.

Grigal, D.F. and L.F. Ohmann. 1992. Carbon storage in upland forests of the lake states. Soil Sci. Soc. Am. J. 56:935-943.

Groninger, J.W., K.H. Johnsen, J.R. Seiler, R.E. Will, D.S. Ellsworth, and C.A. Maier. 1999. Elevated carbon dioxide in the atmosphere: what might it mean for loblolly pine plantation forestry? J. For. 97(7):4-10.

Harding, R.B. and E.J. Jokela. 1994. Long-term effects of forest fertilization on site organic matter and nutrients. Soil Sci. Soc. Am. J. 58:216-221.

Hassink, J. 1995. Density fractions of soil macroorganic matter and microbial biomass as predictors of C and N mineralization. Soil Biol. Biochem. 27(8):1099-1108.

Hassink, J. and A.P. Whitmore. 1997. A model of the physical protection of organic matter in soils. Soil Sci. Soc. Am. J. 61:131-139.

Hassink, J., A.P. Whitmore, and J. Kubat. 1997. Size and density fractionation of soil organic matter and physical capacity of soils to protect organic matter. Europ. J. Agron. 7:189-199.

Hayes, M.H.B. 1985. Extraction of humic substances from soil. p. 329-361. *In* G.R. Aiken *et al.* (eds.) Humic substances in soil, sediment and water. Geochemistry, isolation and characterization. Wiley, New York.

Henry, C.L. and R.B. Harrison. 1996. Carbon fraction in compost and compost maturity tests. p. 51-67. *In* F.R. Magdoff *et al.* (eds.) Soil organic matter: analysis and interpretation. SSSA Spec. Pub. No. 46.

Homann, P.S., P. Sollins, H.N. Chappell, and A.G. Stangenberger. 1995. Soil organic carbon in a mountainous, forested region: relation to site characteristics. Soil Sci. Soc. Am. J. 59:1468-1475.

Hontoria, C., J.C. Rodriguez-Murillo, and A. Saa. 1999. Relationships between soil organic carbon and site characteristics in peninsular Spain. Soil Sci. Soc. Am. J. 63:614-621.

Houghton, R.A. and Woodwell, G.M. 1989. Global climate change. Scientific American 260:36-44.

Hu, S., D.C. Coleman, C.R. Carroll, P.F. Hendrix, and M.H. Beare. 1997. Labile soil carbon pools in subtropical forest and agricultural ecosystems as influenced by management practices and vegetation types. Agric. Ecosyst. Environ. 65:69-78.

Huntington, T.G. 1995. Carbon sequestration in an aggrading forest ecosystem in the southern USA. Soil Sci. Soc. Am. J. 59:1459-1467.

Huntington, T.G., D.F. Ryan, and S.P. Hamburg. 1988. Estimating soil nitrogen and carbon pool in a northern hardwood forest ecosystem. Soil Sci. Soc. Am. J. 52:1162-1167.

Hutchinson, M.F. 1995. Documentation for ANUDEM version 4.4. Centre for
    Resources and Environmental Studies, Australian National University,
    Canberra.

Janzen, H.H., C.A. Campbell, S.A. Brandt, G.P. Lafond, and L. Townley-Smith.
    1992. Light fraction organic matter in soils from long-term crop rotations. Soil
    Sci. Soc. Am. J. 56:1799-1806.

Jenny, H. 1941. Factors of soil formation, a system of quantitative pedology.
    McGraw-Hill, New York, 281 p.

Jenson, S.K. 1991. Application of hydrologic information automatically extracted
    from digital elevation modeling. p. 35-48. *In* K.J. Bevens and I.D. Moore (eds.)
    Terrain analysis and distributed modeling in hydrology. Wiley and Son,
    Chichester, NY.

Johnsen, K.H., D. Wear, R. Oren, R.O. Teskey, F. Sanchez, R. Will, J. Butnor, D.
    Markewitz, D. Richter, T. Rials, H.L. Allen, J. Seiler, D. Ellsworth, C. Maier, G.
    Katul, and P.M. Dougherty. 2001. Carbon sequestration and southern pine
    forests. J. For. 99(4):14-21.

Johnson, D.W. 1992. Effects of forest management on soil carbon storage.
    Water Air Soil Pollut. 64:83-120.

Johnson, D.W. and P.S. Curtis. 2001. Effects of forest management on soil C
    and N storage: meta analysis. For. Ecol. Manage. 140:227-238.

Johnson, D.W., J.D. Knoepp, W.T. Swank, J. Shan, L.A. Morris, D.H. Van Lear, and P.R. Kapeluck. 2002. Effects of forest management on soil carbon: results of some long-term resampling studies. Environ. Pollut. 116:201-208.

Johnson, M.G. and J.S. Kern. 2003. Quantifying the organic carbon held in forested soils of the United States and Puerto Rico. p. 48-72. *In* Kimble *et al.* (eds.) The potential of U.S. forest soils to sequester carbon and mitigate the greenhouse effect. CRC Press, Boca Raton, FL.

Kern, J.S. 1994. Spatial pattern of soil carbon in the contiguous United States. Soil Sci. Soc. Am. J. 58:439-455.

Kern, J.S. and M.G. Johnson. 1993. Conservation tillage impacts on national soil and atmospheric carbon levels. Soil Sci. Soc. Am. J. 57:200-210.

Knoepp, J.D. and W.T Swank. 1997. Forest management effects on surface soil carbon and nitrogen. Soil Sci. Soc. Am. J. 61:928-935.

Laiho, R., F. Sanchez, A. Tiarks, P.M. Dougherty, C.C. Trettin. 2003. Impacts of intensive forestry on early rotation trends in site carbon pools in the southeastern US. For. Ecol. Manage. 177-189.

Lehmann, J., M. da Silva Crave, and W. Zech. 2001. Organic matter stabilization in a Xanthic Ferralsol of the central Amazon as affected by single trees: chemical characterization of density, aggregates, and particle size fractions. Geoderma 99:147-168.

Magdoff, F. 1996. Soil organic matter fractions and implications for interpreting organic matter tests. p. 11-20. *In* Magdoff *et al.* (eds.) Soil organic matter interpretation. SSSA Special Pub. 46. Soil Sci. Soc. Am., Madison, WI.

Magid, J. A. Gorissen and K.E. Giller. 1996. In search of the elusive "active" fraction of soil organic matter: three size-density fractionation methods for tracing the fate of homogeneously ^{14}C-labeled plant materials. Soil Biol. Biochem. 28(1):89-99.

McBratney, A.B., I.O.A. Odeh, T.F.A. Bishop, M.S. Dunbar, and T.M. Shatar. 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97:293-327.

McKenzie, N.J. and M.P. Austin. 1993. A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation. Geoderma 57:329-355.

McSweeney, K.M., P.E. Gessler, B. Slater, R.D. Hammer, J.C. Bell, and G.W. Petersen. 1994. Towards a new framework for modeling the soil-landscape continuum. *In* Amundson, R.G. (eds.) Factors of soil formation: A fiftieth anniversary retrospective. SSSA Spec. Publ. 33. Soil Sci. Soc. Am., Madison, WI. 127-145.

Meijboom, F.W., J. Hassink, and M. Van Noorwijk. 1995. Density fractionation of soil macroorganic matter using silica suspensions. Soil Biol. Biochem. 27(8):1109-1111.

Moore, I.D., P.E. Gessler, G.A. Nielsen, and G.A. Peterson. 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57:443-452.

Moore, I.D., R.B. Grayson, and A.R. Ladson. 1991. Digital terrain modeling: a review of hydrological, geomorphological and biological applications. Hydrological Processes 5:3-30.

Murray, B.C., S.P. Prisley, R.A. Birdsey, and R.N. Sampson. 2000. Carbon sinks in the Kyoto protocol: potential relevance for US forests. J. For. 98(9):6-11.

Musselman, R.C. and D.G. Fox. 1991. A review of the role of temperate forests in the global CO_2 balance. J. Air Waste Manage. Assoc. 41:798-807.

Nelson, D.W. and L.E. Sommers. 1996. Total carbon, organic carbon, and organic matter. p. 961. *In* Page *et al.* (eds.) Methods of soil analysis. Part 3 Chemical Methods. ASA, CSSA, and SSSA, Madison, WI.

Neter, J., W. Wasserman and M.H. Kutner. 1989. Applied linear regression models. 2nd ed. Richard D. Irwin, Homewood.

O'Connell, D.A., P.J. Ryan, N.J. McKenzie and A.J. Ringrose-Voase. 2000. Quantitative site and soil descriptors to improve the utility of forest soil surveys. For. Ecol. Manage. 138:107-122.

Odeh, I.O.A., A.B. McBratney, and D.J. Chittleborough. 1994. Spatial Prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma 63(3-4):197-214.

Olsson, B.A., H. Staaf, H. Lundkvist, J. Bengtsson, and K. Rosen. 1996. Carbon and nitrogen in coniferous forest soils after clear-felling and harvests of different intensity. For. Ecol. Manage. 82:19-32.

Parker, J.L., I.J. Fernandez, L.E. Rustad, and S.A. Norton. 2002. Soil organic matter fractions in experimental forested watersheds. Water Air Soil Pollut. 138:101-121.

Paul, K.I., P.J. Polglase, J.G. Nyakuengam, and P.K. Khanna. 2002. Change in soil carbon following afforestation. For. Ecol. Manage. 168:241-257.

Paustian, K., E. Levine, W.M. Post, and I.M. Ryzhova. 1997. The use of models to integrate information and understanding of soil C at the regional scale. Geoderma 79:227-260.

Polglase, P.J., E.J. Jokela, and N.B. Comerford. 1992. Phosphorus, nitrogen and carbon fractions in litter and soil of southern pine plantations. Soil Sci. Soc. Am. J. 56:566-572.

Post, W. M., and J. Pastor 1996. LINKAGES -An individual based forest ecosystem model. Climatic Change 34:253–261.

Post. W.M., W.R. Emanuel, P.J. Zinke, and A.G. Stangenberger. 1982. Soil carbon and world life zones. Nature 298:156-159.

Richter, D.D., D. Markewitz, S.E. Trumbore, and C.G. Wells. 1999. Rapid accumulation and turnover of soil carbon in a re-establishing forest. Nature 400:56-58.

Richter, D.D., D. Markewitz, C.G. Wells, H.L. Allen, J.K. Dunscome, K. Harrison, P.R. Heine, A. Stuanes, B. Urrego, and G. Bonani. 1995. Carbon cycling in a loblolly pine forest: Implication for the missing sink and for the concept of soil. *In* W.W. McFee and J.M. Kelly (eds.) Carbon forms and functions in forest soils. Soil Sci. Soc. Am. Madison, WI

Ryan, P.J., N.J. McKenzie, D. O'Connell. A.N. Loughhead, P.M. Leppert, D. Jacquier, and L. Ashton. 2000. Integrating forest soils information across scales: spatial prediction of soil properties under Australian forests. For. Ecol. Manage. 138:139-157.

Ruark, G.A. and S.J. Zarnoch. 1992. Soil carbon, nitrogen, and fine root biomass sampling in a pine stand. Soil Sci. Soc. Am. J. 56:1945-1950.

Sanchez, F.G. 1997. Soil organic matter and soil productivity: searching for the missing link. In R.A. Mickler and S. Fox (eds.) The productivity and sustainability of southern forest ecosystems in a changing environment. Springer-Verlag, New York, NY.

Sanchez, F.G. and R.J. Eaton. 2001. Sequestering carbon and improving soils: benefits of mulching and incorporating forest slash. J. For. 99(1):32-36

Schimel, D.S., D.C. Coleman, and K.A. Horton. 1985. Soil organic matter dynamics in paired rangeland and crop toposequences in North Dakota. Geoderma 36:201-214.

Schimel, D.S. and C.S. Porter. 1995. Process modeling and spatial extrapolation. p. 358-383. *In* P.A. Matson and R.C. Harriss (eds.) Biogenic trace gases:

Measuring emissions from soil and water. Blackwell Sci. Ltd., Cambridge, MA.

Schlesinger, W.H. 1977. Carbon balance in terrestrial detritus. Annu. Rev. Ecol. Syst. 8:51-81.

Skjemstad, J.O., L.J. Janik, and J.A. Taylor. 1998. Non-living soil organic matter: what do we know about it? Aust. J. Exp. Agric. 38:667-680.

Sollins, P. G. Spycher, and C.A. Glassman. 1984. Net nitrogen mineralization from light- and heavy-fraction forest soil organic matter. Soil Biol. Biochem. 16:31-37.

Sollins, P., G. Spycher, and C. Topik. 1983. Processes of soil organic-matter accretion at a mudflow chronosequence, Mt. Shasta, California. Ecology 64(5): 1273-1282.

Spycher, G., P. Sollins, and S. Rose. 1983. Carbon and nitrogen in the light fraction of a forest soil: vertical distribution and seasonal patterns. Soil Sci. 135(2):79-87.

Strickland, T.C. and P. Sollins. 1987. Improved method for separating light- and heavy-fraction organic material from soil. Soil Sci. Soc. Am. J. 51:1390-1393.

Sundquist, E. 1993. The global carbon dioxide budget. Science 259:934-941.

Swift, R.S. 1996. Organic matter characterization. 1996. p. 1011-1069. *In* Page *et al.* (eds.) Methods of soil analysis. Part 3 Chemical Methods. ASA, CSSA, and SSSA, Madison, WI.

Tans, P.P., I.Y. Fung, and T. Takahashi. 1990. Observational constraints on the global atmospheric CO_2 budget. Science 247:1431-1438.

Theodorou, C. 1990. Nitrogen transformation on particle size fractions from a second rotation pine forest soil. Commun. Soil Sci. Plant Anal. 21:407-413.

Thieken, A.H., A. Lücke, B. Diekkrüger, and O. Richter. 1999. Scaling input data by GIS for hydrological modeling. Hydrol. Process. 13:611-630.

Thompson, J.A. and J.C. Bell. 1996. Color index for identifying hydric soil conditions for seasonally saturated mollisols in Minnesota. Soil Sci. Soc. Am. J. 60:1979-1988.

Thompson, J.A., J.C. Bell, and C.A. Butler. 1997. Quantitative soil-landscape modeling for estimating the areal extent of hydromorphic soils. Soil Sci. Soc. Am. J. 61:971-980.

Thompson, J.A., J.C. Bell, and C.A. Butler. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma 100:67-89.

Torn, M.S., S.E. Trumbore, O.A. Chadwick, P.M. Vitousek, and D.M. Hendricks. 1997. Mineral control of soil carbon storage and turnover. Nature 389:170-173.

Trettin, C.C., D.W. Johnson, and D.E. Todd, Jr. 1999. Forest nutrient and carbon pools at Walker Branch Watershed: changes during a 21-year period. Soil Sci. Soc. Am. J. 63:1436-1448.

Turner, D.P., G.J. Koerper, M.E. Harmon, and J.J. Lee. 1995. A carbon budget for forest of the conterminous United States. Ecol. Applic. 5:421-436.

Turner, J., M. Lambert. 2000. Change in organic carbon in forest plantation soils in eastern Australia. For. Ecol. Manage. 133:231-247.

Van Dam, D., E. Veldkamp, and N. Van Breeman. 1997. Soil organic carbon dynamics: variability with depth in forested and deforested soils under pasture in Costa Rica. Biogeochem. 39:343-375.

Van Lear, D.H., P.R. Kapeluck, and M.M. Parker. 1995. Distribution of carbon in a piedmont soil as affected by loblolly pine management. *In* W.W. McFee and J.M. Kelly (eds.) Carbon forms and functions in forest soils. Soil Sci. Soc. Am. Madison, WI

Varallyay, G.Y. 1990. Influence of climate change on soil moisture regime, texture, structure, and erosion. P 39-50 In H.W. Scharpenseel, M. Schomaker, and A. Ayoub (eds.) Soils on a warmer earth. Elsevier, NY.

Webster, R. 1994. The development of pedometrics. Geoderma 62:2-15.

Whalen, J.K., P.J. Bottomley, and D.D. Myrold. 2000. Carbon and nitrogen mineralization from light- and heavy-fraction addition to soil. Soil Biol. Biochem. 32:1345-1352.

White, S.A. and Y. Wang. 2002. Utilizing DEMs derived from LIDAR data to analyze morphological change in the North Carolina coastline. Remote Sense. Environ. 5809:1-9.

Wilson, J.P. and J.C. Gallant. 2000. Digital terrain analysis. p. 1-28. *In* Wilson
and Gallant (eds.) Terrain analysis: principles and applications. John Wiley
and Sons, New York, NY.

Wilson, J.P., D.J. Spangrud, G.A. Nielsen, J.S. Jacobsen, and D.A. Tyler. 1998.
Global positioning system sampling intensity and pattern effects on computed
topographic attributes. Soil Sci. Soc. Am. J. 62:1410-1417.

Wood, C.W., R.J. Mitchell, B.R. Zutter, and C.L. Lin. 1992. Loblolly pine plant
community effects in soil carbon and nitrogen. Soil Sci. 154(5):410-419.

Woolard, J.W. and J.D. Colby. 2002. Spatial characterization, resolution, and
volumetric change of coastal dunes using airborne LIDAR: Cape Hatteras,
North Carolina. Geomorphology 48:269-287.

Xu, Y.J. and S.P. Prisley. 2000. Linking STATSGO and FIA data for spatial
analyses of land carbon densities. The 3rd USDA Forest Service Southern
Forest GIS Conference.

Young, J.L. and G. Spycher. 1979. Water dispersible soil organic-mineral
particles: carbon and nitrogen distribution. Soil Sci. Soc. Am. J. 43:324-328.

Zinke, P.J. and A.G. Stangenberger. 2000. Elemental storage of forest soil from
local to global scales. For. Ecol. Manage. 138:159-165.
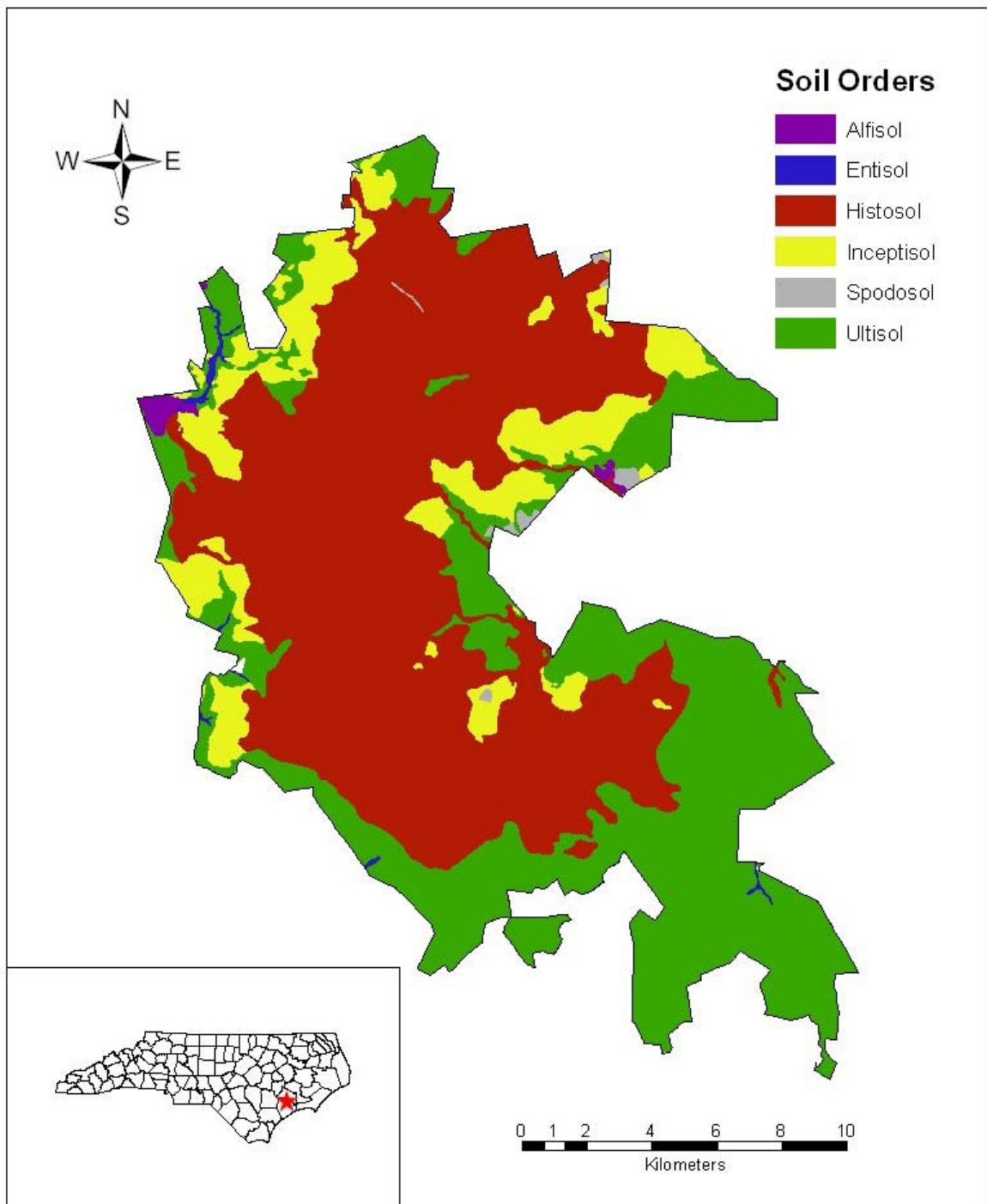
Figure 1. Delineation of soil orders within Hofmann Forest, Jones and Onslow
        Counties, North Carolina.
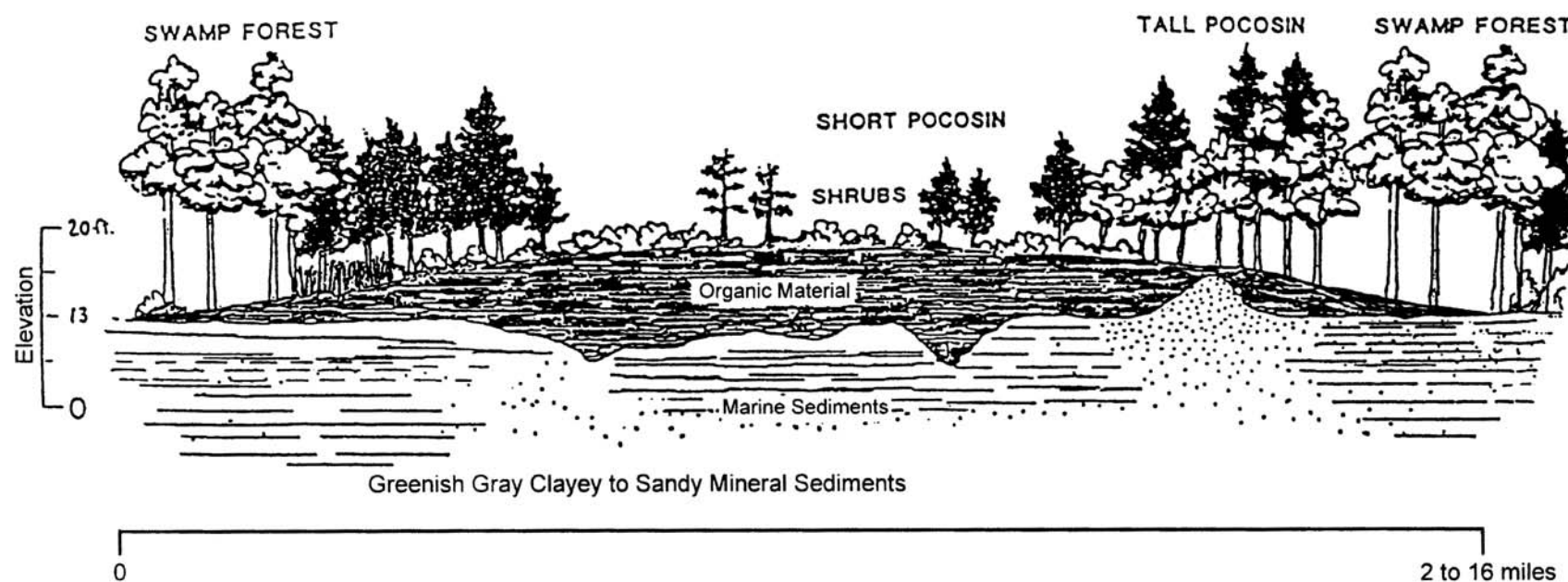
Figure 2. Cross-section of a pocosin illustrating the relationships between pedology, topography and vegetation. From J.A. Gagnon, Jr. Unpublished figure.

1

2

3

4

5

Total LiDAR dataset

6

7

Test data set
50% of original

Validation LiDAR dataset
50%of original

8

9

25% of test data

10

10% of test data

11

5% of test data

12

1% of test data

13
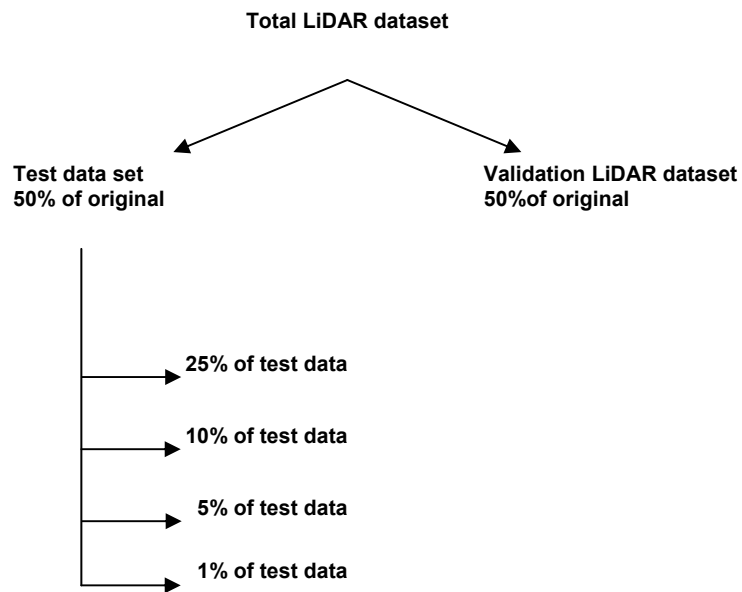
Figure 3. Schematic of LiDAR dataset reduction methodology.

14

15

16

17

18

1 **Table 1**. Potential primary and secondary terrain attributes that may be used for
2   soil landscape modeling.

Attribute	Description	Hydrologic Significance	Source
Elevation (Z), m	Elevation above mean sea level	Proximity to surrounding landscape	Moore *et al.* 1991; Wilson and Gallant, 1996; Florinsky *et al.* 2002
Slope gradient (S), %	Gradient between adjacent grid cells	Velocity of substance flow	Moore *et al.* 1991; Wilson and Gallant, 1996; Florinsky *et al.* 2002
Aspect (ψ), ˚	Slope azimuth	Direction of substance flow	Moore *et al.* 1991; Wilson and Gallant, 1996; Florinsky *et al.* 2002
Specific Catchment Area (A_s), $m^2\ m^{-1}$	Area draining to the catchment outlet, contributing upslope area	Runoff volume	Moore *et al.* 1993; Wilson and Gallant, 1996; Florinsky *et al.* 2002
Flow length (L_f), m	Maximum distance of water flow to a point in catchment	Erosion rates, sediment yields, time of flow	Moore *et al.* 1991
Profile Curvature (K_p), m^{-1}	Slope profile curvature	Relative deceleration and acceleration of substance flows, deposition rates	Moore *et al.* 1991; Wilson and Gallant, 1996; Florinsky *et al.* 2002; Ryan *et al.* 2000
Contour Curvature (K_c), m^{-1}	Contour curvature	Convergence and divergence of substance flows, soil water content	Moore *et al.* 1991; Wilson and Gallant, 1996; Florinsky *et al.* 2002; Ryan *et al.* 2000
Tangential Curvature (K_t), m^{-1}			Wilson and Gallant, 1996
Dispersal Area (D_a), m^2	Area down slope from a short length of contour	Soil drainage rates	Moore *et al.* 1991; Ryan *et al.* 2000
Linear Distance to Stream (L_{ds}), m	Distance from cell location to nearest stream	Soil drainage potential	
Topographic Wetness Index (ω)	Area of flow accumulation, specific catchment area to slope ratio	Extent of soil wetness	Moore *et al.* 1991; Wilson and Gallant, 1996; Florinsky *et al.* 2002; Ryan *et al.* 2000

3

1

CHAPTER 2 LIDAR DATA DENSITY AND LINEAR INTERPOLATOR

INFLUENCE ON ELEVATION PREDICTION ESTIMATES

4

5

6

7

1	**INTRODUCTION**

2	Digital elevation models (DEM) traditionally have been derived from

3	photogrammetric techniques and ground elevation surveys (U.S. Geological

4	Survey (USGS) 1998). Our dependence on DEM for spatial modeling has

5	contributed to a growing need for digital elevation data that provides a better

6	representation of the earth's surface than those derived from small-scale

7	photogrammetric sources (Priestnall *et al.* 2000; Bishop and McBratney 2002;

8	White and Wang 2003). The need for high resolution, high accuracy elevation

9	data for purposes of landscape-scale modeling has resulted in the application of

10	various remote sensing technologies and platforms for digital elevation

11	acquisition including interferometric synthetic aperture radar (IFSAR) (Lloyd and

12	Atkinson 2002; Hodgson *et al.* 2003), light detecting and ranging (LIDAR) (Hill *et*

13	*al.* 2000; Lloyd and Atkinson 2002; White and Wang 2003) and remote sensing

14	platforms (e.g. ASTER [Advanced Spaceborne Thermal Emission and Reflection

15	Radiometer]; IRS-P6 [Indian Remote Sensing satellite]).

16	Increasingly, LIDAR is a source of highly precise, highly accurate

17	elevation data used to produce quality DEM (Lohr 1998; Wehr and Lohr 1999;

18	Lefsky *et al.* 2002). The principles of LIDAR are well documented and readily

19	accepted as a source of elevation data of known accuracy (Huising *et al.* 1998;

20	Lohr 1998; Axelsson 1999; Wehr and Lohr 1999). Although LIDAR allows

21	elevation sampling at exceptionally high densities (>1000 points ha^{-1} in some

1   cases), rasterization of elevation points entails a high degree of uncertainty

2   (Lloyd and Atkinson 2002).

3   Interpolation of irregular spaced LIDAR data sets is needed to develop

4   DEM (Lloyd and Atkinson 2002) to provide better representation of land surfaces

5   and improve spatial modeling. Generally, interpolation is completed using any

6   number of techniques that fall under the broad category of linear interpolators.

7   There are two main groupings of linear interpolation: deterministic and

8   geostatistical. Deterministic interpolation techniques like inverse distance

9   weighted (IDW) create surfaces from measured points (Watson 1992) but do not

10   take into account a model of the spatial processes within the data. Geostatistical

11   interpolation methods such as those imposed by kriging utilize the statistical

12   properties (spatial correlation) of the measured points (Cressie 1993).

13   Comparisons between interpolation methods have been inconclusive, with some

14   research indicating that kriging procedures perform best (Laslett and McBratney,

15   1990; Weber and Englund 1994; Laslett 1994; Phillips *et al.* 1997), while other

16   research indicates splines or IDW interpolations were as good or better (Weber

17   and Englund 1994; Gallichand and Marcotte 1993; Brus *et al.* 1996; Declercq

18   1996). Lloyd and Atkinson (2002) suggested that in many cases a simple IDW

19   approach would suffice but as sample spacing increases, more robust kriging

20   approaches (*viz.* ordinary kriging [OK]) may be required.

21   Inverse distance weighted assumes that each measured point has a local

22   influence that diminishes with distance. Points closer to the prediction location

1   exert greater weights than those farther away. Wood and Fisher (1993) indicated

2   that IDW is best suited for interpolation of sample elevation points that are

3   regularly spaced. Weights in the IDW methodology take the form w(d) = 1 / d^x

4   where w(d) is the weight at distance d that is determined by the inverse of the

5   distance at a user defined power x (Wilson and Gallant 2000).

6   Ordinary kriging relies on the spatial correlation structure of the data to

7   determine weighting values. This is a more rigorous approach to modeling than

8   IDW, as correlation between data points determines estimate values at

9   unsampled points (Cressie 1993). As with IDW, the closest measured values

10   usually have the most influence. However, OK weights are derived from a

11   theoretical semivariogram fitted to empirical data developed by exploring the

12   spatial structure of the data within a given domain (Cressie 1993; Kitanidis 1997).

13   The semivariogram model is based on a statistical approximation of the truth

14   represented by the empirical data. Ordinary kriging takes the form Z(s) = µ(s) +

15   ε(s) where Z(s) is the variable of interest, decomposed into a deterministic trend

16   µ(s) and a random, autocorrelated error in the form of ε(s) (Kitanidis 1997).

17   Advantages of using IDW for interpolation of elevation data include its

18   simplicity of underlying principles, the lack of explicit model parameterization, the

19   speed in calculation, and realistic results produced for many types of data,

20   including elevation (Lloyd and Atkinson 2002). Disadvantages of IDW are that

21   weighting functions may introduce ambiguity, and the interpolation may be

22   affected by non-Gaussian distributions of observational data points given that

1 equal weights are assigned to data points despite spatial clustering. Kriging

2 overcomes many shortcomings of most traditional linear interpolation methods.

3 Kriging is considered the optimal interpolator in the sense that prediction

4 estimates are unbiased and have known minimum variances (Cressie 1993;

5 Kitanidis 1997). Additionally, kriging provides standard error (SE) confidence

6 estimates that help identify unreliable predictions. Kriging weights depend not

7 only on the distances between observational points and estimation locations but

8 on the mutual distances among observational points as well (Cressie 1993).

9 Weaknesses of kriging are that original data points are seldom honored (Kitanidis

10 1997) and it is often difficult to determine whether a particular semivariogram

11 estimate is in fact a true estimator of the spatial correlation within the data set

12 domain (Cressie 1993; Kitanidis 1997).

13   Both methodologies have been shown to be effective in interpolating

14 LIDAR data sets (Lloyd and Atkinson 2002). However, LIDAR data set size is

15 often prohibitively large, resulting in extensive computational requirements for

16 kriging interpolations. While an abundance of data has traditionally been

17 positively viewed, typical LIDAR data sets may contain hundreds or thousands of

18 georeferenced returns ha^{-1} (Lohr 1998) and may impose serious computing

19 constraints. Because of the copious numbers of LIDAR spot elevations returned

20 on an areal basis, even for average field-scaled study areas (~10 ha or >10,000

21 points), the effects of data density reduction on interpolation errors are

22 noteworthy. With a reduction in data, a more manageably and operationally

1   sized elevation data set is possible. Despite substantial data reduction, LIDAR

2   data sets contain a far greater number of elevation reference points than USGS

3   photogrammetrically derived raster DEM.

4   A statistical evaluation is warranted to determine the interpolation

5   technique(s) that maintains landscape morphology yet requires a manageable

6   data set for computer-based operations. The objective of this study was to

7   evaluate interpolation techniques and to analyze the effects of spot elevation

8   data density on the statistical validity of model predictions. No raster DEM were

9   created in this statistical review. Rather, cross-validation and validation with an

10   independent data set were used to assess the prediction errors of IDW and OK

11   linear interpolation techniques. The specific objectives of the study were to (i)

12   perform a statistical analysis of the spatial correlation structure of LIDAR data

13   sets; (ii) determine the effects of data density reduction on prediction estimates

14   through cross-validation and independent validation; and (iii) determine the

15   impacts of IDW and OK linear interpolations on prediction estimates through

16   cross-validation and independent validation.

17   **STUDY AREA**

18   A series of 10 replicate 1000-ha LIDAR data tiles (figure 1) was acquired

19   from the North Carolina Floodplain Mapping Program

20   (http://www.ncfloodmaps.com). Data densities for the LIDAR returns ranged

21   from roughly 100 to 300 points ha^{-1}, with a mean of 183 points ha^{-1}. Tiles used for

22   the study represented LIDAR-acquired digital spot elevation data from forestland

1 in the Lower Coastal Plain (LCP) of eastern North Carolina. Selection criteria

2 focused on maintaining similar landscape morphology and landuse. Elevation

3 ranges varied across the tiles, however a concerted effort was made to select

4 tiles with maximum elevations of 30 m and with similar elevation ranges (figure

5 2). Digital orthophoto quarter quadrangles (DOQQ) were acquired for each tile

6 and were used to verify similar land cover. Each tile was estimated to contain a

7 minimum of 80% forested land cover.

8 Prior to public release of LIDAR data sets, the North Carolina Floodplain

9 Mapping Program processed raw LIDAR data using proprietary-based algorithms

10 to remove artifacts of vegetation, water bodies, and manmade objects. The

11 resulting end products were bare-earth elevation data sets. Processing resulted

12 in a landuse-dependent gradient in the density of bare-earth LIDAR returns, with

13 agriculture lands, shrub lands, and urban lands having much higher point

14 densities than forestlands. All LIDAR data was reviewed and analyzed to assess

15 the quality of the data. The statistics for the combined land cover and the trends

16 for each specific land cover type were reviewed and data that fell outside of the

17 20-25 cm root mean square error (RMSE) criteria were removed from the data

18 sets (http://www.ncfloodmaps.com).

19 **SPATIAL ANALYSIS – SEMIVARIOGRAM**

20 An exploratory analysis of the spatial structure of the LIDAR data set was

21 performed. The global exploratory analysis was performed on the total LIDAR

22 data sets for each of the 10 tiles. The initial step in the exploratory analysis was

1 to determine an appropriate lag interval for producing the semivariograms. Initial

2 lag intervals were established by a criteria based on the minimum number of data

3 points required for model parameterization. The minimum points criteria was

4 arbitrarily set so that the shortest acceptable lag contained a minimum of 300

5 elevation points. Lag intervals were determined with the variogram procedure

6 (proc variogram) executed under the no variogram option of SAS 8.0 (SAS

7 Institute 1999).

8   Large-scale spatial trends in each data set were removed to eliminate any

9 linear trends. A modeled surface was best-fit and used to remove the large-scale

10 trend and establish a set of residuals. Detrending operations were completed

11 using regression procedures (proc reg) of SAS 8.0 (SAS Institute 1999). First,

12 second and third order polynomials were calculated and best-fit using a

13 maximization of R^2 and adjusted-R^2 for each tile. Residual components derived

14 from detrending procedures were used for fitting and development of empirical

15 semivariograms.

16   Weighted nonlinear least squares (WNLS) analysis was used to

17 parameterize variograms to provide best-fit theoretical semivariograms with the

18 empirical data (SAS Institute 2000). Variogram parameters were estimated using

19 the nonlinear least squares procedure (proc nlin) of SAS 8.0 (SAS Institute

20 1999). Under the assumption that the standard deviation of the sum of squares

21 error term is not constant over all sum of squares residual values, WNLS was

22 used to determine variogram parameters that produce the lowest sum of squares

1   error. The WNLS handled regression situations in which the data points were of

2   varying quality but operated on the assumption that the weights were known.

3   Weighted least squares modeled the behavior of the random errors and

4   optimized the weighted fitting criterion to find the parameter estimates that fit

5   each observation to the final parameter estimates.

6   Omni-directional semivariograms were produced for each tile to analyze

7   global trends in the data. Theoretical spherical semivariograms were fitted based

8   on parameter estimates from WNLS procedures. Selection of the best-fit

9   theoretical model was based on comparisons between spherical, exponential,

10   and Gaussian models. Theoretical models were selected by determining which

11   model produced the lowest standard error estimates for sill, nugget, and range

12   estimates. Four-directional semivariograms were produced to evaluate global

13   anisotropic tendencies within each LIDAR data set. Semivariograms along the

14   four directions (0º, 45º, 90º and 135º) were evaluated for similarities in order to

15   assure isotropic variance along directional axes.

16   **DATA REDUCTION**

17   The initial step was to sequentially reduce data density through random

18   selection of a predetermined percentage of the original LIDAR data set. Data set

19   reduction was performed using the Geostatistical Analyst module of ArcGIS 8.1

20   (ESRI 2003). Total LIDAR data sets were initially randomly reduced by 50%.

21   One half of each total data set was used as a training data set and a source of

22   subsequent reduced data sets while the remaining half was used as an

1   independent validation data set for prediction error analysis. The 50% data set

2   selected for subsequent reduction was used to produce a series of data sets

3   representing a selected percentage of the original data sets. Reduction resulted

4   in additional data sets that represented 25%, 10%, 5%, and 1% of the original

5   LIDAR tile (figure 3). The reduction process was repeated for each of the 10

6   LIDAR tiles.

7   **LINEAR INTERPOLATION – IDW AND OK**

8   A statistical evaluation procedure was performed to facilitate a comparison

9   of IDW and OK linear interpolation techniques. Model evaluation, data analysis,

10   model parameterization and elevation prediction and standardized error analysis

11   was performed in the Geostatistical Analyst module of ArcGIS 8.1 (ESRI 2003).

12   Parameters gained from WNLS exploratory spatial analysis were used to

13   populate nugget, range and sill parameters for the OK models. Inverse distance

14   weighted power parameter was set to a power of 2. Both OK and IDW were

15   constrained to isotropic neighborhoods of 30 m and $100 > n > 15$ nearest

16   neighbors were used to predict locational elevations. The search radius was

17   flexible in that it always allowed for a neighborhood of at least 15 elevation

18   points.

19   Prediction and standardized error analysis of both interpolation methods

20   were evaluated by cross-validation and validation using the Geostatistical Analyst

21   module of ArcGIS 8.1 (ESRI 2003). The process was repeated for each density

22   level of each LIDAR tile. Error statistics of cross-validation and independent

1 validation were used to provide a means of evaluating the effects of linear

2 interpolation techniques and data density reduction. Error estimations for all 10

3 tiles were compiled following completion of cross-validation and validation

4 procedures.

5 Cross-validation provided an approach to review potential biases in

6 interpolation methods and data densities. This method, often referred to as the

7 jackknife technique, was used to generate confidence limits on a elevation

8 estimate by iteratively re-estimating the elevation with each of the observations

9 held out; the jackknifed estimates were the mean of these pseudovalue

10 estimates (Crowley 1992). Estimates identified potential bias issues within the

11 data by evaluation of estimation errors. However, cross-validation lacked the

12 ability to determine which interpolator or data density provided the most accurate

13 estimates. Independent validation provided a better approach to evaluate the

14 accuracy of interpolation estimates. Training models were used to predict

15 elevation values for an independent location with a LIDAR estimated elevation.

16 Estimates were then evaluated for accuracy of prediction for the independent

17 location's elevation. This approach allowed for an analysis of the effects of

18 interpolation techniques and data densities on the accuracy of the elevation

19 estimate.

20 Semivariograms for each data density were parameterized with values

21 established by spatial analysis of the original 100% data set. The parameter

22 values assumed a stationary spatial structure across the range of randomly

67

1   reduced data sets. Therefore, semivariograms developed from the original 100%

2   data set were assumed to represent the spatial correlation structure of each of

3   the reduced data sets. The 50% reduced independent validation data set was

4   used for comparison with each training data set density to ensure comparable

5   results across the range of LIDAR densities.

6   For IDW and OK, comparisons were made between mean error (ME),

7   RMSE. For OK only, SE, standardized mean error (SME), and root mean square

8   standardized error (RMSStdE) were used for comparison across the range of

9   data densities. Standardized errors are only generated for OK. This suite of

10   errors was selected for comparison because they are customary error reports

11   generated in the Geostatistical Analyst module of ArcGIS 8.1 (ESRI 2003).

12   **RESULTS AND DISCUSSION**

13   *Global Spatial Exploratory Analysis*

14   Exploratory analysis provided estimates of the spatial structure of the data

15   sets needed for parameterization of interpolation models. Histograms and

16   normality plots indicated that the data were normally distributed across the

17   varying elevation densities. For better understanding of the spatial correlation

18   structure of the data sets and to look for large-scale anisotropic tendencies, four-

19   directional semivariograms were produced with similar lags and maximum lag

20   distances as the omni-directional semivariograms. A visual analysis indicated

21   the data was isotropic. A four-directional semivariogram at a shorter maximum

22   lag distance of approximately 30 m was used to confirm isotropic conditions.

68

1   Fitting procedures were generally successful with theoretical spherical

2   semivariogram models, however some LIDAR tiles had more linearly shaped

3   spatial correlation structures. Exponential and Gaussian models most often

4   yielded semivariograms that fit empirical data poorly or failed to meet

5   convergence criteria. Alternative models were attempted in each case, but

6   parameterization procedures yielded unrealistic nugget, range and sill estimates.

7   Evaluation of nugget-to-sill ratios indicated different levels of spatial

8   dependencies among tiles. Lower spatial correlation within some tile may

9   possibly be accounted for with additional detrending of the data sets. Detrending

10   depended on our *a priori* knowledge of the landscape along with some

11   subjectivity to determine when maximum trend removal was met. Each data set

12   was detrending using third order polynomial regression models. A best-fit

13   semivariogram was produced for each of the ten LIDAR tiles using a theoretical

14   spherical model and the WNLS procedure (figure 4). Inconsistencies in

15   semivariograms between tiles may be an artifact of minor spatial clustering

16   resulting from processing procedures to provide only bare earth returns.

17   ***Cross-Validation and Validation***

18   Comparison of cross-validation errors between OK and IDW indicated that

19   within a given data density, models were unbiased in their predictions. Errors

20   between IDW and OK were remarkably similar, yet IDW had slightly greater ME

21   and tended to have higher RMSE (table 1). The results indicated no obvious

22   advantage in using OK over IDW. However, we must remember that cross-

1 validation only provides information about model bias and not about the accuracy

2 of the predictions.

3   Both cross-validation and validation indicated that errors increased as

4 data density decreased (table 1). This trend was expected because lower

5 density tiles had larger point spacing. The larger distance between neighbors

6 negatively impacted the models ability to estimate an elevation for a given

7 location.

8   Independent validation was used to compare the differences in accuracy

9 of predictions between OK and IDW and examine the effects of data reduction.

10 By calculating SE, SME, and RMSStdE during OK operations, we were able to

11 better address the variability of our predictions. Unfortunately interpolation with

12 IDW does not allow for the calculation of error estimates, however, evaluation of

13 ME and RMSE between OK and IDW generated from independent validation

14 indicated no advantage of using OK over IDW (table 1). Both ME and RMSE

15 followed similar trends for OK and IDW across the density ranges. The

16 maximum difference in RMSE between OK and IDW was less than 2 cm, with

17 IDW having the higher RMSE.

18   Our conclusions of no discernable difference between OK and IDW for

19 interpolation of LIDAR data sets were in opposition to the results of Lloyd and

20 Atkinson (2002). Weber and Englund (1994) found that OK, when performed

21 with variograms estimated from the sample data was more robust than IDW

22 methods. The errors indicated that accuracies of predictions of OK and IDW

1   were very similar within a given data density. Mean errors were consistently less

2   than 1 cm and RMSE were less than 36 cm for all densities for both OK and

3   IDW, and were comparable for interpolators within each data density class.

4   Under the assumption of normally distributed data, SE can be used as a

5   confidence interval around the true value. If the average SE were close to the

6   RMSE, then the model was correctly assessing the variability in prediction (SAS

7   Institute 2000; Kitanidis 1997). If the average SE were greater than the RMSE,

8   then there was an overestimation of the variability of predictions. Conversely, if

9   the average SE were less than the RMSE, an underestimation in the variability of

10   predictions was assumed. The error estimates indicated that OK interpolations

11   were overestimating the variability of elevation predictions. At each density level,

12   RMSE were less than SE (table 1). We were able to confirm that OK was

13   overestimating prediction variability by evaluating the ratio of each prediction

14   error to its estimated prediction SE. When ratios approached 1 the prediction SE

15   was assumed to be valid. If the RMSStdE were greater than 1, variability was

16   underestimated in predictions and if the RMSStdE were less than 1, variability

17   was overestimated (ESRI 2003). In all cases within our data set the RMSStdE

18   were less than 1 (table 1).

19   Data reduction procedures resulted in a range of data densities from less

20   than 2 to more than 180 return points ha^{-1} (table 1). Review of independent

21   validation prediction errors indicated that LIDAR data sets from low-relief forested

22   lands have the ability to withstand data reduction. The density of data used in

1 this study was much greater than was needed for interpolation of a specific

2 location. Following independent validation, data sets reduced to 5% of their

3 original size had RMSE only about 8 cm greater than the 50% data sets. Similar

4 results between OK and IDW support the conclusions that there was no obvious

5 advantage in using OK over IDW.

6   Though there was some decline in RMSE between data sets of different

7 densities, a minor increase in predictions errors across the range of data

8 densities may be acceptable given the improved user-friendliness of the smaller

9 LIDAR data sets. Interestingly, even at the lowest density tested, there were still

10 nearly 2 points ha^{-1}, well above the traditional number of reference points used

11 for photogrammetrically derived elevation models (USGS 1998). If LIDAR data

12 set integrity can withstand a reduction of multiple data points, the computing time

13 for producing DEM could be greatly reduced. While we realize that there would

14 be some resistance to elimination of valid points from elevation data sets, any

15 reduction could be potentially beneficial from a user's standpoint. One of the

16 most vexing issues concerning large LIDAR elevation data sets is the

17 computational time involved for performing even simple statistical exercises.

18   **CONCLUSIONS**

19   Preliminary spatial analysis revealed the possibility of data set reduction

20 without increased prediction errors. Statistical evaluations lead to the conclusion

21 that LIDAR data sets can withstand significant data reductions while maintaining

22 adequate accuracy of predictions of elevations. Errors for OK may have been

1 reduced if variograms were parameterized for each data density rather than

2 assuming the variogram from 100% data set was appropriate for each data

3 density. Despite this, statistical analysis indicated that simple straightforward

4 interpolation approaches such as IDW could be sufficient for interpolating

5 irregular spaced LIDAR data sets. Interpolation techniques such as IDW are not

6 computationally intensive and generally provide for quick interpolation of land

7 surfaces. Reduced data set sizes coupled with a rapid interpolation approach

8 such as IDW could reduce computation times considerably.

9 Users of LIDAR who are seeking to reduce computation times must

10 consider several factors prior to reducing data sets. Is random reduction the best

11 method or should a systematic method of point removal be implemented? Data

12 sets may have different tolerances for reduction and some may simply not allow

13 for any data reduction. Users must have an understanding of acceptable errors

14 for a given function of prediction variables. Those looking for highly accurate

15 micro-scale topographic influences may wish to retain all LIDAR spot elevations

16 for interpolation purposes. However, users looking to simply improve on USGS

17 30-m resolution DEM may be willing to except much greater prediction errors and

18 would benefit from a data reduction strategy.

19 A more robust evaluation of other kriging models and methods (i.e.

20 universal kriging, block kriging and point kriging) is needed to adequately

21 address whether other types of kriging may produce better results than IDW. A

22 complete statistical analysis with procedures to evaluate prediction errors would

1   be required. However, without significant improvements in prediction accuracies,

2   the advantages of IDW outweigh those of the less time efficient kriging

3   interpolators. Additionally, kriging requires the user to have a firm grasp of

4   geostatistics in order to properly parameterize the kriging models.

5   Further analysis is needed to provide replicated analysis of density

6   reduction to elucidate possible strategies for reducing LIDAR data sets sizes on

7   other landuses and landscapes. Given that the density bare earth returns is a

8   function of both landscape morphology and landuse, we must stress that our

9   results may only apply to similarly forested, low-relief landscapes. Similar

10   studies across multiple low, medium, and high relief landscapes under various

11   landuses may elucidate various spatial tendencies not seen in our data set. Our

12   conclusions regarding data reduction also raise the question of scale

13   dependency in interpolation of elevations and the potential for determining an

14   optimum regularly spaced grid resolution for a given density within a LIDAR data

15   set. Future research is needed to determine the effects of data reduction of

16   LIDAR data sets on the creation of raster DEM at multiple resolutions.

17   **REFERENCES**

18   Axelsson, P, 1999, Processing of laser scanning data-algorithms and

19   applications. ISPRS Journal of Photogrammetry and Remote Sensing 54:138-

20   147.

21   Bishop, T.F.A. and A.B. McBratney, 2002, Creating field extent digital elevation

22   models for precision agriculture. Precision Agriculture 3:37-46.

74

1   Brus, D. J., J.J. de Gruijter, B.A. Marsman, R. Visschers, A.K. Bregt and A.

2   Breeuwsma, 1996, The performance of spatial interpolation methods and

3   choropleth maps to estimate properties at points: A soil survey case study:

4   Environmetrics 7(1):1–16.

5   Cressie, N.A., 1993, Statistics for spatial data. John Wiley and Sons, New York,

6   NY.

7   Crowley, P.H., 1992, Resampling methods for computation-intensive data-

8   analysis in ecology and evolution. Annual Review of Ecological Systems. 23:

9   405-447.

10 Declercq, F.A.N. 1996. Interpolation methods for scattered sample data:

11   Accuracy, spatial patterns, processing time. Cartography and Geographic

12   Information Science 23(3):128–144.

13 Environmental Systems Research Institute (ESRI), 2003, ARC/INFO User's

14   Guide. Environmental Systems Research Institute, Redlands, CA, USA.

15 Gallichand, J. and D. Marcotte, 1993, Mapping clay content for subsurface

16   drainage in the Nile delta. Geoderma 58 (3/4):165–179.

17 Hill, J. M., Graham, L. E, Henry, R. J., Cotter, D. M., Ding, A. and Young, D.,

18   2000, Wide-area topographic mapping and applications using airborne light

19   detection and ranging (LIDAR) technology. Photogrammetric Engineering and

20   Remote Sensing 66: 908-927.

21 Hodgson, M.E., J.R. Jensen, L. Schmidt, S. Schill and B. Davis, 2003, An

22   evaluation of LiDAR- and IFSAR-derived digital elevation models in leaf-on

1   conditions with USGS level 1 and level 2 DEMs. Remote Sensing of the

2   Environment 84:295-308.

3   Huising, E.J. and L.M. Gomes Pereira, 1998, Errors and accuracy estimates of

4   laser data acquired by various laser scanning systems for topographic

5   applications. ISPRS Journal of Photogrammetry and Remote Sensing 53:245-

6   261.

7   http://www.ncfloodmaps.com. Last updated February 2004.

8   Kitanidis, P.K., 1997, Introduction to geostatistics: applications in hydrogeology.

9   Cambridge Press, Cambridge, United Kingdom.

10   Laslett, G.M., 1994, Kriging and splines: An empirical comparison of their

11   predictive performance in some applications. Journal of the American

12   Statistical Association 89(426):391–409.

13   Laslett, G.M., and A.B. McBratney, 1990, Further comparison of spatial methods

14   for predicting soil pH. Soil Science Society of America Journal 54(6):1553–

15   1558.

16   Lefsky, M.A., W.B. Cohen, G.G. Parker and D.J. Harding, 2002, LiDAR remote

17   sensing for ecosystem studies. Bioscience 52(1):19-30.

18   Lloyd, C.D. and P.M. Atkinson, 2002, Deriving DSMs from LIDAR data with

19   kriging. International Journal of Remote Sensing 23(12):2519-2524.

20   Lohr, U., 1998, Digital elevation models by laser scanning. Photogrammetric

21   Record 16(91):105-109.

1   Phillips, D.L., E.H. Lee, A.A. Herstrom, W.E. Hogsett. and D.T. Tingey, 1997,

2   Use of auxiliary data for spatial interpolation of ozone exposure in

3   southeastern forests. Environmetrics 8(1):43–61.

4   Priestnall, G., J. Jaafar and A. Duncan, 2000, Extracting urban features from

5   LiDAR digital surface models. Computers, Environment and Urban Systems

6   24:65-78.

7   SAS Institute, 1999, SAS OnlineDoc. Version 8.0. SAS Institute, Cary, NC.

8   SAS Institute, 2000, SAS/STAT user's guide. Version 8.01. SAS Institute, Cary,

9   NC.

10   United States Geological Survey (USGS), 1998, Standards for digital elevation

11   models. National Mapping Program, Reston, VA.

12   Watson, D.F., 1992, Contouring-a guide to analysis and display of spatial data.

13   Pergamon, Oxford.

14   Wehr, A. and U. Lohr, 1999, Airborne laser scanning- an introduction and

15   overview. ISPRS Journal of Photogrammetry. 54(2-3):68-82.

16   Weber, D.D. and E.J. Englund, 1994, Evaluation and comparison of spatial

17   interpolators 2. Mathematical Geology 26(5):589-603.

18   White, S.A. and Y. Wang, 2003, Utilizing DEMs derived from LIDAR data to

19   analyze morphological change in the North Carolina coastline. Remote

20   Sensing of the Environment. 85(1):39-47.

21   Wilson, J.P. and J.C. Gallant, 2000, Terrain analysis: principles and applications.

22   John Wiley and Sons, New York, NY.

1   Wise, S., 1998, The Effect of GIS Interpolation Errors on the Use of Digital

2   Elevation Models in Geomorphology, In: Landform Monitoring, Modelling and

3   Analysis, Edited by S. N. Lane, K. S. Richards and J. H. Chandler, John Wiley

4   and Sons, New York, NY. 300 pp.

5   Wood, J.D. and P.F. Fisher, 1993, Assessing interpolation accuracy in elevation

6   models. IEEE Computer Graphics and Applications 13(2):48-56.

7

8
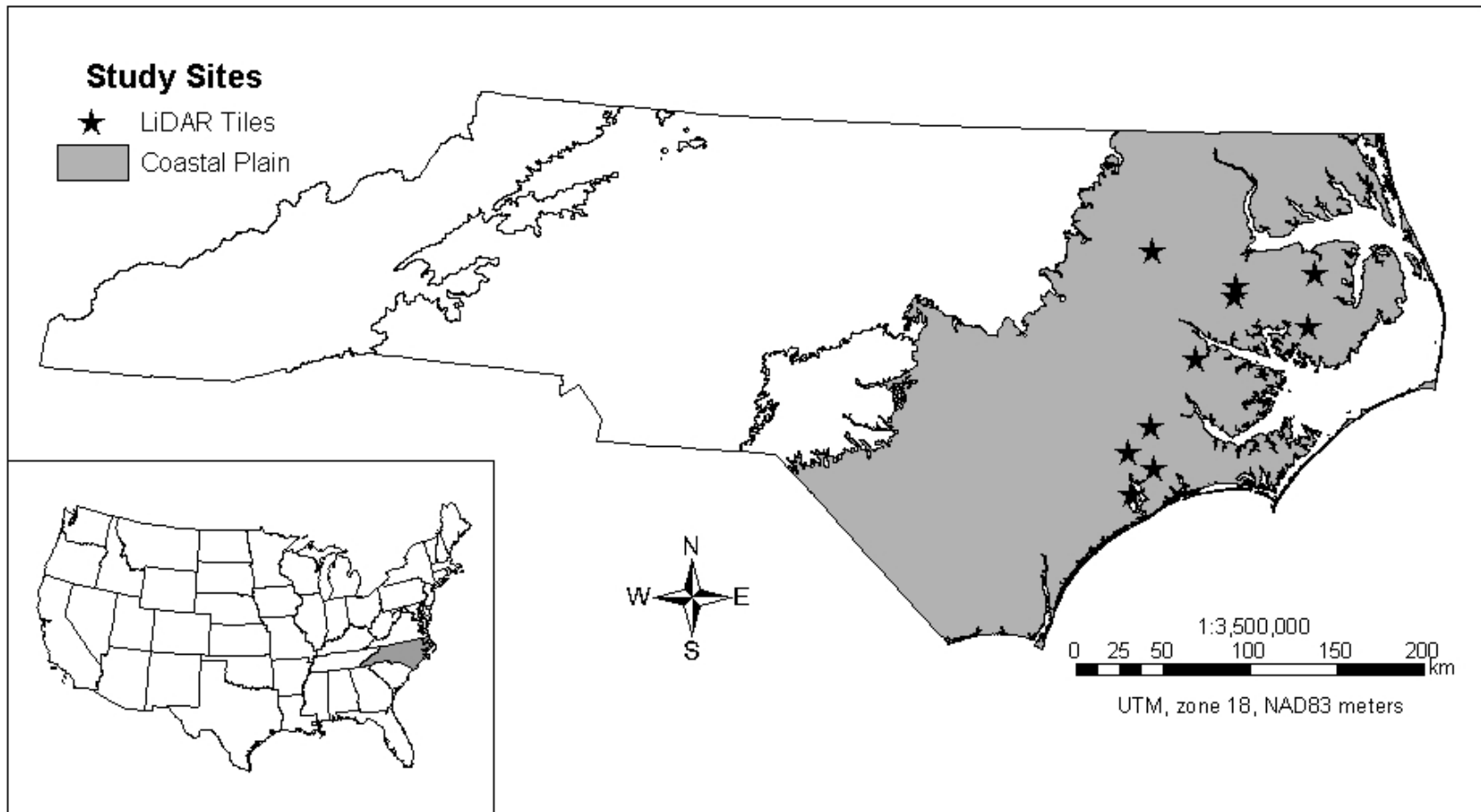
Figure 1. Ten selected LIDAR tiles located on the Lower Coastal Plain of eastern North Carolina, USA. Inset: Location of NC in continental USA.
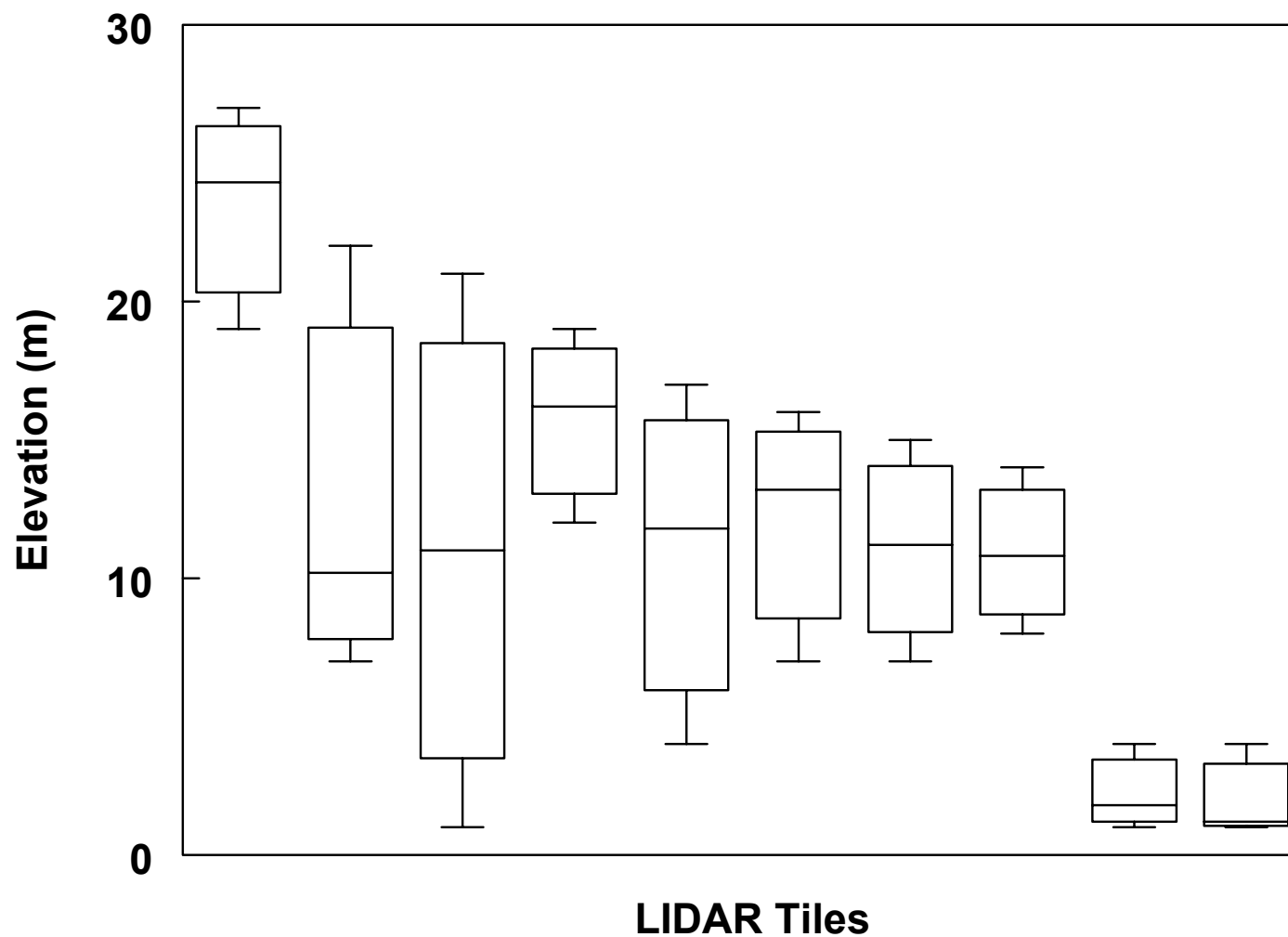
Figure 2. Elevation mean, maximum, and minimum for the 10 forested LIDAR tiles located in the Lower Coastal Plain of eastern North Carolina, USA, arranged from highest to lowest maximum elevation.

a)

100%

50%

25%

10%

5%

1%

N
W—E
S

1:12000

0   250   500   750   1,000
⌐meters
UTM, zone 18, NAD83 meters

b) ┌────LIDAR Tile────┐

50% Training   50% Validation

→ 25% Training
→ 10% Training
→ 5% Training
→ 1% Training
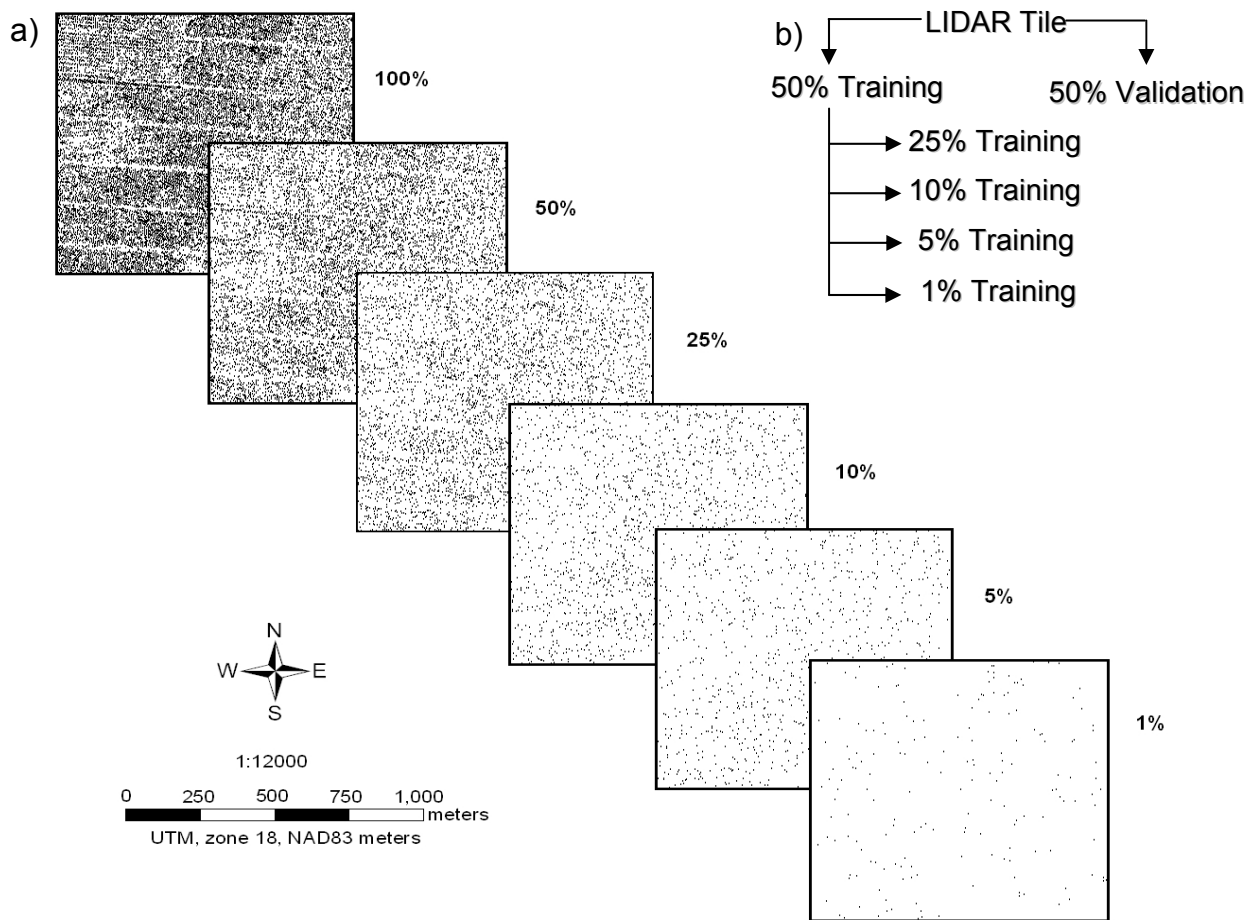
Figure 3. a) Example of spot elevation density gradient following sequential reduction of a LIDAR data sets. Scale was increased to show a sub-region of individual tiles b) Reduction scheme for deriving test and validation LIDAR data sets form the original LIDAR data set.
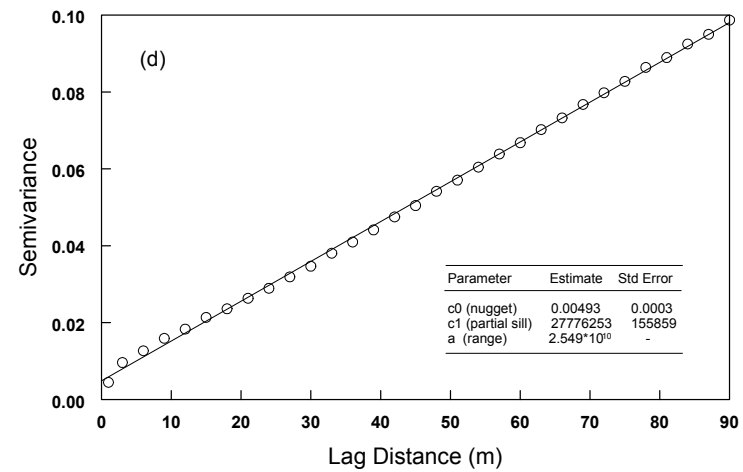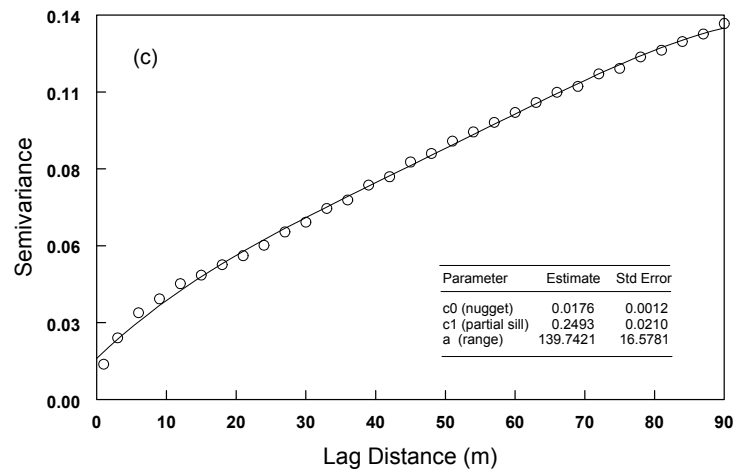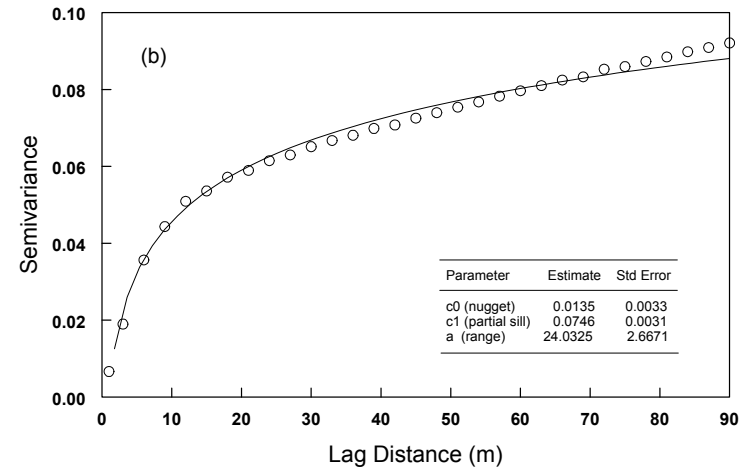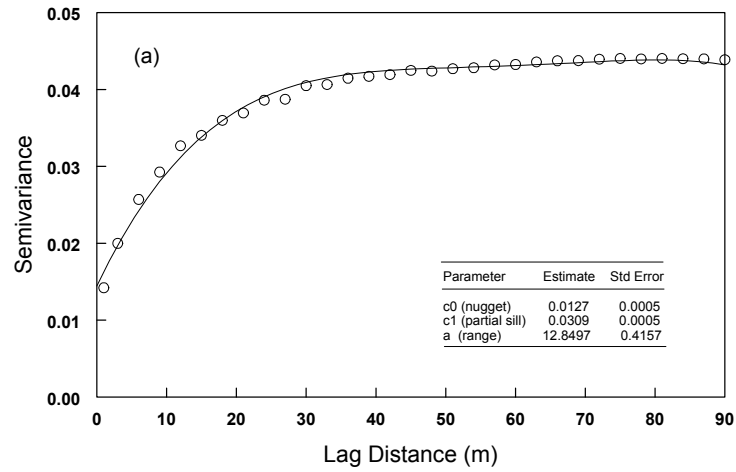
Figure 4. Selected semivariograms from LIDAR tiles fit with theoretical spherical models. Tile a and b illustrate our best fits with spherical models, while tiles c and d had more linear trends fitted with spherical models.

1   **Table 1.** Mean cross-validation and validation prediction errors for 10 forested LIDAR data sets with different densities
2   using IDW^2 and OK interpolation methodologies.

	%	Density (pts ha^{-1})	ME	RMSE	SE m	SME	RMS StdE	%	Density (pts ha^{-1})	ME	RMSE	SE m	SME	RMS StdE
			CROSS VALIDATION							**VALIDATION**				
IDW^2	1	1.80	0.0033	0.3566	-	-	-	1	1.80	-0.0027	0.3576	-	-	-
	5	9.02	0.0032	0.2672	-	-	-	5	9.02	0.0045	0.2648	-	-	-
	10	18.09	0.0024	0.2375	-	-	-	10	18.09	0.0025	0.2374	-	-	-
	25	45.25	0.0018	0.2075	-	-	-	25	45.25	0.0017	0.2080	-	-	-
	50	90.48	0.0016	0.1888	-	-	-	50	90.48	0.0017	0.1891	-	-	-
	100	181.03	0.0015	0.1731	-	-	-	100	181.03	-	-	-	-	-
OK	1	1.80	0.0029	0.3424	0.7564	0.0061	0.8512	1	1.80	-0.0015	0.3397	0.9506	0.0002	0.8575
	5	9.02	0.0017	0.2607	0.4548	0.0045	0.8167	5	9.02	-0.0030	0.2592	0.5957	0.0068	0.7989
	10	18.09	0.0018	0.2556	0.4663	0.0045	0.8150	10	18.09	0.0007	0.2498	0.4682	0.0033	0.8250
	25	45.25	0.0006	0.2000	0.4200	0.0017	0.7893	25	45.25	0.0004	0.2025	0.4160	-0.0006	0.8056
	50	90.48	0.0006	0.1854	0.4567	0.0012	0.7667	50	90.48	0.0010	0.1857	0.4793	0.0016	0.7793
	100	181.03	0.0007	0.1725	0.3713	0.0009	0.7706	100	181.03	-	-	-	-	-

1

CHAPTER 3 HORIZONTAL RESOLUTION AND DATA DENSITY EFFECTS ON

2

REMOTELY SENSED LIDAR-BASED DEM

3

4

INTRODUCTION

1

2   With the emergence of quantitative pedologic measurement and modeling

3 techniques, or pedometrics, in the 1960's (Webster, 1994) soil scientists have sought a

4 more quantitative approach to modeling the spatial distribution of soil properties.

5 Pedometrics resulted in the development of statistical approaches that incorporate

6 surrogate environmental and edaphic explanatory variables and provide estimations of

7 selected soil properties (McBratney *et al.*, 2000). An approach to spatial extrapolation

8 that has found recent use in soil science and geomorphology is soil-landscape

9 modeling. Numerous soil-landscape modeling studies have demonstrated the

10 importance of land surface representation through digital elevation models (DEM) and

11 terrain attributes (McSweeney *et al.*, 1994; Bell *et al.*, 2000; Gessler *et al.*, 2000). Given

12 the role of DEM in spatial modeling, it is important to consider the accuracy of the

13 elevation input data that are used. Data source, vertical precision, and horizontal

14 resolution are crucial in determining accuracy of DEM and derived landscape attributes

15 (Thompson *et al.*, 2001).

16   The horizontal and vertical qualities of a DEM are directly linked to the source of

17 data used for its production. Traditionally, DEM have been derived from

18 photogrammetric techniques. A widely used DEM within the United States (U.S.) has

19 been the 30-m DEM (level 1) produced by the U.S. Geological Survey (USGS). These

20 DEM are produced using stereocorrelation techniques that estimate elevations from

21 relief displacement for areas within the stereomodel. A lattice of elevation points within

22 a pair of stereo models was developed and resampled to a digital 30-m grid to create

1   the DEM. At best, the USGS photogrammetrically derived 30-m level 1 DEM have a

2   vertical precision of 1 m with an accuracy of ≤15 m. Other USGS DEM such as 10-m

3   level 1 DEM have improved resolution and precision, but these DEM are not be readily

4   available for all areas of the U.S.

5   More recently, light detecting and ranging (LIDAR) has been used as a source of

6   elevation data for producing quality DEM (Lohr, 1998; Wehr and Lohr ,1999; Lefsky *et*

7   *al.* 2002). Typical LIDAR data sets may contain hundreds to thousands of

8   georeferenced returns per ha (Lohr, 1998). Federal Emergency Management Agency

9   (FEMA) reported LIDAR accuracies of less than 1 m for both low and medium relief

10   terrain (FEMA, 2003). Application of LIDAR elevation data to produce raster DEM

11   requires interpolation of spot elevation LIDAR points (Lloyd and Atkinson, 2002).

12   Raster DEM created from bare earth LIDAR points can be interpolated to a 0.1m

13   vertical precision at a horizontal resolution most appropriate for its intended use without

14   generalization or resampling techniques.

15   Landscape attribute prediction exhibits a direct dependence on the qualities of

16   the DEM used for surface representation and attribute derivation (Jenson, 1991;

17   Thieken *et al.*, 1999). Thieken *et al.* (1999) and Thompson *et al.* (2001) indicated that

18   DEM resolution contributes to differences in the distribution and representation of

19   landscape attributes. However, Gessler *et al.* (2000) found little difference among

20   landscape models based on a landscape attribute derived from a series of DEM with 2-

21   to 10-m resolutions. Likewise, Chaplot *et al.* (2000) found that 10-m to 30-m DEM

22   generally provided an unbiased prediction of landscape terrain but prediction was

1   influenced as DEM resolution decreased to 50 m. Florinsky and Kuryakova (2000)

2   emphasized that the importance of DEM resolution was highly dependent on the scale

3   of the process modeled, concluding that high resolution (between 2.25- and 3.25-m)

4   DEM were important for modeling processes at the microscale. In general, these

5   studies indicated that attribute value ranges increased and predictive capabilities

6   decreased as DEM grid size increased. The amount of relief on a landscape

7   contributes to the effects of DEM resolution on terrain attributes, with low relief

8   landscapes being less sensitive to resolution impacts. Consequently, large study areas

9   that may incorporate a larger range of relief may require higher resolution DEM.

10   Production of different horizontal resolution DEM with the same vertical precision

11   from the same data source is important for predicting scale dependent environmental

12   variables. The use of LIDAR elevation data sets offers the flexibility needed to produce

13   multiple horizontal resolutions of DEM from the same data source. However, LIDAR

14   data set size is often prohibitively large, resulting in extensive computational

15   requirements for producing DEM. Because of the copious number of LIDAR spot

16   elevations returned on an areal basis, the effects of data density reduction on DEM of

17   various horizontal resolutions is worthy of study, particularly for landscape-scale

18   studies. With a reduction in data, a more manageably and operationally sized elevation

19   data set is possible. Despite substantial data reduction, the resultant LIDAR data sets

20   contain a far greater number of elevation reference points than USGS

21   photogrammetrically derived raster DEM.

1   Depending on the DEM resolution of interest, LIDAR data set spot elevations

2   may be excessively dense. Low resolution (30-m) DEM may require fewer points for

3   producing accurate raster centroid values than medium (10-m) and high resolution (≤10-

4   m) DEM. We hypothesized that given the high density of points within a LIDAR data

5   set, it would be possible to substantially reduce LIDAR data density yet maintain the

6   accuracies of the DEM end products. Furthermore, we hypothesized that resolution of

7   the DEM would play a critical role in determining the level of data reduction that was

8   feasible. The objective of this study was to evaluate the effects of LIDAR data density

9   on the production of DEM at different resolutions. The impetus for reducing LIDAR data

10 sets was to reduce time restraints associated with producing DEM from large LIDAR

11 data sets. The specific objects of this study were to (i) produce a series of DEM at

12 different horizontal resolutions along a LIDAR point density gradient; (ii) compare each

13 DEM produced with different LIDAR data density at a given horizontal resolution to a

14 baseline DEM produced from the highest available LIDAR data density; and (iii)

15 determine the optimum LIDAR point density suitable for producing a DEM at a given

16 horizontal resolution.

17   **MATERIAL AND METHODS**

18 ***Study site***

19   A series of 61 LIDAR tiles (1000 ha) were collected from the North Carolina

20 Flood Mapping Program (NCFMP) to cover the spatial extent of the Hofmann Forest.

21 The Hofmann Forest consists of a 32,500 ha forest ecosystem located in Jones and

22 Onslow Counties of eastern North Carolina, USA (Figure 1) on the low-relief landscape

1   of the Lower Coastal Plain. Greater than 90% of Hofmann Forest is covered in pine

2   plantation and natural pocosin vegetation, with minor inclusions of agricultural lands.

3   Elevations range from 12 to 20 m above mean sea level (Daniels *et al.*, 1977), with the

4   landscape characterized by broad, flat interfluves. In some areas, relief may be as low

5   as 1.5 m elevation difference in 3 to 4 km (Daniels *et al.*, 1999).

6   *Data acquisition and reduction*

7   Prior to public release of LIDAR data sets, the NCFMP processed raw LIDAR

8   data using proprietary algorithms to remove artifacts of vegetation, water bodies, and

9   manmade objects. The resulting end products were irregularly spaced bare earth

10   elevation data sets. Processing by NCFMP resulted in a land cover dependent gradient

11   in the density of bare earth LIDAR returns, with agriculture lands having much higher

12   point densities than forestlands. The statistics for the combined land cover and the

13   trends for each specific land cover type were reviewed, and data that fell outside of a

14   20-25 cm root mean square error (RMSE) criteria were removed from the data sets

15   (http://www.ncfloodmaps.com). Elevations were reported to have 0.1 m vertical

16   precision (http://www.ncfloodmaps.com).

17   *LIDAR Data Reduction*

18   All 61 LIDAR tiles were merged to form a single spot elevation coverage

19   consisting of over 9,000,000 LIDAR points. The initial step was to sequentially reduce

20   data density through random selection of a predetermined percentage of the original

21   LIDAR data set. Data set reduction was performed using the Geostatistical Analyst

22   module of ArcGIS 8.1 (Environmental Systems Research Institute (ESRI) 2003).

1   Reduction resulted in data sets that represented 50%, 25%, 10%, 5%, and 1% of the

2   original LIDAR tile. The reduction process was iterated three times to produce multiple

3   random data sets for each reduction percentage. Triplicate data sets were produced to

4   reduce any bias that may have occurred from using only one random data set for each

5   reduction level.

1 *DEM production*

2   Following completion of replicated data reduction, each density level was used to

3 create a series of raster DEM. Data sets were interpolated to raster DEM using

4 ANUDEM vers. 4.6.3 (Hutchinson, 1995). ANUDEM employs an iterative finite

5 difference interpolation technique that allows for maintenance of a drainage network

6 consistent with the original data and removal of spurious sinks (Hutchinson, 1995).

7 Each density level was used to produce a series of raster DEM at multiple resolutions.

8 Raster DEM of 5 m, 10 m, and 30 m horizontal resolution with 0.1 m vertical precision

9 were created in triplicate for each density level. In addition, a single raster DEM using

10 100% of the LIDAR data was produced for each horizontal resolution. The 5 m, 10 m,

11 and 30 m horizontal resolutions were selected because they are typical resolutions

12 found in a number of other sources of publicly available raster DEM.

13   Within a given horizontal resolution, ARC/INFO vers. 8.3 (ESRI, 2003) was used

14 to create a mean raster DEM for each DEM triplicate produced within a given density

15 level. This process resulted in the creation of mean raster DEM for each LIDAR data

16 density level for 5 m, 10 m, and 30 m horizontal resolutions (Table 1). The mean raster

17 DEM were used to compare against raster DEM created from the total LIDAR data set.

18 Each raster DEM was clipped to the spatial bounds of Hofmann Forest using ARC/INFO

19 (ESRI, 2003).

20 *Statistical Analysis*

21   Coregistered DEM for each resolution allowed for a cell-by-cell comparison

22 across the density sequence. A subsampling routine in ArcGIS Geostatistical Analyst

1   was used to randomly select 10,000 raster cell values to facilitate comparison of mean

2   DEM with the DEM created from the original LIDAR data set. The subsampling routine

3   was implemented to reduced processing time. Each set of 10,000 coregistered

4   subsamples elevation values, with each having a vertical precision 0.1, were exported

5   for statistical analysis.

6   Comparisons were designed to evaluate the effects of LIDAR data density at

7   each of the horizontal resolutions of produced DEM. Analysis was devised to isolate

8   the effects of LIDAR data density on elevation values of produced DEM within a given

9   horizontal resolution. No direct comparisons were made between DEM of different

10 horizontal resolutions. In order to make these comparisons, we assumed that DEM

11 created from the complete LIDAR data set were the "best" DEM. Comparisons of

12 elevation values were always made between the DEM created using the total original

13 data set and the mean DEM created at each selected level of data reduction. The

14 various DEM were compared using paired t-test to determine if the true mean of their

15 differences were equal to zero.

16   Mean differences between the two competing DEM were regressed against

17 mean point densities within individual cells for each DEM resolution (Table 1). Mean

18 point density per cell refers to the number of LIDAR points that would fall within a cell.

19 Specifically, points per grid cell is the mean number of LIDAR points that would be

20 interior to any given grid cell at a set resolution and density if the DEM were overlain

21 with the LIDAR data set. This comparison was made to help elucidate the impacts of

22 LIDAR point density to the calculated centroid value across multiple resolutions.

1   **RESULTS AND DISCUSSION**

2   ***Data density reduction and DEM***

3   The effects of reduced LIDAR data density on resultant DEM differed among

4   horizontal DEM resolutions (Table 2). The 30-m DEM produced using 50%, 25% and

5   10% of the original LIDAR data set were not significantly different from the baseline

6   DEM30$_{100\%}$ DEM (Table 2). However, p-values indicated that the DEM30$_{1\%}$ and

7   DEM30$_{5\%}$compared to DEM30$_{100\%}$ were significantly different (Table 2). Mean

8   differences between 10-m DEM produced using reduced and complete LIDAR data sets

9   were similar to those seen in the 30-m DEM suite. However, only DEM10$_{50\%}$ and

10   DEM10$_{25\%}$ were not statistically different from DEM10$_{100\%}$ (Table 2).   The reduced

11   DEM had 3 to 4 cm higher elevations than those predicted for the original DEM. Similar

12   results were seen when comparing the 5-m DEM. However, in the case of the 5-m

13   DEM, only DEM5$_{50\%}$ was not significantly different from the DEM5$_{100\%}$ (Table 2). The

14   predicted elevations in the reduced DEM were 3 to 7 cm greater than the original DEM.

15   We attribute the differences in the effects of LIDAR data reduction across

16   horizontal resolutions to the number of points used to interpolate each centroid.   As a

17   function of LIDAR data reduction and horizontal resolution, each DEM had a different

18   number of LIDAR points per grid cell (Table 1). The 30-m DEM consistently had a

19   greater number of points per grid cell for interpolating grid cell centroids than other DEM

20   resolutions.

21   ***Mean differences between DEM***

22   Compared to DEM produced from the complete LIDAR data set, mean

23   differences between competing DEM indicated that all reduced density DEM

93

1   overestimated elevations for all horizontal resolutions (Table 2). Within a given

2   horizontal resolution, overestimation increased as data reduction increased. This

3   overestimation is attributed to the reduced number of LIDAR points used to interpolate

4   each grid centroids, and a possible systematic error that may be a result of the

5   ANUDEM algorithm used for interpolation.

6   Given that low resolution DEM have more LIDAR points for interpolation of

7   values on a cell-by-cell basis than high resolution DEM, the lowest mean differences

8   were expected in the 30-m DEM. Specifically, at each level of data density reduction,

9   30-m DEM had a substantially greater number of LIDAR points per cell than did the 10-

10   m and 5-m DEM (Table 1). As expected, the 5-m DEM had the widest range of mean

11   differences, with the maximum difference occurring at the lowest possible density. We

12   anticipated that the 30-m DEM suite would have the narrowest range of mean

13   differences, yet the narrowest range was found in the 10-m DEM suite. No plausible

14   explanation as to the root cause of this can be offered other than that the 30-m DEM

15   may not have accounted for as much variability as the 10-m DEM did in the LIDAR

16   points used to interpolate cell centroid values.

17   The general trend indicated that as mean points per grid cell increased, the mean

18   difference decreased between DEM produced from the original LIDAR data density and

19   DEM produced from each density level (Figure 3). Within a given horizontal resolution,

20   the greater the LIDAR density the greater the similarity to coincident grid cell values of

21   the original DEM. The rate of change in mean differences across the density sequence

22   was similar for each horizontal resolution. The range of LIDAR points interior to each

1   grid cell narrows as DEM resolution decreases. We attribute these trends to large

2   difference in the number of LIDAR points per grid cell between the different horizontal

3   resolutions at each reduction increment. As DEM resolution increases, each grid cell

4   represents more of the micro-scale relief of the landscape. The difference in point

5   density ranges (on a per cell basis) between the different resolutions indicate the need

6   for high-density LIDAR data sets when attempting to model micro-scale elevation

7   differences.

8   ***Data density requirements and DEM resolution***

9   To better understand the required LIDAR point density for producing 5-m, 10-m

10 and 30-m DEM, a comparison of density was made across multiple resolutions. To

11 facilitate this, LIDAR points per cell in the 5-m, 10-m, and 30-m DEM at each density

12 were normalized on an equal area (points ha^{-1}) basis. To determine the upper and

13 lower bounds of required LIDAR density attention was focused on the densities for each

14 resolution where subsequent LIDAR data reduction resulted in significantly different

15 DEM. Specifically, analysis centralized on the point density differences between the

16 10% and 5% for 30-m DEM, 25% and 10% for 10-m DEM, and 50% and 25% for 5-m

17 DEM. These regions were the threshold values of points ha^{-1} required to produce the

18 DEM of a given horizontal resolution that was not significantly different from the DEM

19 produced from the complete LIDAR data set.

20   Plotting the upper and lower bounds of the threshold region provided some

21 indication as to the trends of LIDAR data density across multiple resolutions. From this,

22 an estimation of the minimum density of LIDAR points required for producing DEM at

1 any given resolution between 5 m and 30 m was estimated (Figure 2). Caution should

2 be taken in putting too much confidence in the upper bounds of the threshold values of

3 unevaluated horizontal resolutions. Our use of irregular intervals of density reduction

4 could cause the upper bounds to misrepresent the actual points density needed for

5 producing DEM of a given horizontal resolution. Greater confidence may be put in

6 estimates of the lower bounds of the threshold values because of finer incremental data

7 reduction. For our study area, we are confident that levels of density that are indicated

8 to be insufficient will produce inferior DEM.

9 The interpolation was done by a linear fit between adjacent points. In reality the

10 density function may follow a different trend between DEM of different horizontal

11 resolution. This may lead to some errors in interpolating between the DEM to determine

12 minimum required LIDAR data density for producing a DEM of a given horizontal

13 resolution other than those evaluated in this study. However, we believe the general

14 shape of the threshold data density region is representative of the trends expected in

15 changes in point densities along a horizontal resolution gradient. Further study of finer

16 scaled density reduction and additional horizontal resolutions may help to further refine

17 the shape of the threshold region.

18 **CONCLUSIONS**

19 Statistical comparisons indicated that DEM horizontal resolution influenced the

20 level of reduction that LIDAR data sets could withstand. Differences indicated that

21 when producing 30-m DEM, LIDAR data sets could be reduced to 10% of their original

22 data density without statistically altering the produced DEM. However, differences in

1   10-m DEM indicated that data reduction was only feasible to 25% of the original data

2   set. Reduction below these levels resulted in DEM that were statistically different from

3   DEM produced using the total LIDAR data set. Data reduction was more restrictive on

4   the 5-m DEM and indicated that LIDAR data sets could only be reduced to 50% of their

5   original density without producing statistically different DEM.

6   Our conclusions regarding data reduction also raise the question of scale

7   dependency in interpolation of elevations and the potential for determining an optimum

8   grid resolution for a given LIDAR data set density. Evaluation of the effects of data

9   reduction between 30-m, 10-m, and 5-m horizontal resolutions provided some indication

10   as to the minimum required LIDAR data density to produce a DEM of a given horizontal

11   resolution. However, evaluation of additional horizontal resolutions is required to

12   provide a clearer understanding of the effect of LIDAR data density. Our analysis

13   indicated that the rate of change in mean differences varied by horizontal resolution.

14   Increases in mean differences as a function of data density will be faster when

15   producing high resolution DEM as compared to low resolution DEM. This trend

16   indicates that smaller increments of reduction may be needed to elucidate the true

17   threshold LIDAR data density of high resolution DEM.

18   Users of LIDAR who are seeking to reduce computation times through data

19   reduction must consider several factors prior to reducing data sets. Data sets may have

20   different tolerances for reduction and some may simply allow for minimal data reduction.

21   Users must have an understanding of acceptable errors for a given function of

22   prediction variables. Those looking for highly accurate micro-scale topographic

1   influences will need to retain a higher density of LIDAR spot elevations for interpolation

2   purposes. However, users looking to simply improve on USGS 30-m DEM may be able

3   to substantially reduce LIDAR data sets and yet maintain DEM accuracy.

4   Given that the density of LIDAR elevations points is a function of both

5   landscape morphology and landuse, we must stress that our results may only apply to

6   similarly forested, low-relief landscapes. Similar studies across multiple low, medium,

7   and high relief landscapes under various landuses may elucidate various spatial

8   tendencies not seen in our data set. Further analysis is needed to provide replicated

9   analysis of density reduction to clarify possible strategies for reducing LIDAR data sets

10   sizes on other landuses and landscapes.

11   **REFERENCES**

12   Bell, J.C., D.F. Grigal, and P. Bates. 2000. A soil-terrain model for estimating spatial

13   patterns of soil organic carbon. P. 295-310. In Wilson and Gallant (eds.) Terrain

14   analysis: principles and applications. John Wiley and Sons, New York, NY.

15   Chaplot, V., C. Walter, and P. Curmi. 2000. Improving soil hydromorphy prediction

16   according to DEM resolution and available pedogical data. Geoderma 97:405-422.

17   Daniels, R.B., E.E. Gamble, W.H. Wheeler, and C.S. Holzhey. 1977. The stratigraphy

18   and geomorphology of the Hofmann Forest Pocosin, North Carolina. Soil Sci. Soc.

19   Am. J. 41:1175-1180.

20   Daniels, R.B., S.W. Buol, H.J. Kleiss, and C.A. Ditzler. 1999. Soil systems of North

21   Carolina. Tech. Bull. 314. North Carolina State Univ., Raleigh.

1   Environmental Systems Research Institute (ESRI), 2003, ARC/INFO User's Guide.

2   Environmental Systems Research Institute, Redlands, CA, USA.

3   FEMA (Federal Emergency Management Agency), 2003, FEMA's Flood Hazard

4   Mapping Program: Guidelines and specifications for flood hazard mapping partners.

5   www.FEMA.gov/fhm/gs_main.shtm.

6   Florinsky, I.V. and G.A. Kuryakova. 2000. Determination of grid size for digital terrain

7   modeling in landscape investigations – exemplified by soil moisture distribution at a

8   micro-scale. Int. J. of GIS. 14(8):815-832.

9   Gessler, P.E., O.A. Chadwick, F. Chamran, L. Althouse, and K. Holmes. 2000. Modeling

10   soil-landscape and ecosystem properties using terrain attributes. Soil Sci. Soc. Am.

11   J. 64:2046-2056.

12   http://www.ncfloodmaps.com. Last updated February 2004.

13   Hutchinson, M.F. 1995. Documentation for ANUDEM version 4.4. Centre for Resources

14   and Environmental Studies, Australian National University, Canberra.

15   Jenson, S.K. 1991. Application of hydrologic information automatically extracted from

16   digital elevation modeling. p. 35-48. *In* K.J. Bevens and I.D. Moore (eds.) Terrain

17   analysis and distributed modeling in hydrology. Wiley and Son, Chichester, NY.

18   Lefsky, M.A., W.B. Cohen, G.G. Parker and D.J. Harding. 2002. LiDAR remote sensing

19   for ecosystem studies. Bioscience 52(1):19-30.

20   Lloyd, C.D. and P.M. Atkinson. 2002. Deriving DSMs from LiDAR data with kriging. Int.

21   J. Remote Sensing 23(12):2519-2524.

1   Lohr, U. 1998. Digital elevation models by laser scanning. Photogramm. Rec.

2   16(91):105-109.

3   McBratney, A.B., I.O.A. Odeh, T.F.A. Bishop, M.S. Dunbar, and T.M. Shatar. 2000. An

4   overview of pedometric techniques for use in soil survey. Geoderma 97:293-327.

5   McSweeney, K.M., P.E. Gessler, B. Slater, R.D. Hammer, J.C. Bell, and G.W. Petersen.

6   1994. Towards a new framework for modeling the soil-landscape continuum. *In*

7   Amundson, R.G. (eds.) Factors of soil formation: A fiftieth anniversary retrospective.

8   SSSA Spec. Publ. 33. Soil Sci. Soc. Am., Madison, WI. 127-145.

9   Thieken, A.H., A. Lücke, B. Diekkrüger, and O. Richter. 1999. Scaling input data by GIS

10   for hydrological modeling. Hydrol. Process. 13:611-630.

11   Thompson, J.A., J.C. Bell, and C.A. Butler. 2001. Digital elevation model resolution:

12   effects on terrain attribute calculation and quantitative soil-landscape modeling.

13   Geoderma 100:67-89.

14   United States Geological Survey (USGS), 1998, Standards for digital elevation models.

15   National Mapping Program, Reston, VA.

16   Webster, R. 1994. The development of pedometrics. Geoderma 62:2-15.

17   Wehr, A. and U. Lohr. 1999. Airborne laser scanning- an introduction and overview.

18   ISPRS J. Photogramm. Remote Sens. 54(2-3):68-82.

Figure 1. Ten selected LIDAR tiles located on the Lower Coastal Plain of eastern North Carolina, USA. Inset: Location of NC in continental USA.
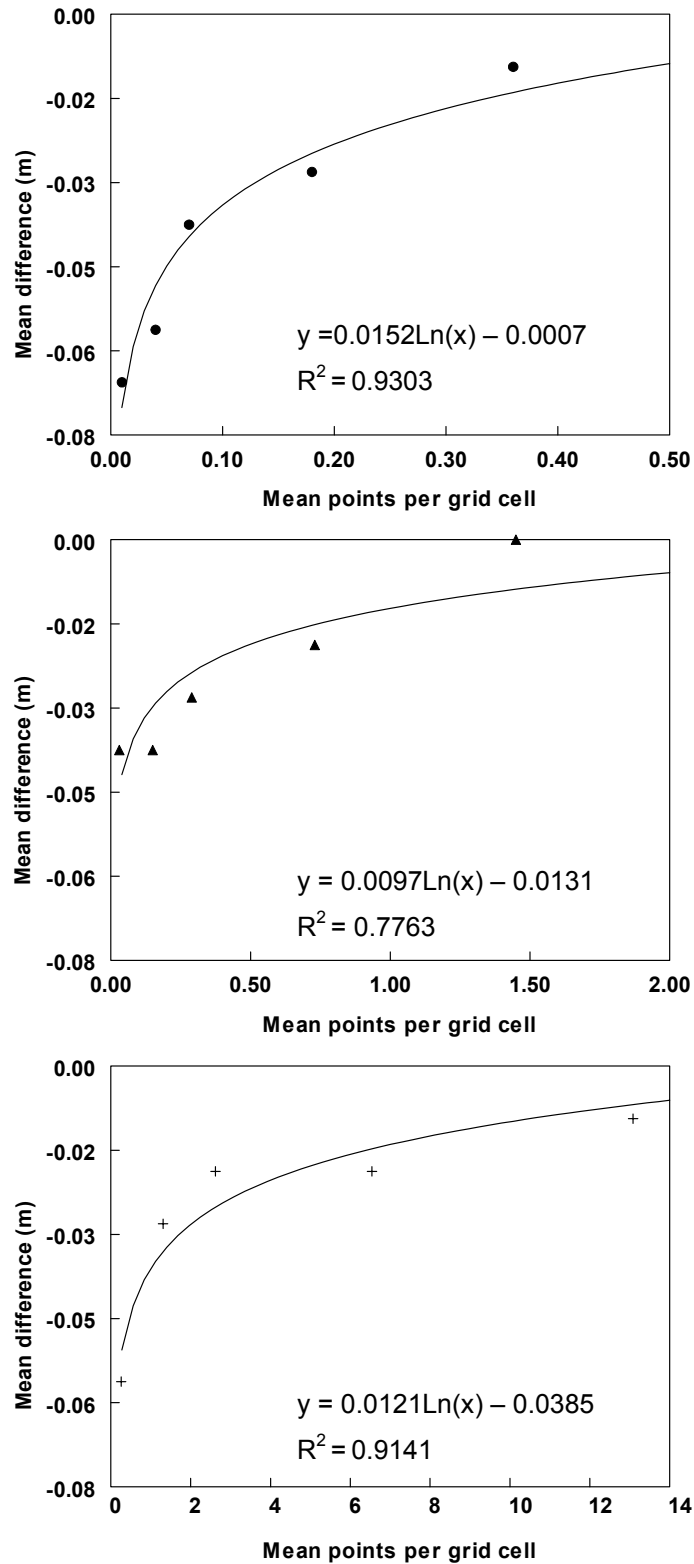
Figure 2.   The effects of mean LIDAR point density per grid cell on the mean difference between coincident points of competing DEM subsamples for (a) 30-m DEM, (b) 10-m DEM, and (c) 5-DEM.
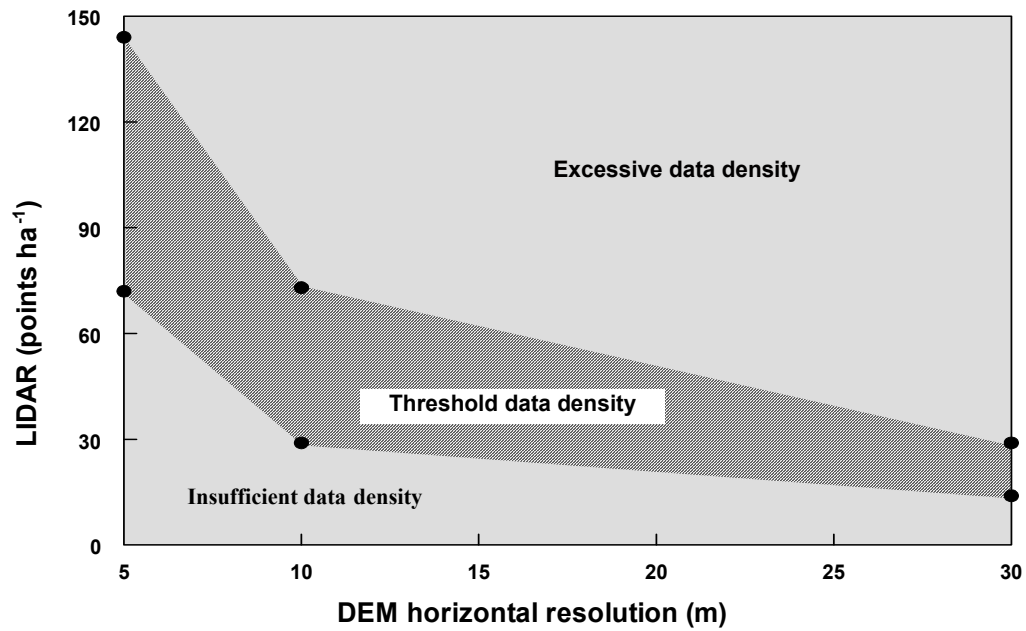
Figure 3. LIDAR points per cell in the 5-m, 10-m, and 30-m DEM for each data density level standardized on an equal area (points ha^{-1}) basis. The threshold region represents the densities for each resolution at which subsequent LIDAR data reduction resulted in significantly different DEM.

Table 1. Designation of 30 m, 10 m, and 5 m DEM produced from each data density level. Designation of mean indicates that DEM elevations were the mean elevation of three DEM produced from different randomly reduced LIDAR data at each data reduction level. Points per grid cell is the mean number of LIDAR points that would be interior to any given grid cell at a set resolution and density.

% original data set	30 m		10 m		5 m	
	Designation	Points per grid cell	Designation	Points per grid cell	Designation	Points per grid cell
100	$DEM30_{100\%}$	26.16	$DEM10_{100\%}$	2.90	$DEM5_{100\%}$	0.72
50	$DEM30_{50\%}$	13.08	$DEM10_{50\%}$	1.45	$DEM5_{50\%}$	0.36
25	$DEM30_{25\%}$	6.54	$DEM10_{25\%}$	0.73	$DEM5_{25\%}$	0.18
10	$DEM30_{10\%}$	2.62	$DEM10_{10\%}$	0.29	$DEM5_{10\%}$	0.07
5	$DEM30_{5\%}$	1.31	$DEM10_{5\%}$	0.15	$DEM5_{5\%}$	0.04
1	$DEM30_{1\%}$	0.26	$DEM10_{1\%}$	0.03	$DEM5_{1\%}$	0.01

Table 2. Differences in elevation values and paired t-test p-values for comparison between raster DEM at three resolutions (30m, 10m, and 5m) produced from the total 100% LIDAR data set and reduced density (50%, 25%, 10%, 5%, and 1%) data sets. The DEM were evaluated as the difference between $TOTAL_{100\%}$ DEM and MEAN DEM elevations for each density.

% original data set	30 m Mean difference TOTAL-MEAN (m)	p-value	10 m Mean difference TOTAL-MEAN (m)	p-value	5 m Mean difference TOTAL-MEAN (m)	p-value
50	-0.01	0.4379	-0.01	0.7680	-0.01	0.1824
25	-0.02	0.1095	-0.02	0.0552	-0.03	0.0145
10	-0.02	0.1830	-0.03	0.0117	-0.04	0.0007
5	-0.03	0.0441	-0.04	0.0086	-0.06	0.0006
1	-0.06	0.0013	-0.04	0.0047	-0.07	0.0001

CHAPTER 4 SPATIAL PREDICTION OF FOREST SOIL CARBON: SPATIAL MODELING AND GEOSTATISTICAL APPROACHES

INTRODUCTION

Soils constitute the major terrestrial carbon (C) reservoir, 1400-1600 Pg (10^{15} g) C globally (Falloon 1998; Sundquist, 1993), approximately three to five times the amount of C contained in terrestrial biomass (Brady and Weil, 2000; Houghton and Woodwell, 1989). Historically, most estimates of soil organic C (SOC) were based on means extrapolated from broad categories of soils and vegetation on a regional scale (Kern, 1990; Post *et al.* 1982). Better analysis and forecast of spatial patterns of SOC is important for sustainable land management (Florinsky *et al.*, 2002) and for potentially formulating strategies for offsetting global C emissions.

Spatial Modeling

Although we are beginning to understand patterns of SOC storage at the site or hillslope scale (Gessler *et al.* 2000), we need better methods to scale our findings to larger landscapes. Mental models developed by soil scientist based on landscape attributes, vegetation, hydrology, and other environmental variables have long been integrated in soil science for soil mapping purposes. However, these methods result in qualitative models that produce broad schemes that attempt to encompass the soil continuum and seek to provide simplistic classification regimes (Cook *et al.*, 1996). With the emergence of quantitative pedologic measurement and modeling techniques, or pedometrics, in the 1960's (Webster, 1994) soil scientists have sought a more quantitative approach to modeling the spatial distribution of soil properties. Pedometrics

resulted in the development of statistical based approaches that incorporate

surrogate environmental and edaphic explanatory variables and provide

estimations of selected soil properties (McBratney *et al.*, 2000). Several

approaches have historically been applied to quantitatively predict soil properties

on various scales.

Geostatistical spatial models have been developed that integrate standard

measures of variables within discrete land units. These "measure and multiply"

(Schimel and Potter, 1995) models provide a coarse estimation of selected soil

properties but often lack the ability to predict soil properties on scales of 1:50,000

to 1:100,000 needed for intensive land management (McKenzie and Ryan,

1999). Conversely, a spatial extrapolation approach referred to as "paint by

numbers" (Schimel and Potter, 1995) integrates a series of independent soil and

environmental variable classes with a known relationship to the dependent soil

variable. Thus, discrete classes with defined combinations of explanatory

variables are formed for model parameterization.

An approach to spatial extrapolation that has found recent use in soil

science and geomorphology is soil-landscape modeling. Soil-landscape

modeling is an approach to analyzing soil variability in response to exogenous

environmental variables related to topographic and hydrologic parameters

(McSweeney *et al.*, 1994; Paustian *et al.*, 1997; Thompson *et al.*, 2001).

McSweeney *et al.* (1994) incorporated three stages for soil-landscape modeling:

(i) physiographic representation through DEM and terrain attributes; (ii)

georeferenced training data with information about soil properties; and (iii) development and validation of explicit quantitative models. The approach provided a hierarchal regime of dissimilarly scaled variables for soil-landscape modeling. This is an important concept of landscape models because it allows for multi-resolution modeling of the soil properties. Furthermore, soil-landscape models provide quantification of soil properties through proxy variables and are not intended to provide a process-level understanding of individual soil properties within these models.

Soil-landscape models use discrete land units of similar vegetation, soils or ecological zones to guide representative sampling strategies to help integrate process dynamics on the landscape. Most models are founded in Jenny's (1941) "Factors of Soil Formation" by $S = f$ (cl, o, r, p, t) where S is a selected soil property as a function of climate (cl), organisms (o), relief (r), parent material (p), and time (t). Soil-landscape models normally assume that at the county (10 km^2) or regional (10 km^3) scale, (Ryan et al., 2000) variability within cl, p and t are controlled across the study site (Paustian et al., 1997). Sampling strategies are designed to control for those factors that vary across the area of interest. Thus, the driver of soil-landscape modeling is that variation in topography and vegetation provide responsive proxy variables that can be utilized for prediction of soil properties.

Topographic-based spatial models derived using GIS have been utilized for spatial predictions of soil properties (Gessler et al., 2000; Ryan et al. 2000;

Moore *et al.*, 1993), including forest SOC pools (McKenzie and Austin, 1993). A number of soil-landscape modeling techniques that use readily available geomorphic and pedologic based environmental explanatory variables to quantitatively predict spatial patterns of soil properties have been developed (McSweeney *et al.*, 1994; Odeh *et al.* 1994, McKenzie and Ryan, 1999; Gessler *et al.*, 2000; Florinsky *et al.*, 2002). Studies using soil-landscape spatial models have been able to predict and quantify specific soil properties such as A-horizon depth (Moore *et al.*, 1993; McKenzie and Ryan, 1999; Gessler *et al.*, 2000), organic matter content (Moore *et al.*, 1993), and total SOC (Arrouays *et al.*, 1998; McKenzie and Ryan, 1999; Gessler *et al.*, 2000; Ryan *et al.*, 2000). These studies were able to explain 40 to 85% of the variability in the predicted soil properties.

Scaling is a serious issue in the study of C cycling in terrestrial ecosystems (Schimel and Potter, 1995). One underlying problem is that the factors that control soil variability occur across a range of scales. Proximal factors (e.g. pH, soil moisture) that influence SOC contribute to the variability on a much smaller scale. Distal factors (*viz.* Jenny's (1941) factors) vary over a much larger scale. Integration of multiple attributes at varying scales imposes a problem for developing models that predictably explain SOC variability across a landscape. Moreover, there is an inherent discord between the scale at which soil property dynamics occur (microscale), the scale at which they are measured (mesoscale) and the scale at which they are modeled (macroscale). With each

increase in coarseness of scale, the variability, and therefore the uncertainty of the prediction increases (Kern, 1994; Gessler *et al.*, 2000; Ryan *et al.*, 2000).

DEM Quality – Resolution, Accuracy and Precision

Landscape attribute prediction exhibits a direct dependency on the qualities of the DEM used for surface representation and attribute derivation (Jenson, 1991; Thieken *et al.*, 1999). Thieken *et al.* (1999) and Thompson *et al.* (2001) indicated that DEM resolution contributes to differences in the distribution and representation of landscape attributes. Gessler *et al.* (2000) found little difference among landscape models based on a landscape attribute derived from a series of DEM with 2- to 10-m resolutions. Likewise, Chaplot *et al.* (2000) found that 10- to 30-m DEM generally provided an unbiased prediction of landscape terrain but prediction was influenced as DEM resolution increased to 50-m. Florinsky and Kuryakova (2000) emphasized that DEM resolution was highly dependent on the scale of the process modeled, concluding that high resolution (between 2.25 and 3.25 m) DEM were important for modeling processes at the microscale. In general, these studies indicated that attribute value ranges increased and predictive capabilities decreased as DEM grid size increased. The amount of relief on a landscape contributes to the effects of DEM resolution on terrain attributes, with low relief landscapes being less sensitive to resolution impacts. Consequently, large study areas that may incorporate a larger range of relief may require higher resolution DEM to capture break-lines and ridges that control water movements.

Thompson *et al.* (2001) found statistical differences in landscape attributes (specific catchment area, compound topographic index) when comparing DEM of 1 m and 0.1 m vertical precisions. Paired data revealed that lower precision (1 m) DEM had higher slope gradients and lower values for specific catchment area and compound topographic index (Thompson *et al.*, 2001). Thieken *et al.* (1999) concluded that changes in vertical precision, particularly in low relief landscapes, affected individual cell values for terrain attributes such as slope, specific catchment area and compound topographic index, but did not affect the cumulative distribution of theses attributes. When precision was held constant, Thompson *et al.* (2001) indicated that there was a dependency on horizontal resolution, with low resolution DEM creating smoother transitions between adjacent cells than did high resolution DEM. However, Thieken *et al.* (1999) found that low vertical precision often results in a "stair stepped" appearance. Similarly, Thompson *et al.* (2001) found that decreased vertical precision created a greater segregation of slope values that included a large number of zero slope areas and steeply sloped areas. A possible solution to this may be found in calculating attributes over greater distances rather than by using only adjacent cells (Thieken *et al.*, 1999).

The horizontal and vertical qualities of a DEM are directly linked to the source of data used for its production. Traditionally, DEM have been derived from photogrammetric techniques (includes contour mapping) and ground elevation surveys, with more recent application of remotely sensed elevation data

acquired through interferometric synthetic aperture radar (IFSAR), light detecting

and ranging (LiDAR) or similar technologies. The most widely used DEM within

the United States have historically been the 30-m DEM (level 1) produced by the

U.S. Geological Survey (USGS). A more recent approach has been the use of

radar- or laser-based remotely sensed elevation data for model derivation.

Application of remotely sensed data requires digital spot elevation data

interpolation into raster-based DEM. Landscape modeling studies have made

use of a large variety of sources of elevation data for terrain representation. The

need for high-resolution, high accuracy elevation data for purposes of landscape-

scale modeling has resulted in the application of Light Detecting and Ranging

(LiDAR) technology is a source of data used to produce high-resolution, high-

accuracy DEM.

Modeling SOC

Prediction of SOC and other soil properties is dependent upon the

selection of pedologically important proxy landscape attributes and soil properties

for use in explicit spatial models (Gessler *et al.*, 2000). A study was conducted

to develop spatial models to predict total SOC using selected pedologically

important topographic variables. Spatial models of SOC were developed and

tested around two hypotheses: (i) that spatial autocorrelation of SOC exist in low-

relief landscapes; and (ii) that spatial patterns of SOC on a watershed scale are

predictable by models based on pedological relationships displayed by

topographic variation. The specific objectives of this study were: (i) to quantify

spatial autocorrelation of SOC patterns predicted using geostatistical models; (ii) to utilize landscape attributes to develop and validate an explicit, quantitative, and spatially realistic model of SOC for a 32,500 ha forest ecosystem; (iii) to quantify total SOC storage in Hofmann Forest.

MATERIALS AND METHODS

Study Site

Investigation into the spatial distribution of SOC occured on a 32,500 ha forest ecosystem located entirely within the bounds of Hofmann Forest. Hofmann Forest is located in Jones and Onslow Counties of eastern North Carolina, USA (Figure 1) and lies on the Lower Coastal Plain (LCP) Wicomico and Talbot morphostratigraphic units of the mid-Atlantic seaboard. Located in the temperate climate zone, the study area is characterized by warm summers and mild winters with a mean summer temperature of 25°C and a mean winter temperature of 7°C. Mean annual precipitation is 1400 mm with a large portion of the rainfall received in late summer. Elevations range from 12 to 20 m above mean sea level (Daniels *et al.*, 1977), with the landscape characterized by broad, flat interfluves. In some areas, relief may be as low as 1.5 m elevation difference in 3 to 4 km (Daniels *et al.*, 1999).

Soils

Soils of the Hofmann Forest were derived from surficial marine sediments of the Wicomico and Talbot morphostratigraphic units, alluvial deposits, and organic deposits on low-relief interfluves. The soils are predominately poorly to

114

somewhat-poorly drained Saprists, Aquults and Aquepts. Organic soils dominate Hofmann Forest, representing nearly 24,000 ha (Daniels *et al.*, 1977). The Hofmann Forest landscape has poorly drained organic soils occurring on broad low-relief interfluves and better-drained soils in close proximity to drainages. Mineral soils fringe the broad interstream divides and are typified by deep water tables with light surface and subsurface horizons. Daniels *et al.* (1977) provides a detailed description of the stratigraphy, geomorphology and pedological units with the Hofmann Forest.

The unique hydrological conditions within Hofmann Forest appeared beneficial to the accumulation of C within interfluve areas. Daniels *et al.* (1977) indicated an influence of distances to nearest major drainage on organic material accumulation within Hofmann Forest. The wide spacing and low slope between natural drains inhibit lateral water movement, thus creating large partially saturated regions within the centers of the forest. These areas, commonly referred to as pocosins, are often characterized by substantial accumulation of organic materials.

Pocosin Vegetation and Forest Management

The Hofmann Forest contains a large diversity of vegetation, both as a result of natural regeneration and intensive forest management. Natural palustrine wetland plant communities dominate the pocosin area with the presence of pond pine (*Pinus serotina*), redbay (*Persea borbonia*), loblolly bay (*Gordonia lasianthus*), sweetbay (*Magnolia virginiana*), bamboo (*Arundo donax*),

115

gallberry (*Ilex glabra*), and many others. A major portion of the land surrounding the pocosin includes managed pineland dominated by loblolly, slash and longleaf pine (*Pinus palustris*) plantations. Other vegetation land use/land classification in Hofmann Forest includes bottomland hardwood, hardwood flats, headwater swamps, non-managed pine flats, swamp forests, agricultural fields and pine savannas (unpublished data).

Moderate to intensively managed pinelands occupy nearly 15,000 ha of Hofmann Forest. Drainage networks have been installed throughout most of the managed pine plantation lands with minor drainage networks on approximately 100 m spacing between adjacent drains. Pine plantations of loblolly pine (*Pinus taeda*) in Hofmann Forest range in age from 0 to 67 yrs, fairly distributed across the age groups of 0-5, 5-10, 10-15, 15-20, 20-25, 25-30, 30-35, 35-40 yrs. However, the 20-25 yrs age group was by far the largest, representing over 3000 ha (20% of total plantations).   Harvesting is performed on a commercial contract basis. Harvest regimes are site specific and may include whole-tree and sawlog-only harvests. Digital information regarding tree volume removals and other details of the harvest are acquired and incorporated into a GIS database for Hofmann Forest. Other information including soils, vegetation, land use/land classification, and age class are digitally georeferenced and integrated into the database.

Sample Stratification Regime

A stratified random sampling scheme for collection of SOC samples was used to minimize variability in physical and chemical differences between samples, while maintaining adequate coverage of the watershed. Variation may be reduced when sampling on a stratification scheme based on *a priori* understanding of factors that potentially influence the prediction variable (Ryan *et al.*, 2000; Turner and Lambert, 2000; Zinke and Stangenberger, 2000). The stratified sampling regime was based on three criteria: (i) plantation pine versus natural pocosin plant communities, (ii) age groupings within pine plantations, and (iii) distance from major streams.

Hofmann Forest was divided into two major vegetation classes (plantation and pocosin) for initial stratification. In the LCP, vegetation and land use are often indicative of underlying soil conditions. These two vegetation groups were selected based on their influence on C storage and their dominant areal extent within Hofmann Forest. Pocosin areas were anticipated to represent areas of greatest SOC accumulation. Plantation areas are suspected to occupy drier, less organic soils or have less SOC as a result of anthropogenic activities. Plantation and pocosin vegetation classes were subsequently stratified into 4 groups representing distances of 0-2000m, 2000-3000m, 3000-4000m and 4000^{+}m to major natural drainages. It was suspected that distance to natural drainage was influential in the genesis of soils within Hofmann Forest. Previous research in Hofmann Forest by Daniels *et al.* (1977) indicated that morphological

differences occurred on the broad flat interfluves as result of distance from major streams and drains.

A further stratification of pine plantations was based on plantation age. Numerous studies indicate high variability of SOC pools within different aged pine stands (Huntington, 1995; Richter *et al.*, 1995; Van Lear *et al.*, 1995; Trettin *et al.*, 1999; Paul *et al.*, 2002). Age groupings included 0-5, 5-15, 15-25, 25-35 and 35[+] yr. This strategy was to ensure proper stratification based on stand establishment (0-5yr), initial stand closure, thinning, and fertilization (5-15yr), late rotation thins or other silvicultural activities (15-25yr), stands scheduled for harvest (25-35yr) and mature older plantations (35[+]yr).

2.5 Soil Sampling and SOC Analysis

Soil samples were collected at 190 georeferenced locations throughout the study site for chemical analysis and determination of soil bulk density. Four sub-samples were taken at roughly 7.5 m at approximate cardinal directions from each prescribed geo-referenced sampling location. Sub-samples on bedded plantation sites were oriented to provide 2 inter-bed and 2 intra-bed samples. Soils were collected in 3 cm butyrate plastic liners with a stainless steel soil recovery probe with slide hammer attachment (JMC Soils-ESP soil sampler).

Intact volumetric soil samples were collected to a depth of approximately 1 m from the surface, including the organic horizon. Samples were stored indoors at 16 to 20°C in the plastic butyrate liners during the sample collection period. Subsequently, soil samples were extruded from the plastic butyrate liners and

dried in a forced-air oven at approximately 40˚C for 72h. Bulk density were

measured on oven-dried samples by standard technique using the cylindrical

volume and soil mass within a given depth increment. Bulk density represented

an average of the 4 sub-samples or the total number of uninterrupted cores from

a given location. Soil samples were composited with the other sub-samples from

each sampling location. Total C of whole soil samples was determined by dry

combustion in a Perkin-Elmer Series II 2400 CNH analyzer (Nelson and

Sommers, 1996; Bremner, 1996).

Total SOC was calculated as kg m^{-2} for each sample collected. Summary

statistics including mean, maximum, minimum, and standard deviation were

calculated by land use and as a total for all cores collected using S-Plus® 6.0.

Geostatistics and Spatial Autocorrelation

Core samples collected from 75% of the georeferenced sampling locations

were used to develop a geostatistical model of SOC across the Hofmann Forest

landscape. Ordinary kriging (OK) was selected purely as a way to preliminarily

assess the possibility of spatial modeling of SOC using landscape and land use

attributes. Model evaluation, data analysis, model parameterization and SOC

prediction was performed in the Geostatistical Analyst module of ArcGIS 8.1

(ESRI, 2003). Semivariogram nugget, range, and sill parameters for the OK

model were calculated. Development data for the OK model was constrained to

isotropic neighborhoods of 1000 m and $100 > n > 15$ nearest neighbors were

used to predict SOC. The search radius was flexible in that it always allowed for

a neighborhood of at least 15 SOC cores.

Spatial autocorrelation was employed to evaluate systematic spatial

variation in estimates of SOC. Methods for investigating spatial autocorrelation

include the statistical coefficients Moran's *I* and Geary's *c* which provided an

indication of the type and degree of spatial autocorrelation present in the SOC

data set. Both indices allow using a single value to describe the spatial

distribution of features. They can be used to determine the degree of adjustment

necessary when modeling the phenomena. Moran's *I* is produced by

standardizing the spatial auto covariance by the variance of the data using a

measure of the connectivity of the data. Geary's *c* uses the sum of squared

differences between pairs of data values as its measure of covariation. These

measures were used to evaluate whether there was a spatial trend in the SOC

data for Hofmann Forest.

DEM and Terrain Attributes

A series of 61 LIDAR tiles (1000 ha) were collected from the North

Carolina Flood Mapping Program (NCFMP) to cover the spatial extent of the

Hofmann Forest. Prior to public release of LIDAR data sets, the NCFMP

processed raw LIDAR data using proprietary algorithms to remove artifacts of

vegetation, water bodies, and manmade objects. The resulting end products

were irregularly spaced bare earth elevation data sets. Processing by NCFMP

resulted in a land cover dependent gradient in the density of bare earth LIDAR

returns, with agriculture lands having much higher point densities than forestlands. The statistics for the combined land cover and the trends for each specific land cover type were reviewed, and data that fell outside of a 20-25 cm root mean square error (RMSE) criteria were removed from the data sets (http://www.ncfloodmaps.com). Elevations were reported to have 0.1 m vertical precision (http://www.ncfloodmaps.com).

All 61 LIDAR tiles were merged to form a single spot elevation coverage consisting of over 9,000,000 LIDAR points. Data sets were interpolated to raster DEM using ANUDEM 4.6.3 (Hutchinson, 1995). ANUDEM employs an iterative finite difference interpolation technique that allows for maintenance of a drainage network consistent with the original data and removal of spurious sinks (Hutchinson, 1995). Raster DEM of 10 m, 30 m, and 100 m horizontal resolution with 0.1 m vertical precision were created.

Selected terrain attributes were calculated (Table 1), with parameters developed using each of the three resolutions; 10 m, 30 m, and 100m. This method was used to address the effects of scale across the various landscape attributes. Terrain attributes were designed to provide quantitative parameters indicative of landform shape, connectivity, and adjacency that control external landscape geomorphology and represent hydrologic tendencies (Gessler *et al.*, 2000). Primary attributes such as elevation and slope gradient were derived directly from DEM, whereas secondary attributes are derived from combinations of primary attributes (Moore *et al.*, 1991). Secondary attributes serve as

surrogates for complex hydrological, geomorphological and pedological processes (Moore *et al.*, 1991).

2.8 Statistical Analysis – Landscape Modeling

Seventy-five percent (n = 143, randomly selected) of the geo-referenced sampling locations were used as model training data while the remaining 25% (n = 47, randomly selected) of the sites served as a validation dataset for a prediction error analysis. Correlation analysis was performed to evaluate prediction variables and to address attributes with strong interdependent correlation with multiple variables. Stepwise multiple regression methods related the target variable (SOC) to explanatory prediction variables. Regression modeling was completed using S-Plus® 6.0 stepwise regression procedures (Neter *et al.*, 1989). Forward stepwise regression successively added variables to the prediction model that exceed the established partial *F*-test statistic threshold, yet only added variables that improve the overall predictive capabilities of the model (Neter *et al.*, 1989; SAS Institute, 2000). Models were evaluated by comparison of R^2 and residual standard error (RSE).

Two separate spatial models were developed to predict SOC in Hofmann Forest. One model was developed using the entire suite of available topographic and land use data, while the second model was developed using only topographic attributes. This process was to allow for a better understand of the spatial and anthropogenic influence on SOC in Hofmann Forest. Both spatial models were implemented into the GIS and displayed on a raster basis. Map

algebra using the topographic and land use coverages was employed to provide quantification with raster units.

Predictive models were evaluated against the remaining 25% data set established for model testing. Mean predicative error (ME), root mean square error (RMSE) were calculated based on the difference between each model and the validation data set. The end result were a predictive map of areal quantities of SOC with known error estimates and predictive capabilities.

<div align="center">

RESULTS AND DISCUSSION

</div>

SOC Analysis

Samples collected in Hofmann Forest were characterized by land use as well as by combined land use classes (Table 2). Data represented all SOC cores collected during field sampling. Several studies have indicated that a reasonable estimate for total SOC in southeastern forest ecosystems should range from 6-20 kg C m^{-2} for mineral soils and around 80 kg C m^{-2} for organic soils to a depth of 1 m (Johnson and Kern, 2003; Birdsey and Lewis, 2003; Garten *et al.*, 1999). Our SOC measures from soil cores were substantially greater than previously reported values. This was attributed to the overall depth and extent of organic soils within Hofmann Forest. Additionally, the SOC values had a large range in maximum and minimum values and standard deviation and indicate a large amount of variability in the 190 SOC cores. The variability of SOC was influential in the amount of variability explained by the explicit spatial models.

Geostatistical Modeling of SOC

Prediction of SOC using a strictly geostatistical approach provided a method of evaluating the spatial distribution of SOC within Hofmann Forest. Prediction of SOC by OK was completed on 30 m resolution grid for Hofmann Forest (Figure 2a).

Spatial trends in SOC predicted by OK indicated that spatial modeling might be possible for Hofmann Forest. To better understand the spatial trends and the level of spatial autocorrelation in the SOC data, two methods of spatial autocorrelation analysis were employed. A Moran's I coefficient of 0.984 for the OK models indicated that there was a spatial trend in the SOC data. Furthermore, a Geary's c of 0.001 provided additional evidence of spatial autocorrelation of the SOC data. Based on the spatial autocorrelation of the SOC data (Moran's I = 0.984, Geary's c = 0.001), its was believed that by using landscape and land use attributes that an explicit spatial model of SOC was possible for Hofmann Forest.

Landscape Models

Correlation coefficients were calculated for the suite of terrain and land use attributes for Hofmann Forest (Table 3). Several attributes were identified as possible explanatory variables. Of initial interest were L_{binary}, SI, Z_{10m}, Z_{30m}, Z_{100m}, $K_{c(10m)}$, $K_{c(30m)}$, and $K_{p(30m)}$. Review of interdependent correlations revealed minimal interference between most of the attributes that had the greatest potential for inclusion in the explicit spatial models. However, the correlation

coefficient between L_{binary} and SI ($r = -0.41$) indicated a possible interfence between these two variables.

Multiple regression performed on the landscape and land use attributes resulted in the selection of variables for inclusion in the spatial model (Table 4). A model (model = LandTopo) was developed to include both landscape and land use attributes. A separate model (model = Topo) was developed that included only landscape attributes. These models included two landscape variables in common (K_c, and $K_{p()}$) but were at different resolutions. LandTopo included L_{binary} and SI variables that were of concern because of the interdependent relationships determined by r. This interaction may have had some influence on the predictive capability of the model.

Despite low R^2 for the best models for LandTopo and Topo, statistical analysis provided an explicit spatial model of SOC for Hofmann Forest (Table 5; Figure 2b,c). The best variables included in both LandTopo and Topo models of SOC provided limited ability to explain the variability of SOC in Hofmann Forest. Even with possible interaction effects between L_{binary} and SI, these variables were not eliminated from the LandTopo model. It was understood that by including these variables some additional statistical limitations in predicting SOC were likely included. However, based on the amount of unexplained variability in SOC by the LandTopo model, these limitations were acceptable.

It was anticipated that landscape variables at multiple resolution would have provided additional ability to model SOC variability. Each model contains

variables of different spatial resolution that did in fact increase the ability of the models to predict SOC. This was anticipated because of the multiple scales at which soil forming process interact. It was believed that by not including landscape variables at multiple resolutions we would have further limited the success of the explicit spatial models to predict SOC in Hofmann Forest

Based on preliminary examination of soil survey information, a topographic influence on spatial patterns of SOC was expected to be model with greater efficiency. Daniels *et al.* (1977) indicated an influence of distances to nearest major drainage on organic material accumulation within Hofmann Forest. Landscape attributes that were anticipated to help explain SOC variability proved to be ineffective in predicting SOC across Hofmann Forest. Of greatest interest was the exclusion of L_{ds} from both models. Prior to development of explicit spatial models, it was anticipated that L_{ds} would help to explain a major portion of the SOC variability. Possibly, other methods of evaluating or describing the affect of distance from major drainages in Hofmann Forest may have provided a better estimation of SOC.

The low relief landscape that encompasses Hofmann Forest greatly contributed to our inability to successfully model SOC. The unique hydrological conditions of pocosins pose a serious challenge to model development using landscape and land use attributes alone. Landscape attributes such as S, K_p, K_c, and K_t that have been successfully included other explicit spatial models (Gessler et al., 2000; Thompson et al., 2001) failed to adequately explain the variability of

126

SOC in Hofmann Forest. Possible inclusion of other yet undeveloped landscape variables more adept at addressing low relief edaphic characteristics may have improved the models ability to explain SOC variability.

Validation and Comparison of SOC models

LandTopo, Topo and OK spatial models were implemented in the Hofmann Forest GIS to provide an estimation of the spatial distribution of SOC. Estimated SOC was compared to the validation data set (n= 47) to exhibit the inability of the explicit spatial models to explain SOC variability in Hofmann Forest. Additionally, validation provided a means of comparing the explicit spatial models to SOC estimates from the OK geostatistical approach.

Coincident grid cells for each SOC model were compared to the validation data set (Table 6). The RMSE between each model and the validation data set indicated that estimates of SOC by OK were three-fold better than estimates by either LandTopo or Topo. The OK model tended to slightly overestimate SOC, while LandTopo and Topo both slightly underestimated SOC compared to the empirical data. Given the low R^2 of both LandTopo and Topo, the large RMSE from validation was not unrealistic from what was anticipated.

The estimates of SOC produced using the geostatistical OK model were much more clustered and presented a more homogeneous distribution of SOC within Hofmann Forests (Figure 2a). Estimates formed concentric bands of SOC that decreased in SOC along the fringes of Hofmann Forest. This was similar to

the trends explained by Daniels *et al.* (1977) and more closely resembled our initial ideas of the spatial trends of SOC within Hofmann Forest.

LandTopo and Topo were more discretized in their estimates of SOC across Hofmann Forest (Figure 2b,c). Rather than producing homogenous blocks of SOC similar to the OK model, estimates were much more heterogeneous across Hofmann Forest. LandTopo had a greater range of SOC values (0 – 350 kg m^{-2} C) than the estimates produced by Topo (0 – 300 kg m^{-2} C). LandTopo provided some separation between pocosin and plantation land uses as was expected due to the inclusion of both L$_{binary}$ and SI land use attributes. Additionally, LandTopo indicated low SOC along the outer rim of Hofmann Forest, similar to the conditions described by Daniels *et al.* (1977). Spatial patterns of SOC in the Topo model were along an east-west gradient across Hofmann Forest. The Topo model failed to exhibit any decreased levels of SOC along the fringes of the Forest. Rather, SOC were concentrated along the eastern edge.

Given the visual and statistical differences between the geostatistical OK SOC models and LandTopo and Topo models, we anticipated there would be a large difference in the total SOC for Hofmann Forest. However, SOC estimates on an areal basis for the resulted in similar values for all models. The OK model predicted a total SOC store of 40.2 Gt C in the top 1 m. LandTopo and Topo had similar predictions of 37.2 Gt C for LandTopo and 37.4 Gt C for Topo. All three models estimated SOC stores at a much higher inventory than was anticipated.

The SOC storage values were a result of high SOC values in the original training data used for developing each model.

CONCLUSIONS

Spatial trends in SOC exist in Hofmann Forest. Geostatistical approaches and spatial autocorrelation coefficients (Moran's I and Geary's c) provided evidence of these spatial trends in SOC. However, we failed to adequately explain the variability in SOC using explicit spatial models. Certainly, we can attribute our lack of ability to model SOC to the highly variable nature of the SOC collected in the Hofmann Forest.

We had anticipated that our use of high-resolution, high-accuracy LiDAR-based DEM would improve our ability to model SOC on the low relief landscape of Hofmann Forest. Despite the use LiDAR-based DEM, multi-scaled landscape attributes and the inclusion of land uses variables in our models, we were unsuccessful in our attempts to model SOC. We believe that a multi-scaled approach provided some help in explaining the variability of SOC but to what extent remains unknown because of our overall poor modeling results. Our models produced landscape scaled SOC stores above what we had expected, yet only additional sampling and fine tuning of our explicit spatial models will provide insight into the extent of error in our SOC estimations.

Low relief landscapes pose a serious problem for producing explicit spatial models of SOC. These models may be inhibited by including only landscape and land use attributes, however we believe that these are likely the most influential

parameters affecting the development of soils in low relief landscapes. One

possible explanation is that we may be missing a yet undetermined landscape or

land use attribute that vastly improve our ability to capture the variability of SOC

in these low relief landscapes. Future exploration of additional landscape and

land use variables that correlate to edaphic properties on low relief landscapes

should be a continued focus for spatial modelers if we hope to provide spatial

estimates of landscape scaled SOC stores. Pedologist will need to continue to

seek innovative approaches to modeling soil process on low-relief landscapes.

LITERATURE CITED

Arrouays, D., and P. Pelissier. 1993. Modeling carbon storage profiles in

temperate forest humic loamy soils of France. Soil Sci. 157(3):185-192.

Birdsey, R.A. and G.M. Lewis. 2003. Current and historical trends in use,

management, and disturbance of U.S. forestlands. p. 15-33. *In* Kimble *et al.*

(eds.) The potential of U.S. forest soils to sequester carbon and mitigate the

greenhouse effect. CRC Press, Boca Raton, FL.

Brady, N.C. and R.R. Weil. 2000. Elements of the nature and properties of soils.

Prentice Hall, Upper Saddle River, NJ.

Bremner, J.M. 1996. Nitrogen-Total. p. 1048. *In* Page *et al.* 1996. Methods of soil

analysis. Part 3: Chemical Methods. ASA, CSSA, and SSSA, Madison, WI.

Chaplot, V., C. Walter, and P. Curmi. 2000. Improving soil hydromorphy

prediction according to DEM resolution and available pedogical data.

Geoderma 97:405-422.

Cook, S.E., R.J. Corner, G. Grealish, P.E. Gessler, and C.J. Chartes. 1996. A rule-based system to map soil properties. Soil Sci. Soc. Am. J. 60:1893-1900.

Daniels, R.B., E.E. Gamble, W.H. Wheeler, and C.S. Holzhey. 1977. The stratigraphy and geomorphology of the Hofmann Forest Pocosin, North Carolina. Soil Sci. Soc. Am. J. 41:1175-1180.

Daniels, R.B., S.W. Buol, H.J. Kleiss, and C.A. Ditzler. 1999. Soil systems of North Carolina. Tech. Bull. 314. North Carolina State Univ., Raleigh.

Falloon, P.D., P. Smith, J. U. Smith, J. Szabo, K. Coleman, and S. Marshall. 1998. Regional estimates of carbon sequestration potential: linking the Rothamsted Carbon Model to GIS databases. Biol. Fert. Soils. 27:236-241.

Florinsky, I.V., R.G. Eilers, G.R. Manning, and L.G. Fuller. 2002. Prediction of soil properties by digital terrain modeling. Environ. Model. Soft. 17:295-311.

Florinsky, I.V. and G.A. Kuryakova. 2000. Determination of grid size for digital terrain modeling in landscape investigations – exemplified by soil moisture distribution at a micro-scale. Int. J. of GIS. 14(8):815-832.

Garten, C.T., W.M. Post, P.J. Hanson, and L.W. Cooper. 1999. Forest soil carbon inventories and dynamics along an elevation gradient in the southern Appalachian Mountains. Biogeochem. 45:115-145.

Gessler, P.E., O.A. Chadwick, F. Chamran, L. Althouse, and K. Holmes. 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Sci. Soc. Am. J. 64:2046-2056.

Houghton, R.A. and Woodwell, G.M. 1989. Global climate change. Scientific American 260:36-44.

Huntington, T.G. 1995. Carbon sequestration in an aggrading forest ecosystem in the southern USA. Soil Sci. Soc. Am. J. 59:1459-1467.

Hutchinson, M.F. 1995. Documentation for ANUDEM version 4.4. Centre for Resources and Environmental Studies, Australian National University, Canberra.

Jenny, H. 1941. Factors of soil formation, a system of quantitative pedology. McGraw-Hill, New York, 281 p.

Jenson, S.K. 1991. Application of hydrologic information automatically extracted from digital elevation modeling. p. 35-48. *In* K.J. Bevens and I.D. Moore (eds.) Terrain analysis and distributed modeling in hydrology. Wiley and Son, Chichester, NY.

Johnson, M.G. and J.S. Kern. 2003. Quantifying the organic carbon held in forested soils of the United States and Puerto Rico. p. 48-72. *In* Kimble *et al.* (eds.) The potential of U.S. forest soils to sequester carbon and mitigate the greenhouse effect. CRC Press, Boca Raton, FL.

Kern, J.S. 1994. Spatial pattern of soil carbon in the contiguous United States. Soil Sci. Soc. Am. J. 58:439-455.

McBratney, A.B., I.O.A. Odeh, T.F.A. Bishop, M.S. Dunbar, and T.M. Shatar. 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97:293-327.

McKenzie, N.J. and M.P. Austin. 1993. A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation. Geoderma 57:329-355.

McSweeney, K.M., P.E. Gessler, B. Slater, R.D. Hammer, J.C. Bell, and G.W. Petersen. 1994. Towards a new framework for modeling the soil-landscape continuum. *In* Amundson, R.G. (eds.) Factors of soil formation: A fiftieth anniversary retrospective. SSSA Spec. Publ. 33. Soil Sci. Soc. Am., Madison, WI. 127-145.

Moore, I.D., P.E. Gessler, G.A. Nielsen, and G.A. Peterson. 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57:443-452.

Moore, I.D., R.B. Grayson, and A.R. Ladson. 1991. Digital terrain modeling: a review of hydrological, geomorphological and biological applications. Hydrological Processes 5:3-30.

Nelson, D.W. and L.E. Sommers. 1996. Total carbon, organic carbon, and organic matter. p. 961. *In* Page *et al.* (eds.) Methods of soil analysis. Part 3 Chemical Methods. ASA, CSSA, and SSSA, Madison, WI.

Neter, J., W. Wasserman and M.H. Kutner. 1989. Applied linear regression models. 2nd ed. Richard D. Irwin, Homewood.

Odeh, I.O.A., A.B. McBratney, and D.J. Chittleborough. 1994. Spatial Prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma 63(3-4):197-214.

Paul, K.I., P.J. Polglase, J.G. Nyakuengam, and P.K. Khanna. 2002. Change in soil carbon following afforestation. For. Ecol. Manage. 168:241-257.

Paustian, K., E. Levine, W.M. Post, and I.M. Ryzhova. 1997. The use of models to integrate information and understanding of soil C at the regional scale. Geoderma 79:227-260.

Post. W.M., W.R. Emanuel, P.J. Zinke, and A.G. Stangenberger. 1982. Soil carbon and world life zones. Nature 298:156-159.

Richter, D.D., D. Markewitz, C.G. Wells, H.L. Allen, J.K. Dunscome, K. Harrison, P.R. Heine, A. Stuanes, B. Urrego, and G. Bonani. 1995. Carbon cycling in a loblolly pine forest: Implication for the missing sink and for the concept of soil. *In* W.W. McFee and J.M. Kelly (eds.) Carbon forms and functions in forest soils. Soil Sci. Soc. Am. Madison, WI

Ryan, P.J., N.J. McKenzie, D. O'Connell. A.N. Loughhead, P.M. Leppert, D. Jacquier, and L. Ashton. 2000. Integrating forest soils information across scales: spatial prediction of soil properties under Australian forests. For. Ecol. Manage. 138:139-157.

SAS Institute, 2000, SAS/STAT user's guide. Version 8.01. SAS Institute, Cary, NC.

Schimel, D.S. and C.S. Porter. 1995. Process modeling and spatial extrapolation. p. 358-383. *In* P.A. Matson and R.C. Harriss (eds.) Biogenic trace gases: Measuring emissions from soil and water. Blackwell Sci. Ltd., Cambridge, MA.

Sundquist, E. 1993. The global carbon dioxide budget. Science 259:934-941.

Thieken, A.H., A. Lücke, B. Diekkrüger, and O. Richter. 1999. Scaling input data by GIS for hydrological modeling. Hydrol. Process. 13:611-630.

Thompson, J.A., J.C. Bell, and C.A. Butler. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma 100:67-89.

Trettin, C.C., D.W. Johnson, and D.E. Todd, Jr. 1999. Forest nutrient and carbon pools at Walker Branch Watershed: changes during a 21-year period. Soil Sci. Soc. Am. J. 63:1436-1448.

Turner, J., M. Lambert. 2000. Change in organic carbon in forest plantation soils in eastern Australia. For. Ecol. Manage. 133:231-247.

Van Lear, D.H., P.R. Kapeluck, and M.M. Parker. 1995. Distribution of carbon in a piedmont soil as affected by loblolly pine management. In W.W. McFee and J.M. Kelly (eds.) Carbon forms and functions in forest soils. Soil Sci. Soc. Am. Madison, WI

Webster, R. 1994. The development of pedometrics. Geoderma 62:2-15.

Zinke, P.J. and A.G. Stangenberger. 2000. Elemental storage of forest soil from local to global scales. For. Ecol. Manage. 138:159-165.

Figure 1. Ten selected LIDAR tiles located on the Lower Coastal Plain of eastern North Carolina, USA. Inset: Location of NC in continental USA
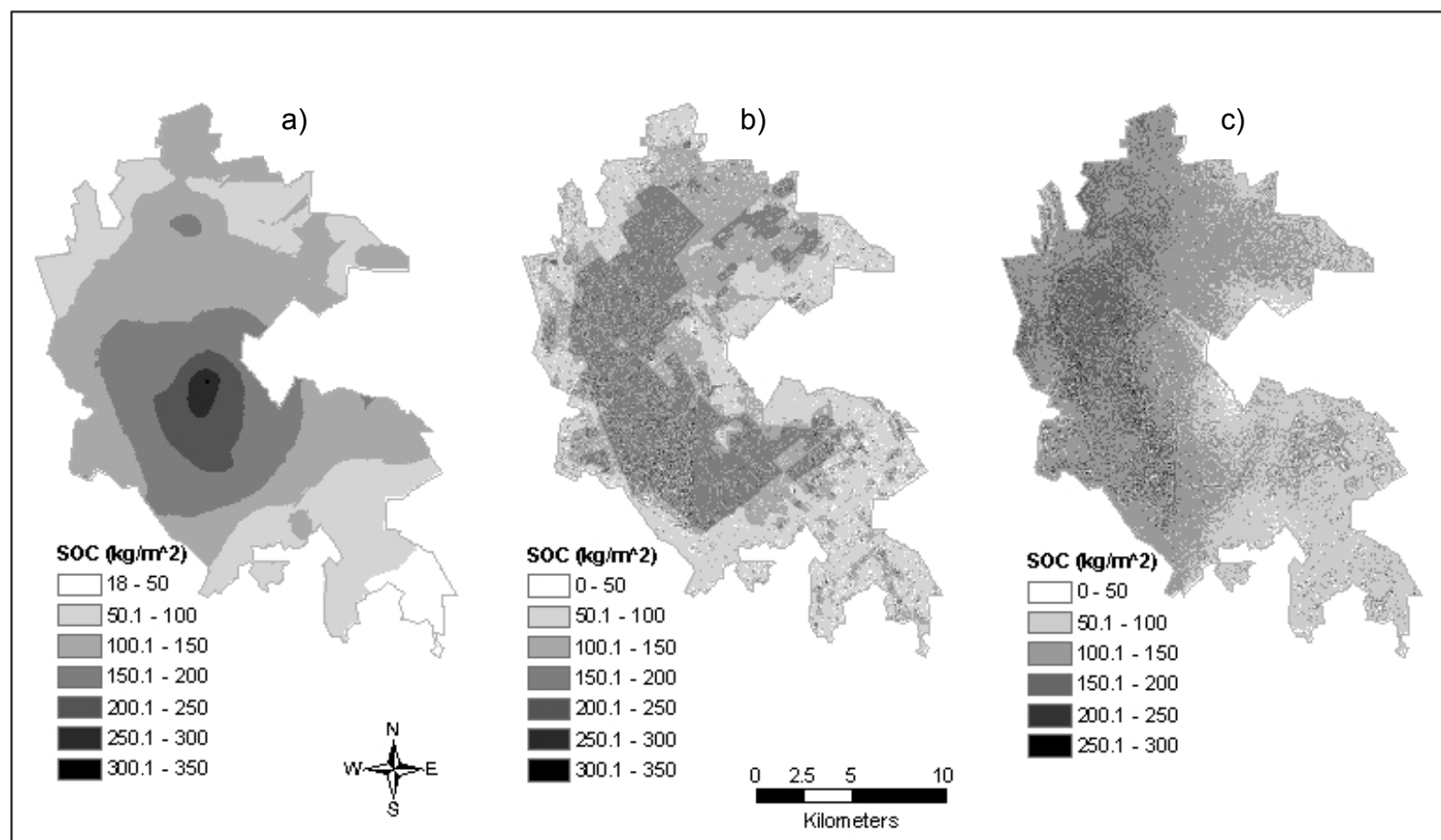
Figure 2. Spatial distribution of SOC in the Hofmann Forest calculated from (a) OK geostatistical model, (b) LandTopo explicit spatial model; and (c) Topo explicit spatial model.

1 **Table 1**. Developed primary and secondary terrain attributes and land use
2   parameters of Hofmann Forest used for soil landscape modeling.
3

Attribute or Parameter	Description
*Elevation (Z), m	Elevation above mean sea level
*Slope gradient (S), %	Gradient between adjacent grid cells
*Profile Curvature (K_p), m^{-1}	Slope profile curvature
*Contour Curvature (K_c), m^{-1}	Contour or plan curvature
*Tangential Curvature (K_t), m^{-1}	Transitional curvature along a tangent not orthogonal to other curvature measures
*Linear Distance to Stream (L_{ds}), m	Distance from cell location to nearest stream
Site Index (SI)	Collective influence of soil factors contributing to the growth of vegetation
Land use (L_{binary})	A binary descriptor given for land use of either pocosin (1) or plantation forestry (0)

4 *Terrain attributes developed at multiple resolution of 10m, 30m and 100m
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

138

Table 2. Summary statistics for SOC cores to a depth of 1m collected in Hofmann Forest.

Land use	Mean	Maximum	Minimum	Std. Deviation
		C (kg m^{-2})		
Pocosin	151.4	390.0	13.5	72.5
Plantation	91.7	754.3	11.2	89.6
Total	123.2	754.3	11.2	86.2

Table 3. Correlation coefficients (*r*) of selected landscape and land use attributes used in development of explicit spatial models of SOC in Hofmann Forest.

Dependent Variables	Terrain Attributes								
	SOC	L_{binary}	SI	Z_{10}	Z_{30}	$K_{c(10)}$	$K_{c(30)}$	$K_{p(30)}$	$K_{p(100)}$
SOC	1.00	0.33	-0.32	0.28	0.25	0.10	0.12	-0.12	0.10
L_{binary}	0.33	1.00	-0.42	0.39	0.39	-0.10	0.08	0.06	0.02
SI	-0.32	-0.42	1.00	-0.38	-0.56	-0.08	-0.15	0.10	-0.09
Z_{10}	0.28	0.39	-0.38	1.00	0.59	0.14	0.02	0.00	0.02
Z_{30}	0.25	0.39	-0.56	0.59	1.00	0.15	0.15	-0.09	-0.03
$K_{c(10)}$	0.10	-0.10	-0.08	0.14	0.15	1.00	-0.02	0.02	0.02
$K_{c(30)}$	0.12	0.09	-0.15	0.02	0.15	-0.02	1.00	-0.66	-0.18
$K_{p(30)}$	-0.12	0.06	0.10	0.00	-0.09	0.02	-0.66	1.00	0.13
$K_{p(100)}$	0.10	0.02	-0.09	0.02	-0.03	0.02	-0.18	0.13	1.00

1 **Table 4.** Multiple linear regression model components developed using backward

Coefficients	Value	SE	t-value	Pr(>|t|)
		LandTopo		
Intercept	146.2	26.6	5.5	0.000
L_{binary}	50.0	15.9	3.1	0.002
SI	-0.5	0.2	-2.2	0.034
K_p (30m)	-792.5	517.6	-1.5	0.128
K_c (10m)	114.8	76.7	1.5	0.137
		Topo		
Intercept	-150.6	81.9	-1.8	0.068
Z (10m)	17.6	5.3	3.3	0.001
K_c (30m)	996.7	592.9	1.7	0.095
K_p (100m)	6582.1	4516.7	1.5	0.147

2 stepwise regression procedures in S-plus.
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

1 **Table 5.** Best-fit multiple linear regression models for the prediction of soil C in
2 Hofmann Forest.
3

Name	Model	R^2	RSE
LandTopo	$C = 50(L_{binary}) - 0.5(SI) - 792.5(K_{p(30m)}) + 114.8(K_{c(10m)}) + 146.2$	0.18	84.3
Topo	$C = 17.6(Z_{(10m)}) + 996.7(K_{c(30m)}) + 6582.2(K_{p(100m)}) - 150.6$	0.10	87.7

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

1 **Table 6.** Validation parameters from comparison of geostatistical and explicit
2 spatial models to georeferenced testing locations with known soil C.

Model	RMSE	ME	Std. Deviation
OK	127.6	4.2	25.7
LandTopo	439.1	-4.0	63.3
Topo	480.0	-3.5	69.2

3
4
5
6
7
8

1

CHAPTER 5. FINAL CONCLUSIONS OF SOIL CARBON, LIGHT DETECTING

AND RANGING (LIDAR) AND SPATIAL MODELING AND GEOSTATISTICS

4

1   The results of these studies have brought to lght a number of issues that

2   of particular concern for pedologist and our ability to model soil properties,

3   specifically SOC. In order to produce highly precise, highly accurate DEM we

4   need to continue to resolve computing and statistical constraints that currently

5   prevent the use of LiDAR data as a source of digital elevation data. Currently,

6   LiDAR based DEM are readily available as raster data sets but the methods by

7   which the original bare earth data were interpolated are not disclosed.

8   Interpolation of the bare earth data is vexingly demanding on both time and

9   computer resources. Further work on methods developed in this research may

10   prove fruitful if the data reduction strategies can be replicated on landscapes of

11   various relief. The low relief landscapes of eastern North Carolina served as a

12   reasonable starting point for developing statistically valid data reduction

13   strategies. Further work is need in fine-tuning the level of data reduction possible

14   across a range of DEM of different resolutions. Smaller incremental analysis in

15   both the level of data reduction and the resolution of the DEM produced needs to

16   be completed before firm conclusion can be drawn with regards to the

17   appropriate level of data density required for producing a DEM of any given

18   resolution.

19   Production of explicit spatial models of SOC will continue to be of great

20   interest as more and more scientist focus on balancing the global C budget.

21   Focus will no doubt continue to reside on terrestrial stores, with an emphasis of

22   SOC in temperate forests in the US. Spatial models developed in this study

1   provide some indication as to the complexity of biogeochemical interactions that

2   drive SOC stores. A number of study have successfully model SOC on smaller

3   landscapes and hillslopes, however, modeling of SOC on flat landscapes will

4   continue to pose serious problems to models based on topographic attributes

5   alone. Innovative think with regard to developing new landscape descriptors is

6   needed if pedoligist hope to produce valid models of SOC on these low-relief

7   landscapes. Future research efforts should make a concerted effort to address

8   landscape attributes that deal with the unique hydrological variables the drive flux

9   of SOC in low relief landscapes. Additionally, more attention may be need to

10   devise sampling schemes that help control the variability of SOC samples taken

11   in the field. Better strategies that address other methods of sample stratification

12   should be addressed. Our models were limited in their abilities to adequately

13   explain SOC variability in Hofmann Forest. Possibly, more scrutiny in the

14   stratifying sampling locations may have enhanced our ability to model SOC.

15   Concern about sampling locations is raised primarily because our models

16   produced similar total SOC estimates to those estimated by geostatistical

17   approaches despite the low R^2 values of our models.

18