Estimating aboveground biomass in interior Alaska: statistical methods for coupling remotely sensed data with field observations to improve precision

Chad Babcock

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Dr. Hans-Erik Andersen, Chair

Dr. L. Monika Moskal

Dr. Andrew Finley

Program Authorized to Offer Degree:
School of Environmental and Forest Sciences

University of Washington

Abstract

Estimating aboveground biomass in interior Alaska: statistical methods for coupling
remotely sensed data with field observations to improve precision

Chad Babcock

Chair of the Supervisory Committee:
Dr. Hans-Erik Andersen
United States Forest Service

The United State Forest Service's (USFS) Forest Inventory and Analysis (FIA) program is charged by Congress to monitor the nation's forestlands—survey growth, mortality and composition, among other aspects related to the status and trends of United States (US) forests. To accomplish this, the FIA repeatedly inventories a massive network of sample plots that spans most of the country's forests. Although the FIA's permanent forest inventory network is the largest in the world, it does not represent the entirety of the nation's forested landscape. Interior Alaska contains approximately 15 percent of the nation's total forestland but, until recently, has not been included in the FIA's monitoring efforts. Being one of the most vulnerable ecosystems to climate change in the US, it is pivotal to begin monitoring the dynamics of interior Alaska's forested landscape.

The vastness of the remote Alaskan landscape makes thorough field-only inventories prohibitively expensive—we need new and innovative methods to track forest dynamics in Alaska. The USFS and NASA's Carbon Monitoring System jointly funded a pilot project in 2013 titled *Monitoring Forest Carbon Stocks in Interior Alaska* to examine the potential of airborne and spaceborne remote sensing technologies to augment sparse collections of field data to obtain carbon estimates with acceptable levels of precision. This study will yield first-ever regional estimates of carbon stocks for interior Alaska's Tanana Valley lever-

aging FIA inventory data—an important initial step in the process of developing a forest monitoring system for Alaska's wilderness.

The research summarized in the following dissertation examines several statistical methods for coupling field and remotely sensed information to estimate aboveground biomass and carbon stocks that can potentially be used for tracking forest carbon dynamics of interior Alaska. Both design- and model-based inferential paradigms are considered. This research is strongly focused on the appropriate characterization of uncertainty in the form of standard errors and confidence intervals. Results form this research will help decision makers meet the challenges of environmental change by providing statistically rigorous methodologies that can estimate interior Alaskan carbon stocks. Having reliable and transparent estimates of carbon and biomass stocks can inform carbon management choices about the role Alaskan forestlands play in the global carbon system.

TABLE OF CONTENTS

Page

LIST OF FIGURES

ACKNOWLEDGMENTS

A countless number of individuals have contributed to my education and research, including teachers, family and friends. The road to my PhD has been long and full of twists and turns. I found that, along the way, I could always rely on support and guidance from those around me.

I'll start by thanking my committee chair, Dr. Hans-Erik Andersen. Through his advisement over the past three years, I was introduced to new avenues of research that challenged me to look at situations differently. As a result, he turned me into a more well-rounded researcher with greater appreciation for the many ways problems can be approached.

In addition to the invaluable guidance concerning research, Dr. Monika Moskal helped me understand what it takes to survive and thrive in academia beyond simply doing research. Monika pushed me to better connect and work with a wider research audience. Through her, I learned that effective science communication is an integral part of the research process.

I want to thank other current and former PhD committee members including Dr. Miles Logsdon, Dr. Christian Torgersen, Dr. Peter Hoff, Dr. David Butman and Dr. John Skalski. Conversations I've had and insights I've gained from these individuals shaped my research. I'd also like to thank Dr. Bruce Cook and other researchers at NASA, Dr. Douglas Morton and Dr. Ross Nelson, for providing me the opportunity to pursue interesting science. I couldn't imagine a more productive, exciting and fun group of scientists to work with. I'd like to thank all the students, postdocs and visiting scientists who I've interacted with in the Remote Sensing and Geospatial Analysis Lab (RSGAL) over the past three years. And I'd like to thank Bob McGaughey for his assistance, especially with computing. Coming to the University of Washington as an open source and Linux junky, I found it extremely helpful

to have a remote sensing computing specialist across the hall who was always eager to help.

Since high school I knew that I wanted to be a scientist and conduct research. What I didn't understand, however, was how I might make that a reality. Nearing the end of an exceptionally fun and messy undergraduate experience I cut off the dreadlocks and met with Dr. Andrew Finley. He agreed to advise me for an independent study, marking the beginning of a nine year long mentorship experience that I don't think either one of us saw coming. Andy has taught me more about academic life than can be mentioned. All of my academic achievements thus far stem from his guidance. Andy and his family also saw my family and I through many of life's experiences over the past nine years, including the birth of my son Oren and daughter Emilia. I thank them for all their support through the several life changing events my family has gone through in the past decade.

The inspiration to choose academia as a career originated in my high school physics class. Mr. Evele's energizing approach to education and eagerness to help cement difficult concepts in the classroom woke up a deep passion in me to pursue knowledge just for the sake of it. Mr. Evele taught me to follow my interests and following my interests has led me down the most rewarding life path I could imagine. I want to thank Mr. Evele for his encouragement and showing me how rewarding learning about science can be.

Lia is hands down the most tolerant, supportive and understanding human being I've ever met. And she continues to prove it every day—from listening to me rant about whatever I may have been dealing with scientifically, computationally or socially nearly every evening to providing insights I never would have considered if not for her. I know that without Lia, I never would have earned a PhD. Beginning the day we started dating, Lia has been continually providing encouragement and support. She is always the person I go to first for guidance. For the past twelve years, Lia stood alongside me through every victory and defeat that led to my graduation and I thank her for all her support.

I would also like to thank my mother and father for everything. I don't think anyone can fully appreciate the effort it takes for great parents to raise a child until they are given the opportunity to become a parent themselves. The birth of my son made me immediately aware of how hard my parents worked to provide the opportunities I was given. I thank my mother for providing a safe and stable home. By example, my mom has also instilled in me the value of perseverance. She has shown me that hard work pays off and to not give up if you believe in something. I thank my father for his unwavering support of all my pursuits. Whenever I asked him for help, he did so without hesitation. I model my parenting style based on the example they set for me and I hope that I can be at least as good a parent to Oren and Emilia as they are to me.

I'll end by thanking my kids. Oren and I got to graduate together this summer—him from preschool and I with a PhD. It seems fitting for him to start his education by entering kindergarten the same time I finish mine. Oren's energy, focus and creativity I think are his greatest strengths. I'm so excited to see the direction his life takes. Oren is the only person who actually begged me to go to my defense and I'm pretty sure that no one will ever show that much desire to hear me ramble on about trees and statistics again. I thank him for his support. Millie is super smart. At nine months I can already see that she's going to excel. She's incredibly determined and is constantly observing and testing. I want to thank Millie for sleeping through the night. Having a child at the end of your PhD is adventurous to say the least and if not for Millie's easy-going personality, graduating could have been much more difficult. I hope I can show her and Oren the same level of support as has been shown to me by my teachers, family and friends. Thank you everyone who helped me along the way to earning a PhD. There were many and I'm failing to mention most.

Chapter 1

INTRODUCTION

1.1 National Forest Inventory in the United States

In conjunction with establishing the statutory basis for the management of forests part of the National Forest System, the *Organic Administration Act of 1897* laid out initial provisions for the inventory of national forestlands (Organic Administration Act, 1897). After the passing of *McSweeny-McNary Act of 1928*, the United States Forest Service (USFS) established regional forest survey projects to conduct state-wide inventories (McSweeny-McNary Act, 1928). By the 1960's, each state in the Lower 48 had completed at least one forest survey, with more heavily forested states having conducted multiple state-wide inventories. The *Agriculture Research, Extension and Reform Act of 1998* mandated the annual inventory of forestlands throughout the United States (Agricultural Research, Extension and Education Reform Act, 1998). This Act requires nationally consistent measurement of a core set of forest variables and for the publication of state-level reports every five years. In response to the increased demands for reliable forest inventory and monitoring information expressed in *Agriculture Research, Extension and Reform Act of 1998*, the USFS significantly overhauled its Forest Inventory and Analysis (FIA) program to increase capacity to collect forest inventory information at a national scale, publish data in a timely manner and provide state- and national-level forestland analyses useful to decision makers.

As the nation's forest census, the FIA is responsible for reporting on the status and trends of United States (US) forest resources. To meet this end, the FIA has established a large network of permanent inventory plots than spans forestlands of the US. At plot locations, a suite of vegetation and soil measurements are collected on a five to ten year interval and serve as the basis for FIA reporting on forest area, species distributions, forest health, growth, mortality, forest biomass/carbon stocking and harvest removals among many other aspects of the nation's forests.

The FIA inventory network is the largest plot-based forest survey in the world; currently covering most US forestlands. However, more than 450 000 km^2 of forests in interior Alaska (15% of US forestland) are not presently included in the FIA program, as these remote regions are difficult and expensive to monitor with standard field methods. Most of the region is inaccessible from the existing road network. The remoteness of interior Alaska, combined with its vast size, means that traditional ground sampling for forest inventory is prohibitively expensive. To monitor forests of interior Alaska, new inventory approaches are needed; potentially leveraging ancillary information sources, e.g., remote sensing.

1.2 Forest biomass/carbon stocking in interior Alaska

The interior Alaskan landscape is part of the tiaga-tundra ecotone and contains vast tracts of boreal forests. Over the past few decades, this biome has experienced the most surface warming of any region on Earth (IPCC, 2013). As a result, the boreal environment has been changing rapidly, including loss of permafrost, upslope migration of treelines, moisture-stress related decline in spruce tree growth, increases in pest infestation and a rise in the prevalence and severity of wildfires (Soja et al., 2007; Camill, 2005; Chapin et al., 2000). The rapid climatic shifts in ecosystem processes call into question the tiaga-tundra ecotone's continued ability to act as a global carbon sink (Hadden & Grelle, 2016; Harden et al., 1997).

The threat of climate change has increased demand in reducing the high level of uncertainty surrounding terrestrial carbon stocks and fluxes in order to craft government policies to mitigate climate change (Bradshaw & Warkentin, 2015). Pan et al. (2011) offers estimates that the global boreal forest landscape has stored 499 teragrams (Tg) of carbon each year between 2000 and 2007 but the uncertainty associated with this estimate is ± 83 Tg. Given this high uncertainty and the fact that the tiaga forests of interior Alaska are not even included in this *global* estimate underscores the need for reliable estimates of forest biomass/carbon stocks for interior Alaska from a carbon monitoring perspective.

1.3 Potential for lidar to aid forest inventory in interior Alaska

The use of light detection and ranging (lidar) to gather information on the three-dimensional structure of forests has increased rapidly over the past decade (Wulder et al., 2013). A multitude of studies show strong relationships between metrics derived from lidar and forest biophysical properties (Asner et al., 2009; Babcock et al., 2013; Erdody & Moskal, 2010; Finley et al., 2014b, 2017; Lim et al., 2003; Næsset, 2004, 2011). These findings have encouraged considerable investment in the development of more capable sensors, the collection of data over larger extents and the development of algorithms/models to extract useful information from point clouds (Hyyppä et al., 2008; Kashani et al., 2015; Wulder et al., 2012b; Zhen et al., 2016).

In boreal landscapes, the use of lidar to aid forest inventory is of particular interest due to the high costs associated with field data collection (Lutz et al., 2008). It is often thought that sparse collections of field data could potentially be augmented with cheaply acquired lidar information to improve inventory precision (Wulder et al., 2012b). The collection of lidar over large spatial extents, e.g., interior Alaska, proves problematic however. From an airborne platform, it is financially impossible to collect wall-to-wall coverage lidar data over expansive spatial domains (Andersen et al., 2011; Bolton et al., 2013; Nelson, 2010; Nelson et al., 2004). Further, currently proposed spaceborne lidar systems are not designed to procure wall-to-wall coverage information. Rather, these campaigns are only optimized to collect data for relatively narrow bands along the orbital tracts of the sensors' host satellite (GEDI, 2014; ICESat-2, 2015a). In order to glean any additional information provided by sampled remote sensing data in a statistically rigorous manner, estimation frameworks that can accommodate incomplete coverage auxiliary information are necessary.

The USFS and NASA's Carbon Monitoring System jointly funded a pilot project in 2013 titled *Monitoring Forest Carbon Stocks in Interior Alaska* to examine the potential of airborne and spaceborne remote sensing technologies, including lidar, to augment sparse collections of field data to obtain forest biomass estimates with acceptable levels of precision

(Pattison et al., In prep). This study will yield first-ever regional estimates of carbon stocks for interior Alaska's Tanana Valley—an important initial step in the process of developing a forest monitoring system for Alaska's wilderness.

1.4 Dissertation summary

Chapter 2 presents results of a simulation study comparing design-based, model-assisted and model-based approaches to quantify aboveground biomass in interior Alaska. A synthetic population was derived using real biomass field samples and associated lidar remote sensing data collected over the Tetlin National Wildlife Refuge. For evaluation criteria, estimation efficiency, bias of total aboveground biomass point estimates and nominal versus actual repeated sampling coverage was considered. Results indicated that model-assisted and model-based approaches can successfully leverage lidar ancillary data to improve total aboveground biomass estimation in interior Alaska and produce uncertainty intervals with adequate coverage. Both the two-stage model-assisted and coregionalization model-based estimators proved applicable under a lidar strip sampling scenario and produced similar uncertainty estimates. This simulation study also provided evidence that spatial models fitted within a Bayesian framework produce prediction uncertainty estimates with better empirical coverage than in a frequentist framework due to the ability to account for spatial covariance parameter uncertainty via prior specification.

Simulation study results provided evidence that a coregionalization framework estimated within a Bayesian paradigm of inference could produce estimates of total aboveground biomass with interpretable measures of uncertainty. Chapter 3 was intended to explore more deeply the possibility of this modeling approach to improve estimation precision and facilitate mapping of aboveground biomass density. The goal of this research was to develop and examine the performance of a geostatistical coregionalization modeling approach for combining field inventory measurements, strip samples of airborne lidar and Landsat-based remote sensing data products to predict aboveground biomass in interior Alaska's Tanana Valley. The proposed modeling strategy facilitates pixel-level mapping of aboveground biomass density predictions across the entire spatial domain. Additionally, the coregionalization framework allows for statistically sound estimation of total aboveground biomass for arbitrary areal units within the study area—a key advance to support diverse management objectives in interior Alaska. This research focuses on appropriate characterization of prediction uncertainty in the form of posterior predictive coverage intervals and standard deviations. Using the framework detailed here, it is possible to quantify estimation uncertainty for any spatial extent, ranging from pixel-level predictions of aboveground biomass density to estimates of aboveground biomass stocks for the full domain. The lidar-informed coregionalization models consistently outperformed their counterpart lidar-free models in terms of point-level predictive performance and total aboveground biomass precision. Additionally, the inclusion of Landsat-derived forest cover as a covariate further improved estimation precision in regions with lower lidar sampling intensity. Our findings also demonstrate that model-based approaches that do not explicitly account for residual spatial dependence can grossly underestimate uncertainty, resulting in falsely precise estimates of aboveground biomass. On the other hand, in a geostatistical setting, residual spatial structure can be modeled within a Bayesian hierarchical framework to obtain statistically defensible assessments of uncertainty for aboveground biomass estimates. The inferential capabilities of aboveground biomass posterior predictive distribution (PPD) products extend beyond simply mapping aboveground biomass density. PPD products can provide insight regarding drivers of biomass heterogeneity in boreal forests, including permafrost and fire, highlighting the range of potential applications for Bayesian geostatistical methods to integrate field, airborne and satellite data.

The methods explored in Chapters 2 and 3 are applicable when field observations and lidar data are collected in the same season. Given the FIA's rotating panel design for the collection of field measurements, we need statistical modeling strategies to account for temporal misalignment between field observations and lidar data. Chapter 4 explores a possible solution to the temporal misalignment problem using a Bayesian geostatistical approach. We move to Penobscot Experimental Forest near Bradley, Maine where a spatio-temporally rich set field observations is available. In Chapter 4, we present and examine a novel modeling approach to improve prediction of aboveground biomass and estimate aboveground biomass growth using lidar data. The proposed model accommodates temporal misalignment between field measurements and remotely sensed data by including multiple time-indexed measurements at plot locations to estimate aboveground biomass growth. We pursue a Bayesian modeling framework that allows for appropriately complex parameter associations and uncertainty propagation through to prediction. Specifically, we identify a space-varying coefficients model to predict and map aboveground biomass and its associated growth simultaneously. The proposed model is assessed using lidar data acquired from NASA Goddard's LiDAR, Hyper-spectral & Thermal imager and field inventory data from the Penobscot Experimental Forest in Bradley, Maine. The proposed model outperformed the time-invariant counterpart models in predictive performance as indicated by a substantial reduction in root mean squared error. The proposed model adequately accounts for temporal misalignment through the estimation of forest aboveground biomass growth and accommodates residual spatial dependence. Results from this analysis suggest that future aboveground biomass models informed using remotely sensed data, such as lidar, may be improved by adapting traditional modeling frameworks to account for temporal misalignment and spatial dependence using random effects.

The development of statistically rigorous approaches to monitor forestlands in interior Alaska is critically important. As the effects of climate change on this vulnerable ecosystem become increasingly apparent, the need for reliable assessments of how the boreal environment is adapting intensifies. The aim of this dissertation is to provide novel insights about cost-effective, remote sensing-assisted forest parameter estimation approaches applicable to remote boreal regions, such as interior Alaska. Results from the research endeavors presented here can aid in the further development of forest monitoring procedures for Alaska's vast forested landscape.

Chapter 2

ON DESIGN-BASED, MODEL-ASSISTED AND MODEL-BASED ESTIMATION OF TOTAL ABOVEGROUND BIOMASS IN INTERIOR ALASKA

Abstract: We used simulated sampling from a realistic forest aboveground biomass (AGB) synthetic population to compare design-based, model-assisted and model-based methods for estimating total AGB in interior Alaska. The synthetic population was derived using real AGB field samples and associated lidar remote sensing data collected over the Tetlin National Wildlife Refuge. For evaluation criteria we considered estimation efficiency, bias of total AGB point estimates and nominal versus actual repeated sampling coverage. Results indicate that model-assisted and model-based approaches can successfully leverage lidar ancillary data to improve total AGB estimation in interior Alaska and produce uncertainty intervals with adequate coverage. Both the two-stage model-assisted and coregionalization model-based estimators proved applicable under a lidar strip sampling scenario and produced similar uncertainty estimates. This simulation study also provides evidence that spatial models fitted within a Bayesian framework produce prediction uncertainty estimates with better empirical coverage than in a frequentist framework due to the ability to account for spatial covariance parameter uncertainty via prior specification.

2.1 Introduction

Remote sensing information, e.g., lidar, radar and spectral data, offers the potential to improve our understanding about the status and trends of forest resources in isolated boreal landscapes (Park et al., 2016; Praks et al., 2007; Neigh et al., 2013). Concerns related to climate change have raised interest in the ability to monitor forest aboverground biomass

(AGB) and carbon dynamics in the boreal ecotone, considering that global warming is projected to have major impacts in this region (Balshi et al., 2009; Gauthier et al., 2015). By increasing estimation precision, remote sensing-assisted AGB inventory of tiaga-tundra regions can lead to greater confidence about the role subartic forests play in the global carbon cycle (Lutz et al., 2008; Margolis et al., 2015).

Many studies have showcased the effectiveness of lidar remote sensing to aid in AGB estimation (Babcock et al., 2015, 2016; Næsset, 2011). Lidar's ability to capture the vertical structure of forests has made it one of the most effective remote sensing technologies to aid AGB accounting (Zolkos et al., 2013). With either a complete census or sufficiently large sample of lidar data, subsequent AGB inventory precision can potentially be vastly increased with little additional cost (Wulder et al., 2012b,a). Increases in AGB estimation efficiency due to the incorporation of lidar can potentially allow for reductions in costly field sample sizes while retaining acceptable levels of precision (Gregoire et al., 2011; Junttila et al., 2013; Junttila & Laine, 2017).

National forest inventory programs, including the United States Forest Service's (USFS) Forest Inventory and Analysis (FIA) program, typically rely on design-based (classical) sampling and estimation practices to gain insight about forest inventory parameters, e.g., AGB (Bechtold & Patterson, 2005; Gschwantner et al., 2016; Haywood & Stone, 2017). Design-based approaches are attractive due to their low reliance on distributional assumptions and asymptotic guarantees with respect to estimation bias (Gregoire, 1998; Ver Hoef, 2002). Given a sample of observations with known probabilities of inclusion, unbiased estimators for means or totals and associated variances can be derived, e.g., using the Horvitz-Thompson method (Horvitz & Thompson, 1952).

Design-based forest inventory becomes challenging when the incorporation of auxiliary information, e.g., lidar data, is considered. Concerns include (1) difficulties associated with the use of models to relate field and ancillary data; (2) the inability to collect full spatial coverage high quality remote sensing information for large isolated landscapes; and (3) challenges related to implementing probabilistic field (remote sensing) data sampling. Many avenues have been explored that attempt to overcome these hurdles. There is a growing body of research supporting the use of model-assisted estimation in remote sensing-aided forest inventory (Strunk et al., 2012; Kangas et al., 2016). These approaches allow the use of models to leverage relationships between field and remote sensing information while preserving favorable design-based statistical properties (Ståhl et al., 2016). Research has been published detailing the use of model-assisted estimation for a variety of remote sensing data collection scenarios including wall-to-wall and flight-line strip sampling schemes (Næsset et al., 2011; Gregoire et al., 2011; Saarela et al., 2015). Several model-based approaches have also been proposed to relate remote sensing and forest inventory datasets to improve estimation precision (Saarela et al., 2016; Ståhl et al., 2011).

Model-assisted and model-based estimation frameworks both use models, e.g., linear regressions, to relate field and ancillary data. The two approaches diverge in how estimation uncertainty is characterized (Särndal et al., 1978). Model-assisted estimation techniques assess uncertainty via analyzing model residuals using design-based variance estimators (Särndal et al., 1992). Therefore, as in design-based estimation, field (remote sensing) sample locations need to be obtained probabilistically for assurances regarding estimation bias or confidence interval coverage. Model-based approaches do not require probabilistic sampling of population units, but also do not guarantee unbiased estimation of means or totals in the same way design-based or model-assisted approaches do (Schreuder et al., 2001). Model-based inference also requires one to make assumptions about the distribution of the data to be modeled when interpreting uncertainty estimates and it is often impossible to verify that these assumptions are sufficiently met (Koch & Gillings, 1983; Sterba, 2009). However, the increased flexibility of model-based inference—afforded, in part, by relaxing the probabilistic sampling requirement—has fueled research on the use of this inferential paradigm in survey work (Gelman, 2007; Magnussen, 2015; Snijders & Bosker, 2012).

Within the model-based paradigm, both frequentist and Bayesian statistical modes of inference have received attention in the forest inventory literature (Kauranne et al., 2017; Junttila et al., 2008; Babcock et al., 2013; McRoberts et al., 2005; Chen et al., 2016). McRoberts et al. (2013) compared model-assisted and model-based approaches, fitted using frequentist methods, to estimate mean growing stock volume and found that the tested model-based approach produced lower standard errors than the design-based and model-assisted estimators they examined. Babcock et al. (2015) explored Bayesian spatial model-based approaches to map AGB density with associated uncertainty and found that models with space varying intercepts produced better predictions than non-spatial counterpart models at three of the four study sites.

Given the multitude of potential approaches for estimating AGB with uncertainty, all stemming from differing modes of statistical inference, it is important to assess the tradeoffs between strategies. To examine the statistical properties of design-based, model-assisted and model-based candidate approaches for AGB estimation in boreal landscapes, we conduct a simulation study using a realistic forest AGB and lidar synthetic population. The synthetic population is derived leveraging AGB field samples and lidar remote sensing data collected over the Tetlin National Wildlife Refuge (TNWR) in interior Alaska in 2014 as part of the Tanana Inventory Unit (TIU) pilot project (Pattison et al., In prep). For evaluation criteria we consider estimation efficiency, repeated sampling bias of total AGB point estimates and nominal versus actual uncertainty interval coverage. Specifically, we examine three classes of total AGB estimators: 1) design-based/model-assisted; 2) non-spatial model-based; and 3) spatial model-based. In addition, we estimate parameters of the non-spatial and spatial model-based estimators using frequentist and Bayesian modes of inference. We examine these classes of estimators under two different lidar data acquisition scenarios: 1) no lidar collected (**Case A**); and 2) wall-to-wall lidar acquired (**Case B**). Additionally, we evaluate a model-assisted approach for a flight-line strip lidar sampling scenario (**Case C**). We repeat the Monte Carlo sampling analysis under four different plot spacings to gain insight about the role of field sample size on estimation efficiency and confidence interval coverage. Computational limitations preclude repeated sampling evaluation of the candidate spatial model-based approaches applicable to the **Case C** sampling scheme. Due to this, we instead estimate total AGB using the real field and lidar data collected over TNWR using the

Bayesian coregionalization spatial model-based approach and the two-stage model-assisted estimator to compare total AGB and associated uncertainty estimates.

This study is intended to provide insight regarding the applicability of design-based, model-assisted and model-based approaches to estimate total AGB in boreal landscapes under various data collection schemes. Conclusions resulting from this analysis can inform decisions about how subartic forestlands are inventoried by offering insight about the effects of field plot spacing, lidar sampling intensity and the suitability of estimation procedures from differing statistical paradigms.

2.2 Synthetic population

2.2.1 Tetlin National Wildlife Refuge

Tetlin National Wildlife Refuge (TNWR) is nearly 300 000 ha in size with lowland areas characterized by extensive wetlands and poorly drained soils. Wet upland sites are home to black spruce (*Picea mariana*) forests whereas drier landscapes favor white spruce (*Picea glauca*). Deciduous species including quaking aspen (*Populus tremuloides*), paper birch (*Betula papyrifera*) and balsam poplar (*Populus balsamifera*) persist on well-drained south-facing slopes. Shrub vegetation consisting of willow (*Salix spp.*), alder (*Alnus spp.*) and dwarf birch (*Betula spp.*) can be found in lowland areas around water bodies (U.S. Fish and Wildlife Service, 2016).

2.2.2 Field and remote sensing data

At TNWR, a spatially-balanced systematic design was implemented using a tessellation of hexagons covering the study site. The hexagons were approximately 12 141 ha (30 000 acres) in size. In the summer of 2014, FIA field crews established standard FIA plots at the center of each hexagon complete with four subplots; one at the midpoint and three additional subplots extended radially approximately 36 meters (m) at 0°, 120° and 240°. Each subplot has an approximate 7.3 m radius and an area of 168.11 m^2. Field measurements were taken using an augmented FIA inventory design (Bechtold & Patterson, 2005). Pattison et al. (In prep) provide detailed field protocols for the TIU pilot project. AGB for individual trees on subplots were tabulated using the Component Ratio Method described in Woodall et al. (2015). AGB for all trees on a subplot with breast height diameters 2.5 cm and larger were scaled to megagrams per ha (Mg/ha) and summed to obtain subplot-level AGB density.

Lidar data were collected using a flight-line strip sampling approach with Goddard's LiDAR, Hyperspectral, and Thermal (G-LiHT) imager (Cook et al., 2013). G-LiHT is a portable multi-sensor system that can be mounted to a fixed wing aircraft. G-LiHT's on-board laser altimeter (VQ-480, Riegl Laser Measurement Systems, Horn, Austria) provides an effective measurement rate of up to 150 kHz along a 60° swath perpendicular to the flight direction using a 1 550 nm laser. At a nominal flying altitude of 335 m, laser pulse footprints have an approximate 10 cm diameter. The instrument is capable of producing up to 8 returns per pulse. Point cloud information was summarized to a 13 × 13 m pixel size to approximate field subplot areas. Over each grid cell, percentile heights were calculated at 10 percent intervals ranging from 10 percent to 100 percent. Maximum height (100^{th} percentile height) relative densities were also calculated at 10 equal width intervals. Identical lidar metrics were obtained using point clouds extracted over each subplot. G-LiHT data for the study area are available online (G-LiHT, 2016).

2.2.3 Synthetic Population generation

Metrics generated using the lidar strip samples collected over TNWR were resampled to produce a realistic synthetic wall-to-wall population of lidar derived variables. Due to the high degree of multicollinearity between lidar metrics an orthogonalization approach was performed which allowed for resampling of the lidar data acquired while preserving the inherent relationships between the different metrics. After orthogonalization, exponential spatial autocovariance functions were fit to each of the orthogonal lidar variables. The fitted spatial variance parameters (nugget, partial sill and range) were then used to simulate a multivariate normal spatial random field covering the target domain for each variable. The uncorrelated lidar metrics were then resampled with replacement. Coordinate locations for the resampled lidar variables were assigned according the rank order of the spatial random field realization. After all orthogonal lidar metrics were resampled and assigned to a location, the resampled set was transformed back to the basis of the original lidar data. The synthetic population percentile height metrics for each pixel were then reordered to ensure that 10^{th} percentile height $< 20^{th}$ percentile height $< \ldots < 100^{th}$ percentile height. The relative densities for each pixel in the synthetic population were also normalized to sum to one. This process helped to transfer the correlations between variables and the second order spatial structure of the lidar dataset to the synthetic population. The mathematical details associated with the synthetic population generation process are given in the appendix.

The synthetic population was simulated at a 13 x 13 m spatial resolution so pixel areas approximate FIA subplot size (168.2 m^2). Figs. 2.1 and 2.2 show statistical summaries of the synthetic and real datasets. It can be seen that the marginal distributions of 90^{th} percentile height ($P90$) and AGB in the synthetic dataset closely resemble that of the real data. We also see that the second order spatial structure of the real $P90$ metric is well represented in the synthetic population. The images of the lidar correlation matrix for the real and synthetic data are similar, indicating that the inherent correlations between lidar metrics transferred effectively to the synthetic population.

2.3 Simulated sampling

For all sampling scenarios, AGB samples of the synthetic population were collected using a procedure that mimics the process used for plot selection in the TIU project detailed in Section 2.2.2. Center points for pixel clusters, designed to represent the typical FIA subplot layout, were selected using a systematic sampling scheme by selecting the centroids of a randomly placed hexagonal grid. Fig. 2.3a shows an example pixel cluster selection. At each pixel cluster location, four pixels were selected to be part of the simulated sample. An example pixel cluster configuration is shown in the inset image of Fig. 2.3a. The midpoints the central and northern pixels are separated by 39 m (two pixels). The pixels to the south

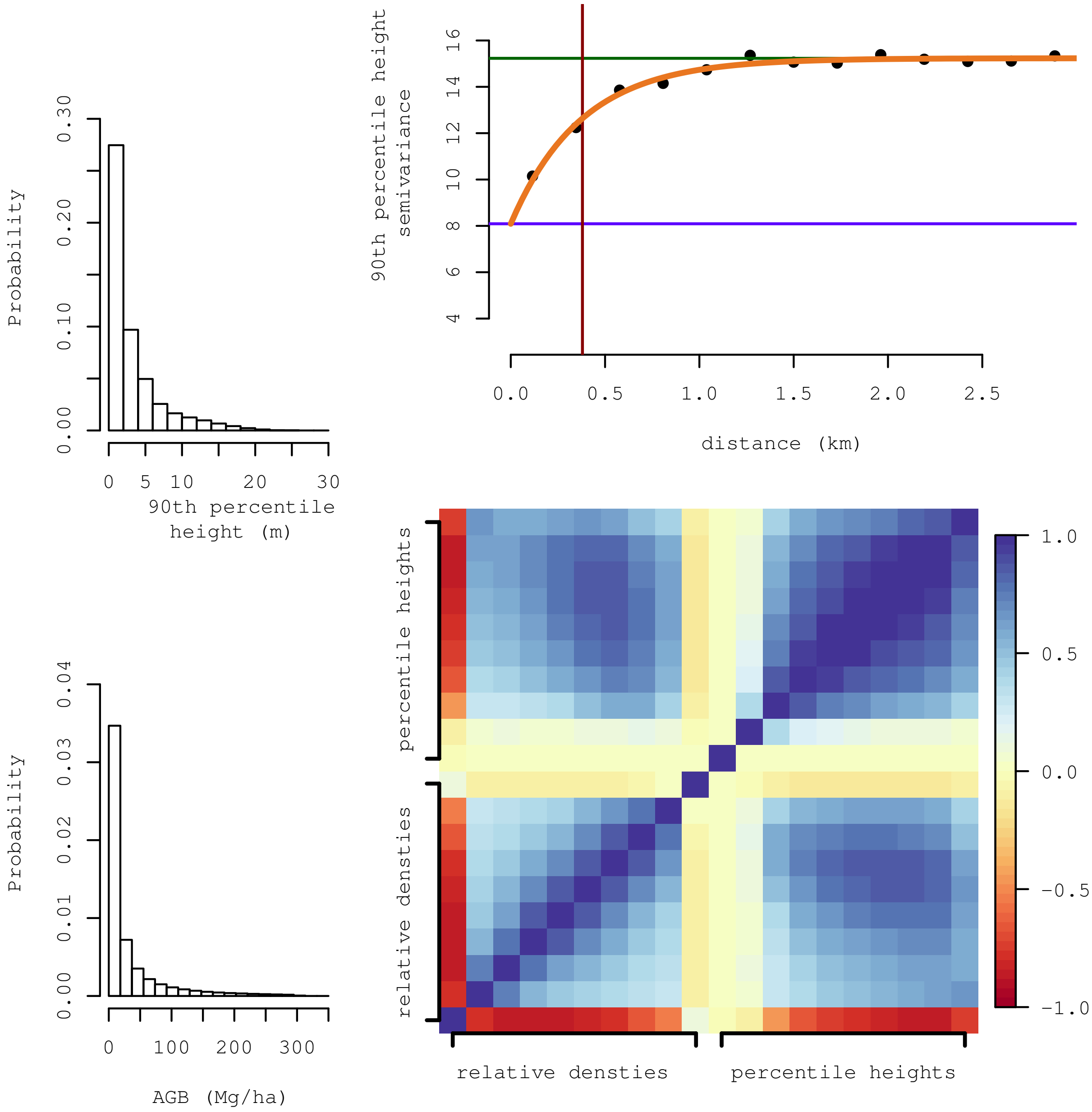


Fig. 2.1: Summary figures showing statistical characteristics of synthetic population. The semivariogram plot (top right) depicts the second order spatial dependence structure of the synthetic 90^{th} percentile height lidar metric. The green, blue and red lines portray the partial sill, nugget and range estimates respectively. An image of the correlation matrix is shown in the bottom right image. The top-left and bottom-left histograms show the marginal distribution of 90^{th} percentile height and aboveground biomass (AGB) respectively.

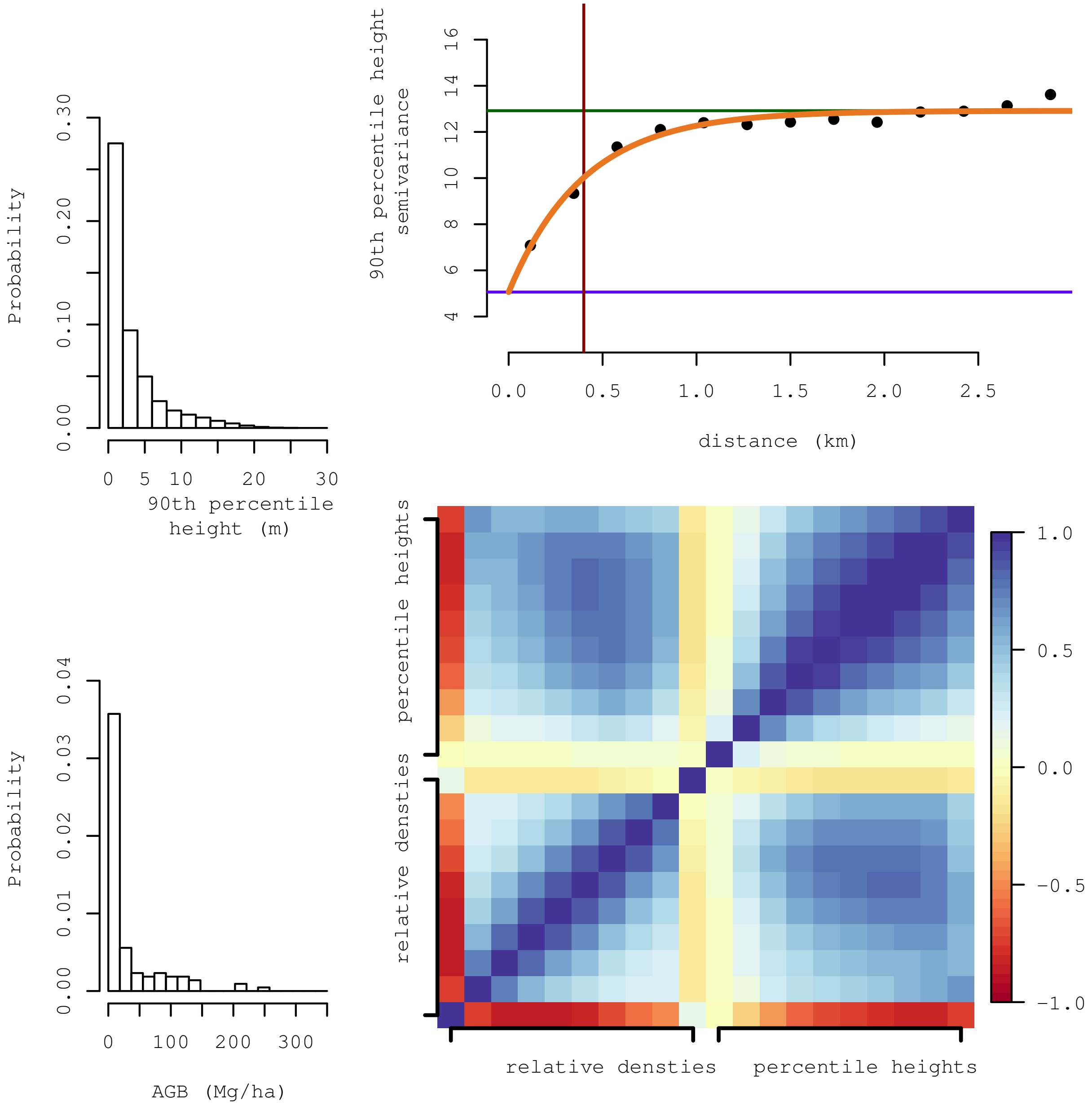


Fig. 2.2: Summary figures showing statistical characteristics of data collected of Tetlin National Wildlife Refuge as part of the Tanana Inventory Unit pilot project. The semivariogram plot (top right) depicts the second order spatial dependence structure of observed 90^{th} percentile height. The green, blue and red lines portray the partial sill, nugget and range estimates respectively. An image of the correlation matrix is shown in the bottom right image. The top-left and bottom-left histograms show the marginal distribution of 90^{th} percentile height and aboveground biomass (AGB) respectively.

east and south west are extended approximately 37 m from the central pixel (separated by one pixel along the diagonal). Simulated samples were selected using four different pixel cluster spacings; 10, 7.5, 5 and 2.5 km. The average number of pixel clusters selected for the repeat samples was 33, 59, 137 and 547 for 10, 7.5, 5 and 2.5 km pixel cluster spacings respectively.

2.4 Model definition

We define a general spatial model for the prediction of AGB density (Mg/ha) as,

$$y(\boldsymbol{s}) = \boldsymbol{x}(\boldsymbol{s})'\boldsymbol{\beta} + w(\boldsymbol{s}) + \epsilon(\boldsymbol{s}), \tag{2.1}$$

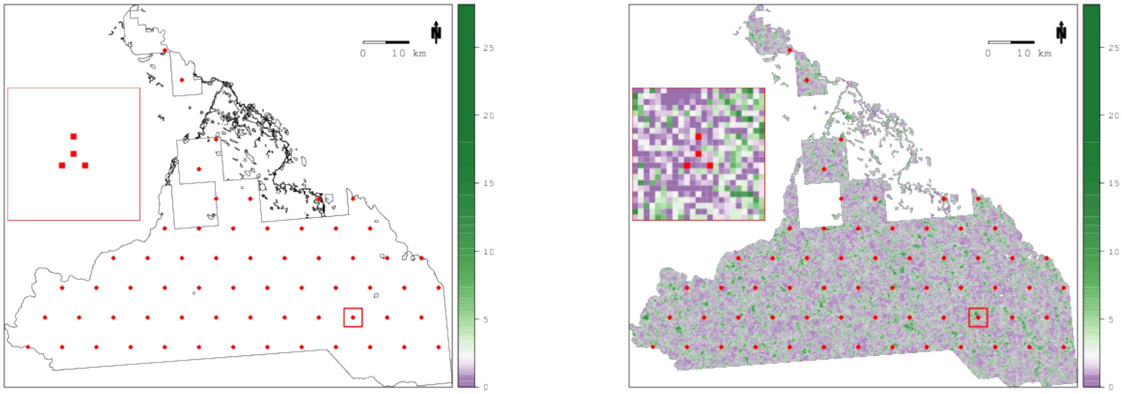
where $y(\boldsymbol{s})$ is AGB density at location $\boldsymbol{s} \in \mathcal{D} \subseteq \mathbb{R}^2$ and $\mathcal{D}$ is the spatial domain. The vector $\boldsymbol{x}(\boldsymbol{s})$ is $p \times 1$ and composed of a 1 and $p-1$ lidar covariates. The term $\boldsymbol{x}(\boldsymbol{s})'\boldsymbol{\beta}$ models the linear relationship between the outcome variable, $y(\boldsymbol{s})$, and lidar variables contained in $\boldsymbol{x}(\boldsymbol{s})$ through a vector of regression coefficients, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})'$. The $w(\boldsymbol{s})$ term absorbs any spatially structured variability in $y(\boldsymbol{s})$ after accounting for $\boldsymbol{x}(\boldsymbol{s})'\boldsymbol{\beta}$ and $\epsilon(\boldsymbol{s})$ captures residual non-spatial variability.

We superimpose a two-dimensional square lattice over the continuous domain $\mathcal{D}$ and define $\boldsymbol{S} = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N\}$ as the collection of coordinates for the lattice nodes. The set $\boldsymbol{S}$ constitutes a finite population of N units completely tessellating domain $\mathcal{D}$. Let $\boldsymbol{y} = (y(\boldsymbol{s}_1), y(\boldsymbol{s}_2), \ldots, y(\boldsymbol{s}_N))'$ and $\boldsymbol{X} = (\boldsymbol{x}(\boldsymbol{s}_1), \boldsymbol{x}(\boldsymbol{s}_2), \ldots, \boldsymbol{x}(\boldsymbol{s}_N))'$ be the collection of AGB densities and covariates for the finite population defined by the superimposed lattice, respectively. The continuous population model defined in (2.1) can be expressed in terms of the finite lattice as,

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{w} + \boldsymbol{\epsilon}, \tag{2.2}$$

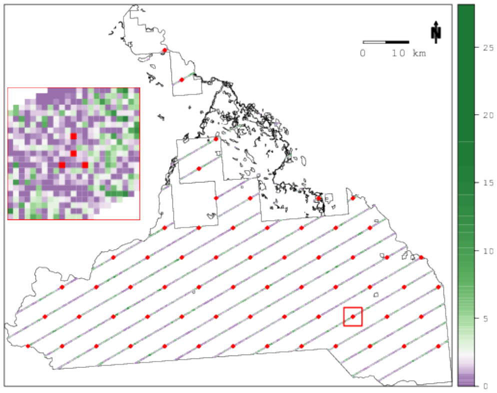
where $\boldsymbol{w} = (w(\boldsymbol{s}_1), w(\boldsymbol{s}_2), \ldots, w(\boldsymbol{s}_N))'$ and $\boldsymbol{\epsilon} = (\epsilon(\boldsymbol{s}_1), \epsilon(\boldsymbol{s}_2), \ldots, \epsilon(\boldsymbol{s}_N))'$. Total AGB for domain $\mathcal{D}$ is then given by,

$$Y = |\mathcal{D}| \left(\sum_{i=1}^{N} \frac{y(\boldsymbol{s}_i)}{N} \right), \tag{2.3}$$



(a) Case A: No lidar

(b) Case B: Wall-to-wall lidar



(c) Case C: Strip sampled lidar

Fig. 2.3: Example figures depicting the three sampling scenarios tested. Subfigures show sampled plots on a hexagonal grid in red. The boundary of Tetlin National Wildlife Refuge (TNWR) is shown in black. The red square boundary within TNWR outlines the boundary of the inset image, shown to the left of TNWR. In the inset image, it can be seen that each plot location consists of four subplot pixels selected to approximate FIA design. Each pixel is 13 x 13 m which approximates the area of a typical FIA subplot (7.3 m radius). Each subfigure displays a map of the 90^{th} percentile height data (m) used to aid estimation. Subfigure (a) shows the Case A scenario where no lidar data was sampled. Subfigure (b) shows the Case B scenario where wall-to-wall lidar information was used. Subfigure (c) shows the Case C scenario where lidar strip samples were acquired.

where $|\mathcal{D}|$ is the area of $\mathcal{D}$ measured in ha.

Given a probability sample $\boldsymbol{y}_s = (y(\boldsymbol{s}_1), y(\boldsymbol{s}_2), \ldots, y(\boldsymbol{s}_n))'$, where n is the sample size, inference about Y can be made within a design- or model-based statistical paradigm. The following sections detail inferential approaches from both paradigms under three different lidar covariate availability scenarios; (**Case A**) no lidar information is observed; (**Case B**) $P90$ is observed for all population units; and (**Case C**) $P90$ is observed for a sample of h flight-line strips. Fig. 2.3 shows examples of the three lidar availability scenarios.

2.5 Model-based inference under the Case A and Case B lidar availability scenarios

Under a model-based perspective, the modeler assumes $\boldsymbol{y}$ to be a realization from a random process, hence $\boldsymbol{y}_s$ is treated as a random vector. The source of randomness is provided by the model, which describes the probability distribution of $\boldsymbol{y}_s$. This requires the modeler to assign probability distributions for any stochastic components in the super-population model (2.2). In Sections 2.5.1 and 2.5.2 we detail the estimation procedure for the frequentist and Bayesian candidate spatial models, respectively.

2.5.1 Frequentist inference

For the frequentist model-based approaches, we only treat $\boldsymbol{w}$ and $\boldsymbol{\epsilon}$ in (2.2) as stochastic vectors. Defining $\boldsymbol{\gamma} = \boldsymbol{w} + \boldsymbol{\epsilon}$, we assign $\boldsymbol{\gamma}$ to follow a multivariate normal distribution, $\boldsymbol{\gamma} \sim MVN(\mathbf{0}, \boldsymbol{C})$. The spatial covariance matrix

$$\boldsymbol{C} = \sigma^2 \boldsymbol{R}(\phi) + \tau^2 \boldsymbol{I}, \tag{2.4}$$

where $(\boldsymbol{R}(\phi))_{ij} = \rho(\phi; d_{ij})$. $d_{ij} = ||\boldsymbol{s}_i - \boldsymbol{s}_j||$ is the euclidean distance between locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ and ρ is a valid spatial correlation function. Here, we assume $\rho(\phi; d_{ij}) = \exp(-\phi d_{ij})$, which is a matérn-class spatial correlation function proven to be strictly positive definite when observation locations are distinct; hence $\boldsymbol{C}$ is invertible (?). $\boldsymbol{C}, \boldsymbol{R}(\phi)$ and the identity matrix $\boldsymbol{I}$ are all $N \times N$ matrices. The variable σ^2 captures the portion of the overall variability in $\boldsymbol{\gamma}$ that is spatially structured while τ^2 absorbs the corresponding non-spatial component. The spatial decay parameter, ϕ, controls how far-reaching the spatial dependence structure is.

To fit model (2.2) we need to estimate $\boldsymbol{\beta}, \phi, \sigma^2$ and τ^2. Given the sample set $\boldsymbol{y}_s$ and associated covariate matrix $\boldsymbol{X}_s$, restricted maximum likelihood (REML) can be used to estimate the parameters of the autocovariance matrix, $\boldsymbol{C}$, i.e., ϕ, σ^2 and τ^2. For this analysis, we used the `SpatialTools` package written for the `R` statistical computing environment (French, 2015). Once REML estimates of ϕ, σ^2 and τ^2 are obtained, they can be used to populate the elements of $\boldsymbol{C}$. Generalized least squares (GLS) estimates for $\boldsymbol{\beta}$ are obtained using,

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}_s' \hat{\boldsymbol{C}}_s^{-1} \boldsymbol{X}_s)^{-1} \boldsymbol{X}_s' \hat{\boldsymbol{C}}_s^{-1} \boldsymbol{y}_s, \tag{2.5}$$

where $\hat{\boldsymbol{C}}_s$ is the estimated spatial covariance matrix between the sample locations, i.e., $(\hat{\boldsymbol{C}}_s)_{ij} = \hat{\sigma}^2(\exp(-\hat{\phi} d_{ij})) + (\hat{\tau}^2 \boldsymbol{I})_{ij}$, where $\hat{\sigma}^2$, $\hat{\phi}$ and $\hat{\tau}^2$ are REML estimates of σ^2, ϕ and τ^2, respectively. Note that we are treating $\hat{\phi}, \hat{\sigma}^2$ and $\hat{\tau}^2$ as *plug-in* estimates and assuming these parameters to be known *without error*. This is a common assumption implied when fitting spatial models within a frequentist framework and means that uncertainty about the spatial covariance parameters is not accounted for in subsequent inference using this approach. The expected block-kriging prediction of total AGB for domain $\mathcal{D}$ is given by

$$\hat{Y} = |\mathcal{D}| \sum_{i=1}^{N} \left(\boldsymbol{x}(\boldsymbol{s}_i)' \hat{\boldsymbol{\beta}} + (\hat{\boldsymbol{c}}(\boldsymbol{s}_i)' \hat{\boldsymbol{C}}_s^{-1})' (\boldsymbol{y}_s - \boldsymbol{X}_s \hat{\boldsymbol{\beta}}) \right) / N \tag{2.6}$$

where $\hat{\boldsymbol{c}}(\boldsymbol{s}_i)$ is an $n \times 1$ vector of estimated covariances between the observed outcomes ($\boldsymbol{y}_s$) and the unvisited location ($y(\boldsymbol{s}_i)$). The variance estimate associated with $\hat{Y}$ is given by

$$\begin{aligned} \widehat{Var}(\hat{Y}) = & |\mathcal{D}|^2 \sum_{i=1}^{N} ((\hat{\sigma}^2 + \hat{\tau}^2) - (\boldsymbol{c}(\boldsymbol{s}_i)' \hat{\boldsymbol{C}}_s^{-1} \boldsymbol{c}(\boldsymbol{s}_i)) + \\ & (\boldsymbol{x}(\boldsymbol{s}_i) - \boldsymbol{X}_s' \hat{\boldsymbol{C}}_s^{-1} \boldsymbol{c}(\boldsymbol{s}_i))' (\boldsymbol{X}_s' \hat{\boldsymbol{C}}_s^{-1} \boldsymbol{X}_s)^{-1} (\boldsymbol{x}(\boldsymbol{s}_i) - \boldsymbol{X}_s' \hat{\boldsymbol{C}}_s^{-1} \boldsymbol{c}(\boldsymbol{s}_i))) / N. \end{aligned} \tag{2.7}$$

A 95 percent uncertainty interval is constructed as,

$$\hat{Y}_{CI} = \left(\hat{Y} - \Phi^{-1}(1 - .95/2) \sqrt{\widehat{Var}(\hat{Y})}, \hat{Y} + \Phi^{-1}(1 - .95/2) \sqrt{\widehat{Var}(\hat{Y})} \right), \tag{2.8}$$

where $\Phi^{-1}(\cdot)$ is the quantile function of the standard normal distribution.

2.5.2 Bayesian inference

For the Bayesian model-based approach, we treat all unknown variables stochastically. In a Bayesian context, model (2.2) can be viewed as a hierarchical model. Given the observed sample, $\boldsymbol{y}_s$, the posterior distribution of the unknown model parameters is,

$$p(\boldsymbol{\beta}, \phi, \sigma^2, \tau^2, \boldsymbol{w}_s | \boldsymbol{y}_s) \propto p(\boldsymbol{y}_s | \boldsymbol{\beta}, \tau^2, \boldsymbol{w}_s) p(\boldsymbol{w}_s | \phi, \sigma^2) p(\boldsymbol{\beta}) p(\phi) p(\sigma^2) p(\tau^2). \tag{2.9}$$

We set $p(\boldsymbol{y}_s | \boldsymbol{\beta}, \tau^2, \boldsymbol{w}_s)$ to be $MVN(\boldsymbol{X}_s \boldsymbol{\beta} + \boldsymbol{w}_s, \tau^2 \boldsymbol{I}_{n \times n})$. The distribution $p(\boldsymbol{w}_s | \phi, \sigma^2)$ is set to $MVN(\boldsymbol{0}, \sigma^2 \boldsymbol{R}_s(\phi))$, where $\boldsymbol{R}_s(\phi))$ is the spatial correlation matrix for the sample. A vague prior distribution for $\boldsymbol{\beta}$, $p(\boldsymbol{\beta})$, is set as $MVN(\boldsymbol{0}, 10\,000 \boldsymbol{I}_{p \times p})$. The prior $p(\phi)$ is set to a uniform distribution with support over the geographic range of the simulated population to be sufficiently non-informative. The priors $p(\sigma^2)$ and $p(\tau^2)$, are set to $IG(2, \hat{\sigma}^2)$ and $IG(2, \hat{\tau}^2)$, where $IG(a, b)$ is the inverse-gamma distribution and $\hat{\sigma}^2$ and $\hat{\tau}^2$ are REML estimates for σ^2 and τ^2, respectively. Setting the IG parameter a to 2 results in a prior distribution with a central tendency equal to b and infinite variance, establishing these priors as vague. We choose to use sample-based inference by constructing a Markov chain Monte Carlo (MCMC) algorithm as described in Banerjee et al. (2014) to sample from $p(\boldsymbol{\beta}, \phi, \sigma^2, \tau^2, \boldsymbol{w}_s | \boldsymbol{y}_s)$. Models estimated using Bayesian methods were fitted using the `spBayes` package written for the `R` statistical computing environment (Finley et al., 2015a).

The posterior predictive distribution of the spatial random effect, $\boldsymbol{w}$, is given by,

$$p(\boldsymbol{w} | \boldsymbol{y}_s) \propto \int p(\boldsymbol{w} | \boldsymbol{w}_s, \boldsymbol{\Omega}, \boldsymbol{y}_s) p(\boldsymbol{w}_s | \boldsymbol{\Omega}, \boldsymbol{y}_s) p(\boldsymbol{\Omega} | \boldsymbol{y}_s) \, \mathrm{d}\boldsymbol{\Omega} \boldsymbol{w}_s, \tag{2.10}$$

where $\boldsymbol{\Omega} = \{\boldsymbol{\beta}, \phi, \sigma^2, \tau^2\}$. The integral in (2.10) does not need to be evaluated explicitly. Given L posterior samples from $p(\boldsymbol{\beta}, \phi, \sigma^2, \tau^2, \boldsymbol{w}_s | \boldsymbol{y}_s)$, this distribution can be obtained via composition sampling by drawing $\boldsymbol{w}^{(l)}$ for each l from $p(\boldsymbol{w} | \boldsymbol{w}_s, \boldsymbol{\Omega}, \boldsymbol{y}_s)$, where this last distribution is derived as a conditional distribution from a multivariate normal and, hence, is multivariate normal. The realizations at the unobserved locations are conditionally independent of the observed outcomes given the realizations over the observed locations and the model parameters. In other words, $p(\boldsymbol{w}|\boldsymbol{w}_s, \boldsymbol{\Omega}, \boldsymbol{y}_s) = p(\boldsymbol{w}|\boldsymbol{w}_s, \boldsymbol{\Omega})$, which is a multivariate normal distribution with mean and variance given by

$$\mathrm{E}\left[\boldsymbol{w}|\boldsymbol{w}_s, \boldsymbol{\Omega}\right] = \boldsymbol{R}(\phi)'\boldsymbol{R}_s(\phi)^{-1}\boldsymbol{w}_s \tag{2.11}$$

and

$$\mathrm{Var}\left[\boldsymbol{w}|\boldsymbol{w}_s, \boldsymbol{\Omega}\right] = \sigma^2\left\{\boldsymbol{R}_s(\phi) - \boldsymbol{R}^*(\phi)'\boldsymbol{R}_s(\phi)^{-1}\boldsymbol{R}^*(\phi)\right\}, \tag{2.12}$$

where $\boldsymbol{R}^*(\phi)$ is a $n \times N$ matrix with $\boldsymbol{R}^*_{ij}(\phi)$ holding the spatial correlation between observed sample i and prediction location j. Then, given $\boldsymbol{x}(\boldsymbol{s}_i)$, samples from the posterior predictive distribution of AGB density at location $\boldsymbol{s}_i$, $y(\boldsymbol{s}_i)^{(l)}$, are drawn from $N(\boldsymbol{x}(\boldsymbol{s}_i)\boldsymbol{\beta}^{(l)} + \boldsymbol{w}(\boldsymbol{s}_i)^{(l)}, \tau^{2(l)})$ for $l = 1, 2, \ldots L$. Samples from the posterior predictive distribution of total AGB are then obtained via composition sampling like so,

$$Y^{(l)} = |\mathcal{D}|\sum_{i=1}^{L} y(\boldsymbol{s}_i)^{(l)}/N. \tag{2.13}$$

A point estimate for total AGB is given by,

$$\hat{Y} = \sum_{l=1}^{L} Y^{(l)}/L. \tag{2.14}$$

The variance of MCMC samples from Y serves as an estimate for $Var(\hat{Y})$ and calculated as,

$$\widehat{Var}(\hat{Y}) = \frac{\sum_{l=1}^{L}(Y^{(l)} - \hat{Y})^2}{L}. \tag{2.15}$$

To generate a 95 percent uncertainty interval, $\hat{Y}_{CI}$, we calculate the 2.5 percent and 97.5 percent quantile breaks for the set of sample draws from posterior predictive distribution of Y.

2.6 Bayesian model-based inference under the Case C lidar availability scenario

A bivariate extension to model (2.1) to jointly predict AGB and lidar information contained in $\boldsymbol{x}(\boldsymbol{s}_i)$, $x(\boldsymbol{s}_i)$ can be formulated as,

$$\begin{bmatrix} y(\boldsymbol{s}_i) \\ x(\boldsymbol{s}_i) \end{bmatrix} = \begin{bmatrix} \beta_{0y} \\ \beta_{0x} \end{bmatrix} + \boldsymbol{A} \begin{bmatrix} u_y(\boldsymbol{s}_i) \\ u_x(\boldsymbol{s}_i) \end{bmatrix} + \begin{bmatrix} \epsilon_y(\boldsymbol{s}_i) \\ \epsilon_x(\boldsymbol{s}_i) \end{bmatrix} \tag{2.16}$$

and is a spatial model applicable to the **Case C** sampling scenario. The mean component of $y(\boldsymbol{s}_i)$ and $x(\boldsymbol{s}_i)$ is modeled through the intercepts, β_{0y} and β_{0x}, respectively. The spatial random effects $u_y(\boldsymbol{s}_i)$ and $u_x(\boldsymbol{s}_i)$ capture the spatially structured components of $y(\boldsymbol{s}_i)$ and $P90(\boldsymbol{s}_i)$, respectively. The non-spatial variability in $y(\boldsymbol{s}_i)$ and $x(\boldsymbol{s}_i)$ is modeled through $\epsilon_y(\boldsymbol{s}_i)$ and $\epsilon_x(\boldsymbol{s}_i)$. The 2×2 matrix $\boldsymbol{K} = \boldsymbol{A}\boldsymbol{A}'$ models the covariance between $u_y(\boldsymbol{s}_i)$ and $u_x(\boldsymbol{s}_i)$. By treating $x(\boldsymbol{s}_i)$ as an additional response variable, rather than a covariate as in model (2.1), we no longer require x to be observed at all lattice locations, meaning that flight-line samples of x can be leveraged to potentially improve estimates of Y using this framework. Given that the number of observations $x(\boldsymbol{s}_i)$ is much larger than $y(\boldsymbol{s}_i)$, if the spatial covariance between $y(\boldsymbol{s}_i)$ and $x(\boldsymbol{s}_i)$ (off-diagonal elements of $\boldsymbol{K}$) is large, we should expect more precise estimation of Y.

We point the interested reader to Banerjee et al. (2014, pp.278-283) and (Babcock et al., 2017) for further details regarding this model specification and Bayesian approaches to inference, including appropriate vague prior specifications. Priors for the model(3.5) were set similarly to those found in Babcock et al. (2017). Note that, within a Bayesian inferential paradigm, once a sufficient number of samples from the posterior distribution of Y is obtained, $\hat{Y}$, $\widehat{Var}(\hat{Y})$ and $\hat{Y}_{CI}$ can be garnered similarly to the procedure described in Section 2.5.2.

Due to the increased complexity of this model, many more MCMC samples are needed to ensure convergence, which makes the approach much more computationally intensive. This increased computational complexity precludes the testing of this framework in the simulated sampling analysis on a typical high-performance workstation. Therefore, we only use this modeling approach on the real data for comparison with the model-assisted estimator applicable to **Case C** detailed in Section 2.8.

2.7 Design-based inference under the Case A and Case B lidar availability scenarios

Under design-based inference, $y(\boldsymbol{s}_i)$ is treated as a fixed quantity and randomness is introduced through the probabilistic selection of sample locations. For the design-based approaches examined here, it is assumed that pixel cluster locations were selected via simple random sampling. Under these conditions, a design-based estimate of total AGB can be obtained using,

$$\hat{Y} = |\mathcal{D}| \left(\sum_{i=1}^{N} \frac{\hat{y}(\boldsymbol{s}_i)}{N} + \sum_{i=1}^{n_c} \frac{\hat{\epsilon}_c(\boldsymbol{s}_i)}{n_c} \right), \tag{2.17}$$

where $\hat{y}(\boldsymbol{s}_i) = \boldsymbol{x}(\boldsymbol{s}_i)'\hat{\boldsymbol{\beta}}$ and $\hat{\epsilon}_c(\boldsymbol{s}_i)$ is the pixel cluster average residual centered on location $\boldsymbol{s}_i$ and n_c is the number of pixel clusters sampled, e.g., $n_c = n/4$ in the study. The variance of $\hat{Y}$ is estimated as,

$$\widehat{Var}(\hat{Y}) = |\mathcal{D}|^2 \left(\frac{\sum_{i=1}^{n_c}(\hat{\epsilon}_c(\boldsymbol{s}_i) - \sum_{i=1}^{n_c}(\hat{\epsilon}_c(\boldsymbol{s}_i)/n_c))^2}{n_c(n_c-1)} \right). \tag{2.18}$$

Using $\hat{Y}$ and $\widehat{Var}(\hat{Y})$ garnered from (2.17) and (2.18), $\hat{Y}_{CI}$ is obtained using (2.8).

2.8 Design-based inference under the Case C lidar availability scenario

For the **Case C** lidar availability scenario we envision a two-stage sampling design with flight-line strips constituting the primary sampling unit and pixel clusters comprising the secondary sampling unit with simple random sampling assumed at both stages. Extensions to this approach, including adaptations for unequal inclusion probabilities, post-stratification and finite sample size correction are provided in (Ringvall et al., 2016). A two-stage design-based estimate of total AGB is,

$$\hat{Y} = N \sum_{i=1}^{h} \frac{\hat{Y}_i}{M_i}, \tag{2.19}$$

where M_i is the number of pixels in strip i and h is the number of strips sampled. The within-strip total AGB estimate is given by,

$$\hat{Y}_i = |\mathcal{B}_i| \left(\sum_{\boldsymbol{s}_j \in \mathcal{B}_i} \frac{\hat{y}(\boldsymbol{s}_j)}{M_i} + \sum_{\hat{\epsilon}_c(\boldsymbol{s}_j) \in \mathcal{B}_i} \frac{\hat{\epsilon}_c(\boldsymbol{s}_j)}{m_i} \right), \tag{2.20}$$

where $\mathcal{B}_i$ is the spatial domain of strip i with $|\mathcal{B}_i|$ being its area measured in hectares and m_i is the number of pixel clusters in strip i. A variance estimate for total AGB is obtained using,

$$\widehat{Var}(\hat{Y}) = \frac{|\mathcal{D}|^2}{\widehat{|\mathcal{D}|}^2} \left(H^2 \frac{s^2}{h} + \frac{H}{h} \sum_{i=1}^{h} \widehat{Var}(\hat{Y}_i) \right), \tag{2.21}$$

where

$$\widehat{Var}(\hat{Y}_i) = |\mathcal{B}_i|^2 \left(\frac{\sum\limits_{\hat{\epsilon}_c(\boldsymbol{s}_j) \in \mathcal{B}_i} (\hat{\epsilon}_c(\boldsymbol{s}_j) - \sum\limits_{\hat{\epsilon}_c(\boldsymbol{s}_j) \in \mathcal{B}_i} (\hat{\epsilon}_c(\boldsymbol{s}_j))/m_i)^2}{m_i(m_i - 1)} \right), \tag{2.22}$$

$$s^2 = \frac{1}{h-1} \sum_{i=1}^{h} \left(\hat{Y}_i - \frac{\sum_{i=1}^{h} \hat{Y}_i}{\sum_{i=1}^{h} |\mathcal{B}_i|} |\mathcal{B}_i| \right)^2, \tag{2.23}$$

$\widehat{|\mathcal{D}|} = (H/h) \sum_{i=1}^{h} |\mathcal{B}_i|$ and H is the number of overlapping flight-line strip domains required to completely tessellate $\mathcal{D}$. Using $\hat{Y}$ and $\widehat{Var}(\hat{Y})$ garnered from (2.19) and (2.21), $\hat{Y}_{CI}$ is obtained using (2.8).

2.9 Candidate estimators and repeat sampling assessment

Section 2.5 describes the frequentist and Bayesian *spatial* model-based approaches presented in Table 2.1 and 2.2. The *non-spatial* candidates presented in Tables 2.1 and 2.2 are derived as special cases of the *spatial* model-based estimators, where σ^2 (spatial variance) and ϕ (spatial range) are forced to equal zero. The *design-based/model-assisted* approaches use the non-spatial special case as well and estimate the regression coefficients, $\boldsymbol{\beta}$, using ordinary least squares (OLS), noting that the GLS $\boldsymbol{\beta}$ estimation formula shown in (2.5) becomes OLS when $\hat{C}_s = \hat{tau}^2 \boldsymbol{I}_{n \times n}$.

For the **Case A** lidar availability scenario $\boldsymbol{x}(\boldsymbol{s}_i) = (1)$. Observing that the general *design-based/model-assisted* detailed in Section 2.7 reduces to a simple random sampling estimator, in Tables 2.1 and 2.4, we label this approach *SRS*. For the **Case B** scenario $\boldsymbol{x}(\boldsymbol{s}_i) = (1, P90(\boldsymbol{s}_i))'$ and the *design-based/model-assisted* is label *MA* in Table 2.2. The *design-based/model-assisted* estimator for the **Case B** scenario uses $\boldsymbol{x}(\boldsymbol{s}_i) = (1, P90(\boldsymbol{s}_i))'$ and is labeled *TSMA* in Tables 2.3 and 2.4. The Bayesian estimator detailed in Section 2.6 was used to estimate total AGB for the real data and results are presented in Table 2.4 (labeled as *coregionalization*). For the *coregionalization* estimator, $x(\boldsymbol{s}_i) = P90(\boldsymbol{s}_i)$.

To assess estimation efficiency, bias and uncertainty interval coverage for the candidate estimators (aside from the *coregionalization* approach), 1 000 sample sets were drawn from the synthetic population according to Section 2.3 at four different sampling intensities; 10, 7.5, 5 and 2.5 km plot separation distances. For each sample set, the candidate estimators were used to generate a point estimate of total AGB ($\hat{Y}$), obtain a variance for the estimate ($\widehat{var}(\hat{Y})$) and calculate a 95 percent uncertainty interval $\hat{Y}_{CI}$. The estimates were then compared to the true total AGB for the synthetic population (Y).

2.10 Results

*2.10.1 Comparing design-based and model-assisted estimators for **Cases A**, **B** and **C***

Tables 2.1, 2.2 and 2.3 show simulation results for sampling **Cases A**, **B** and **C**, respectively. The design-based and model-assisted total AGB estimators exhibited negligible bias for their respective sampling scenarios regardless of pixel cluster separation distance, evidenced by near-zero empirical repeat sampling bias for the *SRS*, *MA* and *TSMA* approaches (Tables 2.1, 2.2 and 2.3). The *SRS*, *MA* and *TSMA* estimators produced 95 percent uncertainty intervals with adequate empirical coverage for the 2.5 and 5 km pixel cluster spacings under the 3 lidar availability cases examined. However, results from the 10 and 7.5 km pixel cluster spacings showed 95 percent empirical uncertainty interval coverage to be lower than 95 percent. This is likely due to the low sample sizes induced by setting pixel cluster separation distances to greater than 5 km.

For the pixel cluster spacings considered, the average variance estimate of the repeated samples for the *MA* approach under **Case B** was approximately half that of the *SRS* estimator. This indicates that the lidar ancillary data provided substantial improvement in terms of estimation efficiency. Comparing the **Case B** *MA* and **Case C** *TSMA* approaches, we see *TSMA* average variance estimates were only slightly larger than *MA* for all pixel cluster spacings. This indicates that leveraging wall-to-wall lidar provided only minimal improvement over using strip samples when estimating total AGB for the synthetic population.

*2.10.2 Comparing model-based and design-based approaches for **Case A***

For the **Case A** sampling scenario, the *non-spatial* approaches produced lower variance estimates than their counterpart *spatial* estimators using both frequentist and Bayesian estimation strategies. We also see that the *non-spatial* models had empirical coverage percentages further below the nominal 95 percent than corresponding *spatial* estimators. This result indicates that the *spatial* estimators were better able to characterize uncertainty for the **Case A** sampling scenario. This result is not surprising, considering that the synthetic population was simulated to preserve inherent second-order spatial structure seen in the real data and that the *spatial* estimators were formulated to attempt to account for this extraneous structure.

It is evident from Table 2.1 that the model-based approaches estimated using Bayesian inference leveraging vague prior information rendered higher total AGB variance estimates than the model-based approaches estimated using frequentist techniques. The Bayesian models also produced uncertainty intervals with better repeated sampling coverage. This illustrates the effect of accounting for variance parameter uncertainty through the specification of appropriately vague prior information afforded by the Bayesian approach to estimation used in this study. Treating the estimated spatial variance parameters as *known without error*, as in the frequentist approaches, resulted in underestimation of total AGB uncertainty.

As with the *design-based/model-assisted* approaches, none of the model-based strategies produced uncertainty intervals exhibiting adequate coverage for the 10 and 7.5 km pixel cluster separation distances, suggesting that low sample size is still of concern. The *model-*

Table 2.1: Simulated sampling results for **Case A** lidar availability scenario.

pixel cluster spacing	Candidate estimator	$\overline{\sqrt{\widehat{Var}(\hat{Y})}}$	$\overline{(Y-\hat{Y})/Y}$	$\hat{Y}_{CI}$ empirical coverage
10 km	*Design-based/model-assisted*			
	SRS	1.711	0.53%	90.0%
	Frequentist model-based			
	non-spatial	1.273	0.52%	80.3%
	spatial	1.486	0.56%	83.9%
	Bayesian model-based			
	non-spatial	1.313	0.55%	81.6%
	spatial	1.657	0.49%	90.1%
7.5 km	*Design-based/model-assisted*			
	SRS	1.299	0.00%	92.7%
	Frequentist model-based			
	non-spatial	0.959	-0.02%	83.8%
	spatial	1.128	-0.01%	88.2%
	Bayesian model-based			
	non-spatial	1.106	-0.02%	85.7%
	spatial	1.274	-0.01%	92.4%
5 km	*Design-based/model-assisted*			
	SRS	0.862	0.32%	94.5%
	Frequentist model-based			
	non-spatial	0.636	0.30%	85.1%
	spatial	0.732	0.31%	89.2%
	Bayesian model-based			
	non-spatial	0.702	0.31%	87.7%
	spatial	0.870	0.32%	94.9%
2.5 km	*Design-based/model-assisted*			
	SRS	0.433	0.04%	94.2%
	Frequentist model-based			
	non-spatial	0.319	0.02%	83.5%
	spatial	0.433	0.02%	94.3%
	Bayesian model-based			
	non-spatial	0.436	0.02%	94.2%
	spatial	0.496	0.01%	96.6%

Table 2.2: Simulated sampling results for **Case B** lidar availability scenario.

pixel cluster spacing	Candidate estimator	$\overline{\sqrt{\widehat{Var}(\hat{Y})}}$	$\overline{(Y-\hat{Y})/Y}$	$\hat{Y}_{CI}$ empirical coverage
10 km	*Design-based/model-assisted*			
	MA	0.700 Tg	0.39%	88.4%
	Frequentist model-based			
	non-spatial	0.724 Tg	-1.30%	91.5%
	spatial	0.727 Tg	-1.10%	91.4%
	Bayesian model-based			
	non-spatial	0.749 Tg	-1.10%	92.0%
	spatial	0.742 Tg	-1.30%	92.2%
7.5 km	*Design-based/model-assisted*			
	MA	0.551 Tg	0.11%	91.2%
	Frequentist model-based			
	non-spatial	0.550 Tg	-1.59%	91.3%
	spatial	0.552 Tg	-1.38%	91.7%
	Bayesian model-based			
	non-spatial	0.578 Tg	-1.39%	93.1%
	spatial	0.576 Tg	-1.59%	92.6%
5 km	*Design-based/model-assisted*			
	MA	0.369 Tg	0.36%	93.2%
	Frequentist model-based			
	non-spatial	0.368 Tg	-1.35%	93.0%
	spatial	0.368 Tg	-1.13%	92.9%
	Bayesian model-based			
	non-spatial	0.407 Tg	-1.14%	94.9%
	spatial	0.406 Tg	-1.45%	95.3%
2.5 km	*Design-based/model-assisted*			
	MA	0.189 Tg	0.00%	94.2%
	Frequentist model-based			
	non-spatial	0.186 Tg	-1.69%	86.9%
	spatial	0.186 Tg	-1.47%	89.0%
	Bayesian model-based			
	non-spatial	0.255 Tg	-1.58%	96.4%
	spatial	0.254 Tg	-1.69%	96.0%

Table 2.3: Simulated sampling results for **Case C** lidar availability scenario.

pixel cluster spacing	Candidate estimator	$\overline{\sqrt{\widehat{Var}(\hat{Y})}}$	$\overline{(Y-\hat{Y})/Y}$	$\hat{Y}_{CI}$ empirical coverage
10 km	*Design-based/model-assisted*			
	TSMA	0.798 Tg	0.66%	88.6%
7.5 km	*Design-based/model-assisted*			
	TSMA	0.615 Tg	0.25%	89.7%
5 km	*Design-based/model-assisted*			
	TSMA	0.427 Tg	0.42%	95.0%
2.5 km	*Design-based/model-assisted*			
	TSMA	0.227 Tg	-0.09%	97.3%

based approaches nominally have sample sizes four times larger than the *design-based/model-assisted* approaches because individual pixels are treated as distinct observations. The fact that the *model-based* approaches still underestimated variance suggests that the *effective sample size* was still too low at the 10 and 7.5 km pixel cluster separation distances. For the *non-spatial* estimators, assuming individual pixels to be completely independent information sources was inappropriate and, when the spatial dependence between the pixels was explicitly accounted for in the *spatial* estimators, variance estimates for total AGB increased, thereby producing uncertainty intervals with better coverage.

The *SRS* estimator produced uncertainty intervals with better empirical coverage than the frequentist *spatial* and *non-spatial* approaches for all but the 2.5 km separation distance. At the 2.5 km pixel cluster spacing, the *SRS* and frequentist *spatial* estimators performed similarly. The Bayesian *spatial* model-based estimator showed conservative coverage at the 2.5 km sampling intensity. At the 10, 7.5 and 5 km distances, the *SRS* and Bayesian *spatial* estimators produced comparable variance and empirical coverage statistics.

2.10.3 Comparing model-based and model-assisted approaches for ***Case B***

The model-based approaches used for the **Case B** sampling scenario showed more bias than for the corresponding **Case A** estimators, though the absolute values of all bias metrics were less than 2 percent, suggesting that estimation bias was still quite minimal. However, the *MA* estimator for **Case B** clearly demonstrated less bias than the candidate model-based estimators.

Unlike in the **Case A** lidar availability scenario, the *non-spatial* and counterpart *spatial* estimators produced similar total AGB variance estimates and subsequent uncertainty interval coverages. This suggests that the $P90$ covariate explained substantial amounts of the spatially structured variability in AGB density for the synthetic population. As in **Case A**, the Bayesian estimators showed higher variance on average and better uncertainty interval coverage than the frequentist model-based approaches. Interestingly, we also see that the Bayesian model-based approaches showed improved empirical coverage, though slight, compared to *SRS* for the 10, 7.5 and 5 km sampling scenarios. As with **Case A** the Bayesian approaches showed conservative coverage at the 2.5 km pixel cluster separation distance.

2.10.4 Comparing coregionalization and two-stage model-assisted estimators for real data

The *TSMA* and *coregionalization* approaches both estimated $\sqrt{\widehat{\mathrm{Var}}\,(\hat{Y})}$ to be drastically lower than for *SRS*, indicating that the use of strip sampled lidar vastly improved precision of total AGB estimates. There is no evidence that either the *TSMA* or *coregionalization* approaches produced markedly different estimates of uncertainty although the *coregionalization* variance estimate was slightly higher. This result is consistent with the results of the simulation study under **Case B** where the Bayesian *spatial* approach tended to estimate higher variances than the *MA* approach, leading to uncertainty intervals with slightly better empirical coverage for 10, 7.5 and 5 km pixel cluster separation distances.

Table 2.4: Candidate estimator results for real data

Candidate estimator	$\hat{Y}$	$\sqrt{\widehat{Var}(\hat{Y})}$	Y_{CI}
Design-based/model-assisted			
SRS	9.28 Tg	2.250 Tg	(4.87, 13.69)
TSMA	9.00 Tg	1.072 Tg	(6.90, 11.10)
Bayesian model-based			
coregionalization	9.89 Tg	1.094 Tg	(7.75, 12.03)

2.11 Discussion

Results from this study indicate that model-assisted and model-based approaches can successfully leverage lidar ancillary data to improve total AGB estimation in interior Alaska and produce uncertainty intervals with adequate coverage. It appears that the *TSMA* estimator examined here can be useful for estimating total AGB, and potentially other forest inventory variables, with increased precision over the current *SRS* approach implemented by the FIA. Concerning model-based approaches to estimation, this study exemplified the importance of specifying model frameworks that can account for spatially structured variability not accounted for by covariates. The simulation study results suggest that the spatial models examined here may provide an effective means to account for this extraneous variability.

It is quite well understood, and this study provides additional evidence to support, that uncertainty intervals produced from spatial models fitted using frequentist methods, e.g., regression kriging, do not, in general, exhibit adequate empirical coverage due to the treatment of spatial covariance parameters as *known without error* (Luna & Young, 2003; Oliveira & Kone, 2015). Results from this study show that a Bayesian approach to spatial model fitting can help to alleviate this issue by allowing the modeler to specify vague prior information to represent a lack of knowledge about the *true* covariance parameter values, thereby propagating uncertainty pertaining to the spatial random field, $\boldsymbol{w}$, through to prediction.

The results from this simulation study suggest that the *TSMA* approach provides total AGB estimates nearly as precise as *MA*, arguing that full-coverage lidar information is unlikely to provide substantial benefit over flight-line strip samples. Considering the cost savings incurred by collecting only samples of lidar information rather than full spatial coverage, for forest inventory purposes a flight-line sampling approach to collect lidar in interior Alaska is likely adequate to produce total AGB estimates for the entire region with acceptable levels of precision. However, with a model-based approach, wall-to-wall lidar provides a mechanism for effectively producing maps of AGB density which can be useful for understanding the spatial distribution of AGB within the study area. With strip-sampled lidar, mapping AGB is not possible when treating lidar information as a covariate, or only produces global mean predictions away from flight-line strips when lidar is jointly predicted within a coregionalization framework. In future research we plan to consider the incorporation readily available wall-to-wall Landsat composites as covariates in the coregionalization framework to potentially provide increased predictive power between the lidar strip coverage areas.

The current plot spacing considered for interior Alaska inventory by the FIA is 10 km. It is clear from the simulation results that, at this spacing, none of the approaches examined here are capable of producing uncertainty intervals with good empirical coverage for TNWR. Though not tested here, it is likely that, for the *SRS* and *MA* estimators used in the **Case A** and **Case B** lidar availability scenarios, uncertainty intervals exhibiting good empirical coverage could have been produced using a t distribution quantile function with appropriate degrees of freedom at the 10 km spacing. However, the *TSMA* approach having different sample sizes within stages, there is no clear choice for degrees of freedom when selecting a t distribution quantile function. Though the 10 km field plot spacing proves problematic for estimating total AGB in TNWR, field sample sizes should not be a concern when estimating total AGB for the TIU. The projected number of FIA plots to be measured in the TIU is well over 1 000. The results here indicated that the *TSMA* estimator produces conservative variance estimates when field sample sizes average over 500 plots as in the 2.5 km spacing scenario. If interested in estimating total AGB for TNWR, it may be preferable to adopt the coregionalization approach, fit the model using all FIA plots for the TIU, then generate a *small area* estimate for TNWR.

Overall, this study shows that design-based, model-assisted and model-based approaches to AGB estimation in interior Alaska can be useful. Both model-assisted and model-based estimators, when appropriately structured, can effectively be used to dramatically improve total AGB estimation precision, thereby allowing the number of costly field sample sizes to be reduced in remote boreal regions. However, a minimum field sample size must be met in order to produce uncertainty intervals with intended coverage.

APPENDIX

2.A Synthetic population generation process

Let $\boldsymbol{A}$ be the $n \times p$ lidar variable matrix where n is the number of lidar pixels observed and p is the number of lidar metrics derived for each pixel. For this analysis, $p = 20$ since the lidar metric set consists of 10 percentile height metrics and 10 relative density measures. Let $\widetilde{\boldsymbol{A}}$ be the standardized scores for $\boldsymbol{A}$ calculated as follows,

$$\widetilde{\boldsymbol{A}}_{ij} = \frac{\boldsymbol{A}_{ij} - \mu_j}{\delta_j}, \tag{2.24}$$

where μ_j and δ_j are the mean and standard deviation of the j-th column of $\boldsymbol{A}$, respectively. A correlation matrix $\boldsymbol{\rho}$ is constructed for $\boldsymbol{A}$ by,

$$\boldsymbol{\rho} = \frac{\widetilde{\boldsymbol{A}}'\widetilde{\boldsymbol{A}}}{n-1}. \tag{2.25}$$

Let $\boldsymbol{\rho} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}'$ be the eigen decomposition of $\boldsymbol{\rho}$ with the diagonal elements of $\boldsymbol{\Lambda}$ $(p \times p)$ holding the eigenvalues of $\boldsymbol{\rho}$ in decreasing order (off diagonal elements are zero) and $\boldsymbol{P}$ $(p \times p)$ holding the corresponding eigenvectors along the columns. $\boldsymbol{P}'$ is the matrix transpose of $\boldsymbol{P}$. $\boldsymbol{P}$ is a rotation matrix that can be used to project $\widetilde{\boldsymbol{A}}$ to an orthonormal basis. Let $\boldsymbol{A}^* = \widetilde{\boldsymbol{A}}\boldsymbol{P}$ be the transformed standardized scores of the lidar variable matrix. The columns of $\boldsymbol{A}^*$ are now uncorrelated.

Let $\boldsymbol{S}$ be the $n \times 2$ matrix of pixel center coordinates for the variables in $\boldsymbol{A}^*$, i.e., $\boldsymbol{S}_i$ holds the coordinates for the lidar row vector $\boldsymbol{A}^*_i$. The exponential spatial autocovariance function,

$$C(||\boldsymbol{S}_k - \boldsymbol{S}_l||) = \begin{cases} \sigma_j^2 \exp(-\phi_j ||\boldsymbol{S}_k - \boldsymbol{S}_l||) & \text{if } ||\boldsymbol{S}_k - \boldsymbol{S}_l|| > 0 \\ \sigma_j^2 + \tau_j^2 & \text{if } ||\boldsymbol{S}_k - \boldsymbol{S}_l|| = 0 \end{cases}, \tag{2.26}$$

was fit using a restricted maximum likelihood (REML) approach to $\boldsymbol{A}^*_j$ to estimate spatial range (ϕ_j), spatial variance (σ_j^2) and nugget variance (τ_j^2). The ϕ_j, σ_j^2 and τ_j^2 estimates were then used to generate an $m \times m$ spatial covariance matrix $\boldsymbol{\Sigma}^j$ where m is the number of pixels for the synthetic population,

$$\boldsymbol{\Sigma}^j_{kl} = \begin{cases} \sigma^2_j \exp(-\phi_j ||\boldsymbol{T}_k - \boldsymbol{T}_l||) & \text{if } ||\boldsymbol{T}_k - \boldsymbol{T}_l|| > 0 \\ \sigma^2_j + \tau^2_j & \text{if } ||\boldsymbol{T}_k - \boldsymbol{T}_l|| = 0 \end{cases}, \tag{2.27}$$

and $\boldsymbol{T}$ is the $m \times 2$ matrix of pixel center coordinates for the synthetic population. Let $\mathbf{R}_j$ ($m \times 1$) be one realization from a zero-centered multivariate normal spatial random field drawn from $MVN(\mathbf{0}, \boldsymbol{\Sigma}^j)$. Let $\boldsymbol{A}^{*r}_j$ ($m \times 1$) be m resampled (with replacement) values from $\boldsymbol{A}^*_j$ sorted according to the rank order of $\boldsymbol{R}_j$. Let $\widetilde{\boldsymbol{A}}^r = \boldsymbol{A}^{*r}\boldsymbol{P}^{-1}$ be the rotation of $\boldsymbol{A}^{*r}$ back to the basis of $\widetilde{\boldsymbol{A}}$. Let $\boldsymbol{A}^r_j$ be m resampled (with replacement) values from $\boldsymbol{A}_j$ sorted according to the rank order of $\widetilde{\boldsymbol{A}}^r_j$. For each row of $\boldsymbol{A}^r$, percentile height variable entries were sorted to ensure that 10^{th} percentile height $<$ 20^{th} percentile height $< ... <$ 100^{th} percentile height. Also, for each row of $\boldsymbol{A}^r$, relative density measures were normalized to sum to one.

Let $\boldsymbol{A}^*_p$ be the subset $\boldsymbol{A}^*$ over subplot locations and let $\boldsymbol{y}$ be the corresponding vector of field measured AGB values. The following linear model,

$$\sqrt{\boldsymbol{y}} = \alpha + \boldsymbol{A}^*_p\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2.28}$$

was fit using ordinary least squares (OLS) to obtain estimates for α, $\boldsymbol{\beta}$ and the variance of $\boldsymbol{\epsilon}$ (σ^2_ϵ). α is the model intercept and $\boldsymbol{\beta}$ is a column vector of regression slope coefficients. $\boldsymbol{\epsilon}$ is a residual error vector often assumed to be independent and identically distributed normally, i.e., $MVN(\mathbf{0}, \sigma^2_\epsilon \boldsymbol{I}_{n\times n})$. Let $\sqrt{\boldsymbol{y}_r}$ be one realization from $MVN(\alpha + \boldsymbol{A}^*_r\boldsymbol{\beta}, \sigma^2_\epsilon \boldsymbol{I}_{m\times m})$. $\boldsymbol{y}_r$ is now the AGB component of the synthetic population.

Chapter 3

GEOSTATISTICAL ESTIMATION OF FOREST BIOMASS IN INTERIOR ALASKA COMBINING LANDSAT-DERIVED TREE COVER, SAMPLED AIRBORNE LIDAR AND FIELD OBSERVATIONS

Abstract: Lidar data provide critical information on the three-dimensional structure of forests. However, collecting wall-to-wall laser altimetry data at regional and global scales is cost prohibitive. As a result, studies employing lidar data for large area estimation typically collect data via strip sampling; leaving large swaths of the forest domain unmeasured by the instrument. The goal of this research was to develop and examine the performance of a geostatistical coregionalization modeling approach for combining field inventory measurements, strip samples of airborne lidar and Landsat-based remote sensing data products to predict aboveground biomass (AGB) in interior Alaska's Tanana Valley. The proposed modeling strategy facilitates pixel-level mapping of AGB density predictions across the entire spatial domain. Additionally, the coregionalization framework allows for statistically sound estimation of total AGB for arbitrary areal units within the study area—a key advance to support diverse management objectives in interior Alaska. This research focuses on appropriate characterization of prediction uncertainty in the form of posterior predictive coverage intervals and standard deviations. Using the framework detailed here, it is possible to quantify estimation uncertainty for any spatial extent, ranging from pixel-level predictions of AGB density to estimates of AGB stocks for the full domain. The lidar-informed coregionalization models consistently outperformed their counterpart lidar-free models in terms of point-level predictive performance and total AGB precision. Additionally, the inclusion of Landsat-derived forest cover as a covariate further improved estimation precision in regions with lower lidar sampling intensity. Our findings also demonstrate that model-based

approaches that do not explicitly account for residual spatial dependence can grossly underestimate uncertainty, resulting in falsely precise estimates of AGB. On the other hand, in a geostatistical setting, residual spatial structure can be modeled within a Bayesian hierarchical framework to obtain statistically defensible assessments of uncertainty for AGB estimates. The inferential capabilities of AGB posterior predictive distribution (PPD) products extend beyond simply mapping AGB density. PPD products can provide insight regarding drivers of AGB heterogeneity in boreal forests, including permafrost and fire, highlighting the range of potential applications for Bayesian geostatistical methods to integrate field, airborne and satellite data.

3.1 Introduction

Coupling remotely sensed data with field-based forest measurements via regression frameworks offers the potential to increase the precision of inventory estimates and provides a mechanism for mapping the spatial distribution of forest biophysical properties. A plethora of studies show strong relationships between lidar metrics and forest variables (Asner et al., 2009; Babcock et al., 2013; Finley et al., 2014b, 2017; Lim et al., 2003; Næsset, 2004, 2011). These findings have spurred investment in collecting lidar data for large areas from aircraft and satellites alike. Of particular interest is the use of lidar to assist in the estimation of forest inventory parameters in high-latitude terrestrial ecosystems. From a carbon monitoring perspective, forests in boreal systems may contain large stores of aboveground biomass (AGB) and carbon, but the uncertainty associated with current estimates is extremely high (Bradshaw & Warkentin, 2015; Pan et al., 2011). Understanding that the taiga-tundra ecotone is one of the most vulnerable environmental systems to climate change and that its boreal forests can contribute mightily to the global carbon cycle, methods are needed to begin monitoring forest carbon stocks and fluxes for these systems (Gauthier et al., 2015; Magnani et al., 2007; Neigh et al., 2013).

Current approaches used by the United States Forest Service's (USFS) Forest Inventory and Analysis (FIA) program to quantify AGB and carbon stocks in temperate regions rely on extensive, spatially-balanced field plot probability samples to generate forest inventory estimates with acceptable levels of precision (Bechtold & Patterson, 2005; Woodall et al., 2015). In vast remote landscapes, implementing the estimation techniques used by the FIA in the contiguous United States becomes prohibitively expensive due to the high cost of collecting field inventory data in difficult-to-access boreal regions, e.g., interior Alaska (Barrett & Gray, 2011). A potential solution commonly put forward to reduce the expense of monitoring AGB in boreal forest systems is to augment sparse collections of field samples with remotely sensed auxiliary data (Wulder et al., 2012b). Lidar-derived measures of forest structure tend to be highly correlated with AGB field observations and, thus, are prime candidates to supplement boreal field campaigns. Additionally, passive sensors such as Landsat can be used to derive remote sensing data products correlated with forest AGB (Kumar et al., 2015). Methodologies leveraging relationships between field and lidar can potentially be further improved by incorporating Landsat-based products (Margolis et al., 2015; Pflugmacher et al., 2014; Powell et al., 2010; Zheng et al., 2004).

Here, we address two challenges encountered when attempting to estimate forest AGB for large areas using lidar coupled with other remote sensing information: (1) incomplete spatial coverage of remote sensing data; and (2) prediction uncertainty quantification. Incomplete spatial coverage is a common problem for studies using airborne or spaceborne lidar over sizable study domains (Andersen et al., 2011; Bolton et al., 2013; Nelson, 2010; Nelson et al., 2004). Model-based methodologies used to link field and lidar data to estimate and map AGB typically require laser altimetry information for the entire spatial domain of interest (Babcock et al., 2015, 2016; McRoberts et al., 2013). The expansive nature of boreal systems, make wall-to-wall collections of airborne lidar data financially impossible. Further, currently proposed spaceborne lidar systems are not designed to procure global coverage information. Rather, these campaigns are only optimized to collect data for relatively narrow bands along the orbital tracts of the sensors' host satellite (GEDI, 2014; ICESat-2, 2015a). In order to glean any additional information provided by sampled remote sensing data in a statistically rigorous manner, estimation frameworks that can accommodate incomplete coverage auxiliary information are necessary.

Studies employing model-assisted estimators within a design-based inferential paradigm have been proposed that assume either two-stage or two-phase data collection schemes (Gregoire et al., 2011; Saarela et al., 2015). In these studies, a model is employed to relate field and remote sensing data and residuals are analyzed using design-based variance estimators to approximate error associated with mean or total AGB estimates. By subscribing to a design-based paradigm, one only requires remote sensing data to cover the plot locations, which can be accomplished by collecting flight-line strips of auxiliary information (Nelson et al., 2012). Any model-assisted estimator depends on plots and remote sensing data being established as probability samples, which in practice, is not easy to implement (Särndal et al., 1992, Chap. 8-9). Typically, systematic sampling designs are used to approximate probabilistic ones hoping that uncertainty will be overestimated (Cochran, 1946). Hybrid estimators that mix design- and model-based inferential paradigms have also been proposed, typically assuming a probabilistic selection of flight-line strips and a purposive sample of field plots (Ståhl et al., 2011, 2016).

The second issue examined here is the problem of obtaining useful estimates of uncertainty about forest AGB stocks using model-based statistical procedures—necessary for decision making with imperfect predictions of forest AGB. In design-based estimation frameworks, error is assumed to arise from the sampling design, which can be appropriately characterized when plots are selected probabilistically (Cochran, 1977; Thompson, 2002). In model-based inference, error is attributed to the underlying process by which the response, e.g., AGB, is generated (Gregoire, 1998; Ver Hoef, 2002). Studies attempting to estimate means and totals for areal units using ancillary data within a model-based paradigm need to take care to specify frameworks that reliably accommodate the structure of the data to be modeled. It can be the case that modelers who attempt to use model-based forest inventory estimation approaches posit potentially unrealistic assumptions about the distributional characteristics of model residuals, e.g., independent and identically distributed (*iid*) errors. In a spatial context, it is highly likely the field observations of AGB will be spatially autocorrelated. If the auxiliary information used in the model fails to fully account for the spatially dependent component of the field observations, model-based approximations of AGB uncertainty can be grossly underestimated (Cressie, 1993; Griffith, 2005).

Coregionalization models constructed within a Bayesian hierarchical framework offer a solution to both above-mentioned challenges (Gelfand et al., 2004). This class of multivariate spatial models is designed to predict multiple response variables simultaneously while leveraging spatial cross-correlation structures between residual components of the different responses. If the lidar data is treated as an explanatory variable (used on the right-hand side of the model as in most lidar studies) predictions are only possible where lidar data is available. Within a coregionalization model, the lidar-derived metrics can be treated as additional response variables (moved to the left-hand side) and jointly predicted across the entire landscape while explicitly modeling the spatially co-varying relationship between them (Finley et al., 2014a). A coregionalization framework also allows for the inclusion of wall-to-wall information, e.g., Landsat-derived data, as covariates to assist in the joint prediction of forest AGB and lidar information.

When multivariate coregionalization models are estimated using a Bayesian hierarchical approach, uncertainty occurring at all levels of the model can be propagated through to prediction and subsequent estimation of totals for areal units (Berliner, 1996; Cressie & Wikle, 2011; Gelfand & Smith, 1990; Hobbs & Hooten, 2015). Forms of multivariate spatial prediction models have been in existence since the 1960s, e.g., cokriging (Matheron, 1963). These non-hierarchical implementations, however, struggle to effectively deal with uncertainty associated with spatial covariance parameters, e.g., spatial variances and decays (Diggle & Ribeiro, 2007, Section 7.1.1). Due to increased computational efficiencies gained by ignoring uncertainty in spatial variability, ‘plug-in’ spatial covariance parameters are used in cokriging interpolation routines available in popular GIS software packages. This impedes their use for fully model-based predictive inference by researchers (Schelin & Sjöstedt-de Luna, 2010).

The development of inferential approaches for complex spatial prediction within a statistical framework is an active area of research. In a hierarchical modeling context, coregionalization frameworks can be constructed using random effects that arise from spatially correlated Gaussian processes and partition variability into spatial and non-spatial components (Banerjee et al., 2014; Cressie et al., 2009). When formulated as such, estimation approaches including Restricted Maximum Likelihood (REML) or Markov chain Monte Carlo (MCMC) become possible in frequentist and Bayesian paradigms of statistical model-based inference, respectively (Ver Hoef et al., 2004). There are advantages to choosing a Bayesian hierarchical approach to inference over complementary frequentist methods. Access to the full posterior predictive distribution (PPD), a by-product of Bayesian inference, allows for easy posterior summarization of means or totals with associated uncertainty for the full spatial domain in addition to any sub-domains that may be of interest; even under back-transformation (Stow et al., 2006). Access to PPDs facilitate subsequent, i.e., post model-fitting, analysis to inform ecological or management objectives while accounting for prediction uncertainty. However, these increases in flexibility come with substantial increases in computational demand.

The aim of this study is to develop and examine the performance of a statistical modeling framework that can (1) incorporate partial coverage lidar data and wall-to-wall Landsat products to improve AGB density prediction; and (2) accommodate extraneous residual structure, thereby allowing for more reliable model-based characterizations of uncertainty and improved prediction. We look to the Tanana Inventory Unit (TIU) in interior Alaska to explore the potential for a coregionalization modeling framework to estimate forest AGB stocking by coupling spatially sparse field inventory, incomplete coverage lidar and Landsat-derived tree cover data products in boreal landscapes. The USFS and National Aeronautics and Space Administration (NASA) collaborated on the collection of field inventory data using an augmented FIA sampling design and flight-line samples of lidar data in 2014. Within this study region, four areal domains containing systematic samples of field and lidar data serve as study sites in this analysis. We use in-depth model comparison to identify strengths and limitations of candidate modeling approaches and information sources. These comparisons outline tradeoffs associated with different estimation methods and highlight benefits of coregionalization modeling for large-area estimation of AGB using sampled remote sensing data. We also demonstrate the potential of Bayesian spatial hierarchical models for generating small-area forest parameter estimates and analysis of variability in ecosystem structure via PPD summarization. At one site, generating AGB PPDs at a watershed scale offered insight regarding the drivers of biomass variability in interior Alaska, including permafrost and fire.

3.2 Methods

3.2.1 Study sites

The four study sites explored in this analysis all fall within the TIU in interior Alaska (Fig. 4.1). The Tanana Valley State Forest (TVSF) is a 730 000 hectare (ha) tract of predominately boreal forest stretching along the Tanana river basin. Nearly 90 percent of the TVSF is considered forested and close to 50 percent of all productive forestland in the TIU is contained within the boundary of the TVSF (Alaska Department of Natural Resources, 2016). Tree species typical of taiga forest can be found throughout, including white spruce (*Picea glauca*), black spruce (*Picea mariana*), tamarack (*Larix laricina*), quaking aspen (*Populus tremuloides*) and balsam poplar (*Populus balsamifera*).

Tetlin National Wildlife Refuge (TNWR) is nearly 300 000 ha in size with lowland areas characterized by extensive wetlands and poorly drained soils. Wet upland sites are home to black spruce forests whereas drier landscapes favor white spruce. Deciduous species including quaking aspen, paper birch (*Betula papyrifera*) and balsam poplar persist on well-drained south-facing slopes. Shrub vegetation consisting of willow (*Salix spp.*), alder (*Alnus spp.*) and dwarf birch (*Betula spp.*) can be found in lowland areas around water bodies (U.S. Fish and Wildlife Service, 2016).

Bonanza Creek Experimental Forest (BCEF) is a Long-Term Ecological Research (LTER) site within the TVSF, consisting of vegetation and landforms typical of interior Alaska. The BCEF domain delineated for this study is 21 000 ha and includes a section of the Tanana River floodplain along the southeastern border. The BCEF is a mixture of forest and non-forest vegetation compositions featuring white spruce, black spruce, tamarack, quaking aspen and balsam poplar trees mixed with willow and alder shrubland species (Bonanza Creek LTER, 2016a).

Caribou-Poker Creeks Research Watershed (CPCRW) is an intensively studied basin reserved for hydrological and ecological research (Bonanza Creek LTER, 2016b). CPCRW is approximately 10 600 ha in size and divided into 11 watershed units. Many research initiatives at CPCRW involve vegetation and hydrology comparisons among watershed units (Amatya et al., 2016; Rinehart et al., 2015; Tanaka-Oda et al., 2016). Upland areas are dominated by paper birch and aspen on south-facing slopes. North-facing slopes are largely occupied by black spruce. Patches of alder exist in the understory. Lowland sites are composed of moss and dwarf shrubs. CPCRW has a rich source of associated mapped data products including permafrost polygon layers and fire maps (Chapin & Hollingsworth, 2006; Rieger et al., 1972; Verbyla, 2011). In 2004, the Boundary Fire in interior Alaska burned a significant portion of CPCRW, predominantly along the southeastern border of the research area (Hollingsworth et al., 2013, Fig. 1).

At all sites, AGB density was not predicted over *Water* and *Barren Lands* pixels classified by the National Land Cover Database (Homer et al., 2015). These areas were not considered to be part of the study domains.

3.2.2 Field measurements

At TVSF and TNWR, a spatially-balanced systematic design was implemented using a tessellation of hexagons covering the study areas. The hexagons were approximately 12 141 ha (30 000 acres) in size. These polygons are five times larger than the hexagons used to establish plots in the continental United States. In the summer of 2014, FIA field crews established standard FIA plots at the center of each hexagon complete with four subplots; one at the midpoint and three additional subplots extended radially approximately 36 meters (m) at 0°, 120° and 240°. Each subplot has an approximate 7.3 m radius and an area of 168.11 m^2. Field measurements were taken using an augmented FIA inventory design (Bechtold &

Patterson, 2005). A notable change to the typical protocol included the addition of a second micro-plot to be inventoried in each subplot. Adding a second micro-plot helped to ensure that a sufficient number of small diameter trees were tallied (between 2.5 and 12.7 centimeter (cm) diameter at breast height) (Andersen et al., 2011). Pattison et al. (In prep) provide detailed field protocols for the TIU inventory pilot project. The same sampling protocol was used at BCEF and CPCRW, although at a greatly increased sampling intensity. Field plots at BCEF and CPCRW were each inventoried once in the summer months of 2011, 2012 or 2014. The total number of subplots measured at each site was 263, 123, 292 and

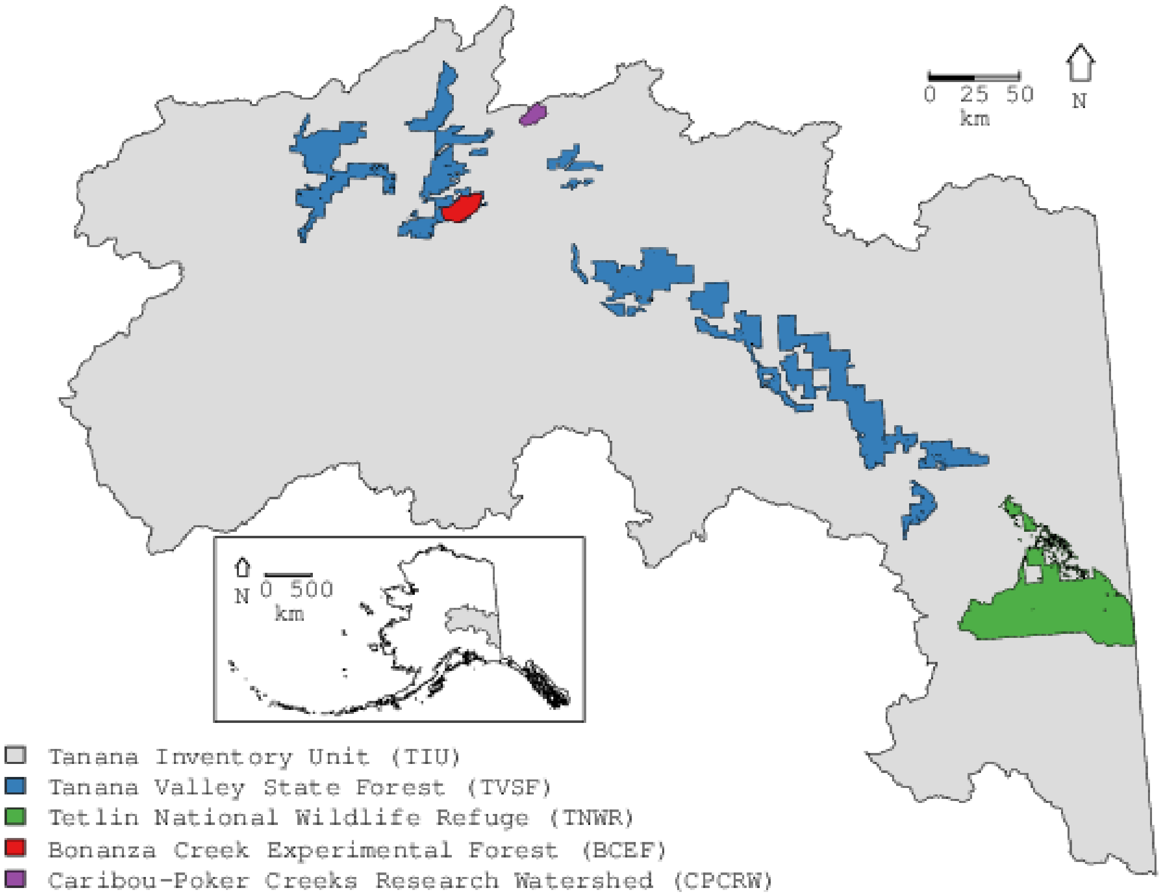


Fig. 3.1: Map showing locations of the four study sites within the Tanana Inventory Unit (TIU).

149 at TVSF, TNWR, BCEF and CPCRW, respectively (note that some subplots at each plot were not measured for various logistical reasons or fell outside the study area boundary making the total number of subplots not wholly divisible by four). AGB for individual trees on subplots were tabulated using the Component Ratio Method described in Woodall et al. (2015). AGB for all trees on a subplot with breast height diameters 2.5 cm and larger were scaled to megagrams per ha (Mg/ha) and summed to obtain subplot-level AGB density. Prior to model-fitting, a square-root transformation was applied to the subplot-level AGB densities to better approximate a Gaussian residual distribution and ensure positive support following back-transformation of predicted values. For this analysis each inventory subplot was treated as a distinct observation. To ease explanation, the term *plot* will be used in subsequent sections to refer to the individual FIA subplots described above.

3.2.3 Remote sensing data

Lidar data were collected using a flight-line strip sampling approach with Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) imager in the summer of 2014 (Cook et al., 2013). G-LiHT is a portable multi-sensor system that can be mounted to a fixed wing aircraft. G-LiHT's on-board laser altimeter (VQ-480, Riegl Laser Measurement Systems, Horn, Austria) provides an effective measurement rate of up to 150 kilohertz along a 60° swath perpendicular to the flight direction using a 1 550 nanometer laser. At a nominal flying altitude of 335 m, laser pulse footprints have an approximate 10 cm diameter. The instrument is capable of producing up to eight returns per pulse. Point cloud information was summarized to a 13 x 13 m pixel size (pixel area equal to 169 m^2) to approximate field plot areas. Over each grid cell, percentile heights were calculated at 10 percent intervals ranging from 10 percent to 100 percent. Maximum height (100^{th} percentile height) relative densities were also calculated at 10 equal width intervals. Additional metrics including point cloud skewness and kurtosis, among others, were calculated for each pixel as well. Identical lidar metrics were obtained using point clouds extracted over each field plot. Exploratory regression model fits indicated that 90^{th} percentile height alone accounted for substantial amounts of variability in square-root transformed AGB at all sites ($0.78 \leq R^2 \leq 0.87$). Due to the high correspondence between 90th percentile height and AGB, it was decided that only this metric will be considered for subsequent modeling efforts. G-LiHT data for the study area are available online (G-LiHT, 2016).

To reduce computational demand associated with fitting the proposed models, the gridded lidar covariate sets were dramatically subsampled. At the four study sites, 0.05 percent of the original lidar pixels were randomly selected. The lidar subset pixels were combined with the 90th percentile height values calculated over the field plots to construct the lidar metric sets for subsequent model-fitting. Specifically, 1 216 of 199 045, 813 of 132 635, 2 011 of 491 621 and 7 069 of 1 374 118 lidar observations were used for model-fitting at BCEF, CPCRW, TNWR and TVSF, respectively. See Section 3.4 for discussion about future research avenues that may lead to increased computational efficiencies, thereby alleviating the need to substantially thin sampled remote sensing datasets to implement the modeling frameworks proposed here.

A Landsat-derived percent tree cover data product was examined to see if additional predictive gains over the use of lidar data alone could be obtained (Hansen et al., 2013). Percent tree cover estimates for peak growing season in 2010 (most recent year available) were produced using a regression tree approach leveraging Landsat 7 ETM+ annual composites. These data are provided by the United States Geological Survey (USGS) on an approximate 30 m lattice covering the entire globe (USGS, 2016). From this dataset, percent tree cover information was extracted over the four study areas. Preliminary model-fitting exercises also examined the use of individual band reflectances from 2014 Landsat 8 OLI composite surface reflectance products. Additionally, various vegetation indices, e.g., Tasseled Cap values and Normalized Difference Vegetation Index, derived from the 2014 Landsat bands were considered. It was found that the 2010 percent tree cover data product exhibited the strongest relationship with AGB. Further, including 2014 Landsat composite bands and indices in addition to 2010 percent tree cover in regressions did not result in appreciable gains in fit performance. For these reasons, only the 2010 percent tree cover metric was considered for this analysis. R^2 values for preliminary models relating square-root transformed AGB density and percent tree cover ranged between 0.25 and 0.55 for the study areas.

3.2.4 Candidate models

To assess the utility of coupling plot-level AGB measurements with sampled airborne lidar and the 2010 Landsat-based tree cover index developed by Hansen et al. (2013), six candidate models were compared. Candidate models were evaluated on fit to observed data, prediction performance and total AGB estimates.

3.2.4.1 Null

The *Null* model is designed to be a baseline regression where no information beyond plot measurements is considered. The *Null* model is

$$y(\boldsymbol{s}) = \beta_{0y} + \epsilon_y(\boldsymbol{s}), \tag{3.1}$$

where $y(\boldsymbol{s})$ is a square-root transformed field-based measurement of AGB at location $\boldsymbol{s} \in \mathcal{D} \subseteq \mathbb{R}^2$, i.e., $\boldsymbol{s}$ is a vector of geographic coordinates, e.g., easting and northing, in the spatial domain $\mathcal{D}$. β_{0y} is an intercept parameter to be estimated. Since the *Null* model contains no additional regression parameters, the estimated value for β_{0y} should approximate the overall mean of square-root transformed AGB from the field plots, i.e., $\sum_{i=1}^{n} y(\boldsymbol{s}_i)/n$, where n is the total number of field plots measured. The residual term $\epsilon_y(\boldsymbol{s})$ captures any departure from the overall mean at site $\boldsymbol{s}$. Imposing a distributional assumption on $\boldsymbol{\epsilon}_y \overset{iid}{\sim} \mathcal{N}(\boldsymbol{0}, \tau_y^2 \boldsymbol{I})$—where $\boldsymbol{\epsilon}_y = (\epsilon_y(\boldsymbol{s}_1), \epsilon_y(\boldsymbol{s}_2), ..., \epsilon_y(\boldsymbol{s}_n))'$, τ_y^2 is a variance parameter to be estimated and $\boldsymbol{I}$ is an n x n identity matrix—allows for model-based statistical inference concerning model parameters and predictions. Any statistical inference, e.g., credible intervals, posterior standard deviations, etc., evaluated for model parameters or predictions can be considered reliable if the distributional assumptions about $\boldsymbol{\epsilon}_y$ are not seriously violated.

3.2.4.2 Tree Cover

The *Tree Cover* model is

$$y(\boldsymbol{s}) = \beta_{0y} + \beta_{1y}x(\boldsymbol{s}) + \epsilon_y(\boldsymbol{s}), \tag{3.2}$$

where $y(\boldsymbol{s})$ and $\epsilon_y(\boldsymbol{s})$ are defined as in model (3.1). The intercept, β_{0y}, and regression slope parameter, β_{1y}, together describe the linear relationship between the Landsat-based tree cover product, $x(\boldsymbol{s})$, and $y(\boldsymbol{s})$.

3.2.4.3 Spatial

Incorporating a spatial random effect into model (3.1) establishes the *Spatial* model and is

$$y(\boldsymbol{s}) = \beta_{0y} + w_y(\boldsymbol{s}) + \epsilon_y(\boldsymbol{s}), \tag{3.3}$$

where $y(\boldsymbol{s})$, β_{0y} and $\epsilon_y(\boldsymbol{s})$ are defined as in model (3.1). Here, $w_y(\boldsymbol{s})$ is modeled as a Gaussian process (GP) with a zero mean and a spatial covariance function that captures the covariance between any pair of locations $\boldsymbol{s}$ and $\boldsymbol{s}^*$, i.e., $w_y(\boldsymbol{s}) \sim GP(0, C(\boldsymbol{s}, \boldsymbol{s}^*; \boldsymbol{\theta}_y))$. We specify $C(\boldsymbol{s}, \boldsymbol{s}^*; \boldsymbol{\theta}_y) = \sigma_y^2 \rho(\boldsymbol{s}, \boldsymbol{s}^*; \phi_y)$ where $\rho(\boldsymbol{s}, \boldsymbol{s}^*; \phi_y)$ is a valid spatial correlation function and $\boldsymbol{\theta}_y = \{\sigma_y^2, \phi_y\}$, where ϕ_y is a correlation decay parameter and $\sigma_y^2 = Var(\boldsymbol{w}_y)$. For this analysis, we assume an exponential correlation function, i.e., $\rho(||\boldsymbol{s} - \boldsymbol{s}^*||; \phi_y) = \exp(-\phi_y ||\boldsymbol{s} - \boldsymbol{s}^*||)$, where $||\boldsymbol{s} - \boldsymbol{s}^*||$ is the Euclidean distance between locations $\boldsymbol{s}$ and $\boldsymbol{s}^*$ in kilometers (km). This specification for the spatial correlation function requires the modeler to assume the spatially structured residual variability to be stationary and isotropic, meaning that spatial variability in the residuals does not depend on location or direction. This is true for the spatial random effects defined in subsequent sections as well. To ease interpretation of the ϕ_y estimates, corresponding effective range estimates, labeled er_y, are presented in the results tables. We define er_y as the distance where the spatial correlation between locations drops to 0.05. See Gelfand et al. (2004) for specifics on how to calculate effective spatial ranges using coregionalization models.

3.2.4.4 Spatial+Tree Cover

Introducing a spatial random effect in model (3.2) defines the *Spatial+Tree Cover* candidate written as

$$y(\boldsymbol{s}) = \beta_{0y} + \beta_{1y}x(\boldsymbol{s}) + w_y(\boldsymbol{s}) + \epsilon_y(\boldsymbol{s}), \tag{3.4}$$

with all terms defined previously.

3.2.4.5 Coregionalization

A bivariate coregionalization model for the joint prediction of square-root transformed AGB and lidar-derived 90^{th} percentile height is

$$\begin{bmatrix} y(\boldsymbol{s}) \\ z(\boldsymbol{s}) \end{bmatrix} = \begin{bmatrix} \beta_{0y} \\ \beta_{0z} \end{bmatrix} + \boldsymbol{A} \begin{bmatrix} u_y(\boldsymbol{s}) \\ u_z(\boldsymbol{s}) \end{bmatrix} + \begin{bmatrix} \epsilon_y(\boldsymbol{s}) \\ \epsilon_z(\boldsymbol{s}) \end{bmatrix}, \tag{3.5}$$

where $y(\boldsymbol{s}), \beta_{0y}$ and $\epsilon_y(\boldsymbol{s})$ are defined previously. $z(\boldsymbol{s})$ is 90^{th} percentile height recorded at location $\boldsymbol{s}$ and, because no other covariates appear in the sub-model for $z(\boldsymbol{s})$, β_{0z} approximates the overall mean of the sampled 90^{th} percentile heights. Here, $u_y(\boldsymbol{s})$ and $u_z(\boldsymbol{s})$ are modeled as GPs with zero means and spatial correlation functions (rather than covariance functions as with $w_y(\boldsymbol{s})$ in *Spatial* and *Spatial+Tree Cover* models), i.e., $u_y(\boldsymbol{s}) \sim GP(0, \rho(\boldsymbol{s}, \boldsymbol{s}^*; \phi_y))$ and $u_z(\boldsymbol{s}) \sim GP(0, \rho(\boldsymbol{s}, \boldsymbol{s}^*; \phi_z))$, where ϕ_y and ϕ_z are spatial decay parameters. We again assume an exponential correlation function for both random effects, i.e., $\rho(||\boldsymbol{s} - \boldsymbol{s}^*||; \phi_y) = \exp(-\phi_y||\boldsymbol{s} - \boldsymbol{s}^*||)$ and $\rho(||\boldsymbol{s} - \boldsymbol{s}^*||; \phi_z) = \exp(-\phi_z||\boldsymbol{s} - \boldsymbol{s}^*||)$. Effective range estimates (er_y and er_z) are again presented in the results tables to facilitate interpretation. The matrix $\boldsymbol{K} = \boldsymbol{A}\boldsymbol{A}'$ models the cross-covariance between the spatial random effects $u_y(\boldsymbol{s})$ and $u_z(\boldsymbol{s})$. The 2 x 2 parameter matrix $\boldsymbol{A}$ is the lower triangular square-root of $\boldsymbol{K}$. The diagonal and off-diagonal elements of $\boldsymbol{K}$ capture the spatial processes variances and covariances, respectively. We assume the errors associated with the lidar sub-model are independent and normally distributed, i.e., $\boldsymbol{\epsilon}_z \overset{iid}{\sim} \mathcal{N}(\boldsymbol{0}, \tau_z^2\boldsymbol{I})$, where $\boldsymbol{\epsilon}_z = (\epsilon_z(\boldsymbol{s}_1), \epsilon_z(\boldsymbol{s}_2), ..., \epsilon_z(\boldsymbol{s}_m))'$ and τ_z^2 is a variance parameter to be estimated.

3.2.4.6 Coregionalization+Tree Cover

The *Coregionalization* model described in Section 3.2.4.5 can be extended to include wall-to-wall Landsat-derived tree cover information as such,

$$\begin{bmatrix} y(\boldsymbol{s}) \\ z(\boldsymbol{s}) \end{bmatrix} = \begin{bmatrix} \beta_{0y} + \beta_{1y} x(\boldsymbol{s}) \\ \beta_{0z} + \beta_{1z} x(\boldsymbol{s}) \end{bmatrix} + \boldsymbol{A} \begin{bmatrix} u_y(\boldsymbol{s}) \\ u_z(\boldsymbol{s}) \end{bmatrix} + \begin{bmatrix} \epsilon_y(\boldsymbol{s}) \\ \epsilon_z(\boldsymbol{s}) \end{bmatrix}, \tag{3.6}$$

where all terms with the exception of β_{1z} have been previously defined. The lidar sub-model intercept, β_{0y}, and regression slope parameter, β_{1y}, together describe the linear relationship between the Landsat-based tree cover product, $x(\boldsymbol{s})$, and $z(\boldsymbol{s})$.

3.2.5 Candidate model parameter estimation

For all six candidate models, a Bayesian paradigm of statistical inference was pursued, which required us to specify prior distributions for all model parameters. Then inference proceeded by sampling from the posterior distribution of the parameters. We set prior distributions to be as non-informative as possible to minimize their influence on posterior inference. For the regression intercept and slope parameters, i.e., β_{0y}, β_{1y}, β_{0z} and β_{1z}, we assumed a $\mathcal{N}(0, 10\,000)$. Any spatial and non-spatial variance components, i.e., σ_y^2, τ_y^2 and τ_z^2, were assigned an inverse Gamma prior $IG(a, b)$. The a hyperparameter was set equal to 2, which results in a prior distribution mean equal to b and infinite variance. The b hyperpriors were determined using preliminary semivariograms fit to the residuals of the non-spatial models (3.1) and (3.2). The spatial decay parameters, ϕ_y and ϕ_z were assigned uniform priors with support over the geographic range of the study areas. The matrix $\boldsymbol{K}$ was assigned an inverse Wishart prior with hyperparameter degrees of freedom set to 2 and diagonal scale matrix equaling $0.1\boldsymbol{I}_2$, where $\boldsymbol{I}_2$ is a 2 x 2 identity matrix. Algorithms for efficient estimation of parameters for all six candidate models are detailed in Banerjee et al. (2014) and Finley et al. (2014a). All six models were fitted using the `spBayes` package written for the `R` statistical computing environment (Finley et al., 2015b).

3.2.6 Pixel-level and areal estimation of aboveground biomass

After collecting a sufficient number of samples from the posterior distribution of a model's parameters via MCMC (following typical sampling and convergence diagnostics in Gelman et al. (2013)), composition sampling was used to obtain samples from the posterior predictive distribution (PPD) of square-root AGB density at all pixel locations within the study area (Banerjee et al., 2014). Then each sample from the PPD was back-transformed (squared) to approximate spatially explicit AGB density PPDs ($\tilde{\boldsymbol{y}} = (\tilde{y}(\boldsymbol{s}_1), \tilde{y}(\boldsymbol{s}_2), \ldots, \tilde{y}(\boldsymbol{s}_{n_0}))'$, where n_0 is the number of prediction units, e.g., pixels). Useful summaries of $\tilde{y}(\boldsymbol{s})$ for each pixel, such as, median, standard deviation, or credible interval width, can be mapped to examine the spatial distribution of AGB density and associated uncertainty. $\tilde{\boldsymbol{y}}$ can also be summarized to estimate average AGB densities for arbitrary areal units. For example, to estimate mean AGB density for the full spatial domain we need to integrate the PPD of AGB density over the entire study region, i.e., $\tilde{y}(\mathcal{D}) = |\mathcal{D}|^{-1} \int_{\mathcal{D}} \tilde{y}(\boldsymbol{s}) \mathrm{d}\boldsymbol{s}$ where $|\mathcal{D}|$ is the area of the domain $\mathcal{D}$. We can approximate this integral via MCMC integration using a fine grain systematic grid of $\tilde{y}(\boldsymbol{s})$ samples, i.e., $\tilde{y}(\mathcal{D}) \approx |\mathcal{D}|^{-1} \sum_{l=1}^{\mathcal{L}} \tilde{y}(\boldsymbol{s}_l)$, where $\mathcal{L}$ is the number of equally-spaced prediction grid points covering domain $\mathcal{D}$. To convert $\tilde{y}(\mathcal{D})$ to a PPD for total AGB, we multiply each sample by $|\mathcal{D}|$. This process also can be used to generate mean AGB density PPDs for sub-domains by integrating pixel-level PPDs over redefined areal blocks, i.e., $|\mathcal{B}|^{-1} \int_{\mathcal{B}} \tilde{y}(\boldsymbol{s}) \mathrm{d}\boldsymbol{s}$, where $\mathcal{B}$ is a sub-region of $\mathcal{D}$. See Banerjee et al. (2014, Chap. 7) for a more thorough discussion on summarizing PPDs over areal units. Total AGB estimates presented in the results tables were obtained by calculating the median of the total AGB PPD for the study area (labeled *Est*). The standard deviation of the total AGB PPD is presented in the results tables as well and serves as a model-based uncertainty estimate for total AGB (labeled *SD*). Relative standard deviations (labeled *RSD* in the results tables) were calculated as $SD/Est * 100\%$.

3.2.7 Cross-validation

To examine pixel-level predictive performance of the six candidate models, for each study site, a 10-fold holdout design was constructed by randomly assigning AGB plot observations to 10 approximately equal size groups. Square-root transformed AGB for each holdout group was sequentially predicted given model parameters estimated using data in the remaining nine groups. 10-fold holdout root mean squared prediction error (*RMSE*) was calculated using back-transformed holdout posterior predicted medians and observed AGB for each model at all four sites. The model with the lowest *RMSE* was considered the *best* pixel-level predictor. The holdout predictions used for *RMSE* calculation were also used to generate the holdout residual semivariograms presented in Fig. 3.1.

3.3 Results

3.3.1 Comparing non-spatial and spatial models

Tables 3.1, 3.2, 3.3 and 3.4 present parameter posterior distribution summaries and prediction accuracy estimates for the six candidate models applied to the BCEF, CPCRW, TNWR and TVSF datasets. At all four sites, results show the *Null* models to be more precise estimators of total AGB (labeled *Est* in Tables 3.1, 3.2, 3.3 and 3.4) than their *Spatial* counterpart models, evidenced by lower total AGB standard deviations (labeled *SD* in Tables 3.1, 3.2, 3.3 and 3.4). However, the *Null* models also show dramatically higher *RMSE* accuracy assessments compared to the *Spatial* models at all sites, suggesting that the models including spatial random effects produced more accurate predictions.

The increased accuracy coupled with apparent decreased precision of the *Spatial* model predictions highlight the role of distributional assumptions concerning the *SD* and *RMSE* estimates. The *RMSE* metric employed here does not rely on distributional assumptions concerning stochastic model components because it only assesses the ability of the PPD median *point estimate* to approximate observed AGB at plot locations—see Efron & Gong (1983) for discussion on robustness of cross-validation strategies to modeling assumptions.

On the other hand, model-based summaries of PPDs, such as, the *SD* precision metric presented in Tables 3.1, 3.2, 3.3 and 3.4, heavily rely on modeling assumptions. Strong violations of distributional assumptions made during model-fitting can lead to unrealistic uncertainty estimates. Looking at the *Null* model holdout residual semivariograms in Fig. 3.1, we see that, at all four sites, there is strong residual spatial autocorrelation. This means that the *Null* models likely violate the *iid* error assumption imposed to conduct model-based inference about any parameters or predictions using standard deviations or other posterior distribution summaries. The *SD* metrics for total AGB resulting from the *Null* models are underestimated because the residuals are falsely assumed to be independent of one another (Griffith, 2005). The *Spatial* model semivariograms, however, show no signs of residual spatial structure, indicating that $w(\boldsymbol{s})$ has effectively absorbed any extraneous spatial variability (Fig. 3.1). Relaxing the *iid* error assumption by incorporating spatial random effects allows for modelers to better interpret model-based uncertainty measures. Total AGB *SD* estimates garnered via the *Spatial* models are more reliable than the *Null* model *SD*s.

A similar phenomenon is observed when comparing the *Tree Cover* and *Spatial+Tree Cover* model predictive performance measures, although less substantial. Comparing the *Null* and *Tree Cover* holdout residual semivariograms in Fig. 3.1 shows that introducing the Landsat-derived tree cover variable absorbs a portion of the spatially structured variability in AGB at all four sites. However, spatial structure still exists in the holdout residuals of the *Tree Cover* models. We argue that, due to positive residual spatial autocorrelation, *SD* estimates are underestimated for the *Tree Cover* models as well. Comparisons between the univariate non-spatial (*Null* and *Tree Cover*) and spatial (*Spatial* and *Spatial+Tree Cover*) models underscore the need to adhere to posited distributional assumptions concerning stochastic model components when attempting to interpret model-dependent measures of predictive performance, including pixel-level and areally integrated PPDs.

Table 3.1: Bonanza Creek Experimental Forest (BCEF) candidate model parameter estimates, prediction accuracy metrics and total aboveground biomass (AGB) estimates. Est = estimated total AGB with associated 95% credible interval in parentheses displayed in teragrams (Tg). SD = posterior predictive standard deviation for AGB total estimate. RSD = relative standard deviation for AGB total estimate ($SD/Est * 100\%$). $RMSE$ = 10-fold holdout root mean squared prediction error accuracy assessment measured in megagrams per hectare (Mg/ha).

		Null	Tree Cover	Spatial	Spatial+Tree Cover	Coregionalization	Coregionalization+Tree Cover
Parameter posterior quantile summaries 50% (2.5%, 97.5%)	β_{0y}	7.07 (6.56, 7.60)	-0.41 (-1.93, 1.10)	7.05 (6.06, 7.81)	1.20 (-0.75, 3.19)	6.65 (5.90, 7.31)	0.97 (-0.56, 2.61)
	β_{1y}	—	0.10 (0.08, 0.12)	—	0.08 (0.05, 0.10)	—	0.08 (0.06, 0.10)
	β_{0z}	—	—	—	—	9.35 (7.81, 11.13)	3.16 (1.81, 4.78)
	β_{1z}	—	—	—	—	—	0.09 (0.08, 0.11)
	τ_y^2	18.79 (15.90, 22.47)	13.56 (11.50, 16.18)	1.01 (0.27, 6.19)	3.43 (1.01, 6.96)	2.36 (1.58, 3.24)	1.81 (0.88, 3.08)
	τ_z^2	—	—	—	—	2.14 (1.54, 2.89)	2.03 (1.36, 2.82)
	σ_y^2	—	—	17.83 (12.42, 23.46)	10.19 (6.34, 14.83)	—	—
	$K[1,1]$	—	—	—	—	20.71 (16.50, 25.57)	13.60 (10.40, 17.98)
	$K[1,2]$	—	—	—	—	25.16 (21.25, 29.50)	14.77 (11.56, 18.66)
	$K[2,2]$	—	—	—	—	41.39 (36.12, 48.35)	29.51 (25.76, 34.92)
	$\text{cor}(\boldsymbol{u}_y, \boldsymbol{u}_z)$	—	—	—	—	0.86 (0.77, 0.92)	0.74 (0.61, 0.86)
	er_y (km)	—	—	0.40 (0.23, 2.88)	0.32 (0.17, 3.03)	0.78 (0.59, 1.01)	0.36 (0.23, 0.54)
	er_z (km)	—	—	—	—	2.37 (1.37, 5.83)	2.15 (1.49, 3.61)
Prediction Acc. and Total Est.	Est (Tg)	1.328 (1.183, 1.495)	1.307 (1.192, 1.446)	1.323 (1.110, 1.582)	1.286 (1.116, 1.485)	1.323 (1.191, 1.484)	1.310 (1.163, 1.469)
	SD (Tg)	0.0809	0.0665	0.1235	0.0938	0.0787	0.0784
	RSD	6.09%	5.09%	9.33%	7.29%	5.95%	5.98%
	RMSE (Mg/ha)	68.58	60.53	47.62	47.44	35.75	36.40

Table 3.2: Caribou-Poker Creeks Research Watershed (CPCRW) candidate model parameter estimates, prediction accuracy metrics and total aboveground biomass (AGB) estimates. *Est* = estimated total AGB with associated 95% credible interval in parentheses displayed in teragrams (Tg). *SD* = posterior predictive standard deviation for AGB total estimate. *RSD* = relative standard deviation for AGB total estimate ($SD/Est * 100\%$). *RMSE* = 10-fold holdout root mean squared prediction error accuracy assessment measured in megagrams per hectare (Mg/ha).

		Null	Tree Cover	Spatial	Spatial+Tree Cover	Coregionalization	Coregionalization+Tree Cover
Parameter posterior quantile summaries 50% (2.5%, 97.5%)	β_{0y}	5.68 (5.13, 6.24)	-0.11 (-1.04, 0.81)	5.65 (4.72, 6.75)	1.00 (-0.35, 2.64)	4.97 (3.92, 5.88)	2.06 (0.76, 3.54)
	β_{1y}	—	0.09 (0.08, 0.11)	—	0.08 (0.05, 0.09)	—	0.05 (0.03, 0.07)
	β_{0z}	—	—	—	—	5.25 (3.14, 7.15)	0.66 (-1.28, 2.24)
	β_{1z}	—	—	—	—	—	0.08 (0.07, 0.10)
	τ_y^2	11.42 (9.16, 14.43)	5.16 (4.13, 6.55)	1.65 (0.76, 2.62)	2.35 (0.80, 3.76)	1.66 (1.22, 2.21)	1.99 (1.25, 2.75)
	τ_z^2	—	—	—	—	1.82 (1.28, 2.52)	2.15 (1.51, 2.91)
	σ_y^2	—	—	8.96 (6.10, 13.58)	2.98 (1.20, 5.41)	—	—
	$K[1,1]$	—	—	—	—	9.63 (6.83, 12.95)	4.86 (2.59, 8.59)
	$K[1,2]$	—	—	—	—	14.03 (10.56, 18.38)	7.10 (4.74, 10.03)
	$K[2,2]$	—	—	—	—	26.32 (20.46, 34.39)	14.60 (11.61, 19.55)
	$\text{cor}(\boldsymbol{u}_y, \boldsymbol{u}_z)$	—	—	—	—	0.89 (0.81, 0.94)	0.85 (0.71, 0.92)
	er_y (km)	—	—	1.28 (0.55, 2.91)	0.33 (0.10, 2.19)	1.96 (1.40, 2.95)	1.20 (0.76, 1.92)
	er_z (km)	—	—	—	—	2.49 (1.66, 6.35)	2.31 (1.32, 6.01)
Prediction Acc. and Total Est.	Est (Tg)	0.464 (0.396, 0.530)	0.462 (0.414, 0.514)	0.457 (0.380, 0.554)	0.444 (0.387, 0.506)	0.413 (0.380, 0.449)	0.403 (0.362, 0.442)
	SD (Tg)	0.0355	0.0250	0.0444	0.0316	0.0173	0.0199
	RSD	7.65%	5.41%	9.73%	7.12%	4.18%	4.94%
	RMSE (Mg/ha)	41.55	27.41	23.88	24.06	20.78	21.19

Table 3.3: Tetlin National Wildlife Refuge (TNWR) candidate model parameter estimates, prediction accuracy metrics and total aboveground biomass (AGB) estimates. *Est* = estimated total AGB with associated 95% credible interval in parentheses displayed in teragrams (Tg). *SD* = posterior predictive standard deviation for AGB total estimate. *RSD* = relative standard deviation for AGB total estimate ($SD/Est * 100\%$). *RMSE* = 10-fold holdout root mean squared prediction error accuracy assessment measured in megagrams per hectare (Mg/ha).

		Null	Tree Cover	Spatial	Spatial+Tree Cover	Coregionalization	Coregionalization+Tree Cover
Parameter posterior quantile summaries 50% (2.5%, 97.5%)	β_{0y}	4.19 (3.46, 4.92)	0.71 (-0.47, 1.92)	4.23 (3.15, 5.40)	1.36 (-0.04, 3.04)	3.87 (3.13, 4.63)	0.98 (-0.05, 1.91)
	β_{1y}	—	0.07 (0.05, 0.10)	—	0.06 (0.04, 0.09)	—	0.07 (0.05, 0.09)
	β_{0z}	—	—	—	—	3.09 (2.38, 3.68)	-0.39 (-1.00, 0.14)
	β_{1z}	—	—	—	—	—	0.08 (0.07, 0.09)
	τ_y^2	14.73 (11.36, 19.61)	10.43 (8.06, 13.80)	0.60 (0.18, 2.28)	1.32 (0.51, 3.80)	0.79 (0.20, 2.02)	0.74 (0.20, 1.62)
	τ_z^2	—	—	—	—	1.96 (1.39, 2.70)	2.24 (1.63, 2.95)
	σ_y^2	—	—	12.90 (8.57, 21.39)	8.59 (5.46, 12.98)	—	—
	$K[1,1]$	—	—	—	—	17.70 (12.69, 24.90)	11.50 (8.12, 15.24)
	$K[1,2]$	—	—	—	—	13.58 (11.27, 16.92)	8.48 (7.05, 10.16)
	$K[2,2]$	—	—	—	—	15.31 (13.57, 17.77)	9.05 (7.80, 10.73)
	$\mathrm{cor}(\boldsymbol{u}_y, \boldsymbol{u}_z)$	—	—	—	—	0.83 (0.75, 0.89)	0.83 (0.76, 0.89)
	er_y (km)	—	—	0.36 (0.19, 0.72)	0.28 (0.15, 0.83)	0.57 (0.44, 0.74)	0.39 (0.30, 0.50)
	er_z (km)	—	—	—	—	5.23 (2.96, 9.71)	5.19 (2.80, 19.17)
Prediction Acc. and Total Est.	Est (Tg)	8.654 (7.091, 11.007)	8.133 (6.716, 10.126)	8.702 (6.189, 12.429)	7.996 (5.688, 10.716)	9.082 (7.662, 11.176)	8.381 (7.183, 9.682)
	SD (Tg)	1.0108	0.861	1.6128	1.1950	0.9482	0.6477
	RSD	11.68%	10.59%	18.53%	14.94%	10.44%	7.73%
	RMSE (Mg/ha)	51.89	44.70	37.37	38.01	34.03	30.89

Table 3.4: Tanana Valley State Forest (TVSF) candidate model parameter estimates, prediction accuracy metrics and total aboveground biomass (AGB) estimates. *Est* = estimated total AGB with associated 95% credible interval in parentheses displayed in teragrams (Tg). *SD* = posterior predictive standard deviation for AGB total estimate. *RSD* = relative standard deviation for AGB total estimate ($SD/Est * 100\%$). *RMSE* = 10-fold holdout root mean squared prediction error accuracy assessment measured in megagrams per hectare (Mg/ha).

		Null	Tree Cover	Spatial	Spatial+Tree Cover	Coregionalization	Coregionalization+Tree Cover
Parameter posterior quantile summaries 50% (2.5%, 97.5%)	β_{0y}	7.10 (6.55, 7.63)	0.48 (-0.97, 1.92)	7.09 (6.19, 8.02)	2.14 (0.24, 4.42)	7.41 (6.74, 8.01)	2.09 (0.74, 3.52)
	β_{1y}	—	0.09 (0.07, 0.11)	—	0.07 (0.04, 0.10)	—	0.08 (0.06, 0.09)
	β_{0z}	—	—	—	—	8.35 (7.68, 9.06)	3.03 (2.50, 3.68)
	β_{1z}	—	—	—	—	—	0.08 (0.07, 0.09)
	τ_y^2	19.53 (16.53, 23.31)	14.47 (12.19, 17.30)	0.91 (0.21, 3.54)	1.47 (0.47, 3.51)	0.99 (0.35, 1.53)	0.68 (0.32, 1.85)
	τ_z^2	—	—	—	—	2.61 (2.03, 3.20)	2.48 (2.05, 2.96)
	σ_y^2	—	—	18.60 (14.04, 24.08)	13.15 (9.67, 18.23)	—	—
	$K[1,1]$	—	—	—	—	23.80 (18.93, 29.98)	17.38 (13.50, 21.67)
	$K[1,2]$	—	—	—	—	22.61 (19.30, 26.37)	16.48 (14.50, 18.90)
	$K[2,2]$	—	—	—	—	40.00 (36.73, 43.89)	27.86 (25.89, 29.63)
	$\mathrm{cor}(\boldsymbol{u}_y, \boldsymbol{u}_z)$	—	—	—	—	0.73 (0.67, 0.80)	0.75 (0.70, 0.81)
	er_y (km)	—	—	0.56 (0.35, 1.28)	0.42 (0.26, 0.87)	0.73 (0.60, 0.91)	0.47 (0.40, 0.58)
	er_z (km)	—	—	—	—	7.10 (5.14, 11.17)	4.49 (3.39, 6.72)
Prediction Acc. and Total Est.	Est (Tg)	49.05 (43.47, 54.73)	47.18 (42.70, 51.86)	49.31 (40.31, 60.10)	46.53 (39.71, 54.81)	55.82 (50.22, 61.59)	52.54 (49.31, 57.12)
	SD (Tg)	2.8404	2.4088	5.0405	4.1593	2.9045	2.0866
	RSD	5.79%	5.11%	10.28%	8.94%	5.20%	3.97%
	RMSE (Mg/ha)	72.37	63.46	42.68	42.15	31.00	28.95

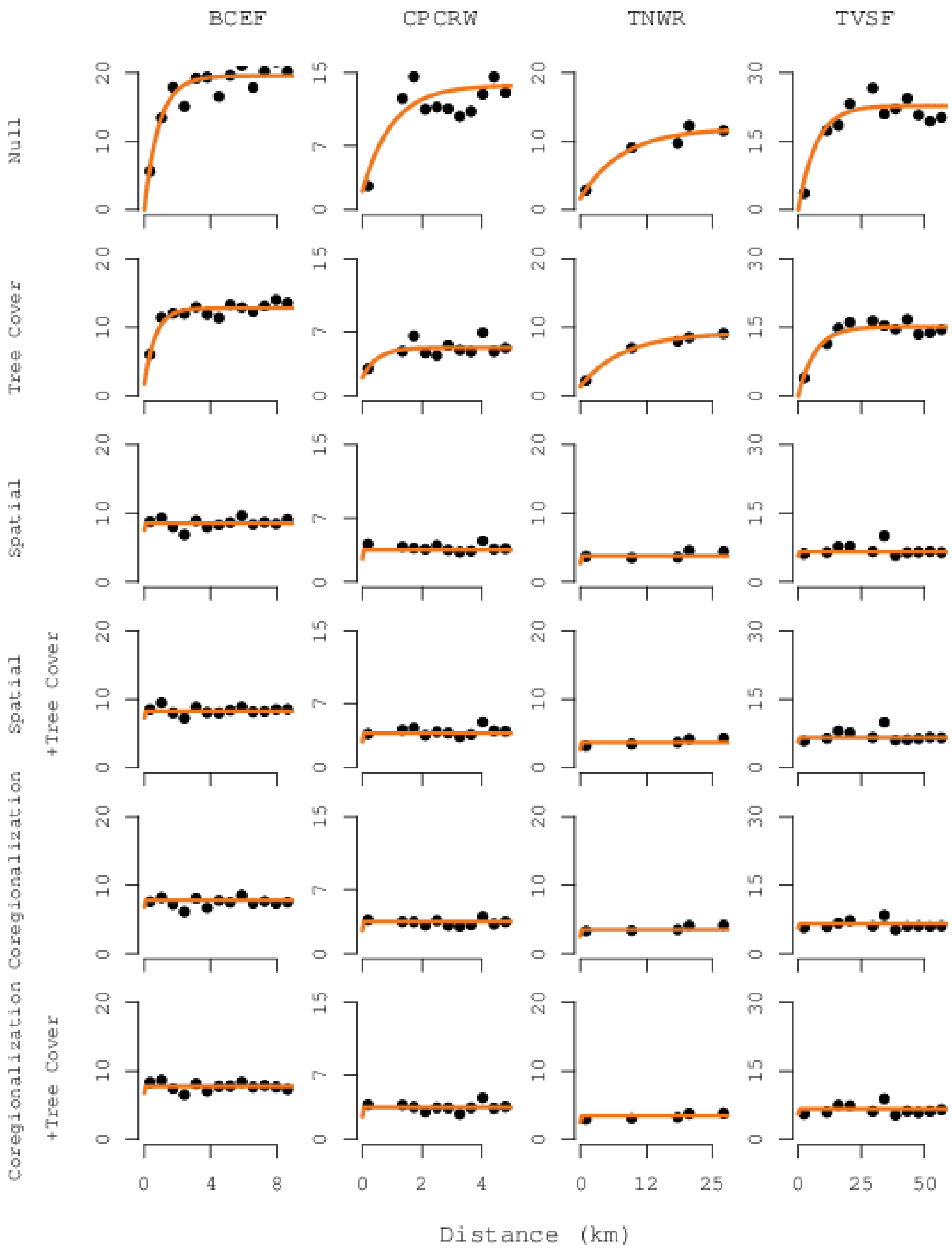


Fig. 3.1: Holdout residual semivariograms for the six candidate models at all four study sites.

3.3.2 Comparing spatial and coregionalization models

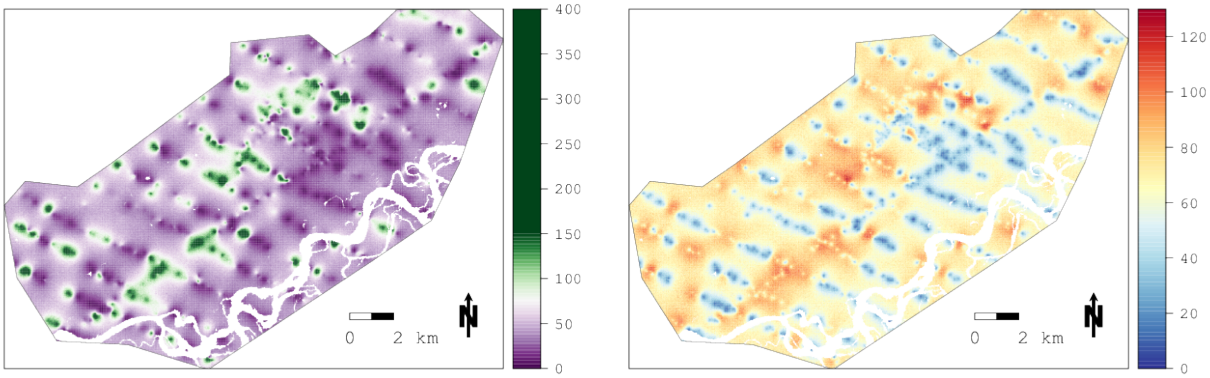
At all four sites, the lidar-informed *Coregionalization* models outperformed their counterpart lidar-free *Spatial* models by producing lower *RMSE* accuracy assessments and total AGB *SD* metrics. Similarly, the *Coregionalization+Tree Cover* models produced better accuracy measures than the *Spatial+Tree Cover* models. Prediction gains resulting from the incorporation of lidar information are not surprising considering that lidar-derived measures of forest height are strongly related to AGB. It is encouraging, however, that this relationship produces appreciable gains in total AGB estimation precision at such low lidar sampling intensities—see Section 3.2.3 for description of the lidar dataset subsampling. Tables 3.1, 3.2, 3.3 and 3.4 show the correlation between the spatial random effects for the AGB and lidar sub-models (labeled $cor(\boldsymbol{u}_y, \boldsymbol{u}_z)$ in Tables 3.1, 3.2, 3.3 and 3.4). Median point estimates of correlation range between 0.73 and 0.89 indicating strong correspondence between square-root transformed AGB density and 90^{th} percentile height. Since the relationship between the AGB and lidar responses is so substantial, considerable improvements in predictive performance and total AGB estimation precision are realized. At all sites, relative standard deviations (labeled *RSD* in Tables 3.1, 3.2, 3.3 and 3.4) for the *Coregionalization* models are nearly half that of the *Spatial* models.

At the sites with sparser field and lidar sampling, i.e., TVSF and TNWR, we see further gains in pixel-level prediction accuracy and total AGB certainty when the Landsat-based tree cover product is included—evidenced by *Coregionalization+Tree Cover* models having lower *RSD* and *RMSE* estimates than counterpart *Coregionalization* frameworks. Examining effective range estimates (labeled er_y and er_z in Tables 3.1, 3.2, 3.3 and 3.4) offers insight concerning this phenomenon. Effective range estimates help us understand how far-reaching the underlying spatial dependence structure is in the responses, i.e., $y(\boldsymbol{s})$ and $z(\boldsymbol{s})$, after accounting for covariates. The longer the effective range, the larger the proximate neighborhood each pixel draws information from. The er_z median point estimates for the TNWR and TVSF *Coregionalization* and *Coregionalization+Tree Cover* models extend less than the

10 km separation distance between lidar flight-lines at those sites (4.49 km $< er_z <$ 7.10 km). However, the effective range estimates for the corresponding models exceed the average flight-line separation distances of $<$ 1.5 km at BCEF and CPCRW (2.15 km $< er_z <$ 2.41 km). On average, pixel-level predictions at BCEF and CPCRW are borrowing more proximate lidar data to inform prediction because there is more available within the effective range of spatial dependence at these sites. Since field plots are also closer together at CPCRW and BCEF compared to TVSF and TNWR, pixel predictions are potentially tapping into more information provided by nearby AGB observations as well. Any AGB explanatory power offered by the Landsat tree cover product is already explained by simply borrowing strength from nearby field and lidar data to inform prediction at BCEF and CPCRW. However, since field and lidar data are much farther apart at TVSF and TNWR, we do see benefit from the wall-to-wall information provided by the Landsat-based tree cover product.

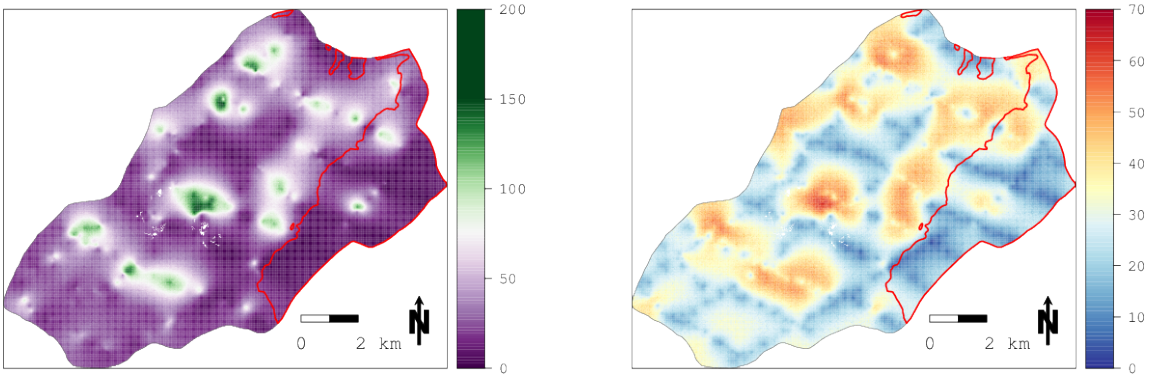
3.3.3 Comparing Coregionalization and Coregionalization+Tree Cover mapped predictions

Figs. 3.2, 3.3, 3.4 and 3.5 show maps of PPD median point estimates and standard deviations for AGB density using the *Coregionalization* and *Coregionalization+Tree Cover* models. The flight-lines where lidar was collected are easily discernible in the *Coregionalization* model prediction maps (Figs. 3.2 and 3.3). We see differentiation of AGB within lidar strips and near field plot locations. As one moves away from regions where information was collected, predicted AGB densities retreat to the global mean. Additionally, pixel-level uncertainty maps show a trend of higher predictive precision in close proximity to field and lidar observations compared to areas farther away. Even though the *Coregionalization* models provide reliable estimates of total AGB and show favorable *RMSE* accuracy assessments, the resulting prediction maps leave something to be desired. Having no information outside the flight-lines, the coregionalization models only rely on borrowing strength from nearby observations to inform prediction between strips. The addition of the wall-to-wall Landsat tree cover product allows for better differentiation of median point predictions outside lidar strips, leading to maps that are potentially more useful for assessing the spatial distribution of AGB across the landscape. At TVSF and TNWR, *RMSE* estimates suggest that the addition of the Landsat tree cover product leads to better map accuracy. However, at BCEF and CPCRW, the inclusion of Landsat tree cover appears to only offer cosmetic adjustments to resulting maps, not actual increases in prediction accuracy. It seems that the intensity of the field and lidar sample at BCEF and CPCRW mixed with constructing models to borrow information from neighboring locations offered more predictive advantage than simply including Landsat tree cover as a covariate.



(a) BCEF: Coregionalization predictions

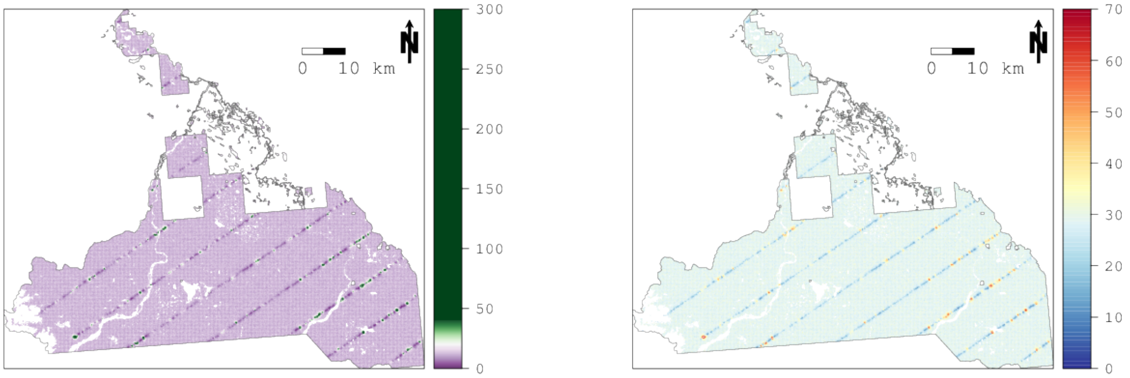
(b) BCEF: Coregionalization prediction standard deviations



(c) CPCRW: Coregionalization predictions

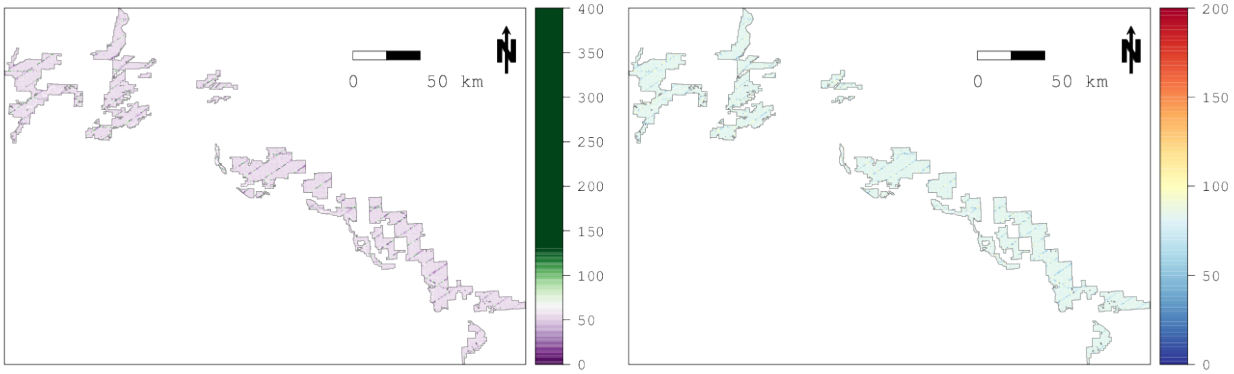
(d) CPCRW: Coregionalization prediction standard deviations

Fig. 3.2: Mapped pixel-level predictions of aboveground biomass density (a and c) and associated standard deviations (b and d) (Mg/ha) using the *Coregionalization* model for Bonanza Creek Experimental Forest (BCEF) and Caribou-Poker Creeks Research Watershed (CPCRW). The red polygon boundary shown in c and d defines the area burned during the 2004 Boundary Fire at CPCRW.



(a) TNWR: Coregionalization predictions

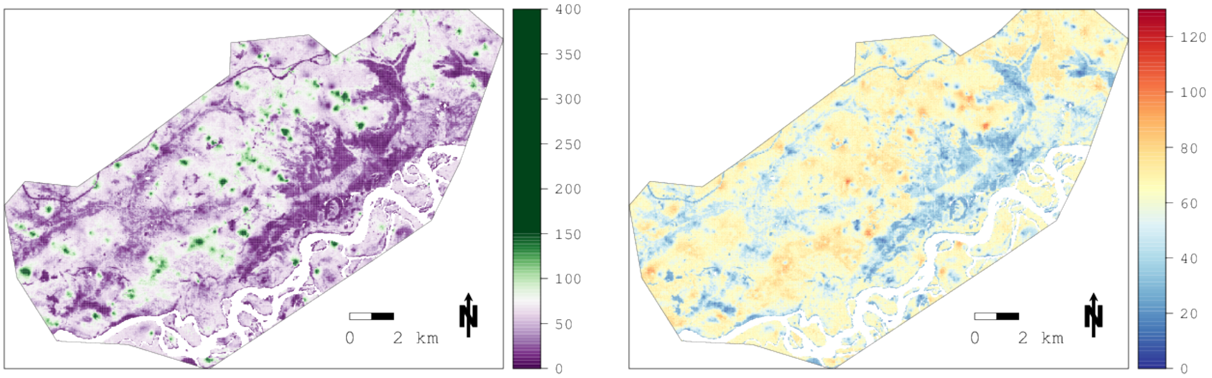
(b) TNWR: Coregionalization prediction standard deviations



(c) TVSF: Coregionalization predictions

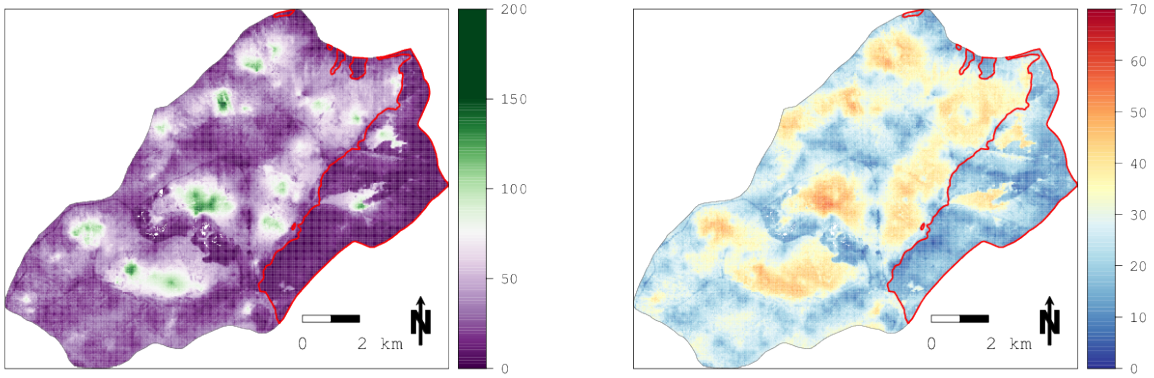
(d) TVSF: Coregionalization prediction standard deviations

Fig. 3.3: Mapped pixel-level predictions of aboveground biomass density (a and c) and associated standard deviations (b and d) (Mg/ha) using the *Coregionalization* model for Tetlin National Wildlife Refuge (TNWR) and Tanana Valley State Forest (TVSF).



(a) BCEF: Coregionalization+Tree Cover predictions

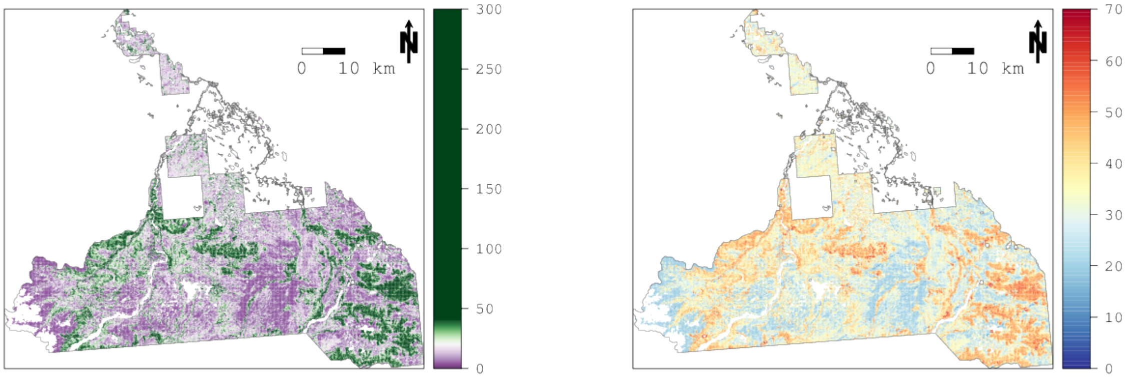
(b) BCEF: Coregionalization+Tree Cover prediction standard deviations



(c) CPCRW: Coregionalization+Tree Cover predictions

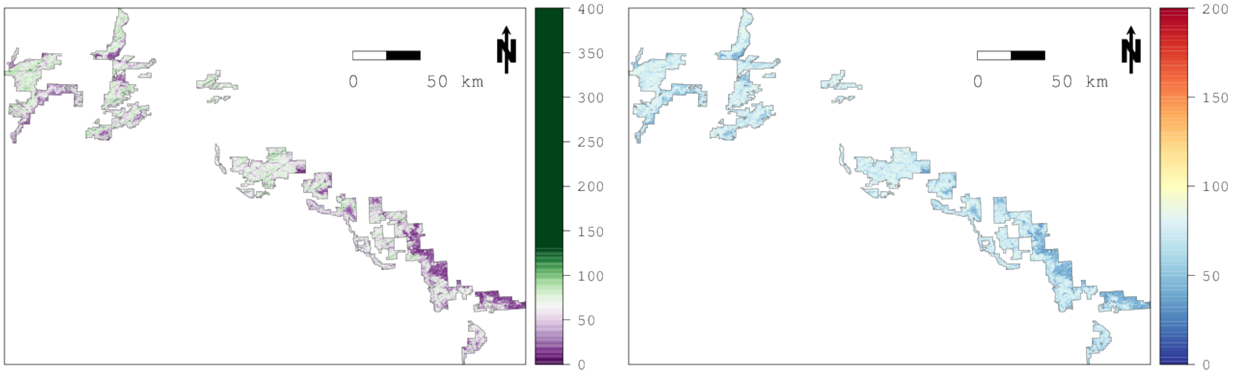
(d) CPCRW: Coregionalization+Tree Cover prediction standard deviations

Fig. 3.4: Mapped pixel-level predictions of aboveground biomass density (a and c) and associated standard deviations (b and d) (Mg/ha) using the *Coregionalization+Tree Cover* model for Bonanza Creek Experimental Forest (BCEF) and Caribou-Poker Creeks Research Watershed (CPCRW). The red polygon boundary shown in c and d defines the area burned during the 2004 Boundary Fire at CPCRW.



(a) TNWR: Coregionalization+Tree Cover predictions

(b) TNWR: Coregionalization+Tree Cover prediction standard deviations



(c) TVSF: Coregionalization+Tree Cover predictions

(d) TVSF: Coregionalization+Tree Cover prediction standard deviations

Fig. 3.5: Mapped pixel-level predictions of aboveground biomass density (a and c) and associated standard deviations (b and d) (Mg/ha) using the *Coregionalization+Tree Cover* model for Tetlin National Wildlife Refuge (TNWR) and Tanana Valley State Forest (TVSF).

3.3.4 Watershed unit AGB estimates at Caribou-Poker Creeks Research Watershed

Fig. 3.6 (left) shows estimated mean AGB densities for the 11 watershed units at CPCRW predicted using the *Coregionalization+Tree Cover* model with a permafrost polygon layer overlaid. Fig. 3.6 (right) shows a scatter plot revealing the relationship between AGB density and permafrost presence at the watershed unit scale. It is evident that average AGB density is negatively correlated with permafrost presence at the watershed unit level. By appropriately summarizing the model's PPD surface, it is possible to obtain an estimate of the correlation between AGB density and permafrost proportion at the watershed unit scale while accounting for uncertainty in the AGB density predictions. By calculating the correlation between the proportion of the watershed unit covered in permafrost with each MCMC sample of average AGB density, we can obtain a PPD summary characterizing their degree of correlation. The median correlation point estimate representing the relationship between permafrost and watershed-unit AGB density is -0.685. The 95 percent upper and lower credible bounds for this estimate are -0.587 and -0.764, respectively. Since this interval does not include zero, we can conclude that permafrost presence is negatively related to AGB density with 95 percent certainty.

We can also explore potential causes for departure from the overall trend between watershed unit-level AGB density and permafrost (dashed line in Fig. 3.6 (right)). We see that the watershed units furthest below the trend-line (points highlighted in red in Fig. 3.6 (right)) correspond to watersheds with the highest proportion of area burned during the 2004 Boundary Fire, suggesting that these areas are still in a state of recovery from fire after a decade. We may expect these two watershed units to continue to accrue AGB until they begin to align with the overall trend between AGB density and permafrost. Fig. 3.6 captures the effect of fire history on watershed-unit AGB variability beyond what would be expected due to variation in permafrost coverage at CPCRW.

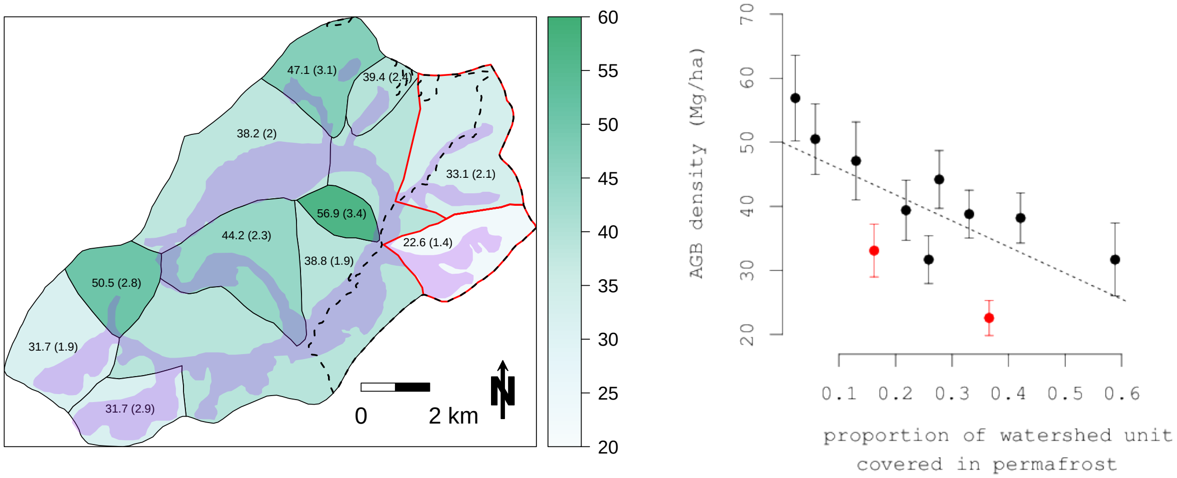


Fig. 3.6: Left figure maps watershed unit-level mean aboveground biomass (AGB) (Mg/ha) with associated standard deviations in parentheses using the *Coregionalization+Tree Cover* model for Caribou-Poker Creeks Research Watershed. Solid polygon boundaries delineate watershed units. Translucent purple polygon identifies area covered in permafrost. The black dashed line delineates the border of the 2004 Boundary Fire. The red outlined watershed units are identified as the two units with the highest proportion of area burned during the 2004 Boundary Fire. The right figure shows a scatter plot highlighting the relationship between the proportion of permafrost in a watershed unit and AGB density. The dashed trend-line shows the general relationship between AGB density and permafrost proportion. The points shown in red correspond to the red outlined watershed units in the left figure.

3.4 Discussion

The goal of this analysis was to develop and test the performance of a statistical modeling framework that can (1) incorporate partial coverage lidar data and wall-to-wall Landsat products to improve AGB density prediction; and (2) accommodate extraneous residual structure, thereby allowing for more reliable model-based characterization of uncertainty and improved prediction. Through model comparison we were able to show that a coregionalization framework can effectively be used to couple sampled lidar and field data to improve pixel-level AGB density prediction accuracy and increase confidence in total AGB estimates. Rigorous model comparison also demonstrated the adverse effects of residual spatial autocorrelation and how including appropriately specified random effects within a Bayesian hierarchical framework can absorb spatial dependence in the response variable not accounted for by the covariates.

An examination of *Coregionalization* and *Coregionalization+Tree Cover* model mapped predictions revealed benefits of including wall-to-wall information to better visualize spatial variability in AGB. We also saw that greater differentiation of AGB median point estimates provided by Landsat-based tree cover data does not necessarily lead to improved pixel-level predictive performance. It can be the case that simply constructing models to leverage proximate observations may improve prediction more than supplementing auxiliary data, especially when sites are intensively sampled and ancillary information is only weakly related to the response variable.

We demonstrated the ability to summarize PPDs from the *Coregionalization+Tree Cover* model to generate small-area estimates of mean AGB density at CPCRW. Small-area estimates are a challenge for traditional estimation methods (Breidenbach & Astrup, 2012; Goerndt et al., 2011), and have also limited the ability to investigate fine-scale spatial heterogeneity within larger management units such as the TVSF or TNWR. In contrast, our results demonstrate how small-area integrated PPDs can be used to examine correlation with other information about sub-domains while accounting for prediction uncertainty. Relating watershed unit-level AGB density to permafrost and fire history polygon layers uncovered possible drivers of AGB variability among watershed units at CPCRW while explicitly accounting for uncertainty of AGB density predictions. This exercise shows the inferential potential of AGB PPD products generated using Bayesian hierarchical spatial models beyond simply mapping AGB density.

Upcoming satellite lidar missions, e.g., ICESat-2 and GEDI LiDAR, are designed to collect data along orbital transects, hence the information provided by these sensors will not be of complete spatial coverage. Results from this study suggest that it may be possible to implement a coregionalization approach to augment the FIA program's network of permanent sample plots over the contiguous United States with sampled space-based lidar to improve forest inventory estimates. Further, Landsat products can be incorporated to potentially improve point-level prediction accuracy and areal inventory estimates at a national scale.

The coregionalization framework detailed here is very computationally intensive. After dramatically subsampling the lidar dataset (described in Section 3.2.3), the TVSF *Coregionalization* model took two days to complete a single MCMC chain of 50 000 iterations on a Linux workstation equipped with an eight-core processor leveraging threaded BLAS and LAPACK C++ libraries (`www.netlib.org/blas` and `www.netlib.org/lapack`). When lidar and/or field observations exceed 10 000, fitting this class of multivariate spatial models within a Bayesian paradigm becomes intractable on typical desktop computers due to computational difficulties related to inverting massive matrices. Recent advances in GP modeling theory and computation have unearthed new approaches for estimating spatial random effects that circumvent the limitations imposed by the necessity to invert large matrices (Datta et al., 2016a,b; Finley et al., 2017). Implementing Nearest-neighbor Gaussian process priors (NNGP) in place of traditional GPs alleviates the need to invert large covariance matrices, allowing modelers to effectively estimate complex Bayesian spatial models using many remote sensing and field observations on standard workstations. In the future, we will investigate the approximation of GPs with NNGPs to allow for the inclusion of more field and remote sensing samples within coregionalization frameworks to further improve pixel-level prediction and areal estimation certainty. We also plan to extend this class of multivariate spatial models to allow for the simultaneous prediction of multiple forest variables, e.g., tree density, basal area and AGB, while leveraging spatial cross-correlations between all responses. By jointly predicting many forest inventory parameters we will be able to preserve the inherent relationships between them and use those intrinsic correlations to aid spatial prediction and subsequent areal estimation of all included forest inventory parameters.

Chapter 4

MODELING FOREST BIOMASS AND GROWTH: COUPLING LONG-TERM INVENTORY AND LIDAR DATA

Abstract: Combining spatially-explicit long-term forest inventory and remotely sensed information from Light Detection and Ranging (lidar) datasets through statistical models can be a powerful tool for predicting and mapping aboveground biomass (AGB) at a range of geographic scales. We present and examine a novel modeling approach to improve prediction of AGB and estimate AGB growth using lidar data. The proposed model accommodates temporal misalignment between field measurements and remotely sensed data—a problem pervasive in such settings—by including multiple time-indexed measurements at plot locations to estimate AGB growth. We pursue a Bayesian modeling framework that allows for appropriately complex parameter associations and uncertainty propagation through to prediction. Specifically, we identify a space-varying coefficients model to predict and map AGB and its associated growth simultaneously. The proposed model is assessed using lidar data acquired from NASA Goddard's LiDAR, Hyper-spectral & Thermal imager and field inventory data from the Penobscot Experimental Forest in Bradley, Maine. The proposed model outperformed the time-invariant counterpart models in predictive performance as indicated by a substantial reduction in root mean squared error. The proposed model adequately accounts for temporal misalignment through the estimation of forest AGB growth and accommodates residual spatial dependence. Results from this analysis suggest that future AGB models informed using remotely sensed data, such as lidar, may be improved by adapting traditional modeling frameworks to account for temporal misalignment and spatial dependence using random effects.

4.1 Introduction

Linking long-term forest inventory with air- and space-borne Light Detection and Ranging (lidar) datasets via regression models offers an attractive approach to mapping forest aboveground biomass (AGB) at stand, regional, continental, and global scales. lidar data have shown great potential for use in estimating spatially explicit forest variables, including AGB, over a range of geographic scales (Asner et al., 2009; Babcock et al., 2013, 2015; Finley et al., 2011; Iqbal et al., 2013; Muss et al., 2011; Næsset, 2011; Neigh et al., 2013). Encouraging results from these and many other studies have spurred massive investment in new lidar sensors, sensor platforms, as well as extensive campaigns to collect field-based calibration data. For example, ICESat-2—planned for launch in 2017—will be equipped with a lidar sensor able to gather data from space at unprecedented spatial resolutions (Abdalati et al., 2010). As currently proposed, ICESat-2 will be a photon-counting sensor capable of recording photon-returns on a $\approx$70 cm footprint (ICESat-2, 2015b). The Global Ecosystem Dynamics Investigation lidar (GEDI) will be an International Space Station mounted system capable of producing 25 m diameter footprint waveforms and is scheduled to be operational in 2018 (GEDI, 2014). One of GEDI's core objectives is to quantify the distribution of AGB at a fine spatial resolution. NASA Goddard's lidar, Hyper-spectral, and Thermal (G-LiHT) imager is an air-borne platform developed, in part, to examine how future space-originating lidar, e.g., ICESat-2, GEDI, or other platforms, may be combined with field-based validation measurements to build predictive models for AGB and other forest variables (Awadallah et al., 2013; Cook et al., 2013).

Along with increasing attention given to development of lidar data acquisition systems, there is continued and expanding interest in acquiring long-term forest inventory datasets for ecosystem monitoring initiatives (Albercht et al., 2012; Phillips et al., 1998; Talbot et al., 2014). Long-term forest inventories including the United States Forest Service's Forest Inventory and Analysis (FIA) program, Long-Term Ecological Research Network (LTER), and National Ecological Observatory Network (NEON) are collecting massive amounts of space and time indexed ecological data (Kampe et al., 2010; Michener et al., 2011; Smith, 2002). These datasets are able to provide the calibration measurements necessary for large-scale AGB model development and mapping using lidar systems (Healey et al., 2012).

Much of the interest in lidar based AGB mapping is to support carbon monitoring, reporting, and verification (MRV) systems, such as defined by the United Nations Programme on Reducing Emissions from Deforestation and Forest Degradation (UN-REDD) and NASA's Carbon Monitoring System (CMS) (Le Toan et al., 2011; Ometto et al., 2014; UN-REDD, 2009; CMS, 2010). There are several challenges to specifying models that yield prediction useful for MRV systems. Here, we consider two common data characteristics that, if ignored in the modeling process, can result in misleading model-based inference about AGB or similar forest variables of interest. The first, and central focus of this study, is temporal misalignment between lidar data acquisition and AGB calibration data measurements. The second is spatial independence among model residuals—a key assumption of regression models. These characteristics are not simple nuisances that need accommodation to ensure valid inference. Explicit development of models that incorporate misalignment and residual spatial dependence can yield improved prediction of forest AGB and growth superior to that of commonly applied models.

In regression models, temporal misalignment occurs when the response variable and covariates (e.g., lidar metrics) are measured at different times. Temporal misalignment is common in settings where permanent sample plots (PSPs)—remeasured on a periodic basis as part of a long-term inventory or monitoring system—are used to calibrate remotely sensed data. For example, Saatchi et al. (2011) use field data from inventories conducted between 1995 and 2005 when building regression models for AGB with lidar data. Zhang et al. (2014) use FIA inventories collected between 2001 and 2007 to predict AGB using lidar data for California. Blackard et al. (2008), trained MODIS based regression models using national inventory data ranging from 1990 to 2003. Gregoire et al. (2011), García et al. (2010), Chen et al. (2012) and Anderson et al. (2008) use data from long-term inventories spanning a three year period to calibrate regression models. Some researchers elect to retain field data from long-term inventories spanning 10+ year intervals (Saatchi et al., 2011; Blackard et al., 2008) to obtain calibration datasets with large numbers of observations, whereas other researchers sacrifice sample size by only using inventories spanning a small time interval (Gregoire et al., 2011; García et al., 2010; Chen et al., 2012; Anderson et al., 2008). Because the FIA (the largest long-term forest inventory network both in spatial extent and number of PSPs) only measures between ten and twenty percent of their PSPs annually, we should expect to continue to see temporal misalignment to be an point of concern when attempting to combine long-term inventory and remote sensing data (Bechtold & Patterson, 2005). As another example, the recently established long-term inventory, NEON, only plans to acquire field inventories from their PSPs every five years, but collect airborne remote sensing data over these plots annually (Kampe et al., 2010; Kao et al., 2012). In order to develop yearly maps of forest structure and prediction uncertainty leveraging the annual lidar acquisitions, NEON will likely need sound statistical modeling strategies that accommodate temporal misalignment.

Regression model parameters (i.e., intercept, slope coefficients associated with lidar metrics, and residual variance) estimated using temporally misaligned data can lead to erroneous model-based inference and prediction especially as forest disturbances operate on temporal time scales rarely aligned with field and remotely sensed inventory programs. For AGB modeling, the worst cases will arise when there is a large time lag between lidar acquisition (covariates) and AGB measurements (response variable). The posited relationship between lidar metrics and response variables will be further weakened if the PSPs are visited across a large time interval (e.g., FIA sample cycle lengths) in concert with rapid changes in AGB (e.g., due to wildfire or wind throw). In both cases, *large* is relative to forest growth rates and disturbance intervals. Strategies to address misalignment include discarding all measurements outside of some *acceptable* time interval around the lidar acquisition date, using growth models to project PSP measurements to align with lidar acquisition date, or simply assume temporal misalignment effects are negligible (Zhang et al., 2014; Blackard et al., 2008; Weiskittel et al., 2011). In many cases, such strategies are less than ideal—due to small resulting sample size, desire to avoid unrealistic assumptions about growth, or need to quantify components of uncertainty in AGB predictions.

Repeated time- and space-indexed measurements within long-term inventory datasets make possible an approach to the misalignment problem that does not discard data and facilitates uncertainty quantification on both AGB and growth. The framework we pursue builds on work by Banerjee & Johnson (2006) and Nathoo (2010) who model spatially varying growth curves using repeated measures. In both works, regression models are specified with intercept and slope coefficients arising from univariate or multivariate spatial Gaussian processes. Here, we also use a space-varying coefficients framework but incorporate lidar covariates to model and map AGB and growth simultaneously. In addition to accommodating temporal misalignment the proposed model captures residual spatial dependence (i.e., spatially structured variability not explained by lidar covariates) via a space-varying intercept (Gelfand et al., 2003; Banerjee et al., 2014). Similar to Banerjee & Johnson (2006) and Nathoo (2010), we use a Bayesian model specification that allows for appropriately complex parameter associations and uncertainty propagation through to prediction.

To better understand how lidar and long-term inventory datasets can be used to map AGB, we look to Penobscot Experimental Forest (PEF) in Bradley, Maine. Using the PEF's long-term inventory database along with G-LiHT data acquired for the area in 2012, we present and examine the proposed modeling approach to improve AGB prediction and estimate growth. The proposed model is compared with two benchmark regressions to examine predictive performance, ability to account for temporal misalignment, and capacity to accommodate spatial dependence.

4.2 Methods

4.2.1 Field data description and preparation

The PEF is a 1 600 hectare (ha) tract of Acadian forest land located in Bradley, Maine (44° 52' N, 68° 38' W) (Fig. 4.1). The average annual temperature and precipitation near Bradley is 6°C and 110 centimeters (cm) respectively. Composition of the PEF is a mix of coniferous and deciduous species including spruce (*Picea spp.*), balsam fir (*Abies balsamea*), red maple (*Acer rubrum*), birch (*Betula spp.*), and aspen (*Populus spp.*), among others. Since the 1950s, the PEF has been subjected to routine management and monitoring initiatives. With over 600 PSPs currently established and scheduled for remeasurement on an approximate ten-year cycle (and immediately pre- and post-harvest), the PEF has a wealth of calibration data available for fitting AGB models using lidar (Brissette & Kenefic, 2008). The PEF is divided into over 50 management units (MU). Within each MU, different silvicultural treatments are implemented, e.g., unregulated harvest, shelterwood, diameter limit cutting, or natural regeneration (Fig. 4.1). MU 32 is a mature, natural area that serves as a reference stand for the PEF with limited harvesting or other management actions conducted since 1954. MU 8 and MU 22 were commercially clearcut in 1984 and 1989, respectively (Fig. 4.1). MU 23 is undergoing a three-stage shelterwood management plan (Sendak et al., 2003). For further details regarding the PEF's MU silvicultural treatments see Hayashi et al. (2014).

For this study we use 2 203 field measurements from 604 PSPs (Fig. 4.2). Only post-disturbance measurements from each PSP were included. Further, only PSPs with at least one post-disturbance measurement recorded between 1999 and 2011 were used. Data for the 33 measurements collected in 2012 were held out for model fit comparison (2012 was the most current data available for the PEF at the time of our study). Each PSP contains three nested sampling areas with radii of approximately 16 m (1/5 acre), 8 m (1/20 acre), and 5 m (1/50 acre), which are used to measure trees of progressively smaller diameter at breast height (dbh) classes (Brissette et al., 2012). For this analysis, only stems with dbh measures greater than or equal to 12.7 cm (5 in) were used to calculate AGB, which precluded trees measured only in the 8 m and 5 m plots. AGB was calculated using dbh based, species specific, allometric equations provided by Jenkins et al. (2003). The response variable used in subsequent regression models was calculated for each PSP measurement date by summing individual tree AGB values, scaling to megagrams per ha (Mg/ha), then applying a square-root transformation to better approximate a Gaussian distribution and ensure positive support following back-transformation of predicted values.

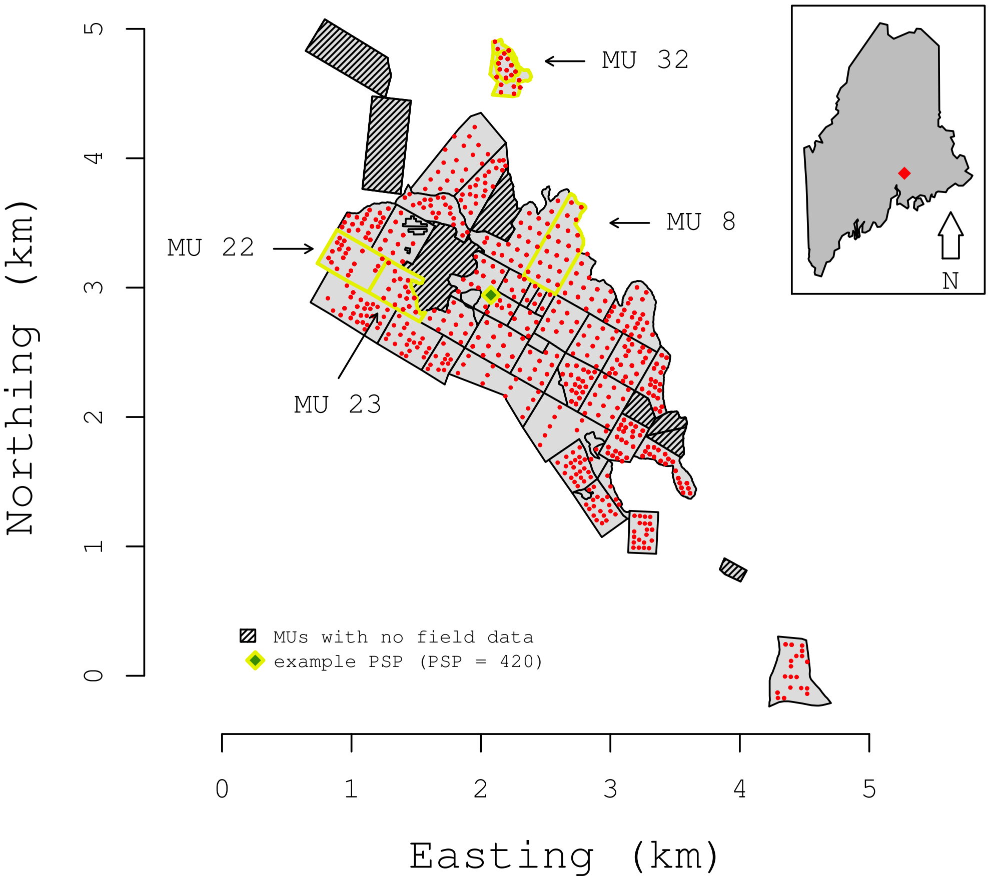


Fig. 4.1: Map of PEF. PSP's highlighted in red. Example PSP (site 420) colored in green. Black polygon boundaries outline different management units (MU) on the PEF. Hashed out polygons identify MUs with no inventory data. Select MUs have been labeled and highlighted in yellow. MU 8 and MU 22 were clear cut in 1984 and 1989 respectively. MU 32 is an old-growth stand that has experienced very limited management activity since 1954. MU 23 is a stand undergoing a three-stage shelterwood harvesting technique. The inset map in the upper right corner shows the location of the PEF with respect to Maine.

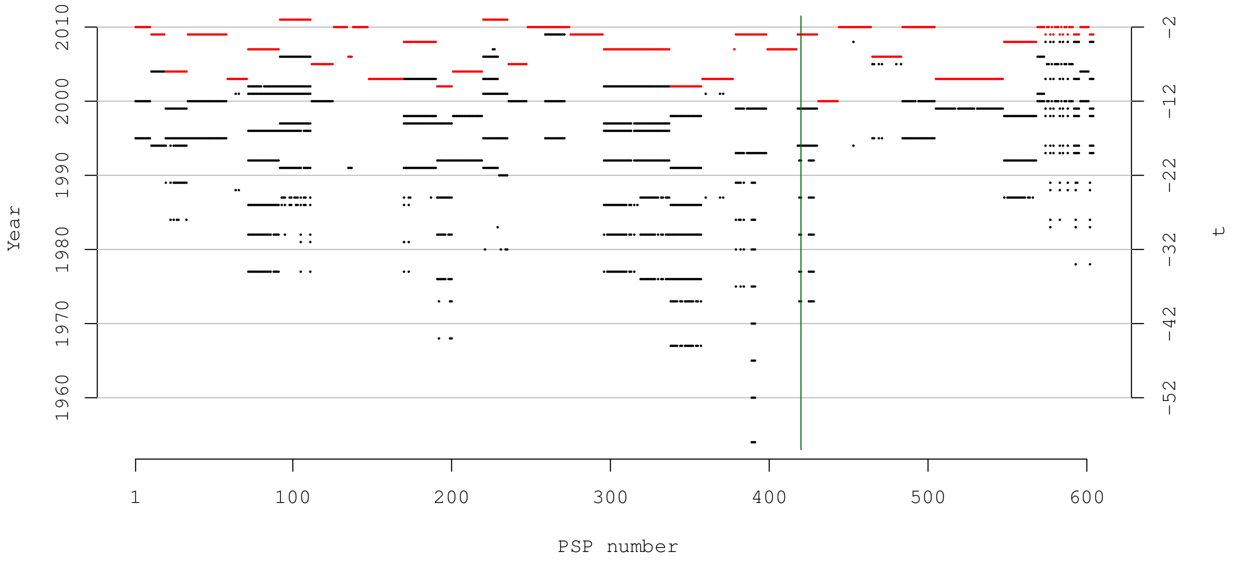


Fig. 4.2: Aboveground biomass remeasurements used in analysis. Most recent remeasurement for each permanent sample plot (PSP) is highlighted in red. Horizontal axis is the PSP identification number. Left vertical axis shows the remeasurement year and the right vertical axis shows the corresponding t value. t is the number of years after lidar collection. Since all remeasurements were conducted before the year of lidar acquisition, all t's are negative. Vertical green line identifies example PSP (site 420, highlighted in Fig. 4.4).

4.2.2 lidar acquisition and preparation

Large footprint waveforms characteristic of upcoming space-based lidar instruments were simulated using discrete, multistop returns from the G-LiHT airborne imager collected over the PEF in 2012. G-LiHT is a portable multi-sensor airborne system developed by NASA Goddard Space Flight Center that simultaneously maps the composition and structure of terrestrial ecosystems. The G-LiHT airborne laser scanner (VQ-480, Riegl Laser Measurement Systems, Horn, Austria) uses a 1 550 nanometer (nm) laser that provides an effective measurement rate of up to 150 kHz along a 60° swath perpendicular to the flight direction. At a nominal flying altitude of 335 meters (m), each laser pulse has a footprint approximately 10 cm in diameter and is capable of producing up to 8 returns.

Pseudo-waveforms were created over the PEF's 604 PSPs by aggregating G-LiHT lidar returns and weighting return heights using a Gaussian shaped 25 m diameter footprint (Blair & Hofton, 1999). The pseudo-waveforms were then used to calculate percentile heights at 5 percent intervals between 5 percent and 100 percent (Fig. 4.3). Principal Components Analysis (PCA) via eigen decomposition of the percentile height variables correlation matrix was used to reduce the dimension of the data set and ensure the variables used as covariates in the subsequent regression analysis were uncorrelated. An assessment of the principal component eigenvalues showed that greater than 85 percent of the variation in the percentile height data was accounted for by the first two eigenvectors (principal components). Hence, scores generated from the first two eigenvectors were used as covariates for model fitting. Pseudo-waveforms were also generated for a grid of 8 283 25×25 m cells over the PEF. PCA scores were calculated for each grid cell and used for subsequent prediction and mapping of AGB and AGB growth.

4.2.3 Model description

Let $Y_t(\boldsymbol{s})$ be the square-root transformed AGB value at location $\boldsymbol{s} \in \mathcal{D}$ and year t centered on the year of lidar acquisition, where $\boldsymbol{s}$ is a vector of location coordinates, e.g., easting and northing, and $\mathcal{D}$ is the forested domain. A model describing AGB growth at site $\boldsymbol{s}$ can be written as,

$$Y_t(\boldsymbol{s}) = \alpha_0(\boldsymbol{s}) + \alpha_1(\boldsymbol{s})t + \epsilon_t(\boldsymbol{s}) \tag{4.1}$$

where $\alpha_0(\boldsymbol{s})$ is the estimated value of $Y_t(\boldsymbol{s})$ when $t = 0$ and $\alpha_1(\boldsymbol{s})$ equals the additive adjustment for every one year change in t. For the example PSP shown in Fig. 4.4, $\alpha_0(\boldsymbol{s})$ and $\alpha_1(\boldsymbol{s})$ correspond to the intercept and slope of the green line respectively. Since $t = 0$ for the year the lidar data was collected, $\alpha_0(\boldsymbol{s})$ can be interpreted as an estimate of AGB at the time of lidar acquisition. Higher values of $\alpha_1(\boldsymbol{s})$ correspond to faster AGB accumulation and lower values of $\alpha_1(\boldsymbol{s})$ suggest slower AGB growth over time. Note the $\boldsymbol{s}$ on α_0 and α_1 identify these as location specific parameters. An independent white-noise process $\epsilon_t(\boldsymbol{s})$

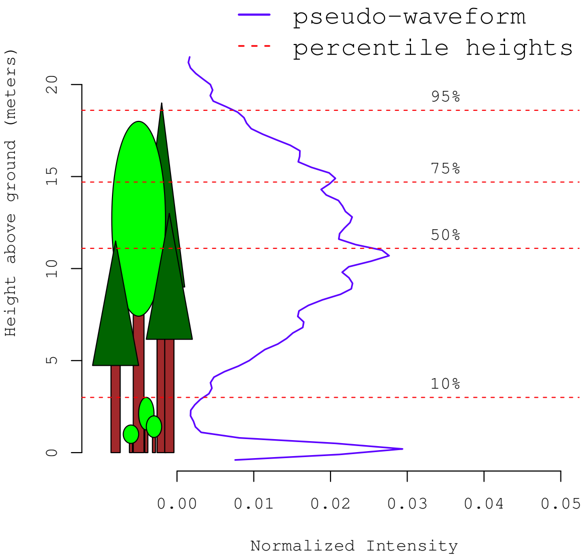


Fig. 4.3: Illustration showing the normalized G-LiHT pseudo-waveform intensity of return energy height profile (blue) along with example percentile height values (red). Return intensity is greater at heights where the forest canopy is encountered. The energy spike at 0 meters signifies energy returning from ground.

captures measurement error in $Y_t(\boldsymbol{s})$ and is modeled as $\mathcal{N}(0, \tau^2)$. This is an attractive choice for modeling long-term inventory and lidar data. Specifying the first stage of our hierarchical design as (4.1) allows the model to acknowledge and account for temporal misalignment between field and remote sensing data through the introduction of $\alpha_1(\boldsymbol{s})t$. Further, this framework allows for the inclusion of multiple time-indexed measurements at one location to inform AGB growth estimation.

Letting $\boldsymbol{\alpha}_0 = (\alpha_0(\boldsymbol{s}_i))_{i=1}^N$ and $\boldsymbol{\alpha}_1 = (\alpha_1(\boldsymbol{s}_i))_{i=1}^N$, where N equals the total number of observed locations (N = 604 PSPs), we write the second level of our hierarchical design as,

$$\boldsymbol{\alpha}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \tag{4.2}$$

$$\boldsymbol{\alpha}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1), \tag{4.3}$$

where, $\boldsymbol{\mu}_0 = (\mu_0(\boldsymbol{s}_i))_{i=1}^N$ and $\boldsymbol{\mu}_1 = (\mu_1(\boldsymbol{s}_i))_{i=1}^N$ constitute the mean components of $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}_1$

Gaussian processes (*GP*), respectively. For our subsequent analysis, we set $\mu_0(\boldsymbol{s}) = \boldsymbol{x}(\boldsymbol{s})\,\boldsymbol{\beta}$, where $\boldsymbol{x}(\boldsymbol{s})$ is a row vector comprised of a 1 and the two lidar PCA scores for site $\boldsymbol{s}$ and $\boldsymbol{\beta}$ is the corresponding column vector of regression parameters, i.e., $\boldsymbol{x}(\boldsymbol{s}) = (1, x_1(\boldsymbol{s}), x_2(\boldsymbol{s}))$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^\top$. We set $\mu_1(\boldsymbol{s}) = \gamma$, where γ is a single regression parameter. Note that this framework can be easily generalized to include any number of covariates to help estimate $\boldsymbol{\alpha}_0$ (AGB) and $\boldsymbol{\alpha}_1$ (AGB growth) simply by redefining $\mu_0(\boldsymbol{s})$ and $\mu_1(\boldsymbol{s})$ accordingly.

The $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ describe the spatial covariance of $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}_1$. We can write (4.2) and (4.3) as

$$\boldsymbol{\alpha}_0 = \boldsymbol{\mu}_0 + \boldsymbol{w}_0, \text{where} \boldsymbol{w}_0 \sim \mathcal{N}(\underline{0}, \boldsymbol{\Sigma}_0) \tag{4.4}$$

$$\boldsymbol{\alpha}_1 = \boldsymbol{\mu}_1 + \boldsymbol{w}_1, \text{where} \boldsymbol{w}_1 \sim \mathcal{N}(\underline{0}, \boldsymbol{\Sigma}_1). \tag{4.5}$$

$\boldsymbol{w}_0 = (w_0(\boldsymbol{s}_i))_{i=1}^N$ and $\boldsymbol{w}_1 = (w_1(\boldsymbol{s}_i))_{i=1}^N$ are zero centered random effects and provide local adjustment with spatial dependence to $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$. Such flexibility is desired because we might expect AGB and growth, i.e., $\boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha}_1$, to have some spatial pattern over a forest domain. If the covariates used to inform the model (i.e., lidar metrics) fail to fully account for the spatial structure in the outcome variable (AGB), $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$ will absorb any unaccounted for spatial structure, thereby reducing spatial dependence among the residuals. Aside from helping to ensure statistical validity, the spatial random effects yield other modeling benefits. For example, we can interpret $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$ as capturing the contribution of unobserved (or unobservable) covariates with spatial pattern. Capturing these latent covariates via the spatial random effects should result in improved fit and predictive performance. The inclusion of $\boldsymbol{w}_1$ provides an additional benefit specific to modeling growth with long-term inventory data. Often at locations, only one post-disturbance inventory will be available—for this analysis 173 PSPs had one post-harvest measurement (Fig. 4.2). By introducing $\boldsymbol{w}_1$, estimation of growth at sites with one measurement becomes possible because the model borrows biomass accumulation information from neighboring sites where multiple inventories may be available.

The spatial random effects are specified as $w_0(\boldsymbol{s}) \sim GP(0, C(\cdot, \cdot, \boldsymbol{\theta}_0))$ and $w_1(\boldsymbol{s}) \sim GP(0, C(\cdot, \cdot, \boldsymbol{\theta}_1))$, where $C(\boldsymbol{s}, \boldsymbol{s}', \boldsymbol{\theta}_0) = Cov(w_0(\boldsymbol{s}), w_0(\boldsymbol{s}'))$ and $C(\boldsymbol{s}, \boldsymbol{s}', \boldsymbol{\theta}_1) = Cov(w_1(\boldsymbol{s}), w_1(\boldsymbol{s}'))$

are functions that model the covariance between any pair of locations $\boldsymbol{s}$ and $\boldsymbol{s}'$. The (i,j)-th elements of the $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ spatial covariance matrices are given by $C(\boldsymbol{s}_i, \boldsymbol{s}_j, \boldsymbol{\theta}_0)$ and $C(\boldsymbol{s}_i, \boldsymbol{s}_j, \boldsymbol{\theta}_1)$, respectively. To ensure $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ are symmetric and positive definite, the spatial covariance functions are defined as

$$C(\boldsymbol{s}, \boldsymbol{s}', \boldsymbol{\theta}_0) = \sigma_0^2 \rho(\boldsymbol{s}, \boldsymbol{s}'; \phi_0)\,, \text{and}\, \boldsymbol{\theta}_0 = \{\sigma_0^2, \phi_0\} \tag{4.6}$$

$$C(\boldsymbol{s}, \boldsymbol{s}', \boldsymbol{\theta}_1) = \sigma_1^2 \rho(\boldsymbol{s}, \boldsymbol{s}'; \phi_1)\,, \text{and}\, \boldsymbol{\theta}_1 = \{\sigma_1^2, \phi_1\}, \tag{4.7}$$

where $\rho(\cdot; \phi_0)$ and $\rho(\cdot; \phi_1)$ are *correlation functions* with ϕ_0 and ϕ_1 determining the rate of correlation decay. The spatial variance parameters σ_0^2 and σ_1^2 equal $Var(w_0(\boldsymbol{s}))$ and $Var(w_1(\boldsymbol{s}))$, respectively. For the subsequent analysis, exponential correlation functions were assumed for $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$, where $\rho(||\boldsymbol{s} - \boldsymbol{s}'||; \phi_0) = \exp(-\phi_0 ||\boldsymbol{s} - \boldsymbol{s}'||)$ and $\rho(||\boldsymbol{s} - \boldsymbol{s}'||; \phi_1) = \exp(-\phi_1 ||\boldsymbol{s} - \boldsymbol{s}'||)$ with $||\boldsymbol{s} - \boldsymbol{s}'||$ being defined as the Euclidean distance separating sites $\boldsymbol{s}$ and $\boldsymbol{s}'$. To ease interpretation of the ϕ estimates, corresponding *effective spatial range* estimates are presented in Table 4.1 and labeled as esr_0 and esr_1. We define *effective spatial range* as the distance (m) where the spatial correlation between locations drops to 0.05.

The data model for (4.1) is constructed as follows. Let $n(\boldsymbol{s})$ be the number of re-measurements at location $\boldsymbol{s}$. Defining the $M \times 1$ column vectors $\boldsymbol{Y} = ((Y_{t_j}(\boldsymbol{s}_i))_{j=1}^{n(\boldsymbol{s}_i)})_{i=1}^{N}$, $\boldsymbol{t} = ((t_j(s))_{j=1}^{n(\boldsymbol{s}_i)})_{i=1}^{N}$, and $\boldsymbol{\epsilon} = ((\epsilon_{t_j}(\boldsymbol{s}_i))_{j=1}^{n(\boldsymbol{s}_i)})_{i=1}^{N}$, where $M = \sum_{i=1}^{N} n(\boldsymbol{s}_i)$, we write the model as,

$$\boldsymbol{Y} = \boldsymbol{Z}\boldsymbol{\alpha}_0 + \boldsymbol{Z}\boldsymbol{\alpha}_1 \boldsymbol{t} + \boldsymbol{\epsilon}, \tag{4.8}$$

where $\boldsymbol{Z}$ is an $M \times N$ block-diagonal matrix with the i-th diagonal element being an $n(\boldsymbol{s}_i) \times 1$ vector of ones.

To complete the Bayesian specification of our hierarchical model, prior distributions are assigned to the parameters and inference proceeds by sampling from the posterior distribution. As customary, we assume $\boldsymbol{\beta}$ and γ follow $\mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}})$ and $\mathcal{N}(\mu_\gamma, \sigma_\gamma^2)$ priors (Gelman et al., 2013). The spatial variance components σ_0^2 and σ_1^2, along with the measurement error τ^2, are assigned inverse-Gamma, $IG(a, b)$, priors. The spatial decay parameters ϕ_0 and ϕ_1

follow a Uniform prior distribution, $Unif(a, b)$, with support over the geographic range of the study area.

Using notation similar to Gelman et al. (2013), we can write the posterior distribution of the parameters as $p(\boldsymbol{\Omega}|\boldsymbol{Y})$, where $\boldsymbol{\Omega} = \{\boldsymbol{\beta}, \gamma, \boldsymbol{w}_0, \boldsymbol{w}_1, \sigma_0^2, \sigma_1^2, \phi_0, \phi_1, \tau^2)$, which is proportional to

$$\begin{gathered} Unif(\phi_0|a_{\phi_0}, b_{\phi_0}) \times Unif(\phi_1|a_{\phi_1}, b_{\phi_1}) \\ \times IG(\sigma_0^2|a_{\sigma_0}, b_{\sigma_0^2}) \times IG(\sigma_1^2|a_{\sigma_1^2}, b_{\sigma_1^2}) \\ \times \mathcal{N}(\boldsymbol{\beta}\,|\boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}}) \times \mathcal{N}(\gamma|\mu_\gamma, \sigma_\gamma^2) \times IG(\tau^2|a_\tau, b_\tau) \\ \times \mathcal{N}(\boldsymbol{w}_0|\underline{0}, \boldsymbol{\Sigma}_0) \times \mathcal{N}(\boldsymbol{w}_1|\underline{0}, \boldsymbol{\Sigma}_1) \times \mathcal{N}(\boldsymbol{Y}|\boldsymbol{Z}\boldsymbol{\alpha}_0 + \boldsymbol{Z}\boldsymbol{\alpha}_1\boldsymbol{t}, \tau^2\boldsymbol{I}_M). \end{gathered} \tag{4.9}$$

Posterior samples of $\boldsymbol{\Omega}$ are collected via Markov chain Monte Carlo (MCMC) algorithms detailed in Banerjee et al. (2014).

The posterior predictive distribution of the spatial random effects $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$ at new (unobserved) locations are given by

$$p(\tilde{\mathbf{w}}_0\,|\boldsymbol{Y}) \propto \int p(\tilde{\mathbf{w}}_0\,|\boldsymbol{w}_0, \boldsymbol{\Omega}_0, \boldsymbol{Y}) p(\boldsymbol{w}_0|\boldsymbol{\Omega}_0, \boldsymbol{Y}) p(\boldsymbol{\Omega}_0|\boldsymbol{Y}) d\boldsymbol{\Omega}_0 \boldsymbol{w}_0 \tag{4.10}$$

$$p(\tilde{\mathbf{w}}_1\,|\boldsymbol{Y}) \propto \int p(\tilde{\mathbf{w}}_1\,|\boldsymbol{w}_1, \boldsymbol{\Omega}_1, \boldsymbol{Y}) p(\boldsymbol{w}_1|\boldsymbol{\Omega}_1, \boldsymbol{Y}) p(\boldsymbol{\Omega}_1|\boldsymbol{Y}) d\boldsymbol{\Omega}_1 \boldsymbol{w}_1, \tag{4.11}$$

where $\tilde{\mathbf{w}}_0 = (w_0(\tilde{\mathbf{s}}_i))_{i=1}^m$ and $\tilde{\mathbf{w}}_1 = (w_1(\tilde{\mathbf{s}}_i))_{i=1}^m$ are the collections over prediction locations $\tilde{\mathbf{s}}$'s, and $\boldsymbol{\Omega}_0$ and $\boldsymbol{\Omega}_1$ include the appropriate subsets of parameters from $\boldsymbol{\Omega}$. Then given the set of lidar covariates at location $\tilde{\mathbf{s}}$, (4.4) and (4.5) can be evaluated to obtain $\alpha_0(\tilde{\mathbf{s}})$ and $\alpha_1(\tilde{\mathbf{s}})$. Finally, the posterior predictive distribution of the response variable, $Y_t(\tilde{\mathbf{s}})$, is obtained from $\mathcal{N}(\alpha_0(\tilde{\mathbf{s}}) + \alpha_1(\tilde{\mathbf{s}})t, \tau^2)$. Samples from the predictive distribution of each component of interest, i.e., $w_0(\tilde{\mathbf{s}})$, $w_1(\tilde{\mathbf{s}})$, $\alpha_0(\tilde{\mathbf{s}})$, $\alpha_1(\tilde{\mathbf{s}})$, and $Y_t(\tilde{\mathbf{s}})$, are obtained via composition sampling, see, e.g., Finley et al. (2015a) and Banerjee et al. (2014) for details. Summary statistics of these posterior predictive distributions, e.g., mean, median, variance, interquartile range and credible intervals, can then be computed and mapped as illustrated in Section 4.3.

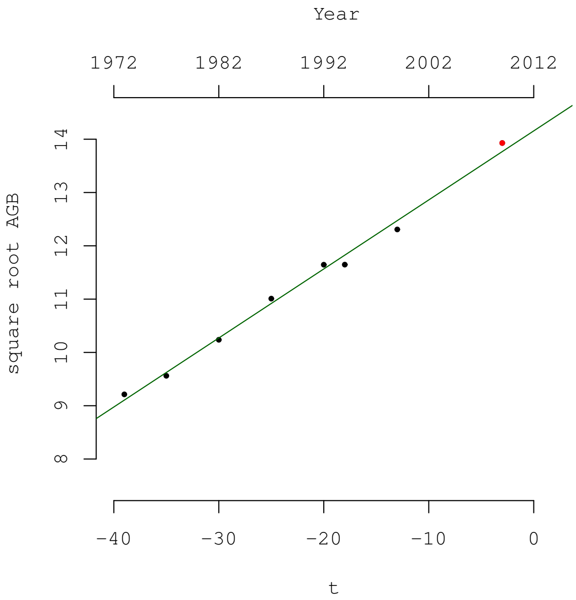


Fig. 4.4: Square root transformed AGB versus time graph for example PSP (site 420, identified with a vertical green line in Fig. 4.2). Upper horizontal axis shows the remeasurement year and the lower horizontal axis shows the corresponding t value. t is the number of years after lidar collection. Since all remeasurements were conducted before the year of lidar acquisition, all t's are negative. Most recent remeasurement highlighted in red.

4.2.4 Model comparison and validation

The modeling framework described in Section 4.2.3 (test model) is compared with the benchmark regression models

$$Y(\boldsymbol{s}) = \boldsymbol{x}(\boldsymbol{s})\,\boldsymbol{\beta} + \epsilon(\boldsymbol{s}) \qquad \text{(benchmark 1)} \tag{4.12}$$

and

$$Y(\boldsymbol{s}) = \boldsymbol{x}(\boldsymbol{s})\,\boldsymbol{\beta} + w_0(\boldsymbol{s}) + \epsilon(\boldsymbol{s}) \text{ (benchmark 2)}, \tag{4.13}$$

where $\boldsymbol{\beta}$ and $w_0(\boldsymbol{s})$ were defined previously. Further, the benchmark models provide a way to evaluate the test model's ability to accommodate temporal misalignment and need to accommodate residual spatial dependence. Specification of the benchmark models (including parameter prior distributions) are defined analogous to the test model, aside from $Y(\boldsymbol{s})$

which is the most recent measurement of square-root transformed AGB for location $\boldsymbol{s}$ (red measurements in Figs. 4.2 and 4.4) and $\epsilon(\boldsymbol{s}) \sim \mathcal{N}(0, \tau^2)$. Again, posterior samples of the benchmark models' parameters and predictions are collected via an MCMC algorithm and composition sampling in the same fashion as the test model. Neither benchmark model acknowledges temporal misalignment between the time lidar and field data were acquired—which is typical of most contemporary modeling exercises that couple lidar and long-term forest inventory data. Because we can only include the most recent measurements at each PSP, the benchmark models lack the ability to estimate AGB growth and hence only estimate AGB. Benchmark model 2 includes a spatial random effect that should absorb any spatially structured variation that may be present in the residuals of benchmark model 1.

To test predictive performance of the three candidate models, a ten-fold holdout set design was constructed by separating the PSPs ($n = 604$) into ten approximately equal-sized groups. Square-root transformed AGB for the holdout groups was sequentially predicted given model parameters estimated using data in the remaining nine groups. Each posterior predictive distribution sample was then back-transformed prior to calculating the AGB posterior predictive distribution mean at each holdout location. Note, because we are working within a Bayesian setting, inference proceeds from the entire back-transformed posterior distribution and hence we avoid issues with back-transformation bias encountered within a frequentest inferential paradigm, see, e.g., Stow et al. (2006) for further discussion. Root mean squared error (RMSE) was calculated using the back-transformed holdout posterior predicted means and observed AGB for each of the three models (labeled *10-fold RMSE* in Table 4.1). The model with the lowest *10-fold RMSE* was considered the "best" predicting model.

To graphically assess the impact of temporal misalignment on 2012 AGB prediction, scatterplots comparing candidate model fitted values from the 33 PSPs with measurements taken in 2012 and the actual 2012 remeasurements were examined (recall, as noted in Section 4.2.1, 2012 data was not used to estimate model parameters and reserved for model comparison).

4.2.5 Posterior inference

Model-based parameter and predictive inference for each of the three candidate models were based on three MCMC chains run for 25 000 iterations. Diffuse hyper-parameters for the $IG(a, b)$, $\mathcal{N}(\boldsymbol{\mu_\beta}, \boldsymbol{\Sigma_\beta})$, and $\mathcal{N}(\mu_\gamma, \sigma^2_\gamma)$ prior distributions were selected to minimize influence on the posterior distribution. The IG a hyper-parameter was set to 2 which results in a prior distribution mean equal to b and infinite variance. The $\mathcal{N}$ hyper-parameters were set to mean zero and variances equal to 10 000. Each MCMC was initiated with different starting values and random seeds. Chain convergence was diagnosed by monitoring mixing and the Gelman-Rubin statistic (Cowles & Carlin, 1996; Gelman & Rubin, 1992). Satisfactory convergence was achieved after 10 000 iterations for all models. Posterior inference was based on a post burn-in sub-sample of 15 000 iterations (5 000 from each chain).

4.3 Results

4.3.1 Parameter estimates and prediction accuracy

Parameter estimates and prediction metrics for the three candidate models, detailed in Sections 4.2.3 and 4.2.4, are given in Table 4.1. Here, we can see the 95 percent credible intervals for β_1 and β_2 do not include zero—the lidar PCA covariates explain a substantial portion of the variability in AGB for all three models. Regarding prediction performance results in Table 4.1, benchmark model 1 yields the largest *10-fold RMSE* (42.88 Mg/ha) among the candidate models. In comparison, the spatial random effects in benchmark model 2 and the test model capture substantial residual spatial structure and hence improve prediction at unobserved locations, i.e., by borrowing information from observed measurements. There is substantial residual spatial dependence and space-varying impact of t, as indicated by long *effective spatial range* estimates (esr_0 and esr_1, noting a maximum inter-location distance between PSP of ∼5.6 km) and relatively large spatial variances (σ^2_0 and σ^2_1). This, along with increased prediction accuracies, is compelling evidence that spatial dependence is present in the residuals of benchmark model 1—suggesting inference about parameters and predictions could be compromised. Indeed, benchmark model 1 prediction is less accurate compared with the other models.

Between the random effects models, the test model provides a substantially lower *10-fold RMSE* compared to the benchmark model 2 (17.52 versus 31.27 Mg/ha). Similarity between the models' β_1 and β_2 estimates suggests that once residual spatial dependence is accommodated, via $\boldsymbol{w}_0$, both models provide similar lidar calibration and any improvement in prediction is due to correction for temporal misalignment.

The test model's increased prediction accuracy, along with higher β_0, suggests benchmark model 2 yields a negative prediction bias. Because the most recent measurement at each PSP occurred before G-LiHT data collection, a negative prediction bias should be expected. Calibrating lidar data to past measures of AGB, without accounting for the temporal mismatch, will result in underestimates of AGB for the year of remote sensing data collection. The difference in *10-fold RMSE* between the two models suggests the negative bias in benchmark model 2 is substantial and correcting for it, as we have in the test model, increases prediction accuracy. Fig. 4.1 graphically demonstrates the negative bias introduced by ignoring the effects temporal misalignment between PSP remeasurement and remote sensing acquisition. We see in Fig. 4.1a that benchmark model 1 fitted values at the 33 PSPs with 2012 remeasurements are substantially lower than actual 2012 AGB. Fig. 4.1b shows that this bias is reduced but still present when a spatial random effect is introduced. Only after temporal misalignment is explicitly accommodated via the introduction of a growth related random effect is the model fit bias effectively removed (Fig. 4.1c).

4.3.2 Predictive uncertainty and precision

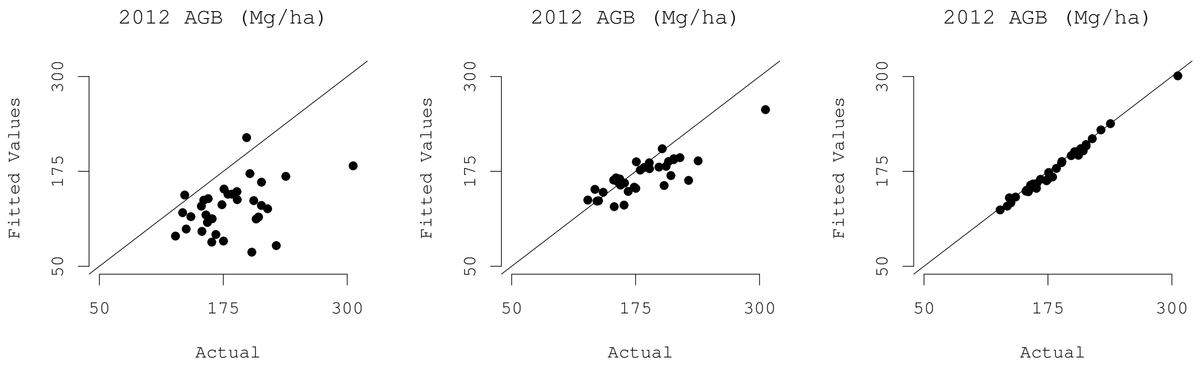
In addition to assessing prediction accuracy, we are interested in candidate models' capacity to provide appropriate measures of prediction uncertainty. To assess the quality of uncertainty quantification, we count the number of holdout set observations that fall within their respective posterior predictive distribution credible intervals, e.g., we can use a credible set that gives a 95% credible interval (Carlin & Louis, 2008, pp. 48), then divide this count by

Table 4.1: Parameter estimates and associated 95% credible intervals for benchmark model 1, 2, and the test model. 10-fold RMSE is presented in Mg/ha units. esr_0 and esr_1 are the *effective spatial ranges*, or the distances (m) at which correlation drops to 0.05, for $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$ respectively.

		Benchmark 1	Benchmark 2	Test Model
Parameter estimates 50% (2.5%, 97.5%)	β_0	9.73 (9.56, 9.90)	9.53 (8.52, 10.51)	10.55 (10.12, 10.98)
	β_1	-0.38 (-0.42, -0.33)	-0.28 (-0.32, -0.23)	-0.27 (-0.31, -0.23)
	β_2	0.72 (0.62, 0.82)	0.38 (0.29, 0.47)	0.34 (0.25, 0.43)
	γ	—	—	0.173 (0.111, 0.23)
	τ^2	4.84 (4.34, 5.45)	1.16 (0.82, 1.53)	0.048 (0.045, 0.052)
	esr_0	—	810.85 (537.45, 1860.61)	283.62 (227.49, 365.23)
	σ_0^2	—	4.10 (2.89, 7.66)	4.19 (3.55, 5.12)
	esr_1	—	—	1 112.12 (693.52, 2501.50)
	σ_1^2	—	—	0.009 (0.006, 0.018)
Prediction metrics	10-fold RMSE	42.88	31.27	17.52
	95% CP	0.94%	0.94%	0.92%
	95% CIW	169.25	117.26	54.21

the total number of observations in the holdout set and multiply by 100 to yield an empirical coverage percentage (CP). We expect the CP to be approximately equal to the chosen credible set, e.g., ~95 percent of the holdout observations should fall within their respective 95 percent credible interval bounds. These CPs are given in Table 4.1 for the 10-fold holdout measurements. The CP results suggest that all models provide appropriate estimates of prediction uncertainty—even benchmark 1 that violates the independent and identically distributed residuals assumption.

CP is useful for identifying misspecified models and issues arising from prediction in extrapolation settings; however, it does not provide information about the precision of prediction. For this we can look at the posterior predictive distribution credible interval width (CIW)—more precise predictions have shorter CIW. Table 4.1 provides the mean CIW calculated using the 95 percent posterior predictive credible interval width for the 10-fold holdout



(a) Benchmark 1 **(b)** Benchmark 2 **(c)** Test Model

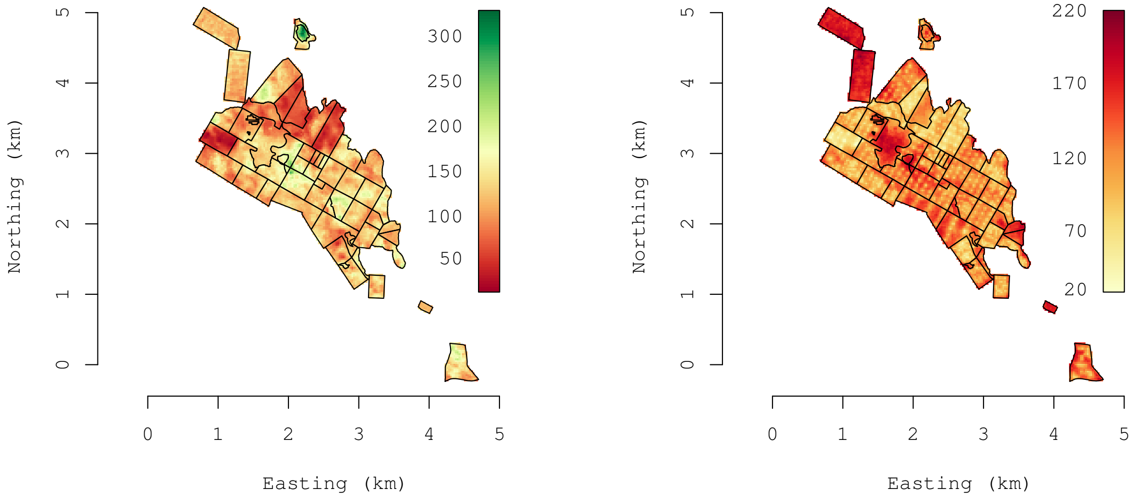
Fig. 4.1: Fitted values versus actual 2012 AGB scatter-plots for the three models. The line on each plot is the one-to-one line.

sets. Here, the 10-fold CIW for the test model is approximately half the width of the benchmark 2 and about 1/3 that of benchmark 1.

4.3.3 Prediction and influence of PSP measurements

The test model was used to generate posterior predictive distributions for 2012 AGB and AGB growth ($\boldsymbol{\alpha}_1$) for each of the 8 283 25×25 m cells over the PEF. The mean of each cell's posterior predictive distribution is mapped in Figs. 4.2a and 4.3a and serve as our best predictions of 2012 AGB and growth (posterior predictive samples were back-transformed prior to calculating summary statistics and subsequent maps). In addition to these *point estimate* maps, the 95% posterior predictive distribution CIW was calculated and mapped for AGB and growth, Figs. 4.2b and 4.3b respectively.

Figs. 4.2b and 4.3b show that CIW for AGB and growth is generally largest in MUs with no inventory data (hashed out MUs in Fig. 4.1). This behavior is expected and clearly illustrates how precision of prediction decreases as we move away from the field data used to calibrate the model. In Fig. 4.2b, we can even identify PSP locations as points of high precision, i.e., narrow CIWs, within MUs. The plot of grid cell AGB CIW versus distance to nearest PSP (Fig. 4.4a) further illustrates the negative relationship between precision of prediction and distance to observed PSPs, i.e., high precision at and near observed PSPs. Additionally, Fig. 4.4a suggests that, on average, AGB CIW increases until distances of ∼200–300 m from PSP locations, after which it levels off. This corresponds to the estimated value for esr_0 for the test model (Table 4.1). Similarly, Fig. 4.4b shows how the precision of predicted $\alpha_1(\boldsymbol{s})$ changes as a function of distance to the nearest PSP. Here, we see the effect PSP measurements have on prediction precision is farther reaching than for AGB (>1 200 m)—also reflected in the esr_1 parameter estimate for the test model (Table 4.1).

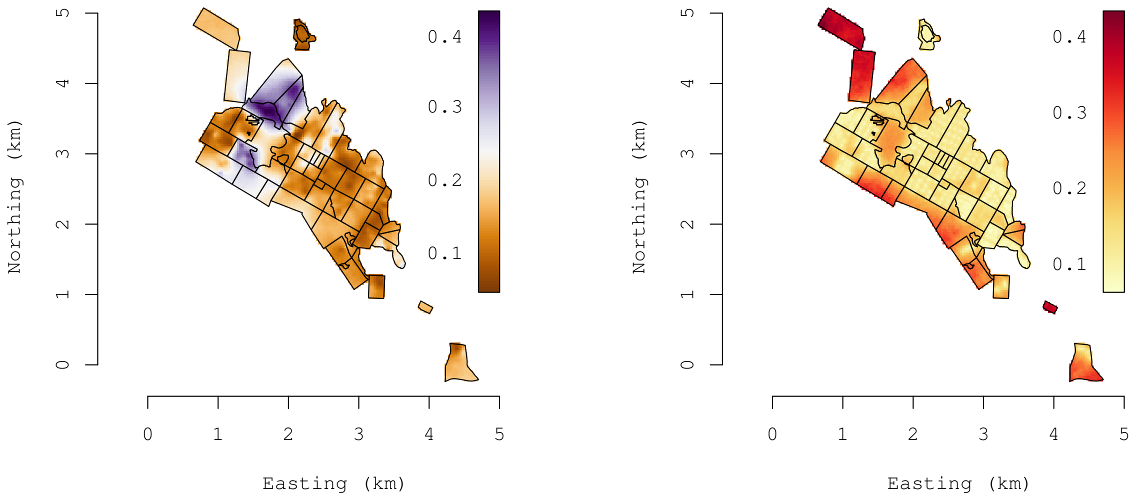


(a) AGB (Mg/ha)

(b) AGB 95% CIW (Mg/ha)

Fig. 4.2: Maps of Predicted 2012 AGB (left) and associated 95 percent credible interval width (CIW) (right) using the test model. Black polygon boundaries outline different management units (MU) on the PEF.

To graphically demonstrate the test model's ability to account for uncertainty resulting from temporal misalignment and the benefit of including multiple measurements at PSPs,

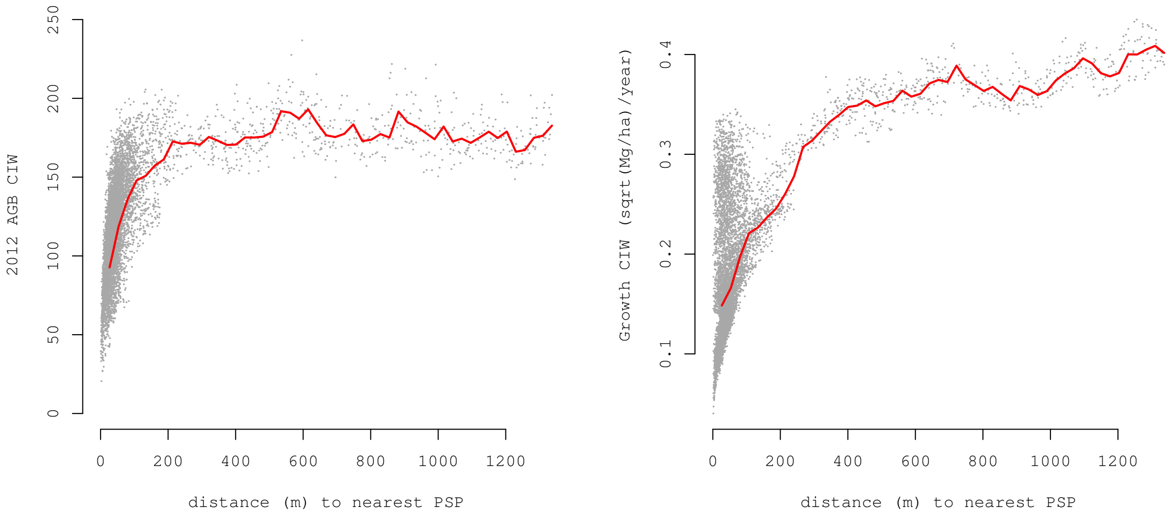


(a) $\boldsymbol{\alpha}_1$

(b) $\boldsymbol{\alpha}_1$ 95% CIW

Fig. 4.3: Maps of predicted $\boldsymbol{\alpha}_1$ (growth) (left) and associated 95 percent credible interval width (CIW) (right) using the test model. Black polygon boundaries outline different management units (MU) on the PEF.

Fig. 4.5 summarizes annual AGB predictions from 1972 to 2012 for 4 example PSPs with associated estimates of uncertainty. We see in all 4 PSPs (especially PSP 120 and 150) that prediction uncertainty increases (wider credible intervals, lower precision) with increasing time since measurement. Figs. 4.5a and 4.5d show that PSPs 420 and 302 have better overall predictions (narrower credible intervals) across the time gradient than PSPs 120 and 150–a direct result of the number of measurements and hence information to inform PSP specific intercept and growth. At PSP 150 (Fig. 4.5c), the single measurement does inform AGB prediction in proximity to the measurement year; however, as expected, the paucity of periodic measurements results in low precision at the time interval extremes. Also, these wide credible intervals suggest that PSP 150 is not close to other PSPs that offer measurements at different values of t to inform prediction. Overall, 2012 prediction of



(a) predicted 2012 AGB

(b) predicted growth

Fig. 4.4: 95% Credible Interval Widths (CIW) for grid cell predictions of 2012 AGB and growth versus distance to nearest permanent sample plot (PSP). The overlayed red trend lines show mean AGB CIW and mean $\alpha_1(\boldsymbol{s})$ CIW values binned for 10 m intervals in (a) and (b), respectively.

AGB at all 4 example PSPs is higher than at the time of the most recent measurement, and growth curve uncertainty is smallest at PSPs with more measurements.

Figs. 4.6a and 4.6c help elucidate the relationship between 2012 AGB and growth prediction uncertainty and the number of measurements at influential PSPs (i.e., nearby PSPs that have the most influence on prediction at unobserved locations.) Fig. 4.6a shows little discernible trend in box-plots summarizing the 8 283 grid cell prediction CIWs subset by number of inventories at nearest PSP, suggesting that 2012 AGB uncertainty is not substantially affected by the number of repeat measurements on proximate PSPs. In contrast, the narrowing of CIWs after ∼2-3 remeasurements at the nearest PSPs seen in Fig. 4.6c underscores the need for such information to define the space-varying slope parameter on t, i.e., AGB growth.

Figs. 4.6b and 4.6d show grid cell prediction CIW box-plots subset by the number of years since last inventory at nearest PSP for 2012 AGB and growth, respectively. Neither figure provides strong evidence that time since remeasurement affects precision of prediction on its own, although Fig. 4.6d might show a slight trend that predictions near PSPs measured in the last four years have higher precision, on average, than predictions near PSPs with older measurements.

4.4 Discussion

4.4.1 Comparisons with other PEF studies

MU 8 and MU 22 were clearcut in 1984 and 1989, respectively (Fig. 4.1). We see the consequence of this in the map of predicted AGB (Fig. 4.2a) where both MU 8 and 22 are predominantly dark red, indicating comparatively low ($<$ 75 Mg/ha) AGB. Hayashi et al. (2014, Tab. 2) found MU 22 to have the lowest stem volume of the MUs they examined. We see slightly higher AGB growth rates in MU 8 compared to MU 22 in Fig. 4.3a.

These differences can be attributed to a time lag in stand *reorganization* (i.e., stand initiation) and *aggradation* (i.e., stem exclusion) phases (Oliver & Larson, 1996). Nyland (2007, pp. 204) points out the transition from the stand *reorganization* to *aggradation* phase, marked by the beginning of AGB accumulation, occurs approximately 20 years after clearcutting in this forest type and region. Since MU 22 and MU 8 were harvested 25 and 30 years ago, respectively, the differences in AGB growth might be attributed to the five-year difference in stand treatments and subsequent transition to the *aggradation* phase. Over the next decade, we expect a rise in AGB growth rates in both stands as they accrue biomass on ever larger trees until such a time that self-thinning processes start to reduce the growth of dominant trees (Oliver & Larson, 1996). MU 23 is undergoing a three-stage shelterwood management system where initial stand entries remove over-mature and undesirable trees to encourage the growth of advance regeneration present at the site. In an empirical study of the silvicultural treatments at the PEF Sendak et al. (2003, Fig. 4) found basal area growth to be fastest in this management system. In support of Sendak's finding, Fig. 4.3a shows high AGB growth rate in MU 23. MU 32 is a mature stand that serves as a reference treatment for the PEF. Only very limited management activity has occurred here since 1954. Fig. 4.2a shows that MU 32 has the highest accumulated AGB of any of the PEF's MUs and Fig. 4.3a indicates that AGB growth in MU 32 is currently very low—characteristics expected of mature, relatively undisturbed, old-growth stands in the *steady-state* phase of development (Nyland, 2007).

4.4.2 Benefits of Proposed Framework

Results in Table 4.1 show the test model detailed in Section 4.2.3 provides greater prediction accuracy and precision compared to the benchmark models. These improvements are a direct result of appropriately modeling residual spatial dependence and capturing the space-varying nature of AGB change by leveraging repeated measurements. Other benefits of the proposed modeling framework are data driven estimates of AGB growth and full uncertainty quantification of parameters and predictions. The ability to estimate AGB and growth together and propagate their uncertainties through to prediction is clearly advantageous from an AGB monitoring and management perspective. As an example, scientists who are part of NASA's CMS initiative, in effort to advance the development of carbon MRV systems, need methods to spatially predict AGB with uncertainty. The modeling framework proposed here aims to help meet this goal.

The performance of the model introduced here provides evidence that temporal misalignment between lidar and field inventory data should be acknowledged during model fitting. The benchmark models' poor performance as compared to the test model, indicates the effects of temporally disjointed AGB and lidar measurements should not be treated as negligible even when the time discrepancies are ten years or less. This has implications for potential national scale mapping of AGB using air- or space-borne lidar and inventory systems, such as FIA. In most western U.S. states, only ten percent of FIA PSPs are measured annually, i.e., we would expect to encounter time disagreements of up to ten years when calibrating lidar with FIA inventory data. Results from this analysis suggest that future continental scale lidar informed AGB models using FIA inventory data may be improved by accounting for temporal misalignment.

One common solution to temporal disagreement between field inventory and remote sensing data is to apply growth models to "project" inventories forward to the time of lidar acquisition. A potential weakness of this approach is that uncertainty in growth model results are not easily propagated through statistical model components to prediction. Rather than attempting to correct for temporal misalignment in a disjoint data preprocessing phase, it is more attractive to specify modeling frameworks capable of acknowledging misalignment explicitly and appropriately propagating resulting uncertainty.

Spatially and temporally explicit measures of uncertainty, a by-product of fitting Bayesian spatio-temporal hierarchical models, can help researchers determine when and where to collect additional inventory data given specific inferential goals. For example, if the goal is to increase growth prediction accuracy and precision at the PEF, Fig. 4.6c and 4.6d suggest remeasurement of PSPs with fewer than three measurements where the most recent measurement is at least four years old. Looking at Figs. 4.6c and 4.6d together suggests the number of remeasurements at a PSP may be more important than time since last inventory in determining growth uncertainty at the PEF. Figs. 4.6a and 4.6b indicate the number and timing of PSP remeasurements may not be a strong determining factor concerning AGB prediction uncertainty. Rather, Fig. 4.4a suggests that collecting new inventories further than 200 m from existing PSPs will improve AGB certainty more than any remeasurement effort. These types of questions cannot be addressed when RMSE alone is considered.

4.4.3 Limitations and indications for future research

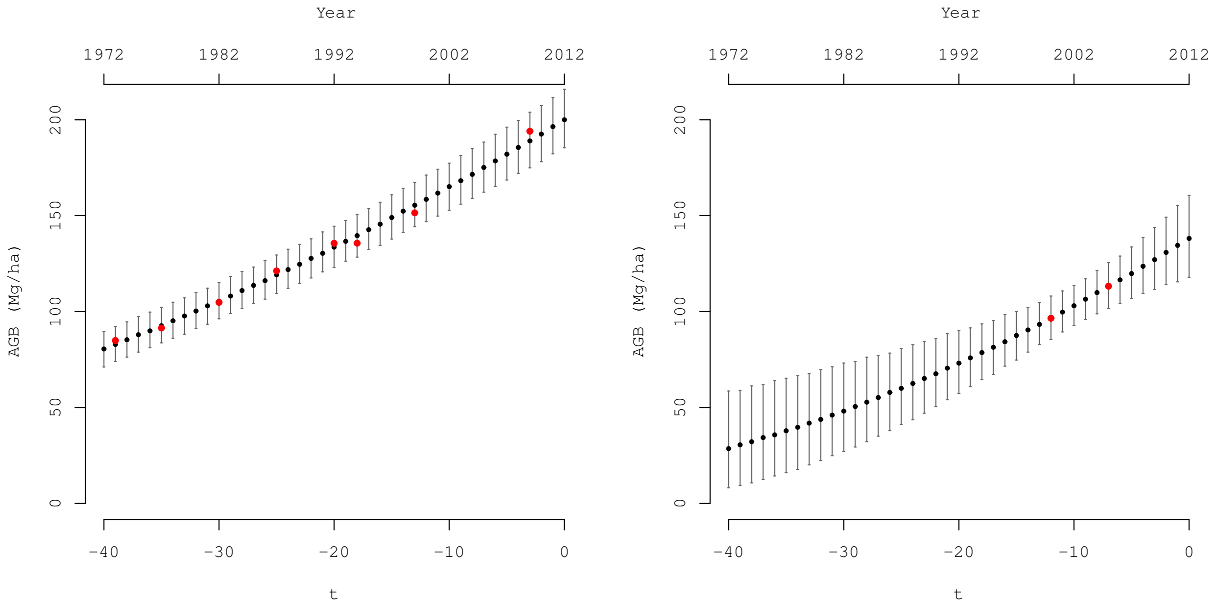
The assumption of a linear increase of AGB over time as imposed by the first stage of the hierarchical model (4.1) is a potential drawback to this specific design's application to other forests. A linear trend was chosen to relate time and square-root transformed AGB as opposed to a sigmoidal growth curve that is commonly put forward for AGB growth rates.

AGB growth on the PEF did not exhibit any strong curvilinear trends (Fig. 4.4 and 4.5) hence a linear trend was considered appropriate. Ongoing management at the PEF is likely keeping much of the forest in an active state of AGB growth. Areas are being harvested (or disturbed in some other fashion) before forest growth begins to slow in the senescence phase. Also, most PSPs had fewer than five post-disturbance measurements, which may not be enough to appropriately characterize a sigmoidal growth curve. Still further, the length of the temporal record at the PEF may not be long enough to expose non-linearity in the system. For datasets where a sigmoidal growth curve can be seen, the framework described here can be adapted. Rather than proposing a linear trend for (4.1), one could consider a non-linear AGB growth model.

The PEF PSP network is spatially dense with many PSPs having a long temporal record of inventories, which is a luxury not common in most forest monitoring systems. In the future, we will extend the proposed modeling framework to large scale sparse PSP networks, e.g., FIA, to estimate and map AGB growth. Previous modeling exercises provide some encouraging results. For example, in an effort to map forest biomass using FIA data over an area in northern Minnesota, Banerjee & Finley (2007) considered spatial regression models with a suite of Landsat derived covariates and residual structure captured using a Gaussian process capable of resolving short- and long-range spatial dependence. Their results showed effective spatial long-range dependence of ∼7.9 km (with 95 percent credible interval from 4.1 to 11.1 km). In a different modeling exercise, aimed at mapping forest biomass over large heterogeneous landscapes (forest/non-forest areas), Finley et al. (2011) estimated residual spatial dependence patterns extending as far as ∼352 km (with 95 percent credible interval from 290 to 429 km) across Michigan's lower peninsula. Datta et al. (2016a) reported similar residual spatial dependence patterns when modeling AGB using FIA data across the conterminous US. All three studies used a model akin to benchmark model 2 and only the most recent FIA PSP measurements. Results from each study showed improved fit and/or holdout set prediction when leveraging residual spatial dependence through estimation of $\boldsymbol{w}_0$. Given the results of our current study, we anticipate including multiple measurements on FIA PSPs might further improve prediction at new locations and fetch additional insight into forest growth rates. Even if there is little residual spatial dependence—either due to effective covariates or simply a high degree of short range variability in the forest variable of interest—then $\boldsymbol{w}_0$ will not improve predictive ability through borrowing information from proximate PSPs. However, this does not negate the useful role of $\boldsymbol{w}_1$ to capture space varying growth rates.

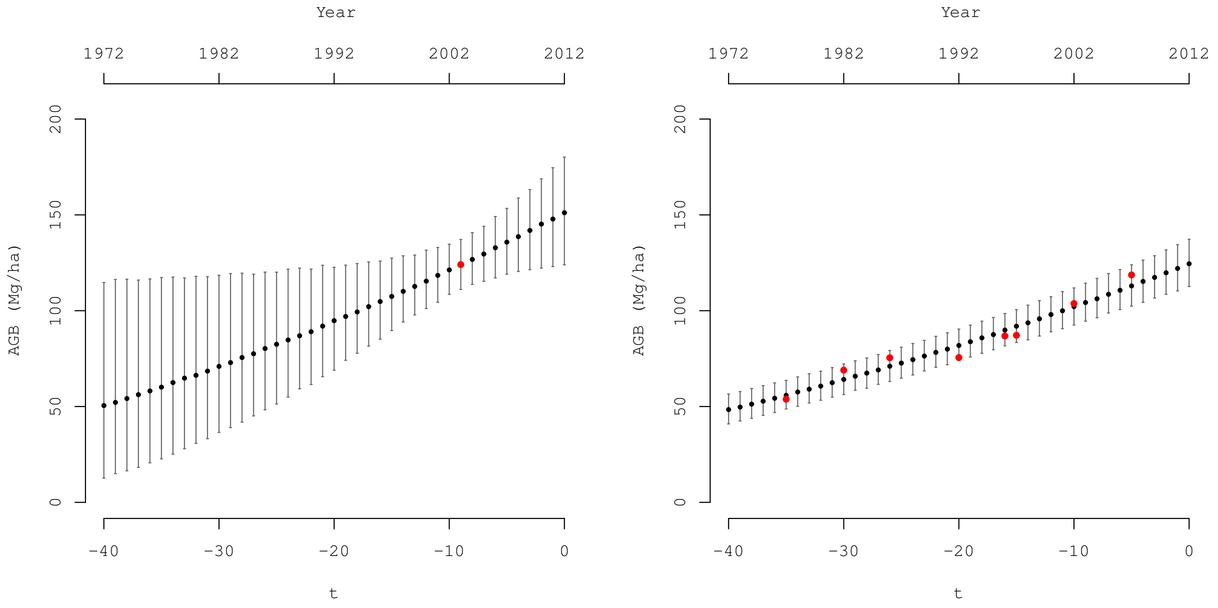
Extensions to the proposed model could incorporate functional forms that use covariates in place of γ and $\boldsymbol{w}_1$. For instance, multi-spectral data may be useful for explaining variability in AGB growth rates across an expansive study region.

As we have demonstrated the benefits and knowledge gaps of adopting our proposed model for estimating forest AGB and associated AGB accumulation, the wider implications should not be understated. Global interest in mitigating the emission of carbon dioxide necessitates accurate estimates of terrestrial AGB and associated sequestration/emission rates. Rigorous combination of spatially sparse and infrequent in-situ measurements of forest AGB with potentially more spatially intense and frequent remote sensing products such as lidar offers the opportunity to refine our monitoring of the earth's terrestrial carbon balance. Beyond reducing uncertainty and informing field campaigns, our proposed model offers the opportunity to rapidly incorporate remotely sensed information following large-scale disturbance events (e.g., wildfires or hurricanes) to inform the distribution of carbon emission to specific disturbances (e.g., Kyoto Protocol removal of non-anthropogenic carbon emissions).



(a) PSP 420

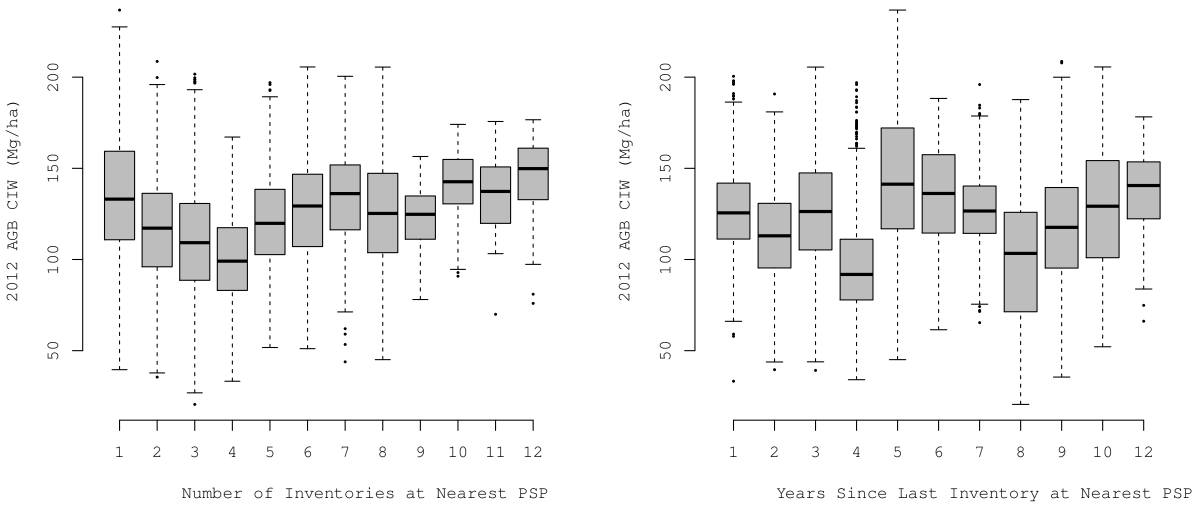
(b) PSP 120



(c) PSP 150

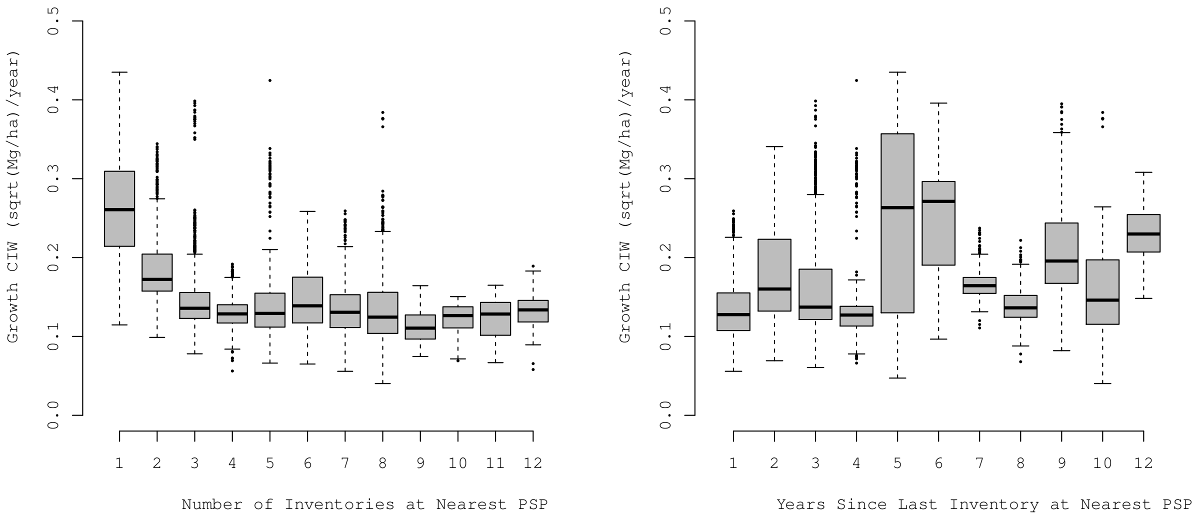
(d) PSP 302

Fig. 4.5: Predictions of AGB from 1972 to 2012 for four example PSPs (black circles). Red circles indicate field inventory measures of AGB. Vertical gray lines highlight the 95 percent credible interval for the AGB predictions.



(a) predicted 2012 AGB

(b) predicted 2012 AGB



(c) predicted growth

(d) predicted growth

Fig. 4.6: Boxplots showing the distributions of 95 percent Credible Interval Widths (CIW) for grid cell predictions of 2012 AGB ((a) and (b)) and productivity ((c) and (d)) versus number of remeasurements ((a) and (c)) and years since last remeasurement ((b) and (d)) at nearest permanent sample plot (PSP).

BIBLIOGRAPHY

Abdalati, W., Zwally, H., Bindschadler, R., Csatho, B., Farrell, S., Fricker, H., Harding, Dand Kwok, R., Lefsky, M., Markus, T., Marshak, A., Neumann, T., Palm, S., Schutz, B., Smith, B., Spinhirne, J., & Webb, C. (2010). The ICESat-2 laser altimetry mission. *Proceedings of the IEEE*, 98, 735–751.

Agricultural Research, Extension and Education Reform Act (1998). Pub. l. 105–185, 112 stat. 523–610.

Alaska Department of Natural Resources (2016). Alaska's state forests. `http://forestry.alaska.gov/stateforests`. Accessed: 10-16-2016.

Albercht, A., Hanewinkel, M., & Bauhus, J. (2012). How does silviculture affect storm damage in forests of south-western Germany? results from empirical modeling based on long-term observations. *European Journal of Forest Resources*, 131, 229–247.

Amatya, D., Campbell, J., Wohlgemuth, P., Elder, K., Sebestyen, S., Johnson, S., Keppeler, E., Adams, M., Caldwell, P., & Misra, D. (2016). Hydrological processes of reference watersheds in experimental forests, USA., in: Amatya, D., Williams, T., Bren, L., & de Jong, C. (Eds.), Forest Hydrology: Processes, Management, and Applications. CABI Publishers, UK, pp. 219–239.

Andersen, H.E., Strunk, J., & Temesgen, H. (2011). Using airborne light detection and ranging as a sampling tool for estimating forest biomass resources in the upper Tanana Valley of interior Alaska. *Western Journal of Applied Forestry*, 26, 157–164.

Anderson, J.E., Plourde, L.C., Martin, M.E., Braswell, B.H., Smith, M.L., Dubayah, R.O., Hofton, M.A., & Blair, J.B. (2008). Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sensing of Environment*, 112, 1856–1870.

Asner, G., Hughes, R., Varga, T., Knapp, D., & Kennedy-Bowdoin, T. (2009). Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems*, 12, 261–278.

Awadallah, A., Abbott, V., Wynne, R., & Nelson, R. (2013). Estimating forest canopy height and biophysical parameters using photon-counting laser altimetry, in: Proc. 13th International Conference on LiDAR Applications for Assessing Forest Ecosystems (SilviLaser 2013), pp. 129–136.

Babcock, C., Finley, A.O., Andersen, H.E., Pattison, R., Cook, B.D., Morton, D.C., Alonzo, M., Nelson, R., Gregoire, T., Ene, L., Gobakken, T., & Næsset, E. (2017). Geostatistical estimation of forest biomass in interior alaska combining landsat-derived tree cover, sampled airborne lidar and field observations. *ArXiv e-prints*, `arXiv:1705.03534`.

Babcock, C., Finley, A.O., Bradford, J.B., Kolka, R., Birdsey, R., & Ryan, M.G. (2015). Lidar based prediction of forest biomass using hierarchical models with spatially varying coefficients. *Remote Sensing of Environment*, 169, 113–127.

Babcock, C., Finley, A.O., Cook, B.D., Weiskittel, A., & Woodall, C.W. (2016). Modeling forest biomass and growth: Coupling long-term inventory and lidar data. *Remote Sensing of Environment*, 182, 1–12.

Babcock, C., Matney, J., Finley, A., Weiskittel, A., & Cook, B. (2013). Multivariate spatial regression models for predicting individual tree structure variables using lidar data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6, 6–14.

Balshi, M.S., McGuire, A.D., Duffy, P., Flannigan, M., Kicklighter, D.W., & Melillo, J. (2009). Vulnerability of carbon storage in north american boreal forests to wildfires during the 21st century. *Global Change Biology*, 15, 1491–1510.

Banerjee, S., Carlin, C., & Gelfand, A. (2014). Hierarchical Modeling and Analysis for Spatial Data. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. second ed., Chapman & Hall/CRC.

Banerjee, S., & Finley, A. (2007). Bayesian multi-resolution modeling for spatially replicated data sets with application to forest biomass data. *Journal of Statistical Planning and Inference*, 137, 3193–3205.

Banerjee, S., & Johnson, G. (2006). Coregionalized single- and multi-resolution spatially-varying growth curve modelling with application to weed growth. *Biometrics*, 61, 617–625.

Barrett, T.M., & Gray, A.N. (2011). Potential of a national monitoring program for forests to assess change in high-latitude ecosystems. *Biological Conservation*, 144, 1285–1294.

Bechtold, W.A., & Patterson, P.L. (2005). The Enhanced Forest Inventory and Analysis Program: National Sampling Design and Estimation Procedures. US Department of Agriculture Forest Service, Southern Research Station Asheville, North Carolina.

Berliner, L.M. (1996). Hierarchical bayesian time series models, in: Maximum Entropy and Bayesian Methods. Springer, pp. 15–22.

Blackard, J., Finco, M., Helmer, E., Holden, G., Hoppus, M., Jacobs, D., Lister, A., Moisen, G., Nelson, M., Riemann, R., et al. (2008). Mapping US forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment*, 112, 1658–1677.

Blair, J., & Hofton, M. (1999). Modeling laser altimeter return waveforms over complex vegetation using high-resolution elevation data. *Geophysical Research Letters*, 26, 2509–2512.

Bolton, D.K., Coops, N.C., & Wulder, M.A. (2013). Measuring forest structure along productivity gradients in the Canadian boreal with small-footprint lidar. *Environmental monitoring and assessment*, 185, 6617–6634.

Bonanza Creek LTER (2016a). Bonanza Creek Experimental Forest. `http://www.lter.uaf.edu/research/study-sites-bcef`. Accessed: 10-16-2016.

Bonanza Creek LTER (2016b). Caribou-Poker Creeks Research Watershed. `http://www.lter.uaf.edu/research/study-sites-cpcrw`. Accessed: 10-16-2016.

Bradshaw, C.J., & Warkentin, I.G. (2015). Global estimates of boreal forest carbon stocks and flux. *Global and Planetary Change*, 128, 24–30.

Breidenbach, J., & Astrup, R. (2012). Small area estimation of forest attributes in the Norwegian national forest inventory. *European Journal of Forest Research*, 131, 1255–1267.

Brissette, J., & Kenefic (2008). Study Plan: Silvicultural Effects on Composition, Structure and Growth of Northern Conifers in the Acadian Forest Region: Revision of the Compartment Management Study on the Penobscot Experimental Forest. USDA Forest Service, Northern Research Station, Newtown Square, PA.

Brissette, J., Kenefic, L., Russel, M., & Puhlick, J. (2012). Overstory Tree and Regeneration Data from the "Silvicultural Effects on Composition, Structure, and Growth" Study at Penobscot Experimental Forest. USDA Forest Service, Northern Research Station, Newtown Square, PA.

Camill, P. (2005). Permafrost thaw accelerates in boreal peatlands during late-20th century climate warming. *Climatic Change*, 68, 135–152.

Carlin, C., & Louis, T. (2008). Bayesian Methods for Data Analysis. Chapman & Hall/CRC Texts in Statistical Science. second ed., Chapman & Hall/CRC.

Chapin, F., & Hollingsworth, J. (2006). GIS file of permafrost distribution in Caribou-Poker Creeks Research Watershed, Bonanza Creek LTER - University of Alaska Fairbanks. `http://www.lter.uaf.edu/data/data-detail/id/121`.

Chapin, F.S., Mcguire, A.D., Randerson, J., Pielke, R., Baldocchi, D., Hobbie, S.E., Roulet, N., Eugster, W., Kasischke, E., Rastetter, E.B., Zimov, S.A., & Running, S.W. (2000). Arctic and boreal ecosystems of western north america as components of the climate system. *Global Change Biology*, 6, 211–223.

Chen, Q., McRoberts, R.E., Wang, C., & Radtke, P.J. (2016). Forest aboveground biomass mapping and estimation across multiple spatial scales using model-based inference. *Remote Sensing of Environment*, 184, 350–360.

Chen, Q., Vaglio Laurin, G., Battles, J.J., & Saah, D. (2012). Integration of airborne lidar and vegetation types derived from aerial photography for mapping aboveground live biomass. *Remote Sensing of Environment*, 121, 108–117.

CMS (2010). NASA carbon monitoring system. `http://carbon.nasa.gov`. Accessed: 8-11-2016.

Cochran, W. (1946). Relative accuracy of systematic and stratified random samples for a certain class of populations. *The Annals of Mathematical Statistics*, 17, 164–177.

Cochran, W. (1977). Sampling Techniques. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, Wiley.

Cook, B., Corp, L., Nelson, R., Middleton, E., Morton, D., McCorkel, J., Masek, J., Ranson, K., Ly, V., & Montesano, P. (2013). NASA Goddard's lidar, hyperspectral and thermal (G-LiHT) airborne imager. *Remote Sensing*, 5, 4045–4066.

Cowles, M., & Carlin, B. (1996). Markov chain monte carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, 91, 883–904.

Cressie, N. (1993). Statistics for Spatial Data. Wiley series in probability and mathematical statistics: Applied probability and statistics, John Wiley & Sons.

Cressie, N., Calder, C.A., Clark, J.S., Ver Hoef, J.M.V., & Wikle, C.K. (2009). Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19, 553–570.

Cressie, N.A.C., & Wikle, C.K. (2011). Statistics for Spatio-Temporal Data. Wiley series in probability and statistics, Hoboken, N.J. Wiley.

Datta, A., Banerjee, S., Finley, A.O., & Gelfand, A.E. (2016a). Hierarchical nearest-neighbor gaussian process models for large geostatistical datasets. *Journal of the American Statistical Association*, 111, 800–812.

Datta, A., Banerjee, S., Finley, A.O., & Gelfand, A.E. (2016b). On nearest-neighbor gaussian process models for massive spatial data. *Wiley Interdisciplinary Reviews: Computational Statistics*, 8, 162–171.

Diggle, P., & Ribeiro, P. (2007). Model-based Geostatistics. Springer Series in Statistics, Springer.

Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, 37, 36–48.

Erdody, T., & Moskal, L. (2010). Fusion of lidar and imagery for estimating forest canopy fuels. *Remote Sensing of Environment*, 114, 725 – 737.

Finley, A., Banerjee, S., & MacFarlane, D. (2011). A hierarchical model for quantifying forest variables over large heterogeneous landscapes with uncertain forest areas. *Journal ot the American Statistical Association*, 106, 31–48.

Finley, A.O., Banerjee, S., & Cook, B.D. (2014a). Bayesian hierarchical models for spatially misaligned data in R. *Methods in Ecology and Evolution*, 5, 514–523.

Finley, A.O., Banerjee, S., & Gelfand, A.E. (2015a). spBayes for large univariate and multivariate point-referenced spatio-temporal data models. *Journal of Statstical Software*, 63.

Finley, A.O., Banerjee, S., & Gelfand, A.E. (2015b). spBayes for large univariate and multivariate point-referenced spatio-temporal data models. *Journal of Statistical Software*, 63, 1–28. URL: `http://www.jstatsoft.org/v63/i13/`.

Finley, A.O., Banerjee, S., Weiskittel, A.R., Babcock, C., & Cook, B.D. (2014b). Dynamic spatial regression models for space-varying forest stand tables. *Environmetrics*, 25, 596–609.

Finley, A.O., Banerjee, S., Zhou, Y., Cook, B.D., & Babcock, C. (2017). Joint hierarchical models for sparsely sampled high-dimensional LiDAR and forest variables. *Remote Sensing of Environment*, 190, 149–161.

Finley, A.O., Datta, A., Cook, B.C., Morton, D.C., Andersen, H.E., & Banerjee, S. (2017). Applying nearest neighbor gaussian processes to massive spatial data sets: Forest canopy height prediction across Tanana Valley Alaska. *ArXiv e-prints*, `arXiv:1702.00434`.

French, J. (2015). SpatialTools: Tools for Spatial Data Analysis. URL: `https://CRAN.R-project.org/package=SpatialTools`. r package version 1.0.2.

G-LiHT (2016). Goddard's lidar hyperspectral and thermal (G-LiHT) imager. `http://www.gliht.gsfc.nasa.gov`. Accessed: 10-16-2016.

García, M., Riaño, D., Chuvieco, E., & Danson, F.M. (2010). Estimating biomass carbon stocks for a mediterranean forest in central Spain using lidar height and intensity data. *Remote Sensing of Environment*, 114, 816–830.

Gauthier, S., Bernier, P., Kuuluvainen, T., Shvidenko, A.Z., & Schepaschenko, D.G. (2015). Boreal forest health and global change. *Science*, 349, 819–822.

GEDI (2014). Global ecosystem dynamics investigation lidar. `http://science.nasa.gov/missions/gedi/`. Accessed: 8-11-2016.

Gelfand, A.E., Kim, H.J., Sirmans, C., & Banerjee, S. (2003). Spatial modeling with spatially varying coefficient processes. *Journal of the American Statistical Association*, 98, 387–396.

Gelfand, A.E., Schmidt, A.M., Banerjee, S., & Sirmans, C. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization. *Test*, 13, 263–312.

Gelfand, A.E., & Smith, A.F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85, 398–409.

Gelman, A. (2007). Struggles with survey weighting and regression modeling. *Statist. Sci.*, 22, 153–164.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2013). Bayesian Data Analysis, 3rd Edition. Chapman & Hall/CRC Texts in Statistical Science, Chapman & Hall/CRC.

Gelman, A., & Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, , 457–472.

Goerndt, M.E., Monleon, V.J., & Temesgen, H. (2011). A comparison of small-area estimation techniques to estimate selected stand attributes using LiDAR-derived auxiliary variables. *Canadian Journal of Forest Research*, 41, 1189–1201.

Gregoire, T.G. (1998). Design-based and model-based inference in survey sampling: Appreciating the difference. *Canadian Journal of Forest Research*, 28, 1429–1447.

Gregoire, T.G., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R., & Holm, S. (2011). Model-assisted estimation of biomass in a lidar sample survey in Hedmark County, Norway. *Canadian Journal of Forest Research*, 41, 83–95.

Griffith, D.A. (2005). Effective geographic sample size in the presence of spatial autocorrelation. *Annals of the Association of American Geographers*, 95, 740–760.

Gschwantner, T., Lanz, A., Vidal, C., Bosela, M., Di Cosmo, L., Fridman, J., Gasparini, P., Kulie̎sis, A., Tomter, S., & Schadauer, K. (2016). Comparison of methods used in european national forest inventories for the estimation of volume increment: towards harmonisation. *Annals of Forest Science*, 73, 807–821.

Hadden, D., & Grelle, A. (2016). Changing temperature response of respiration turns boreal forest from carbon sink into carbon source. *Agricultural and Forest Meteorology*, 223, 30–38.

Hansen, M., Potapov, P., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Thau, D., Stehman, S., Goetz, S., Loveland, T., Kommareddy, A., Egorov, A., Chini, L., Justice, C., & Townshend, J. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 342, 850–853.

Harden, J.W., O'Neill, K.P., Trumbore, S.E., Veldhuis, H., & Stocks, B.J. (1997). Moss and soil contributions to the annual net carbon flux of a maturing boreal forest. *Journal of Geophysical Research: Atmospheres*, 102, 28805–28816.

Hayashi, R., Weiskittel, A., & Sader, S. (2014). Assessing the feasibility of low-density lidar for stand inventory attribute predictions in complex and managed forests of northern Maine, USA. *Forests*, 5, 363–383.

Haywood, A., & Stone, C. (2017). Estimating large area forest carbon stocksa pragmatic design based strategy. *Forests*, 8, 99.

Healey, S., Patterson, P., Saatchi, S., Lefsky, M., Lister, A., & Freeman, E. (2012). A sample design for globally consistent biomass estimation using lidar data from the geoscience laser altimeter system (GLAS). *Carbon Balance and Management*, 7, 10.

Hobbs, N., & Hooten, M. (2015). Bayesian Models: A Statistical Primer for Ecologists. Princeton University Press.

Hollingsworth, T.N., Johnstone, J.F., Bernhardt, E.L., & Chapin III, F.S. (2013). Fire severity filters regeneration traits to shape community assembly in Alaska's boreal forest. *PloS one*, 8, e56033.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J., & Megown, K. (2015). Completion of the 2011 national land cover database for the conterminous United States-representing a decade of land cover change information. *Photogramm. Eng. Remote Sens*, 81, 345–354.

Horvitz, D.G., & Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47, 663–685.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., & Maltamo, M. (2008). Review of methods of smallfootprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, 29, 1339–1366.

ICESat-2 (2015a). Ice, cloud, and land elevation satellite-2. `http://icesat.gsfc.nasa.gov/icesat2`. Accessed: 8-11-2016.

ICESat-2 (2015b). Ice, cloud, and land elevation satellite-2. `http://icesat.gsfc.nasa.gov/`. Accessed: 1-5-2015.

IPCC (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Iqbal, I., Dash, J., Ullah, S., & Ahmad, G. (2013). A novel approach to estimate canopy height using ICESat/GLAS data: A case study in the new forest national park, UK. *International Jorunal of Applied Earth Observation and Geoinformation*, 23, 109–118.

Jenkins, J., Chojnacky, D., Heath, L., & Birdsey, R. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, 49, 12–35.

Junttila, V., Finley, A.O., Bradford, J.B., & Kauranne, T. (2013). Strategies for minimizing sample size for use in airborne lidar-based forest inventory. *Forest Ecology and Management*, 292, 75–85.

Junttila, V., & Laine, M. (2017). Bayesian principal component regression model with spatial effects for forest inventory variables under small field sample size. *Remote Sensing of Environment*, 192, 45–57.

Junttila, V., Maltamo, M., & Kauranne, T. (2008). Sparse bayesian estimation of forest stand characteristics from airborne laser scanning. *Forest Science*, 54, 543–552.

Kampe, T., Johnson, B., Kuester, M., & Keller, M. (2010). NEON: The first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *Journal of Applied Remote Sensing*, 4.

Kangas, A., Myllymäki, M., Gobakken, T., & Næsset, E. (2016). Model-assisted forest inventory with parametric, semiparametric, and nonparametric models. *Canadian Journal of Forest Research*, 46, 855–868.

Kao, R.H., Gibson, C.M., Gallery, R.E., Meier, C.L., Barnett, D.T., Docherty, K.M., Blevins, K.K., Travers, P.D., Azuaje, E., Springer, Y.P., Thibault, K.M., McKenzie, V.J., Keller, M., Alves, L.F., Hinckley, E.L.S., Parnell, J., & Schimel, D. (2012). NEON terrestrial field observations: Designing continental-scale, standardized sampling. *Ecosphere*, 3, 1–17.

Kashani, A.G., Olsen, M.J., Parrish, C.E., & Wilson, N. (2015). A review of lidar radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration. *Sensors*, 15, 28099–28128.

Kauranne, T., Pyankov, S., Junttila, V., Kedrov, A., Tarasov, A., Kuzmin, A., Peuhkurinen, J., Villikka, M., Vartio, V.M., & Sirparanta, S. (2017). Airborne laser scanning based forest inventory: Comparison of experimental results for the perm region, russia and prior results from finland. *Forests*, 8.

Koch, G.G., & Gillings, D.B. (1983). Inference, design-based vs. model-based. *Encyclopedia of statistical sciences*, .

Kumar, L., Sinha, P., Taylor, S., & Alqurashi, A.F. (2015). Review of the use of remote sensing for biomass estimation to support renewable energy generation. *Journal of Applied Remote Sensing*, 9, 097696.

Le Toan, T., Quegan, S., Davidson, M., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H., et al. (2011). The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sensing of Environment*, 115, 2850–2860.

Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. (2003). Lidar remote sensing of forest structure. *Progress in physical geography*, 27, 88–106.

Luna, S.S.D., & Young, A. (2003). The bootstrap and kriging prediction intervals. *Scandinavian Journal of Statistics*, 30, 175–192.

Lutz, D.A., Washington-Allen, R.A., & Shugart, H.H. (2008). Remote sensing of boreal forest biophysical and inventory parameters: a review. *Canadian Journal of Remote Sensing*, 34, S286–S313.

Magnani, F., Mencuccini, M., Borghetti, M., Berbigier, P., Berninger, F., Delzon, S., Grelle, A., Hari, P., Jarvis, P.G., Kolari, P., et al. (2007). The human footprint in the carbon cycle of temperate and boreal forests. *Nature*, 447, 849–851.

Magnussen, S. (2015). Arguments for a model-dependent inference? *Forestry: An International Journal of Forest Research*, 88, 317.

Margolis, H.A., Nelson, R.F., Montesano, P.M., Beaudoin, A., Sun, G., Andersen, H.E., & Wulder, M.A. (2015). Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. *Canadian Journal of Forest Research*, 45, 838–855.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58, 1246–1266.

McRoberts, R.E., Holden, G.R., Nelson, M.D., Liknes, G.C., & Gormanson, D.D. (2005). Using satellite imagery as ancillary data for increasing the precision of estimates for the forest inventory and analysis program of the usda forest service. *Canadian Journal of Forest Research*, 35, 2968–2980.

McRoberts, R.E., Næsset, E., & Gobakken, T. (2013). Inference for lidar-assisted estimation of forest growing stock volume. *Remote Sensing of Environment*, 128, 268–275.

McSweeny-McNary Act (1928). 45 stat. 699–702.

Michener, W., Porter, J., Servilla, M., & Vanderbilt, K. (2011). Long term ecological research and information management. *Ecological Informatics*, 6, 13–24.

Muss, J., Mladenoff, D., & Townsend, P. (2011). A pseudo-waveform technique to assess forest structure using discrete lidar data. *Remote Sensing of Environment*, 115, 824–835.

Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19, 164–179.

Næsset, E. (2011). Estimating above-ground biomass in young forests with airborne laser scanning. *International Journal of Remote Sensing*, 32, 473–501.

Næsset, E., Gobakken, T., Solberg, S., Gregoire, T.G., Nelson, R., Ståhl, G., & Weydahl, D. (2011). Model-assisted regional forest biomass estimation using lidar and insar as auxiliary data: A case study from a boreal forest area. *Remote Sensing of Environment*, 115, 3599–3614.

Nathoo, F.S. (2010). Space-time regression modeling of tree growth using the skew-t distribution. *Environmetrics*, 21, 817–833. URL: `http://dx.doi.org/10.1002/env.1057`, doi:`10.1002/env.1057`.

Neigh, C., Nelson, R., Ranson, K., Margolis, H., Montesano, P., Sun, G., Kharuk, V., Næsset, E., Wulder, M., & Andersen, H. (2013). Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne lidar. *Remote Sensing of Environment*, 137, 274–287.

Nelson, R. (2010). Model effects on GLAS-based regional estimates of forest biomass and carbon. *International Journal of Remote Sensing*, 31, 1359–1372.

Nelson, R., Gobakken, T., Næsset, E., Gregoire, T., Ståhl, G., Holm, S., & Flewelling, J. (2012). Lidar sampling – using an airborne profiler to estimate forest biomass in Hedmark County, Norway. *Remote Sensing of Environment*, 123, 563–578.

Nelson, R., Short, A., & Valenti, M. (2004). Measuring biomass and carbon in Delaware using an airborne profiling lidar. *Scandinavian Journal of Forest Research*, 19, 500–511.

Nyland, R. (2007). Silviculture: Concepts and Applications. McGraw-Hill series in forest resources, Waveland PressInc.

Oliveira, V.D., & Kone, B. (2015). Prediction intervals for integrals of gaussian random fields. *Computational Statistics & Data Analysis*, 83, 37–51.

Oliver, C., & Larson, B. (1996). Forest Stand Dynamics. Biological Resource Managament, Wiley.

Ometto, J.P., Aguiar, A.P., Assis, T., Soler, L., Valle, P., Tejada, G., Lapola, D.M., & Meir, P. (2014). Amazon forest biomass density maps: Tackling the uncertainty in carbon emission estimates. *Climatic Change*, , 1–16.

Organic Administration Act (1897). 30 stat. 36.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., & Hayes, D. (2011). A large and persistent carbon sink in the world's forests. *Science*, 333, 988–993.

Park, T., Ganguly, S., Tømmervik, H., Euskirchen, E.S., Høgda, K.A., Karlsen, S.R., Brovkin, V., Nemani, R.R., & Myneni, R.B. (2016). Changes in growing season duration and productivity of northern vegetation inferred from long-term remote sensing data. *Environmental Research Letters*, 11, 084001.

Pattison, R., Andersen, H.E., Gray, A.N., Schulz, B.K., Smith, K., Jovan, S., & Manies, K.L. (In prep). Forests of the Tanana Valley State Forest and Tetlin National Wildlife Refuge Alaska - results of the 2014 pilot inventory. US Department of Agriculture Forest Service, Pacific Northwest Research Station.

Pflugmacher, D., Cohen, W.B., Kennedy, R.E., & Yang, Z. (2014). Using Landsat-derived disturbance and recovery history and lidar to map forest biomass dynamics. *Remote Sensing of Environment*, 151, 124–137.

Phillips, O.L., Malhi, Y., Higuchi, N., Laurance, W.F., Nez, P.V., Vsquez, R.M., Laurance, S.G., Ferreira, L.V., Stern, M., Brown, S., & Grace, J. (1998). Changes in the carbon balance of tropical forests: Evidence from long-term plots. *Science*, 282, 439–442.

Powell, S.L., Cohen, W.B., Healey, S.P., Kennedy, R.E., Moisen, G.G., Pierce, K.B., & Ohmann, J.L. (2010). Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, 114, 1053–1068.

Praks, J., Kugler, F., Papathanassiou, K.P., Hajnsek, I., & Hallikainen, M. (2007). Height estimation of boreal forest: Interferometric model-based inversion at l- and x-band versus hutscat profiling scatterometer. *IEEE Geoscience and Remote Sensing Letters*, 4, 466–470.

Rieger, S., Furbush, C.E., Schoephorster, D.B., Summerfield Jr, H., & Geiger, L.C. (1972). Soils of the Caribou-Poker Creeks Research Watershed, interior Alaska. Technical Report. DTIC Document.

Rinehart, A.J., Jones Jr, J.B., & Harms, T.K. (2015). Hydrologic and biogeochemical influences on carbon processing in the riparian zone of a subarctic stream. *Freshwater Science*, 34, 222–232.

Ringvall, A.H., Ståhl, G., Ene, L.T., Næsset, E., Gobakken, T., & Gregoire, T.G. (2016). A poststratified ratio estimator for model-assisted biomass estimation in sample-based airborne laser scanning surveys. *Canadian Journal of Forest Research*, 46, 1386–1395.

Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordkvist, K., & Hyyppä, J. (2015). Model-assisted estimation of growing stock volume using different combinations of lidar and Landsat data as auxiliary information. *Remote Sensing of Environment*, 158, 431–440.

Saarela, S., Holm, S., Grafström, A., Schnell, S., Næsset, E., Gregoire, T.G., Nelson, R.F., & Ståhl, G. (2016). Hierarchical model-based inference for forest inventory utilizing three sources of information. *Annals of Forest Science*, 73, 895–910.

Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., et al. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences*, 108, 9899–9904.

Särndal, C., Swensson, B., & Wretman, J. (1992). Model Assisted Survey Sampling. Springer series in statistics, Springer-Verlag.

Särndal, C.E., Thomsen, I., Hoem, J.M., Lindley, D.V., Barndorff-Nielsen, O., & Dalenius, T. (1978). Design-based and model-based inference in survey sampling [with discussion and reply]. *Scandinavian Journal of Statistics*, 5, 27–52.

Schelin, L., & Sjöstedt-de Luna, S. (2010). Kriging prediction intervals based on semiparametric bootstrap. *Mathematical Geosciences*, 42, 985–1000.

Schreuder, H.T., Gregoire, T.G., & Weyer, J.P. (2001). For what applications can probability and non-probability sampling be used? *Environmental Monitoring and Assessment*, 66, 281–291.

Sendak, P.E., Brissette, J.C., & Frank, R.M. (2003). Silviculture affects composition, growth, and yield in mixed northern conifers: 40-year results from the Penobscot Experimental Forest. *Canadian Journal of Forest Research*, 33, 2116–2128.

Smith, W. (2002). Forest inventory and analysis: A national inventory and monitoring program. *Environmental Pollution*, 116, Supplement 1, S233–S242.

Snijders, T., & Bosker, R. (2012). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. SAGE Publications.

Soja, A.J., Tchebakova, N.M., French, N.H., Flannigan, M.D., Shugart, H.H., Stocks, B.J., Sukhinin, A.I., Parfenova, E., Chapin, F.S., & Stackhouse, P.W. (2007). Climate-induced boreal forest change: Predictions versus current observations. *Global and Planetary Change*, 56, 274–296.

Ståhl, G., Holm, S., Gregoire, T., Gobakken, T., Næsset, E., & Nelson, R. (2011). Model-based inference for biomass estimation in a lidar sample survey in Hedmark County, Norway. *Canadian Journal of Forest Research*, 41, 96–107.

Ståhl, G., Saarela, S., Schnell, S., Holm, S., Breidenbach, J., Healey, S.P., Patterson, P.L., Magnussen, S., Næsset, E., McRoberts, R.E., & Gregoire, T.G. (2016). Use of models in large-area forest surveys: Comparing model-assisted, model-based and hybrid estimation. *Forest Ecosystems*, 3, 1–11.

Sterba, S.K. (2009). Alternative model-based and design-based frameworks for inference from samples to populations: From polarization to integration. *Multivariate behavioral research*, 44, 711–740.

Stow, C.A., Reckhow, K.H., & Qian, S.S. (2006). A bayesian approach to retransformation bias in transformed regression. *Ecology*, 87, 1472–1477.

Strunk, J.L., Reutebuch, S.E., Andersen, H.E., Gould, P.J., & McGaughey, R.J. (2012). Model-assisted forest yield estimation with light detection and ranging. *Western Journal of Applied Forestry*, 27, 53–59.

Talbot, J., Lewis, S.L., Lopez-Gonzalez, G., Brienen, R.J., Monteagudo, A., Baker, T.R., Feldpausch, T.R., Malhi, Y., Vanderwel, M., Murakami, A.A., Arroyo, L.P., Chao, K.J., Erwin, T., van der Heijden, G., Keeling, H., Killeen, T., Neill, D., Vargas, P.N., Gutierrez, G.A.P., Pitman, N., Quesada, C.A., Silveira, M., Stropp, J., & Phillips, O.L. (2014). Methods to estimate aboveground wood productivity from long-term forest inventory plots. *Forest Ecology and Management*, 320, 30–38.

Tanaka-Oda, A., Kenzo, T., Toriyama, J., & Matsuura, Y. (2016). Variability in the growth rates and foliage δ 15 N values of black spruce trees across a slope gradient in the Alaskan interior. *Canadian Journal of Forest Research*, .

Thompson, S. (2002). Sampling. Wiley Series in Probability and Statistics, Wiley.

UN-REDD (2009). The UN-REDD programme. `http://www.un-redd.org/`. Accessed: 8-11-2016.

U.S. Fish and Wildlife Service (2016). Tetlin National Wildlife Refuge, Alaska. `https://www.fws.gov/refuge/tetlin/wildlife_and_habitat/habitat.html`. Accessed: 10-16-2016.

USGS (2016). Global tree canopy cover circa 2010. `http://landcover.usgs.gov/glc/`. Accessed: 8-11-2016.

Ver Hoef, J. (2002). Sampling and geostatistics for spatial data. *Ecoscience*, 9, 152–161.

Ver Hoef, J., Cressie, N., & Barry, R.P. (2004). Flexible spatial models for kriging and cokriging using moving averages and the fast fourier transform (FFT). *Journal of Computational and Graphical Statistics*, 13, 265–282.

Verbyla, D. (2011). Statewide 2-km raster of year since last widlfire, Bonanza Creek LTER - University of Alaska Fairbanks. `http://www.lter.uaf.edu/data/data-detail/id/464`.

Weiskittel, A.R., Hann, D.W., Kershaw Jr, J.A., & Vanclay, J.K. (2011). Forest Growth and Yield Modeling. John Wiley & Sons.

Woodall, C.W., Coulston, J.W., Domke, G.M., Walters, B.F., Wear, D.N., Smith, J.E., Andersen, H.E., Clough, B.J., Cohen, W.B., Griffith, D.M., et al. (2015). The US forest carbon accounting framework: Stocks and stock change, 1990-2016, .

Wulder, M., Coops, N., Hudak, A., Morsdorf, F., Nelson, R., Newnham, G., & Vastaranta, M. (2013). Status and prospects for lidar remote sensing of forested ecosystems. *Canadian Journal of Remote Sensing*, 39, S1–S5.

Wulder, M.A., White, J.C., Bater, C.W., Coops, N.C., Hopkinson, C., & Chen, G. (2012a). Lidar plots—a new large-area data collection option: context, concepts, and case study. *Canadian Journal of Remote Sensing*, 38, 600–618.

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., & Gobakken, T. (2012b). Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, 121, 196–209.

Zhang, G., Ganguly, S., Nemani, R.R., White, M.A., Milesi, C., Hashimoto, H., Wang, W., Saatchi, S., Yu, Y., & Myneni, R.B. (2014). Estimation of forest aboveground biomass in California using canopy height and leaf area index estimated from satellite data. *Remote Sensing of Environment*, 151, 44–56.

Zhen, Z., Quackenbush, L.J., & Zhang, L. (2016). Trends in automatic individual tree crown detection and delineationevolution of lidar data. *Remote Sensing*, 8.

Zheng, D., Rademacher, J., Chen, J., Crow, T., Bresee, M., Moine, J.L., & Ryu, S.R. (2004). Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. *Remote Sensing of Environment*, 93, 402–411.

Zolkos, S., Goetz, S., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment*, 128, 289–298.