LiDAR based prediction of forest biomass using hierarchical models with spatially varying coefficients

Chad Babcock^a, Andrew O. Finley^{b,*}, John B. Bradford^c, Randall Kolka^d, Richard Birdsey^e,
Michael G. Ryan^f

^a*School of Environmental and Forest Sciences, University of Washington, Seattle, Washington, U.S.A.*

^b*Departments of Forestry and Geography, Michigan State University, East Lansing, Michigan, U.S.A.*

^c*US Geological Survey, Southwest Biological Science Center, Flagstaff, Arizona, U.S.A.*

^d*USDA Forest Service, Northern Research Station, Grand Rapids, Minnesota, U.S.A*

^e*USDA Forest Service, Northern Global Change Program, Newtown Square, Pennsylvania, U.S.A.*

^f*USDA Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, U.S.A.*

Abstract

Many studies and production inventory systems have shown the utility of coupling covariates derived from Light Detection and Ranging (LiDAR) data with forest variables measured on georeferenced inventory plots through regression models. The objective of this study was to propose and assess the use of a Bayesian hierarchical modeling framework that accommodates both residual spatial dependence and non-stationarity of model covariates through the introduction of spatial random effects. We explored this objective using four forest inventory datasets that are part of the North American Carbon Program, each comprising point-referenced measures of above-ground forest biomass and discrete LiDAR. For each dataset, we considered at least five regression model specifications of varying complexity. Models were assessed based on goodness of fit criteria and predictive performance using a ten-fold cross-validation procedure. Results showed that the addition of spatial random effects to the regression model intercept improved fit and predictive performance in the presence of substantial residual spatial dependence. Additionally, in some cases, allowing either some or all regression slope parameters to vary spatially, via the addition of spatial random effects, further improved model fit and predictive performance. In other instances, models showed improved fit but decreased predictive performance—indicating over-fitting and underscoring the need for cross-validation to assess predictive ability. The proposed Bayesian modeling framework provided access to pixel-level posterior predictive distributions that were useful for uncertainty mapping, diagnosing

*Corresponding author. Tel.: +1 517 432 7219.
Email address: finleya@msu.edu (Andrew O. Finley)
Preprint submitted to Elsevier

June 4, 2015

spatial extrapolation issues, revealing missing model covariates, and discovering locally significant parameters.

Keywords: Bayesian hierarchical models, Markov chain Monte Carlo, Gaussian process, geospatial, LiDAR, forest biomass

1. Introduction

Estimating forest above-ground biomass (AGB), along with other structure variables, using discrete Light Detection and Ranging (LiDAR) data is an active subject in ecological and resource monitoring research. Studies in this area show a strong potential for discrete return LiDAR to be used as a tool for developing spatially explicit estimates of many forest attributes, including AGB, either on its own or in conjunction with other remote sensing technologies (see, e.g., Gonzalez et al., 2010; Sherrill et al., 2008; Lucas et al., 2006; Lim & Treitz, 2004).

Regression models proposed for forest AGB mapping using LiDAR data often do not explicitly accommodate residual spatial dependence, see, e.g, Anderson et al. (2008), Gonzalez et al. (2010), Muss et al. (2011), Tonolli et al. (2011), and Popescu et al. (2011). A non-spatial model can be appropriate if all spatially structured variation in the outcome is accounted for by the covariates used for model fitting; however, this is often an unrealistic assumption when data are spatially indexed. Considering forest attribute data, it is reasonable to expect similar outcomes for neighboring locations. For example, inventory plots close together could comprise like tree species and have comparable stem densities due to similar disturbance histories and environmental conditions. In contrast, inventory plots far apart are less likely to share common composition or structure attributions. A key assumption of regression is that model residuals are not correlated. Given the spatially dependent nature of forest inventory data, the potential for spatially correlated residuals in AGB models is high. Not accounting for residual spatial dependence can result in falsely precise estimates of model parameters and erroneous predictions (Hoeting, 2009). For these reasons, it is important to check for spatially structured residuals when spatial data are used to fit AGB models.

To ensure a model's statistical validity, spatial dependence among residuals can be accommodated by introducing a spatially varying intercept via the addition of an appropriately structured

37 random effect. Including spatial random effects that account for spatial association between ob-
38 served locations can also improve model fit and prediction accuracy, via partitioning error uncer-
39 tainty into a spatial and non-spatial component and borrowing information from observed locations
40 to inform prediction at proximate locations.

41 In addition to a spatially varying intercept, it may be appealing to allow the regression pa-
42 rameters associated with the covariates to vary across the study area. For instance, given the
43 heterogeneity of forest species composition, age classes, and resources (e.g., light, water, soil char-
44 acteristics), a single set of regression parameters might not adequately capture the space varying
45 relationship between forest outcome variables and the covariates, i.e., parameters may have a spa-
46 tially non-stationary relationship with the covariates. In such cases, one might attempt to capture
47 these localized effects by including forest characteristics and/or environmental conditions as covari-
48 ates in the regression model; however, it is not always clear which variables should be included and,
49 in many applied settings, the necessary variables are not available. Rather, localized relationships
50 between the outcome variable and covariates might be more effectively captured by allowing the
51 regression coefficients to vary smoothly over the study area.

52 Such spatially varying coefficient (SVC) models have received some attention in the statistical
53 literature (see, e.g., Gelfand et al., 2004; Finley et al., 2009, 2011), and more recently in applied
54 disciplines (Wheeler & Calder, 2007; Wheeler & Waller, 2009; Finley, 2010). Flexible specification
55 of these models follows the Bayesian paradigm of statistical inference (see, e.g., Carlin & Louis,
56 2000; Gelman et al., 2013; Banerjee et al., 2004), where analysis uses samples, obtained via Markov
57 chain Monte Carlo (MCMC) methods, from the posterior distributions of model parameters.

58 SVC models accommodate residual spatial dependence and parameter non-stationarity differ-
59 ently than more common geostatistical approaches seen in the remote sensing literature. Chen et al.
60 (2012) examined geographically weighted regression (GWR, Fotheringham et al., 2003) alongside
61 other geostatistical approaches for predicting canopy height with remote sensing data and saw im-
62 provements in model fit. In an attempt to accommodate residual spatial dependence and parameter
63 non-stationarity, Hudak et al. (2002) explored kriging and co-kriging geostatistical methods to build
64 prediction models for forest canopy height. They saw improvements in fit using regression kriging

65 and regression co-kriging.

66 Recently, it has been shown that GWR may not be robust to correlation among covariates and
67 produce erroneous results in the presence of complex correlation structures (Wheeler & Waller,
68 2009; Finley et al., 2011). Also, from an inferential standpoint, GWR can be problematic when
69 attempting to draw inference about model parameters and prediction uncertainty. Because GWR
70 lacks a valid underlying probability model, parameter standard error estimates and prediction
71 variance estimates can potentially be unjustifiable. Error maps generated by kriging and co-kriging
72 methods do not account for the uncertainty in the variogram-derived spatial covariance parameters,
73 which is a well-known shortcoming encountered with these geostatistical approaches (Cressie, 1993).
74 In a Bayesian hierarchical framework, it is possible to estimate the spatial covariance parameters
75 within SVC models and, hence, propagate their uncertainties through to the prediction of the
76 outcome variable (Banerjee et al., 2004). Such approaches yield a more statistically defensible map
77 of uncertainty than would otherwise be produced using traditional kriging or GWR methods.

78 In an effort to more fully account for patterns of spatial dependence between AGB and LiDAR
79 covariates, the utility of a Bayesian hierarchical modeling framework that accommodates both
80 residual spatial dependence and non-stationarity of model covariates through the introduction of
81 spatial random effects was assessed. This objective was explored using four forest inventory datasets
82 that are part of the North American Carbon Program (NACP), each comprising point-referenced
83 measures of AGB and discrete LiDAR. For each dataset, a set of regression model specifications
84 of varying complexity was considered. Models were assessed based on fit criteria and predictive
85 performance using a ten-fold cross-validation approach. AGB maps with associated pixel-level
86 prediction uncertainty measures were generated using the model selected as “best” for each site.
87 In cases where the selected model included spatial random effects, additional maps were created
88 to illustrate features of the space-varying coefficients. Results showed, in an exploratory fashion,
89 how uncertainty and space-varying coefficient maps can be used to identify: spatial extrapolation
90 issues; missing covariates, and; locally significant regression coefficients.

91 The remainder of the paper evolves as follows. The motivating data and preprocessing steps are
92 detailed in Section 2. The proposed modeling framework and model assessment are described in

Section 3, followed by analysis results presented in Section 4. Finally, some discussion along with an indication of future work are provided in Section 5.

2. Study Sites

2.1. Fraser Experimental Forest

The Fraser Experimental Forest (FEF) is located in central Colorado (39° 4' N, 105° 52' W) near the town of Fraser. Tree species at FEF consist primarily of *Abies lasiocarpa* Hook. Nutt and *Picea engelmannii* Parry at higher elevations and *Pinus contorta* Bol. at lower elevations. Climate at FEF is characterized by cold and relatively long winters, with mean annual temperature and precipitation of 0°C and 737 mm, respectively. FEF experienced a widespread stand-replacing fire in approximately 1685. Selective clearcuts were conducted at FEF as watershed-scale manipulations in the 1950s. Although FEF is currently experiencing mortality due to mountain pine beetle, the field measurements and LiDAR acquisition for this study were completed prior to beetle infestation.

2.2. Marcell Experimental Forest

Located in northern Minnesota, the Marcell Experimental Forest (MEF) consists of mixed forests that include both upland forests and peatlands. Upland forests are generally dominated by *Populus tremuloides* Michx. and *grandidentata* Michx., but also contain substantial components of *Betula papyrifera* Marshall, *Pinus resinosa* Roezl, *Pinus strobus* L., and *Pinus banksiana* Lamb.. Lowland tree species include *Larix laricina* (Du Roi) K.Koch, *Picea mariana* Britton, Sterns & Poggenb., *Fraxinus nigra* Marshall, and *Thuja occidentalis* L.. Climate at MEF is subhumid continental, with mean annual precipitation of 785 mm, mean annual temperature of 3°C and air temperature extremes from -46°C to 38°C. Forests of the Lake States region experienced widespread logging around the turn of the 20th century, including much of the MEF landscape. Natural disturbances at MEF include windstorms of variable intensity and rare wildfires.

2.3. Glacier Lake Ecosystem Experimental Site

Glacier Lake Ecosystem Experimental Site (GLEES) is located approximately 55 km west of Laramie, Wyoming at 3190 m elevation (41° 22' N, 106° 14' W). GLEES includes a mosaic of trees

119 and alpine meadows; the forest is dominated by *Picea engelmannii* and *Abies lasiocarpa*. GLEES
120 has a mean annual temperature of -2°C and a mean annual precipitation of 1000 mm, primarily
121 as snow. Tree ages at GLEES suggest either a stand-replacing disturbance more than 400 years
122 ago with slow recovery, or a series of smaller disturbances over the last 400 years (Bradford et al.,
123 2008).

124 2.4. Niwot Long Term Ecological Research Site

125 Niwot Long Term Ecological Research Site (NIWOT) is located at 3050 m elevation on the front
126 range of the Rocky Mountains ($40^{\circ} 2' \text{ N}$, $105^{\circ} 33' \text{ W}$), near the town of Nederland, Colorado. Tree
127 species include primarily a mix of *Abies lasiocarpa*, *Picea engelmannii* and *Pinus contorta* with
128 minor components of *Pinus flexilis* A.Murr. and *Populus tremuloides*. Mean annual temperature
129 and precipitation are 4°C and 800 mm, respectively. Disturbance history at NIWOT includes
130 widespread clearcuts between 1900-1910.

131 2.5. Field Data Preparation

132 Field data at each site were collected using methods similar to the Forest Inventory and Analysis
133 style plot design (Bechtold & Patterson, 2005). Each plot location consisted of 4 sub-plots with a
134 radius of 8–10 m (depending on site) with 1 being in the center and 3 others 35 m away from the
135 plot center at 0° , 120° , and 240° . Each site was initially established with 9 (FEF, GLEES, and
136 NIWOT) or 16 (MEF) plots in 2004 located on a regular grid with 250 meters between plot centers
137 in the north-south and east-west directions. Additional plots and/or sub-plots were subsequently
138 added during 2005 and 2006 at each site in locations selected to capture canopy height conditions
139 not adequately represented in the initial gridded plots. FEF, MEF, GLEES and NIWOT contained
140 61, 115, 46 and 62 sub-plots respectively. Figure 1 shows the spatial distribution of the sub-plots
141 inside the LiDAR coverage area. Within each sub-plot, individual tree diameter at breast height,
142 and height measurements for both live and dead trees were taken and used in species-specific
143 allometric equations to estimate AGB (stem, branch and foliage biomass). Additional details about
144 field measurements and allometric equations used for biomass estimation are available in Bradford
145 et al. (2010). The individual tree AGB estimates were totaled for each sub-plot and converted to

146 AGB Mg/ha. The AGB measurements were then square root transformed to better approximate a
147 normal distribution before model fitting.

148 *2.6. LiDAR Data Preparation*

149 Height-above-ground measurements for the first return pulses were calculated by subtracting
150 the point elevations from a digital terrain model constructed from the LiDAR data. LiDAR return
151 height empirical cumulative distribution curves were constructed over each sub-plot and a percentile
152 height dataset was compiled by extracting heights associated with 5% intervals (i.e., 5%–100%) for
153 each curve. These percentile heights served as an initial set of regression covariates. Using the same
154 intervals, a percentile height dataset was constructed for a 20×20 m grid over the LiDAR coverage
155 area and subsequently used to construct AGB prediction maps. The cell size was chosen because
156 it is approximately the same area as the observed sub-plots.

157 Variable selection via dimension reduction was necessary because of high collinearity among
158 the percentile height covariates. We chose principal components analysis (PCA) via eigen (spec-
159 tral) decomposition of the percentile heights correlation matrix to reduce the covariate set. Eigen
160 decomposition is a matrix factorization technique that is useful for characterizing patterns in high-
161 dimensional data (Harville, 1997). Eigen decomposition (or similar orthogonalization techniques) is
162 used widely for data reduction or compression in statistics, signal processing, pattern recognition,
163 remote sensing, and other fields where high-dimensional and/or highly correlated data are encoun-
164 tered (Guanter et al., 2012, 2013; Huang et al., 2014; Wang et al., 2014). It is common to see
165 PCA performed by decomposing the covariance matrix in the remote sensing sciences (Richards,
166 2013). We elected to decompose the correlation matrix due to the natural tendency for upper
167 percentile heights to be more variable than lower percentile heights. Johnson & Wichern (2007)
168 offers a detailed discussion about how standardizing the variables, i.e., using the correlation matrix,
169 reduces the effect of individual covariate variances on principal component loadings. In Section A1,
170 the mathematical procedure for conducting PCA on standardized variables and transforming the
171 gridded data set to the correct basis for prediction is detailed. Criteria for selecting principal com-
172 ponents are also presented in Section A1. For the FEF, NIWOT, and GLEES sites, two principal

173 components were retained and three were selected for MEF.

174 Some grid cell locations contained covariate values that well exceeded the range of values used
 175 for model fitting. This is a result of a relatively small set of observed plots compared to the size and
 176 heterogeneity of the study area. To avoid potential issues arising from model extrapolation (Perrin,
 177 1904), grid cells with values outside the support of the observed covariate values were removed from
 178 the analysis. Water bodies at GLEES and MEF were also excluded from the analysis. Excluded
 179 grid cells are colored black in the subsequent maps.

180 *Note, upon acceptance, inventory plot and LiDAR data described in Section 2 will be made*
 181 *available via the RSE's supplemental material or public data repository, e.g., Dryad. Our hope is*
 182 *the data can serve as a test bed for future method development.*

183 3. Methods

184 3.1. Modeling Framework

185 Geostatistical settings typically assume at location $\mathbf{s} \in \mathcal{D} \subseteq \mathbb{R}^2$, where $\mathbf{s}$ is a vector of geographic
 186 coordinates in domain $\mathcal{D}$, a Gaussian outcome variable $y(\mathbf{s})$ is modeled using the regression model,

$$y(\mathbf{s}) = \mathbf{x}(\mathbf{s})'\boldsymbol{\beta} + \tilde{\mathbf{x}}(\mathbf{s})'\mathbf{w}(\mathbf{s}) + \epsilon(\mathbf{s}). \quad (1)$$

187 Here the linear mean structure that accounts for large scale variation in the outcome is composed of
 188 a $p \times 1$ vector $\mathbf{x}(\mathbf{s})$, comprising an intercept and spatially referenced covariates, and an associated
 189 column vector of regression coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})'$. The $\tilde{\mathbf{x}}(\mathbf{s})$ is a $q \times 1$ vector that
 190 includes the intercept and those covariates from $\mathbf{x}(\mathbf{s})$ whose impact on the outcome is posited
 191 to vary spatially. This space varying impact is captured through the vector of spatial random
 192 effects $\mathbf{w}(\mathbf{s}) = (w_1(\mathbf{s}), w_2(\mathbf{s}), \dots, w_q(\mathbf{s}))'$. Various sub-models are formed by specifying different
 193 combinations of $\tilde{\mathbf{x}}(\mathbf{s})$ and associated $\mathbf{w}(\mathbf{s})$. Customarily, $\epsilon(\mathbf{s})$ is modeled as an independent white-
 194 noise process that captures measurement error or micro-scale variation. With any collection of n
 195 locations, say $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\}$, we assume that the $\epsilon(\mathbf{s}_i)$'s are identically and independently
 196 distributed as $N(0, \tau^2)$, where τ^2 is often called the *nugget*.

For this model, spatial structure is typically introduced through a multivariate Gaussian process (GP) (Cressie, 1993; Wackernagel, 2003; Banerjee et al., 2004), where a *cross-covariance function* explicitly models the covariance of $\mathbf{w}(\mathbf{s})$ both within and among locations. This additional flexibility is attractive in many settings (see, e.g., Gelfand et al., 2004; Banerjee et al., 2011; Finley et al., 2011); however, for the objectives of this study it is not clear the extra effort will fetch substantial advantages. Rather, for simplicity and substantially reduced computational demand, we assumed the elements of $\mathbf{w}(\mathbf{s})$ arise from q independent univariate GPs. Specifically, the process associated with the k -th covariate is $w_k(\mathbf{s}) \sim GP(0, C(\cdot, \cdot; \boldsymbol{\theta}_k))$ where $C(\mathbf{s}, \mathbf{s}^*; \boldsymbol{\theta}_k) = \text{Cov}(w_k(\mathbf{s}), w_k(\mathbf{s}^*))$ is a valid covariance function that models the covariance corresponding to a pair of locations $\mathbf{s}$ and $\mathbf{s}^*$. The process realizations are collected into an $n \times 1$ vector, say $\mathbf{w}_k = (w_k(\mathbf{s}_1), \dots, w_k(\mathbf{s}_n))'$, that follows a multivariate normal distribution $MVN(\mathbf{0}, \Sigma_k)$, where Σ_k is the $n \times n$ covariance matrix of $\mathbf{w}_k$ with (i, j) -th element given by $C(\mathbf{s}_i, \mathbf{s}_j; \boldsymbol{\theta}_k)$. Clearly $C(\mathbf{s}, \mathbf{s}^*; \boldsymbol{\theta}_k)$ cannot be just any function; it must ensure that the resulting Σ_k matrix is symmetric and positive definite. Such functions are known as positive definite functions and are characterized as the characteristic function of a symmetric random variable. Further technical details about positive definite functions can be found in Cressie (1993), Chilès & Delfiner (1999), and Banerjee et al. (2004).

We specify $C(\mathbf{s}, \mathbf{s}^*; \boldsymbol{\theta}_k) = \sigma_k^2 \rho(\mathbf{s}, \mathbf{s}^*; \phi_k)$ with $\boldsymbol{\theta}_k = \{\sigma_k^2, \phi_k\}$, $\rho(\cdot; \phi_k)$ is a valid spatial correlation function, where ϕ_k quantifies the rate of correlation decay and $\text{Var}(w_k) = \sigma_k^2$. For the subsequent analyses we assumed an exponential correlation function, $\rho(\|\mathbf{s} - \mathbf{s}^*\|; \phi_k) = \exp(-\phi_k \|\mathbf{s} - \mathbf{s}^*\|)$, where $\|\mathbf{s} - \mathbf{s}^*\|$ is the Euclidean distance between the locations $\mathbf{s}$ and $\mathbf{s}^*$.

To complete the Bayesian model specification, we assign prior distributions to the model parameters and inference proceeds by sampling from the posterior distribution of the parameters. As customary, we assume $\boldsymbol{\beta}$ follows a $MVN(\boldsymbol{\mu}_\beta, \Sigma_\beta)$ prior with $\boldsymbol{\mu}_\beta = \mathbf{0}$ and $\Sigma_\beta = 10,000 \mathbf{I}_p$, while the spatial variance components σ_k^2 's and the measurement error variance τ^2 are assigned inverse-Gamma, $IG(a, b)$, priors. The process correlation decay parameters ϕ_k 's follow a Uniform, $Unif(a, b)$, with support over the geographic range of the study area.

Using notations similar to Gelman et al. (2013), we can write the posterior distribution of the model parameters as $p(\Omega | \mathbf{y})$, where $\Omega = \{\boldsymbol{\beta}, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_q, \sigma_1^2, \sigma_2^2, \dots, \sigma_q^2, \phi_1, \phi_2, \dots, \phi_q, \tau^2\}$ and

225 $\mathbf{y} = (y(\mathbf{s}_1), \dots, y(\mathbf{s}_n))'$, which is proportional to

$$\begin{aligned} \prod_{k=1}^q \text{Unif}(\phi_k | a_{\phi_k}, b_{\phi_k}) \times \prod_{k=1}^q IG(\sigma_k^2 | a_{\sigma_k}, b_{\sigma_k}) \times N(\boldsymbol{\beta} | \boldsymbol{\mu}_\beta, \Sigma_\beta) \times IG(\tau^2 | a_\tau, b_\tau) \times \\ \times \prod_{k=1}^q N(\mathbf{w}_k | \mathbf{0}, \Sigma_k) \times \prod_{i=1}^n N(y(\mathbf{s}_i) | \mathbf{x}(\mathbf{s}_i)' \boldsymbol{\beta} + \tilde{\mathbf{x}}(\mathbf{s})' \mathbf{w}(\mathbf{s}), \tau^2). \end{aligned} \quad (2)$$

226 An efficient MCMC algorithm for estimation of (2) is obtained by updating $\boldsymbol{\beta}$ from its full
227 conditional then using Metropolis steps for the remaining parameters. Alternatively, the model
228 can be reparametrized such that the spatial random effects $\mathbf{w}$ do not need to be sampled directly
229 (Banerjee et al., 2008). In either case, the MCMC algorithms yield posterior samples of Ω .

230 For predictions, if $\mathcal{S}_0 = \{\mathbf{s}_{0,1}, \mathbf{s}_{0,2}, \dots, \mathbf{s}_{0,m}\}$ is a collection of m new locations, the posterior
231 predictive distribution of the spatial random effects associated with the k -th regression coefficient
232 is given by

$$p(\mathbf{w}_{k,0} | \mathbf{y}) \propto \int p(\mathbf{w}_{k,0} | \mathbf{w}_k, \Omega, \mathbf{y}) p(\mathbf{w}_k | \Omega, \mathbf{y}) p(\Omega | \mathbf{y}) d\Omega \mathbf{w}_k, \quad (3)$$

233 where $\mathbf{w}_{k,0} = (w_k(\mathbf{s}_{0,1}), w_k(\mathbf{s}_{0,2}), \dots, w_k(\mathbf{s}_{0,m}))'$.

234 Because we are using MCMC sample-based inference, i.e., drawing samples from the posterior
235 of Ω , the integral in (3) does not have to be evaluated explicitly. Rather, given L posterior samples
236 for Ω , i.e., $\{\Omega^{(l)}\}_{l=1}^L$, this distribution can be obtained via composition sampling (see, e.g., Banerjee
237 et al. (2008) pg. 132) by first draw $\mathbf{w}_k^{(l)}$ and then draw $\mathbf{w}_{k,0}^{(l)}$ for each l from $p(\mathbf{w}_{k,0} | \mathbf{w}_k^{(l)}, \Omega^{(l)}, \mathbf{y})$,
238 where this last distribution is derived as a conditional distribution from a multivariate normal
239 and, hence, is multivariate normal. More precisely, the process realizations over the new locations
240 are conditionally independent of the observed outcomes given the realizations over the observed
241 locations and the process parameters. In other words, $p(\mathbf{w}_{k,0} | \mathbf{w}_k, \Omega, \mathbf{y}) = p(\mathbf{w}_{k,0} | \mathbf{w}_k, \Omega)$, which is
242 a multivariate normal distribution with mean and variance given by

$$E[\mathbf{w}_{k,0} | \mathbf{w}_k, \Omega] = \text{Cov}(\mathbf{w}_{k,0}, \mathbf{w}_k) \text{Var}^{-1}(\mathbf{w}_k) \mathbf{w}_k = R_0(\phi_k)' R(\phi_k)^{-1} \mathbf{w}_k \quad (4)$$

$$\text{and } \text{Var}[\mathbf{w}_{k,0} | \mathbf{w}_k, \Omega] = \sigma_k^2 \{ R(\phi_k) - R_0(\phi_k)' R(\phi_k)^{-1} R_0(\phi_k) \}, \quad (5)$$

where $R_0(\phi_k)$ is the $n \times m$ matrix with (i, j) -th element given by $\rho(\mathbf{s}_{0,i}, \mathbf{s}_j; \phi_k)$ and $R(\phi_k)$ is the $n \times n$ matrix with (i, j) -th element given by $\rho(\mathbf{s}_i, \mathbf{s}_j; \phi_k)$. This procedure is repeated to generate samples for all $\mathbf{w}_{k,0}$'s. Finally, given a set of covariates at a new location $\mathbf{s}_0$, samples from the posterior predictive distribution of the outcome variable, $y(\mathbf{s}_0)^{(l)}$, are drawn from $N(\mathbf{x}(\mathbf{s}_0)' \boldsymbol{\beta}^{(l)} + \tilde{\mathbf{x}}(\mathbf{s}_0)' \mathbf{w}_0^{(l)}, \tau^2)^{(l)}$ for $l = 1, 2, \dots, L$.

95% credible interval widths (CIW) can be calculated using pixel-level posterior predictive distributions of AGB by measuring the distance between the 2.5% and 97.5% quantile breaks (Figure 2). The 95% CIWs serve as spatially explicit summaries of prediction uncertainty and can be mapped in order to assess model prediction performance spatially.

3.2. Candidate Models

Five candidate models were derived from (1) and include: *non-spatial*, where $\mathbf{w}_k$'s are set to zero; spatially varying intercept (*SVI*), where only the spatial random effects associated with the model intercept are included; the full spatially varying coefficients (*SVC*), where all regression coefficients have associated spatial random effects; *SVC- β_1* , where spatial random effects for the intercept and β_1 are included, and; *SVC- β_2* , where spatial random effects for the intercept and β_2 are included. Due to MEF's additional covariate, four additional candidate models were tested. *SVC- β_3* has spatial random effects associated with β_0 and β_3 ; *SVC- $\beta_1\beta_2$* has spatial random effects associated with β_0 , β_1 , and β_2 ; *SVC- $\beta_1\beta_3$* has spatial random effects associated with β_0 , β_1 , and β_3 , and; *SVC- $\beta_2\beta_3$* has spatial random effects associated with β_0 , β_2 , and β_3 .

Empirical semivariograms fit to the residuals of the *non-spatial* models (Figure 3) were used to help guide hyperprior specification for the candidate models' *IG* and *Unif* priors. Specifically, for the variance parameters the *IG* hyperprior a was set equal to 2, which results in a prior distribution mean equal to b and infinite variance (see, e.g., *IG* definition in Gelman et al. (2013)). Then the b hyperpriors for the models' τ^2 and σ^2 's were set according to the nugget and partial sill of the semivariograms, respectively. The prior for the spatial decay parameters, ϕ 's, was set to *Unif*(0.0006, 3) which, assuming the exponential spatial correlation function, corresponds to support for an *effective spatial range* between ~ 1 -5,000 m. The effective spatial range is defined

270 here as the distance at which the correlation equals 0.05.

271 The MCMC samplers were implemented in C++ and Fortran and leveraged Intel’s Math Ker-
272 nel Library threaded BLAS and LAPACK routines for matrix computations. All analyses were
273 conducted on a Linux workstation using two Intel Nehalem quad-Xeon processors.

274 Three MCMC chains were run for 25,000 iterations each. The most demanding model required
275 ~ 1 hour to complete a single MCMC chain. Convergence was diagnosed using the CODA package
276 in R by monitoring mixing of chains and the Gelman-Rubin statistic (Gelman & Rubin, 1992).
277 Satisfactory convergence was diagnosed within 10,000 iterations for all models. Posterior inference
278 was based on a post burn-in sub-sample of 15,000 iterations—every third sample from the last 15,000
279 iterations of each chain.

280 3.3. Fit and Prediction Accuracy Assessment

281 Model performance was assessed using the popular Deviance Information Criterion (DIC) to
282 rank models in terms of how well they fit the data (Spiegelhalter et al., 2002). This criteria is
283 the sum of the Bayesian deviance (a measure of model goodness of fit) and the effective number
284 of parameters (a penalty for model complexity), D and p_D , respectively. Lower values of DIC
285 indicate better model fit.

286 Predictive performance was assessed using a ten-fold cross-validation approach. Here, the full
287 dataset was split into ten roughly equal sized subsets. Then AGB was predicted for locations within
288 each subset, given parameters estimated from the remaining subsets. Root mean squared prediction
289 error (RMSPE) was then calculated using the observed AGB values and corresponding median of
290 the posterior predictive distribution. Lower RMSPE values signify more accurate predictions.

291 4. Results

292 The experimental design produced 24 combinations of candidate models and sites. Summaries of
293 the candidate models’ parameter estimates, fit, and predictive performance for FEF, MEF, GLEES
294 and NIWOT are provided in Tables 1, 2, 3, and 4. Concerning all sites’ *non-spatial* models, the
295 exclusion of zero from the regression slope parameters’ 95% credible interval suggested the derived

LiDAR covariates explained a substantial portion of variability in AGB, with the exception of NIWOT's β_2 parameter.

For the FEF, GLEES, and NIWOT sites, the *SVI* models produced marginally to moderately lower D , DIC , and $RMSPE$ as compared to the *non-spatial* models (Tables 1, 3, and 4). These results were not surprising given the clear spatial structure seen in the exploratory semivariograms of the *non-spatial* models' residuals. In contrast, the semivariogram of the *non-spatial* model residuals for MEF did not provide strong evidence of any residual spatial correlation (Figure 3 (b)); hence, the *SVI* model produced only slight gains in model fit, along with a decrease in model prediction accuracy, when compared to the *non-spatial* candidate (Table 2). This demonstrated that the addition of a spatial random effect to the MEF model intercept was not only unnecessary, but allowing the intercept to vary spatially caused the model to over-fit the data and reduced predictive performance.

The candidate models' parameter estimates further corroborated the notion that spatial structure was being absorbed by the spatial random effects in the presence of residual spatial autocorrelation (Tables 1, 2, 3, and 4). The ratio $\tau^2/(\tau^2 + \sigma_0^2)$ calculated using estimates from the *SVI* model proved useful for assessing the need for spatial random effects. Lower ratios indicated greater spatial structure in the *non-spatial* models' residuals and, hence, need to include the spatially-varying intercept. This ratio was 0.15, 0.13, and 0.29 for FEF, GLEES, and NIWOT, respectively. Compared to MEF's 0.45 ratio, the *non-spatial* model residuals for FEF, GLEES, and NIWOT contained more spatial structure.

The effective spatial range is another parameter that describes the strength of a spatial component. A long spatial range suggests stronger and farther reaching spatial dependence. MEF's *SVI* spatial range of 29.5 m was much shorter than the *SVI* ranges for FEF, GLEES and NIWOT (Tables 1, 2, 3, and 4). The long spatial ranges and low nugget to total variance ratios suggested the spatial random effects were capturing substantial residual spatial structure and the *non-spatial* models were not appropriate for the FEF, GLEES, and NIWOT datasets. That is, inference drawn from the *non-spatial* model at FEF, GLEES, and NIWOT—particularly uncertainty for parameter point estimates and predictions—should be interpreted with caution due to the correlated residuals

and hence violation of the independent and identically distributed residuals assumption. In contrast, the variance ratio and effective spatial range values suggested there was little spatial structure in the MEF *non-spatial* model residuals—indicating that spatial random effects were not warranted.

The remainder of this section is divided into four subsections, one for each site. In each subsection the “best” model is identified and some site-specific results are explored.

4.1. FEF Results

At FEF, the full *SVC* model produced the lowest D , DIC , and $RMSPE$. The *SVC* model showed an 83% improvement in DIC compared to the *non-spatial* model. There was high correspondence between the observed and fitted values produced with the *SVC* model, whereas the fitted values for the *non-spatial* model were scattered further from the plot data (Figure A2). The credible intervals for the *SVC* fitted values were also much tighter than the *non-spatial* fitted values (Figure A2). There was a 9% improvement in $RMSPE$ when moving from the *non-spatial* to *SVC* model. Most of the predictions produced with the *SVC* model fell within the range of observed AGB (0 – 430 Mg/ha, Figure 4a). Those pixels that exceeded this observed range exhibited high associated uncertainties (Figure 5a). For example, *SVC* model predictions in a small patch in the southeastern corner of the FEF domain showed large uncertainty, perhaps indicating forest structure characteristics in the region are not adequately represented in the plot data used to estimate model parameters. Further, there is a pattern of large uncertainty across the northern region of the FEF, again suggesting a paucity of plot data to characterize the relationship between AGB and the LiDAR covariates.

The mean, or median, posterior predictive surface of $\beta_k(\mathbf{s}) = \beta_k + w_k(\mathbf{s})$ can provide insights into the space-varying relationship between the k -th covariate and AGB. For example, the $\beta_1(\mathbf{s})$ coefficient map showed clear spatial structure and suggested the associated LiDAR covariate explained more variability in AGB in the northern region of the study area, i.e., $\beta_1(\mathbf{s})$ was farther from zero in the north (Figure 6b). After examining the LiDAR derived digital terrain model, it could be seen that $\beta_1(\mathbf{s})$ was related to elevation (Figure 7). This suggests the relationship between AGB and the LiDAR covariate associated with $\beta_1(\mathbf{s})$ changes as a function of elevation—due perhaps

to species composition change or a fertility gradient that influences the relationship between the covariate and stem biomass. Although not further explored here, one could envision adapting this model to include elevation as a covariate to potentially eliminate the need for β_1 to vary spatially.

Compared to $\beta_1(\mathbf{s})$, random effects associated with $\beta_0(\mathbf{s})$ and $\beta_2(\mathbf{s})$ had much shorter effective ranges (Table 1), which made their adjustments less noticeable at the scale shown in Figure 6. By zooming into a region with a high concentration of sub-plots, switching from an equal to quantile interval color classification, and passing the predictions of the $\beta(\mathbf{s})$'s through a smoothing filter, the more subtle spatial adjustments to β_0 and β_2 could be seen. Figures 8a and 8b depict $\beta_0(\mathbf{s})$ and $\beta_2(\mathbf{s})$ at the extents shown by the red outline in Figures 6a and 6c, respectively. The zoomed-in figures show that sub-plot clusters received slightly higher or lower $\beta_0(\mathbf{s})$ and $\beta_2(\mathbf{s})$, confirming that the random effects are capturing local or micro-scale variability. The global estimates of β_0 and β_2 (~ 11.28 and ~ 0.91 respectively) are influenced by their corresponding spatial random effects to better accommodate the non-stationary relationship between the covariates and AGB.

4.2. MEF Results

As noted previously, of the eight spatial models fit to the MEF dataset none out performed the *non-spatial* model with regard to prediction accuracy. Regarding fit to the observed data, only the *SVI* model showed an improvement in *DIC* over the *non-spatial* model (Table 2).

Results from MEF provided an example of how uncertainty maps can aid in identifying spatial extrapolation issues. Figures 4b and 5b show the *non-spatial* model AGB prediction and associated uncertainty for the extent of the LiDAR data. A large number of grid cells in the southwest corner of MEF were removed from the prediction dataset because their covariate values far exceeded those seen in the observed dataset (see Section 2.6). Also, grid cells in this area that were included in the prediction set had high uncertainties. An examination of MEF's aerial photo suggested that the forest in this region has different characteristics than the areas where plot data were collected (Figure 1b). In this instance, having uncertainty maps for MEF helped to diagnose a spatial extrapolation issue that may have otherwise gone unnoticed.

377 4.3. GLEES Results

378 The *SVI* model at GLEES had the lowest *RMSPE* value of the five models (Table 3); however,
 379 all three *SVC* variant models had lower *D* and *DIC* values. Here, the increased flexibility afforded
 380 by the additional spatial random effects in the *SVC* models improved the fit but may have reduced
 381 prediction accuracy due to over-fitting. The *SVI* model exhibited a 33% improvement in *DIC* and
 382 a 4.5% increase in prediction accuracy over the *non-spatial* model. Because prediction is our focus,
 383 the *SVI* model was preferred for the GLEES dataset (Figures 4c and 5c).

384 4.4. NIWOT Results

385 The *SVC*- β_2 model provided the most accurate prediction for the NIWOT dataset. This model
 386 produced a 24% lower *DIC* and 7% lower *RMSPE* value compared to the *non-spatial* model. As
 387 with GLEES, the more complex *SVC* model provided better fit (5% lower than *SVC*- β_2), but this
 388 did not translate into improved prediction accuracy.

389 NIWOT's *non-spatial* model β_2 parameter was statistically not significant, i.e., zero was con-
 390 tained in its 95% credible interval. In traditional regression model selection procedures, parameters
 391 deemed not significant would typically be removed. Here, even though β_2 was globally not signifi-
 392 cant, allowing the parameter to vary across the spatial domain improved model fit and predictive
 393 performance. It can be seen that some clusters of sub-plots received higher estimated values of
 394 $\beta_2(\mathbf{s})$ than others (Figure 10b). This result provided evidence that in some locations, the LiDAR
 395 covariate associated with β_2 was useful for explaining variability in AGB.

396 5. Discussion

397 The usefulness of Bayesian space-varying coefficient models for predicting AGB was assessed
 398 at four forested NACP sites. For three of the four sites, the *non-spatial* model residuals exhibited
 399 enough spatial correlation to warrant the use of a spatial random effect on the intercept. For these
 400 sites, the presence of serial correlations among the residuals suggested the assumptions of the *non*-
 401 *spatial* model were violated, making the validity of any subsequent inference based on parameter

estimates or prediction maps using on the *non-spatial* model questionable. Further, beyond meeting model assumptions, explicitly accommodating the residual dependence via spatially structured random effects fetched improved prediction by borrowing information from proximate observed locations. This was illustrated in the analysis of the FEF, GLEES, and NIWOT datasets, where the *SVI* models marginally outperformed the *non-spatial* models (i.e., providing lower *RMSPE*). A potential reason for seeing only slight prediction accuracy improvements may have been due to the small number of inventory plots. With such small sample sizes it is not possible to obtain precise estimates of spatial covariance parameters. Hence, we might expect some impact on prediction accuracy when the covariance uncertainty is propagated through to the posterior predictive distributions. Also, the sparsity of inventory plots means there is little information to borrow from when making predictions—this is exacerbated by cross-validation, which further reduces the sample size.

For the FEF and NIWOT datasets, allowing regression coefficients associated with LiDAR covariates to vary spatially improved prediction accuracy beyond that achieved by the *SVI* models (i.e., the *SVC* models showed lower *RMSPE*). Although the improvements in *RMSPE* were small, these results suggested the more complex *SVC* models were able to estimate and leverage non-stationary spatial relationships between AGB and the LiDAR covariates. There was no advantage, from a prediction standpoint, for *SVC* models at MEF or GLEES. Given the small number of inventory plots and relatively homogeneous forest characteristics within the NACP sites, it is not too surprising the *SVC* models did not perform consistently better than the simpler models.

In studies based on larger datasets distributed across more heterogeneous domains, substantial improvements in prediction accuracy have been obtained by adding spatial random effects to the model intercept, see, e.g., Babcock et al. (2013), and regression coefficients, see, e.g., Finley et al. (2009), Finley et al. (2011), Jarzyna et al. (2014). In these studies, the *SVC* models were effective because there were strong non-stationary relationships between the covariates and outcome variable—driven by unobserved environmental factors that operated at broad spatial scales. Also, substantially larger sample sizes ($n > 1,000$ in all cases) allowed more precise estimation of the spatial covariance parameters that, in turn, improved the posterior predictive distribution precision and perhaps accuracy.

Study results underscored the difference between goodness of fit metrics, e.g., DIC, and *out of set* prediction performance, assessed here via cross-validation. In three of the four datasets, the exception being FEF, DIC favored a more complex model over the simpler model selected based on RMSPE. The flexibility added by allowing the intercept and regression slope coefficients to vary spatially will generally allow fitted values to more closely approximate the observed data. Unlike an unstructured random effect model, see, e.g., Salas et al. (2010), the Gaussian Process imposes some smoothing that should reduce over-fitting to the observed data. However, the flexibility afforded by a univariate Gaussian Process on each regression coefficient appears to result in some over-fitting and hence reduced prediction performance. Because prediction is the primary goal of many AGB modeling studies it is important to base model selection on metrics that summarize out of sample predictions, e.g., RMSPE, and not goodness of fit.

Uncertainty quantification is an important component of carbon Measurement, Reporting, and Verification systems (MRVs) like those called for by the United Nations Collaborative Programme on Reducing Emissions from Deforestation and Forest Degradation in Developing Countries and actively being developed by NASA’s Carbon Monitoring System Science Team (UN-REDD, 2009; CMS, 2010). Maps like the ones developed in Figure 5 can aid uncertainty analysis by providing a tool to assess the reliability of AGB predictions. As shown with the FEF and MEF results, spatially explicit representations of uncertainty can be used to identify regions of concern regarding spatial extrapolation and highlight areas where more field data may be required.

Mapping space-varying coefficients can shed light on missing covariates. At FEF, examining how the $\beta_1(s)$ parameter varied spatially revealed that elevation may be a missing covariate in the AGB models. If further AGB model refinement was pursued at FEF, the inclusion of elevation as a covariate, either in an additive or multiplicative fashion, should be tested.

Bayesian space-varying coefficient models provide statically valid inference about non-stationary relationships between the covariates and the outcome variable. Results from the NIWOT analysis showed this phenomenon when β_2 was allowed to vary spatially. Examining regression parameters’ spatial impact via the proposed modeling framework can expose locally important covariates, improve overall prediction accuracy, and potentially provide insight into ecological processes.

458 Given the small size of the datasets considered here, the proposed modeling framework was com-
 459 putationally feasible. However, when the datasets consist of several thousand observations, which
 460 is common in forest inventory databases, cubic order matrix operations required for evaluating the
 461 model likelihood make parameter estimation computationally onerous. Therefore, our future work
 462 will focus on exploring strategies for dimension reduction when fitting spatially varying coefficients
 463 models. Further, we will extend the spatially varying coefficients models to accommodate multiple
 464 forest variables of interest.

465 References

466 References

- 467 Anderson, J., Plourde, L., Martin, M., Braswell, B., Smith, M.L., Dubayah, R., Hofton, M., & Blair,
 468 J. (2008). Integrating waveform lidar with hyperspectral imagery for inventory of a northern
 469 temperate forest. *Remote Sensing of Environment*, 112, 1856 – 1870.
- 470 Babcock, C., Matney, J., Finley, A.O., Weiskittel, A., & Cook, B.D. (2013). Multivariate spatial
 471 regression models for predicting individual tree structure variables using lidar data. *Selected*
 472 *Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 6, 6–14.
- 473 Banerjee, S., Carlin, C., & Gelfand, A. (2004). Hierarchical Modeling and Analysis for Spatial Data.
 474 Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Chapman & Hall/CRC.
- 475 Banerjee, S., Finley, A., Waldmann, P., & Ericsson, T. (2011). Hierarchical spatial process models
 476 for multiple traits in large genetic trials. *Journal of the American Statistical Association*, 105,
 477 506 – 521.
- 478 Banerjee, S., Gelfand, A., Finley, A., & Huiyan, S. (2008). Gaussian predictive process models for
 479 large spatial data sets. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*,
 480 70, 825 – 848.

481 Bechtold, W., & Patterson, P. (2005). The enhanced forest inventory and analysis program–national
482 sampling design and estimation procedures. General technical report SRS, USDA Forest Service,
483 Southern Research Station.

484 Bradford, J., Birdsey, R., Joyce, L., & Ryan, M. (2008). Tree age, disturbance history, and carbon
485 stocks and fluxes in subalpine rocky mountain forests. *Global Change Biology*, 14, 2882 – 2897.

486 Bradford, J., Weishampel, P., Smith, M.L., Kolka, R., Birdsey, R., Ollinger, S., & Ryan, M. (2010).
487 Carbon pools and fluxes in small temperate forest landscapes: Variability and implications for
488 sampling design. *Forest Ecology and Management*, 259, 1245 – 1254.

489 Carlin, B., & Louis, T. (2000). Bayes and Empirical Bayes Methods for Data Analysis, Second
490 Edition. Chapman & Hall/CRC Texts in Statistical Science, Taylor & Francis.

491 Chatterjee, S., & Hadi, A. (2006). Regression Analysis by Example, 4th Edition. Wiley Series in
492 Probability and Statistics, John Wiley & Sons.

493 Chen, G., Zhao, K., McDermid, G.J., & Hay, G.J. (2012). The influence of sampling density on
494 geographically weighted regression: a case study using forest canopy height and optical data.
495 *International Journal of Remote Sensing*, 33, 2909–2924.

496 Chilès, J., & Delfiner, P. (1999). Geostatistics: modeling spatial uncertainty. Wiley series in
497 probability and statistics, John Wiley & Sons.

498 CMS (2010). Nasa carbon monitoring system. <http://carbon.nasa.gov>. Accessed: 3-18-2014.

499 Cressie, N. (1993). Statistics for spatial data. Wiley series in probability and mathematical statistics:
500 Applied probability and statistics, John Wiley & Sons.

501 Finley, A. (2010). Comparing spatially-varying coefficients models for analysis of ecological data
502 with non-stationary and anisotropic residual dependence. *Methods in Ecology and Evolution*, 2,
503 143 – 154.

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

507 Finley, A., Banerjee, S., & McRoberts, R. (2009). Hierarchical spatial models for predicting tree
508 species assemblages across large domains. *Annals of Applied Statistics*, 3, 1052 – 1079.

509 Fotheringham, A.S., Brunsdon, C., & Charlton, M. (2003). Geographically weighted regression:
510 the analysis of spatially varying relationships. John Wiley & Sons.

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

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

515 Gelman, A., & Rubin, D. (1992). Inference from iterative simulation using multiple sequences.
516 *Statistical Science*, 7, 457–511.

517 Gonzalez, P., Asner, G., Battles, J., Lefsky, M., Waring, K., & Palace, M. (2010). Forest carbon
518 densities and uncertainties from lidar, quickbird, and field measurements in california. *Remote*
519 *Sensing of Environment*, 114, 1561–1575.

520 Guanter, L., Frankenberg, C., Dudhia, A., Lewis, P., Gómez-Dans, J., Kuze, A., Suto, H., &
521 Grainger, R. (2012). Retrieval and global assessment of terrestrial chlorophyll fluorescence from
522 gosat space measurements. *Remote Sensing of Environment*, 121, 236–251.

523 Guanter, L., Rossini, M., Colombo, R., Meroni, M., Frankenberg, C., Lee, J.E., & Joiner, J. (2013).
524 Using field spectroscopy to assess the potential of statistical approaches for the retrieval of sun-
525 induced chlorophyll fluorescence from ground and space. *Remote Sensing of Environment*, 133,
526 52–61.

527 Harville, D. (1997). Matrix Algebra from a Statistician’s Perspective. Springer.

528 Hoeting, J. (2009). The importance of accounting for spatial and temporal correlation in analyses
529 of ecological data. *Ecological Applications*, 19, 574–577.

530 Huang, B., Song, H., Cui, H., Peng, J., & Xu, Z. (2014). Spatial and spectral image fusion using
531 space matrix factorization. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 1693–
532 1704.

533 Hudak, A.T., Lefsky, M.A., Cohen, W.B., & Berterretche, M. (2002). Integration of lidar and land-
534 sat etm+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment*,
535 82, 397–416.

536 Jarzyna, M.A., Finley, A.O., Porter, W.F., Maurer, B.A., Beier, C.M., & Zuckerberg, B. (2014). Ac-
537 counting for the space-varying nature of the relationships between temporal community turnover
538 and the environment. *Ecography*, .

539 Johnson, R.A., & Wichern, D.W. (2007). Applied multivariate statistical analysis. volume 6. Pear-
540 son Prentice Hall.

541 Lim, K., & Treitz, P. (2004). Estimation of above ground forest biomass from airborne discrete
542 return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest*
543 *Research*, 19, 558–570.

544 Lucas, R., Cronin, N., Lee, A., Moghaddam, A., Witte, C., & Tickle, P. (2006). Empirical relation-
545 ships between airsar backscatter and lidar-derived forest biomass, queensland, australia. *Remote*
546 *Sensing of Environment*, 100, 407–425.

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

549 Perrin, E. (1904). On some dangers of extrapolation. *Biometrika*, 3, 99–103.

550 Popescu, S., Zhao, K., Neuenschwander, A., & Lin, C. (2011). Satellite lidar vs. small footprint
551 airborne lidar: comparing the accuracy of aboveground biomass estimates and forest structure
552 metrics at footprint level. *Remote Sensing of Environment*, 115, 2786–2797.

- 553 Richards, J.A. (2013). Remote sensing digital image analysis. volume 5. Springer.
- 554 Salas, C., Ene, L., Gregoirea, T.G., Næssetb, E., & Gobakkenb, T. (2010). Modelling tree diameter
555 from airborne laser scanning derived variables: A comparison of spatial statistical models. *Remote*
556 *Sensing of Environment*, 114, 1277–1285.
- 557 Sherrill, K., Lefsky, M., Bradford, J., & Ryan, M. (2008). Forest structure estimation and pattern
558 exploration from discrete-return lidar in subalpine forests of the central rockies. *Canadian Journal*
559 *of Forest Research*, 38, 2081–2096.
- 560 Spiegelhalter, D., Best, N., Carlin, B., & van der Linde, A. (2002). Bayesian measures of model
561 complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B*, 64, 583–
562 639.
- 563 Tonolli, S., Dalponte, M., Neteler, M., Rodeghiero, M., Vescovo, L., & Gianelle, D. (2011). Fusion of
564 airborne lidar and satellite multispectral data for the estimation of timber volume in the southern
565 alps. *Remote Sensing of Environment*, 115, 2486–2498.
- 566 UN-REDD (2009). The un-redd programme. <http://www.un-redd.org/>. Accessed: 3-18-2014.
- 567 Wackernagel, H. (2003). Multivariate geostatistics: an introduction with applications, 3rd edition.
568 Springer.
- 569 Wang, C., Yu, W., Wang, R., Deng, Y., & Zhao, F. (2014). Comparison of nonnegative eigen-
570 value decompositions with and without reflection symmetry assumptions. *IEEE Transactions on*
571 *Geoscience and Remote Sensing*, 52, 2278–2287.
- 572 Wheeler, D., & Calder, C. (2007). An assessment of coefficient accuracy in linear regression models
573 with spatially varying coefficients. *Journal of Geographical Systems*, 9, 145–166.
- 574 Wheeler, D., & Waller, L. (2009). Comparing spatially varying coefficient models: a case study
575 examining violent crime rates and their relationships to alcohol outlets and illegal drug arrests.
576 *Journal of Geographical Systems*, 11, 1–22.

Figure captions

Figure 1: Aerial photographs of FEF, MEF, GLEES, and NIWOT with inventory sub-plot locations highlighted in red.

Figure 2: Posterior predicted distribution at an example grid cell location shown in gray. Blue vertical dashed lines indicate the 2.5% and 97.5% quantiles of the posterior predicted distribution, The length of the horizontal red line is the 95% credible interval width (CIW). CIWs for each grid cell are mapped in Figure 5.

Figure 3: Semivariogram models generated using the residuals from the *non-spatial* model estimates for FEF, MEF, GLEES, and NIWOT.

Figure 4: Maps showing predicted AGB means (Mg/ha) for the full LiDAR extent at FEF, MEF, GLEES, and NIWOT. Black squares indicate excluded prediction grid cells (see Section 2.6).

Figure 5: Maps showing predicted AGB 95% credible interval width (Mg/ha) at FEF, MEF, GLEES, and NIWOT. Black squares indicate excluded prediction grid cells (see Section 2.6).

Figure 6: Maps showing the spatially varying coefficients and associated 95% credible interval width (CIW) for the *SVC* model at FEF. Black squares indicate excluded prediction grid cells (see Section 2.6).

Figure 7: Digital terrain model for FEF.

Figure 8: Maps showing the spatially varying coefficients zoomed into the extent outlined in red in Figures 6a and 6c. A quantile interval color classification and smoothing filter are used to highlight the more subtle spatial adjustments of these parameters. The sub-plot locations are identified in green. Black squares indicate excluded prediction grid cells (see Section 2.6).

Figure 9: Maps showing the spatially varying coefficients and associated 95% credible interval width (CIW) for the *SVC- β_2* model at NIWOT. Black squares indicate excluded prediction grid cells (see Section 2.6).

Figure 10: Maps showing the spatially varying coefficients zoomed into the extent outlined in red in Figures 9a and 9c. A quantile interval color classification and smoothing filter are used

604 to highlight the more subtle spatial adjustments of these parameters. The sub-plot locations
605 are identified in green. Black squares indicate excluded prediction grid cells (see Section 2.6).
606 Figure A1: The cumulative sums of the standardized eigenvalues (in decreasing order) for
607 FEF, MEF, GLEES, and NIWOT.

608 Figure A2: Plots showing fitted versus field-observed AGB values for FEF's *non-spatial* and
609 *SVC* models. Vertical Gray error bars depict 95% credible intervals for fitted values.

610 Figure A3: Plots showing fitted versus field-observed AGB values for MEF's *non-spatial* and
611 *SVI* models. Vertical gray error bars depict 95% credible intervals for fitted values.

612 Figure A4: Plots showing fitted versus field-observed AGB values for GLEES's *non-spatial*
613 and *SVI* models. Vertical gray error bars depict 95% credible intervals for fitted values.

614 Figure A5: Plots showing fitted versus field-observed AGB values for NIWOT's *non-spatial*
615 and *SVC- β_2* models. Vertical gray error bars depict 95% credible intervals for fitted values.

Table 1: Candidate model parameter estimates and 95% credible intervals for FEF.

		Non-Spatial		SVI		SVC- β_1		SVC- β_2		SVC	
parameter C.I. 50% (2.5%, 97.5%)	β_0	11.16	(10.49, 11.80)	11.33	(10.44, 12.27)	11.58	(10.88, 12.30)	11.41	(10.69, 12.05)	11.28	(10.69, 11.90)
	β_1	-1.08	(-1.25, -0.92)	-1.06	(-1.23, -0.88)	-1.04	(-1.91, 0.17)	-0.97	(-1.16, -0.78)	-0.81	(-1.76, 0.12)
	β_2	1.10	(0.70, 1.51)	1.01	(0.57, 1.44)	0.77	(0.35, 1.22)	1.46	(0.63, 3.51)	0.91	(0.35, 1.41)
	τ^2	6.49	(4.62, 9.58)	0.89	(0.32, 3.70)	3.38	(1.61, 5.70)	2.36	(0.79, 4.83)	0.07	(0.02, 0.75)
	$3/\phi_0$	—		121.45	(20.05, 284.91)	21.14	(10.29, 241.36)	31.29	(10.37, 1612.90)	20.62	(10.27, 156.90)
	$3/\phi_1$	—		—		2380.95	(160.09, 4838.71)	—		2400.00	(302.72, 4918.03)
	$3/\phi_2$	—		—		—		194.74	(10.82, 2542.37)	21.63	(10.25, 185.87)
	σ_0^2	—		5.18	(2.51, 8.71)	0.48	(0.21, 1.73)	0.53	(0.21, 2.19)	0.48	(0.20, 1.19)
	σ_1^2	—		—		0.66	(0.29, 1.91)	—		0.74	(0.33, 1.95)
	σ_2^2	—		—		—		2.28	(0.80, 7.99)	1.64	(0.89, 3.09)
fit statistics	D	293.21		93.13		248.42		224.71		22.43	
	pD	3.93		38.24		24.17		29.57		25.71	
	DIC	297.14		131.37		272.59		254.28		48.14	
	RMSPE	71.70		69.44		71.16		74.88		65.27	

Table 2: Candidate model parameter estimates and 95% credible intervals for MEF.

	Non-Spatial		Spatial	SVC- β_1		SVC- β_2		SVC- β_3		SVC- $\beta_1\beta_2$		SVC- $\beta_1\beta_3$		SVC- $\beta_2\beta_3$		SVC	
	β_0	9.10 (8.45, 9.73)	9.08 (8.43, 9.73)	9.00 (8.30, 9.69)	9.12 (8.44, 9.80)	9.09 (8.45, 9.73)	9.01 (8.27, 9.75)	9.01 (8.27, 9.75)	8.98 (8.27, 9.69)	9.09 (8.39, 9.79)	9.00 (8.39, 9.79)	8.98 (8.27, 9.69)	9.09 (8.39, 9.79)	9.00 (8.39, 9.79)	8.98 (8.27, 9.69)	9.00 (8.39, 9.79)	8.98 (8.27, 9.69)
50% parameter C.I.	β_1	-1.29 (-1.49, -1.09)	-1.29 (-1.48, -1.09)	-1.23 (-1.50, -0.93)	-1.28 (-1.51, -1.06)	-1.28 (-1.49, -1.06)	-1.22 (-1.51, -0.94)	-1.22 (-1.51, -0.94)	-1.22 (-1.54, -0.84)	-1.27 (-1.51, -1.03)	-1.27 (-1.51, -1.03)	-1.22 (-1.54, -0.84)	-1.27 (-1.51, -1.03)	-1.27 (-1.51, -1.03)	-1.22 (-1.54, -0.84)	-1.21 (-1.52, -0.94)	-1.21 (-1.52, -0.94)
	β_2	0.89 (0.61, 1.15)	0.88 (0.61, 1.15)	0.90 (0.57, 1.23)	0.83 (0.35, 1.29)	0.86 (0.53, 1.18)	0.84 (0.30, 1.38)	0.84 (0.30, 1.38)	0.88 (0.48, 1.25)	0.82 (0.33, 1.30)	0.83 (0.33, 1.30)	0.88 (0.48, 1.25)	0.82 (0.33, 1.30)	0.82 (0.33, 1.30)	0.88 (0.48, 1.25)	0.83 (0.27, 1.38)	0.83 (0.27, 1.38)
	β_3	-1.20 (-1.67, -0.74)	-1.18 (-1.64, -0.71)	-1.26 (-1.84, -0.70)	-1.25 (-1.83, -0.68)	-1.21 (-1.85, -0.60)	-1.34 (-2.03, -0.64)	-1.34 (-2.03, -0.64)	-1.29 (-2.00, -0.60)	-1.29 (-2.00, -0.57)	-1.29 (-2.00, -0.57)	-1.29 (-2.00, -0.60)	-1.29 (-2.00, -0.57)	-1.29 (-2.00, -0.57)	-1.38 (-2.18, -0.60)	-1.38 (-2.18, -0.60)	-1.38 (-2.18, -0.60)
	τ^2	11.56 (8.99, 15.36)	4.98 (1.24, 11.25)	10.51 (7.75, 14.37)	10.15 (7.55, 13.75)	10.30 (7.67, 13.93)	9.52 (6.67, 13.44)	9.52 (6.67, 13.44)	9.27 (6.92, 13.59)	9.85 (7.16, 13.58)	9.85 (7.16, 13.58)	9.85 (7.16, 13.58)	9.85 (7.16, 13.58)	9.85 (7.16, 13.58)	9.10 (6.19, 13.07)	9.10 (6.19, 13.07)	9.10 (6.19, 13.07)
	$3/\phi_0$	—	29.50 (12.04, 192.98)	20.33 (10.26, 284.90)	21.51 (10.28, 607.29)	22.44 (10.29, 928.79)	20.30 (10.29, 332.59)	20.30 (10.29, 332.59)	21.23 (10.29, 417.25)	21.02 (10.29, 572.52)	21.02 (10.29, 572.52)	20.96 (10.27, 466.56)	20.96 (10.27, 466.56)	20.96 (10.27, 466.56)	20.96 (10.27, 466.56)	20.96 (10.27, 466.56)	20.96 (10.27, 466.56)
	$3/\phi_1$	—	—	24.52 (10.32, 1612.90)	—	—	25.30 (10.32, 1363.64)	25.30 (10.32, 1363.64)	32.31 (10.36, 2439.02)	—	—	25.08 (10.31, 594.06)	25.08 (10.31, 594.06)	25.08 (10.31, 594.06)	25.08 (10.31, 594.06)	25.08 (10.31, 594.06)	25.08 (10.31, 594.06)
	$3/\phi_2$	—	—	—	22.33 (10.31, 699.30)	—	21.35 (10.29, 518.13)	21.35 (10.29, 518.13)	—	—	—	22.06 (10.32, 584.80)	22.06 (10.32, 584.80)	22.06 (10.32, 584.80)	22.29 (10.32, 390.12)	22.29 (10.32, 390.12)	22.29 (10.32, 390.12)
	$3/\phi_3$	—	—	—	—	—	19.52 (10.25, 276.50)	—	—	—	—	19.59 (10.26, 386.25)	19.59 (10.26, 386.25)	19.59 (10.26, 386.25)	19.76 (10.27, 250.63)	19.76 (10.27, 250.63)	19.76 (10.27, 250.63)
	σ_6^2	—	5.98 (1.31, 11.37)	0.36 (0.17, 1.02)	0.36 (0.17, 0.90)	0.36 (0.17, 1.01)	0.36 (0.17, 1.01)	0.36 (0.17, 1.01)	0.36 (0.17, 0.96)	0.37 (0.17, 1.05)	0.37 (0.17, 1.05)	0.36 (0.17, 0.92)	0.36 (0.17, 1.05)	0.36 (0.17, 1.05)	0.36 (0.17, 0.92)	0.36 (0.17, 0.92)	0.36 (0.17, 0.92)
	σ_1^2	—	—	0.36 (0.19, 0.69)	—	—	—	0.39 (0.21, 0.80)	0.39 (0.21, 0.80)	0.39 (0.20, 0.80)	0.39 (0.20, 0.80)	—	—	—	0.40 (0.21, 0.79)	0.40 (0.21, 0.79)	0.40 (0.21, 0.79)
	σ_2^2	—	—	—	0.65 (0.30, 1.52)	—	—	0.70 (0.32, 1.71)	0.70 (0.32, 1.71)	—	—	0.64 (0.28, 1.52)	0.64 (0.28, 1.52)	0.64 (0.28, 1.52)	0.70 (0.31, 1.78)	0.70 (0.31, 1.78)	0.70 (0.31, 1.78)
	σ_3^2	—	—	—	—	—	0.99 (0.37, 2.75)	—	—	—	—	1.08 (0.41, 3.23)	0.94 (0.38, 2.77)	0.94 (0.38, 2.77)	1.07 (0.41, 3.10)	1.07 (0.41, 3.10)	1.07 (0.41, 3.10)
fit statistics	D	608.41	501.53	600.15	595.76	597.53	588.21	588.21	588.05	588.05	588.05	588.05	588.05	588.05	588.05	588.05	588.05
	pD	4.95	41.29	28.61	22.85	19.05	41.54	41.54	47.55	47.55	47.55	47.55	47.55	47.55	47.55	47.55	47.55
	DIC	613.37	542.82	628.77	618.61	616.59	628.74	628.74	630.60	630.60	630.60	630.60	630.60	630.60	630.60	630.60	630.60
	RMSPE	66.89	67.46	67.53	70.81	68.23	72.41	72.41	75.28	75.28	75.28	75.28	75.28	75.28	75.28	75.28	75.28

Table 3: Candidate model parameter estimates and 95% credible intervals for GLEES.

		Non-Spatial		SVI		SVC- β_1		SVC- β_2		SVC	
parameter C.I. 50% (2.5%, 97.5%)	β_0	8.27	(7.66, 8.86)	8.40	(7.76, 9.14)	8.31	(7.61, 9.03)	7.94	(7.43, 8.51)	8.06	(7.38, 8.74)
	β_1	-0.62	(-0.79, -0.44)	-0.63	(-0.81, -0.45)	-0.61	(-0.89, -0.31)	-0.61	(-0.77, -0.46)	-0.62	(-0.96, -0.34)
	β_2	0.66	(0.37, 0.96)	0.67	(0.37, 0.99)	0.86	(0.49, 1.23)	0.74	(0.14, 1.29)	0.87	(0.33, 1.45)
	τ^2	4.10	(2.75, 6.63)	0.53	(0.19, 2.01)	0.40	(0.10, 1.67)	0.32	(0.10, 1.22)	0.25	(0.08, 0.95)
	$3/\phi_0$	—		44.11	(12.90, 173.16)	22.85	(10.31, 125.63)	20.96	(10.27, 151.98)	20.19	(10.27, 198.41)
	$3/\phi_1$	—		—		33.88	(10.46, 527.24)	—		29.80	(10.37, 1020.50)
	$3/\phi_2$	—		—		—		58.09	(10.56, 552.49)	28.21	(10.39, 643.78)
	σ_0^2	—		3.45	(1.95, 5.88)	0.98	(0.35, 2.66)	0.71	(0.29, 1.77)	0.53	(0.24, 1.33)
	σ_1^2	—		—		0.44	(0.24, 0.88)	—		0.39	(0.21, 0.85)
	σ_2^2	—		—		—		1.11	(0.54, 2.48)	0.78	(0.35, 1.90)
fit statistics	D	196.30		103.69		90.35		79.59		67.79	
	pD	3.90		29.79		27.58		30.25		33.74	
	DIC	200.20		133.48		117.93		109.85		101.53	
	RMSPE	35.79		34.21		35.63		35.02		35.20	

Table 4: Candidate model parameter estimates and 95% credible intervals for NIWOT.

		Non-Spatial		SVI		SVC- β_1		SVC- β_2		SVC	
parameter C.I. 50% (2.5%, 97.5%)	β_0	12.34	(11.83, 12.85)	12.35	(11.38, 13.18)	12.35	(11.42, 13.01)	12.19	(11.38, 12.85)	12.21	(11.47, 13.06)
	β_1	-0.82	(-0.95, -0.68)	-0.77	(-0.93, -0.60)	-0.84	(-1.17, -0.17)	-0.77	(-0.93, -0.59)	-0.84	(-1.22, -0.29)
	β_2	0.21	(-0.04, 0.49)	0.21	(-0.05, 0.48)	0.31	(-0.01, 0.62)	0.23	(-0.52, 0.78)	0.34	(-0.17, 0.82)
	τ^2	4.01	(2.82, 5.98)	1.29	(0.36, 3.74)	1.09	(0.25, 2.76)	0.99	(0.21, 3.14)	0.66	(0.20, 2.10)
	$3/\phi_0$	—		116.58	(13.64, 1444.64)	60.19	(10.52, 1578.95)	137.24	(11.40, 1167.32)	42.98	(10.43, 1401.87)
	$3/\phi_1$	—		—		50.58	(10.40, 2097.90)	—		77.34	(10.46, 2419.35)
	$3/\phi_2$	—		—		—		45.68	(10.45, 2343.75)	22.29	(10.30, 669.64)
	σ_0^2	—		2.66	(0.78, 5.07)	1.19	(0.34, 3.21)	1.74	(0.41, 3.87)	0.87	(0.30, 2.51)
	σ_1^2	—		—		0.37	(0.20, 1.03)	—		0.38	(0.20, 1.00)
	σ_2^2	—		—		—		0.74	(0.32, 2.03)	0.78	(0.33, 1.90)
fit statistics	D	262.40		189.65		177.34		172.77		150.60	
	pD	3.93		27.65		34.78		28.97		41.52	
	DIC	266.32		217.29		212.12		201.74		192.12	
	RMSPE	50.05		48.27		49.31		46.31		49.40

(a) FEF

(b) MEF

(c) GLEES

(d) NIWOT

Figure 1: Aerial photographs of FEF, MEF, GLEES, and NIWOT with inventory sub-plot locations highlighted in red.

Figure 2: Posterior predicted distribution at an example grid cell location shown in gray. Blue vertical dashed lines indicate the 2.5% and 97.5% quantiles of the posterior predicted distribution, The length of the horizontal red line is the 95% credible interval width (CIW). CIWs for each grid cell are mapped in Figure 5.

(a) FEF

(b) MEF

(c) GLEES

(d) NIWOT

Figure 3: Semivariogram models generated using the residuals from the *non-spatial* model estimates for FEF, MEF, GLEES, and NIWOT.

(a) FEF

(b) MEF

(c) GLEES

(d) NIWOT

Figure 4: Maps showing predicted AGB means (Mg/ha) for the full LiDAR extent at FEF, MEF, GLEES, and NIWOT. Black squares indicate excluded prediction grid cells (see Section 2.6).

(a) FEF

(b) MEF

(c) GLEES

(d) NIWOT

Figure 5: Maps showing predicted AGB 95% credible interval width (Mg/ha) at FEF, MEF, GLEES, and NIWOT. Black squares indicate excluded prediction grid cells (see Section 2.6).

Figure 6: Maps showing the spatially varying coefficients and associated 95% credible interval width (CIW) for the *SVC* model at FEF. Black squares indicate excluded prediction grid cells (see Section 2.6).

Figure 7: Digital terrain model for FEF.

Figure 8: Maps showing the spatially varying coefficients zoomed into the extent outlined in red in Figures 6a and 6c. A quantile interval color classification and smoothing filter are used to highlight the more subtle spatial adjustments of these parameters. The sub-plot locations are identified in green. Black squares indicate excluded prediction grid cells (see Section 2.6).

(a) $\beta_0(\mathbf{s})$

(b) $\beta_0(\mathbf{s})$ 95% CIW

(c) $\beta_2(\mathbf{s})$

(d) $\beta_2(\mathbf{s})$ 95% CIW

Figure 9: Maps showing the spatially varying coefficients and associated 95% credible interval width (CIW) for the *SVC*- β_2 model at NIWOT. Black squares indicate excluded prediction grid cells (see Section 2.6).

Figure 10: Maps showing the spatially varying coefficients zoomed into the extent outlined in red in Figures 9a and 9c. A quantile interval color classification and smoothing filter are used to highlight the more subtle spatial adjustments of these parameters. The sub-plot locations are identified in green. Black squares indicate excluded prediction grid cells (see Section 2.6).