

2017

Joint hierarchical models for sparsely sampled high-dimensional LiDAR and forest variables

Andrew O. Finley
Michigan State University

Sudipto Banerjee
University of California, Los Angeles

Yuzhen Zhou
University of Nebraska - Lincoln, yuzhenzhou@unl.edu

Bruce D. Cook
NASA/Goddard Space Flight Center

Chad Babcock
University of Washington

Follow this and additional works at: <http://digitalcommons.unl.edu/nasapub>

 Part of the [Applied Mathematics Commons](#), [Earth Sciences Commons](#), [Instrumentation Commons](#), and the [Oceanography and Atmospheric Sciences and Meteorology Commons](#)

Finley, Andrew O.; Banerjee, Sudipto; Zhou, Yuzhen; Cook, Bruce D.; and Babcock, Chad, "Joint hierarchical models for sparsely sampled high-dimensional LiDAR and forest variables" (2017). *NASA Publications*. 235.
<http://digitalcommons.unl.edu/nasapub/235>

This Article is brought to you for free and open access by the National Aeronautics and Space Administration at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in NASA Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



Joint hierarchical models for sparsely sampled high-dimensional LiDAR and forest variables



Andrew O. Finley^{a,*}, Sudipto Banerjee^b, Yuzhen Zhou^c, Bruce D. Cook^d, Chad Babcock^e

^aDepartment of Forestry, Michigan State University, East Lansing, MI 48824, United States

^bDepartment of Biostatistics, University of California, Los Angeles, CA, 90095-1772, United States

^cDepartment of Statistics, University of Nebraska-Lincoln, Lincoln, NE, 68583-0963, United States

^dBiospheric Sciences Branch, NASA/Goddard Space Flight Center, Greenbelt, MD 20742, United States

^eSchool of Environmental and Forest Sciences, University of Washington, Seattle, WA 98195, United States

ARTICLE INFO

Article history:

Received 27 June 2016

Received in revised form 22 October 2016

Accepted 5 December 2016

Available online 26 December 2016

Keywords:

Forest biomass

Uncertainty quantification

Functional analysis

Dimension reduction

Predictive process

Hierarchical models

Markov chain Monte Carlo

ABSTRACT

Recent advancements in remote sensing technology, specifically Light Detection and Ranging (LiDAR) sensors, provide the data needed to quantify forest characteristics at a fine spatial resolution over large geographic domains. From an inferential standpoint, there is interest in prediction and interpolation of the often sparsely sampled and spatially misaligned LiDAR signals and forest variables. We propose a fully process-based Bayesian hierarchical model for above ground biomass (AGB) and LiDAR signals. The process-based framework offers richness in inferential capabilities, e.g., inference on the entire underlying processes instead of estimates only at pre-specified points. Key challenges we obviate include misalignment between the AGB observations and LiDAR signals and the high-dimensionality in the model emerging from LiDAR signals in conjunction with the large number of spatial locations. We offer simulation experiments to evaluate our proposed models and also apply them to a challenging dataset comprising LiDAR and spatially coinciding forest inventory variables collected on the Penobscot Experimental Forest (PEF), Maine. Our key substantive contributions include AGB data products with associated measures of uncertainty for the PEF and, more broadly, a methodology that should find use in a variety of current and upcoming forest variable mapping efforts using sparsely sampled remotely sensed high-dimensional data.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Coupling forest inventory with remotely sensed Light Detection and Ranging (LiDAR) datasets using regression models offers an attractive approach to mapping forest variables at stand, regional, continental, and global scales. LiDAR data have shown great potential for use in estimating spatially explicit forest variables over a range of geographic scales (Asner et al., 2009; Babcock et al., 2013; 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 and sensor platforms, as well as extensive campaigns to collect field-based calibration data.

Much of the interest in LiDAR based forest variable 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) (CMS, 2010; Le Toan et al., 2011; Ometto et al., 2014; UN-REDD, 2009). In these, and similar initiatives, AGB is the forest variable of interest because it provides a nearly direct measure of forest carbon (i.e., carbon comprises ~50% of wood biomass, West, 2004). Most efforts to quantify and/or manage forest ecosystem services, e.g., carbon, biodiversity, and water, seek high spatial resolution wall-to-wall data products such as gridded maps with associated measures of uncertainty, e.g., point and associated credible intervals (CIs) at the pixel level. In fact several high profile international initiatives include language concerning the level of spatially explicit acceptable error in total forest carbon estimates, see, e.g., UN-REDD (2009) and UNFCCC (2015).

Many current LiDAR data acquisition campaigns focus on achieving complete coverage at a high spatial resolution over the domain of interest, e.g., resulting in a fine grid with each pixel yielding a high-dimensional LiDAR signal. In practice, a variety of non-statistical approaches are then used to characterize the LiDAR signals—effectively a dimension reduction step, Anderson et al. (2008), Gonzalez et al. (2010), Muss et al. (2011), Tonolli et al. (2011), Popescu and Zhao (2008), and Babcock et al. (2013). These signal characteristics serve as regressors in models where the outcome

* Corresponding author at: Department of Forestry, Michigan State University, East Lansing, MI 48824, United States.

forest variables are measured at a relatively small set of georeferenced forest inventory plots. The regression model is then used to predict the forest outcome variables at all LiDAR pixels across the domain. This approach works well for small-scale forest variable mapping efforts. However, next generation LiDAR acquisition campaigns aimed at mapping and quantifying variables over large spatial extents, such as ICESat-2 (Abdalati et al., 2010; ICESat-2, 2015), Global Ecosystem Dynamics Investigation LiDAR (GEDI) (GEDI, 2014), and NASA Goddard's LiDAR, Hyper-spectral, and Thermal (G-LiHT) imager (Cook et al., 2013; Stockton, 2014), will collect LiDAR data *samples* from the domain of interest, e.g., using transect or cluster designs. The designs specify point-referenced LiDAR sampling across the domain extent and also over forest inventory plot locations (again for regression model calibration). In such settings the primary objective is still delivery of high resolution wall-to-wall predictive maps of forest variables, but also corresponding maps of LiDAR signal predictions at non-sampled locations. Further, to inform future LiDAR collection sampling designs, there is interest in characterizing the spatial dependence of within and, more importantly, among LiDAR signals. This information can help guide LiDAR sampling strategies with the aim to maximize some information gain criterion; see, e.g., Xia et al. (2006), and Mateu and Müller (2012).

We propose a flexible framework to jointly model spatially misaligned LiDAR signals and forest inventory plot outcomes (e.g., AGB) that will *i*) automatically (i.e., no explicit variable selection step) extract information from the high-dimensional LiDAR signals to explain variability in the forest variable of interest, *ii*) estimate spatial dependence among and within LiDAR signals to improve inference and possibility help inform future LiDAR sampling strategies, and *iii*) provide full posterior predictive inference for both LiDAR signals and forest variables at locations where either one or neither of the data sources are available (i.e., wall-to-wall prediction).

Meeting these objectives is particularly challenging for several reasons. From a computational standpoint each LiDAR signal is high-dimensional and the signals as well as the forest inventory plots are observed at a potentially large number of locations. From a model specification standpoint there are several sources of dependence that should be accommodated, including *i*) within and between LiDAR signals, *ii*) between LiDAR signals and spatially proximate forest variable measurements, and *iii*) residual spatial dependence in the signals and forest variables. These dependencies often result from strong vertical and horizontal similarities in forest structure caused by past management and/or natural disturbances.

Our primary methodological contribution is the development of a modeling framework for high-dimensional misaligned data. Given the rich inference we seek (see preceding paragraph), our Bayesian hierarchical framework jointly models LiDAR signals and forest variables as a random process using latent Gaussian processes (GPs). This considerably enhances the computational burden of fitting them to datasets with a large number of spatial locations. The costs are exacerbated further by even a modest number of heights at which the LiDAR signal is observed. We achieve dimension reduction through bias-adjusted reduced-rank representations of the joint LiDAR-AGB process.

The manuscript is organized as follows. Section 2 provides an overview of the motivating dataset that comprises G-LiHT LiDAR and AGB measured at forest inventory plots on the Penobscot Experimental Forest (PEF) in Bradley, Maine. Section 3 describes the proposed hierarchical model for the joint LiDAR-AGB process. The details on Bayesian prediction and implementation are given in the Supplemental material. Section 4 offers an analysis of a synthetic dataset and PEF analysis. Finally, Section 5 concludes the manuscript with a brief summary and pointers toward future work.

2. Data

The PEF is a 1600 ha tract of Acadian forest located in Bradley, Maine (44° 52' N, 68° 38' W). The forest is divided into over 50 management units (MU)—delineated as black polygons in Fig. 1 (a)—that received management and monitoring since the 1950s (Sendak et al., 2003). Within each MU, different silvicultural treatments are implemented, e.g., unregulated harvest, shelterwood, diameter limit cutting, or natural regeneration. Following procedures described in Finley et al. (2014), AGB (Mg/ha) was calculated for each of 451 permanent sample plots (PSPs) across the PEF, shown as point symbols in Fig. 1 (a). The underlying surface in Fig. 1 (a) was generated by passing the point-referenced AGB through a deterministic surface interpolator. Due to MU specific harvesting and subsequent regrowth cycles, the surface exhibits patterns of spatial dependence with relatively strong homogeneity within MUs. For example, MU U7B—highlighted in Fig. 1 (a)—received a shelterwood harvest in 1978 with a final overstory harvest in 2003. This silvicultural treatment results in a MU with relatively young trees and even-aged composition with low AGB (indicated by a lighter surface color in Fig. 1 (a)). In contrast to U7B, C12 is characterized by older and larger trees, but also greater

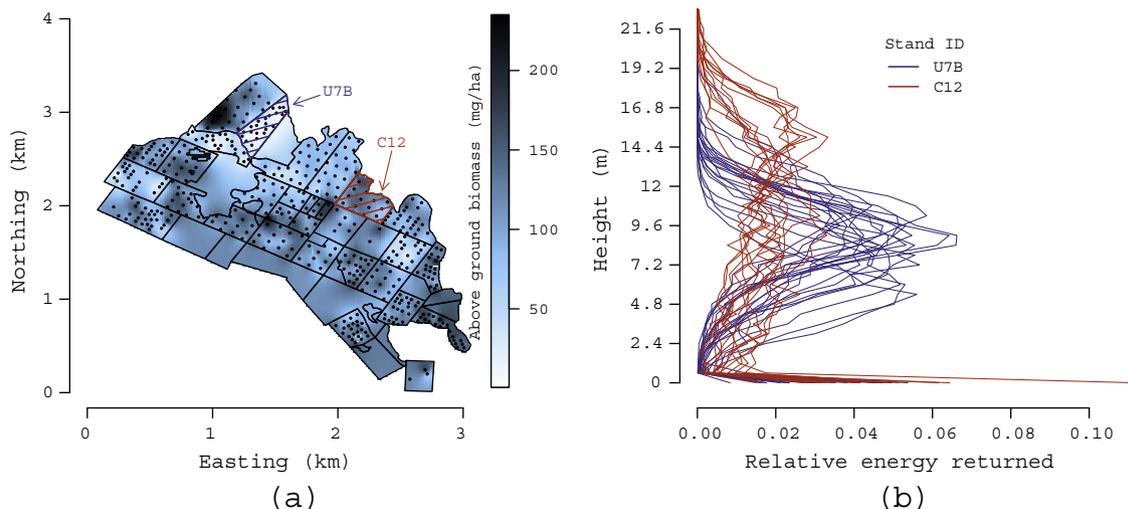


Fig. 1. (a) Penobscot Experimental Forest, Maine, with management units and forest inventory plot locations delineated as polygons and points, respectively. (b) G-LiHT LiDAR signals observed at forest inventory plots highlighted in (a).

vertical and horizontal forest structure complexity due to repeated selection harvests that aim to concentrate growth on economically desirable trees. Sendak et al. (2003) and Hayashi et al. (2014) provide additional silvicultural treatment details.

Large footprint waveforms, characteristic of space-based LiDAR sensors, were calculated using discrete multistop returns from a 2013 PEF G-LiHT data acquisition campaign (Cook et al., 2013). As noted in Section 1, 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 laser scanner (VQ-480, Riegl Laser Measurement Systems, Horn, Austria) uses a 1550 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 m, each laser pulse has a footprint approximately 10 cm in diameter and is capable of producing up to 8 returns. Following data processing methods in Blair and Hofton (1999), G-LiHT produced 26,286 georeferenced pseudo-waveform LiDAR signals across the PEF with 451 of these spatially coinciding with the observed PSPs. Each pseudo-waveform covers a 15 m diameter footprint with a signal comprising 113 values between 0 and 33.9 m above the ground. A signal value is the amount of energy returned to the sensor from a given height divided by the total energy emitted by the sensor over the footprint (additional details are given in Section 4.2.1). The signal can be used to characterize the vertical distribution of forest structure within the footprint. Signals corresponding to PSPs within the MUs highlighted in Fig. 1 (a) are shown in Fig. 1 (b). Here, U7B's even-aged and structurally homogeneous composition is apparent in the signals' consistent peak at ~8 m—corresponding to the densest layer of the forest canopy—and minimal energy returns above ~17 m—corresponding to maximum forest canopy height. In contrast, C12's signals are characterized by non-zero values at greater heights—reflecting the prevalence of taller trees—and greater vertical distribution of energy returns—indicative of a vertically complex forest structure resulting from the MU's silvicultural treatments. The relative energy distribution in the signal does not exactly portray the vertical distribution of vegetation because dense overstory may act to reduce the amount of energy available to characterize lower canopy structures. Therefore, if inferential interest is in the vertical distribution of leaf area density, then we would want to transform the signal energy returns to account for decreasing transmittance of energy through the canopy, see, e.g., MacArthur and Horn (1969) for theoretical motivation for such transformations and Stark et al. (2015) for a recent application. Our focus is on modeling the observed signal and glean information from signal characteristics to explain variability in AGB. It is not clear that applying a MacArthur-Horn transformation (MacArthur and Horn, 1969) to the signal data would fetch improved inference about AGB, and hence we do not pursue such methods here. We do, however, identify these topics as potential extensions to our proposed modeling framework, see Section 5.

3. Models

We envision AGB as a continuous spatial process $\{y(\mathbf{s}) : \mathbf{s} \in \mathcal{D} \subset \mathbb{R}^2\}$ measured over a finite collection of PSP's $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{n_s}\} \subset \mathcal{D}$, where \mathcal{D} is the domain of interest. LiDAR signals are also assumed to arise as the partial realizations of a process $\{z(\ell) : \ell \in \mathcal{D} \times \mathcal{H}\}$, where $\ell = (s, x)$ is a space-height coordinate, $z(\ell)$ represents the LiDAR signal's relative energy return at spatial location \mathbf{s} and height x , and \mathcal{H} is the compact interval $[0, M]$ representing the range of possible heights. The LiDAR signals are also measured at the PSP's in \mathcal{S} and heights $\mathcal{X} = \{x_1, x_2, \dots, x_{n_x}\} \subset \mathcal{H}$. We will assume that $\mathcal{L} = \{\ell_1, \ell_2, \dots, \ell_n\}$ is a complete enumeration of space-height coordinates at which the LiDAR signals have been measured. Each ℓ_i will correspond to a unique ordered pair

(\mathbf{s}_j, x_k) , where $\mathbf{s}_j \in \mathcal{S}$ and $x_k \in \mathcal{X}$. If the measurements are balanced across space and height, i.e., every PSP has measured the LiDAR signal at each of the points in \mathcal{X} , then there will be $n = n_s n_x$ measurements. This, however, need not be assumed for the subsequent development.

3.1. Model for $z(\ell)$

We write the LiDAR signal as

$$z(\ell) = \mu_z(\ell; \boldsymbol{\beta}_z) + u(\ell) + \epsilon_z(\ell), \quad (1)$$

where $\mu_z(\ell; \boldsymbol{\beta}_z)$ is a mean function capturing large-scale variation, $u(\ell)$ is an underlying zero-centered stochastic process over $\mathcal{D} \times \mathcal{H}$ that characterizes spatial dependence, and $\epsilon_z(\ell) \stackrel{\text{ind}}{\sim} N(0, \tau_z^2(x))$ models random disturbances at finer scales, at least part of which is attributed to measurement error. The variance of this fine scale disturbance is assumed to remain invariant over the locations, but depends upon the height x at which the signal is measured.

We assume that $u(\ell)$ is a zero-centered Gaussian process over $\mathcal{D} \times \mathcal{H}$ with a covariance function $C_u(\ell, \ell'; \boldsymbol{\theta}_u) := \text{Cov}[u(\ell), u(\ell')]$. This function must ensure that the resulting variance-covariance matrix corresponding to realizations of the process over any finite subset of $\mathcal{D} \times \mathcal{H}$ is positive definite. A natural class of such functions is that of spatiotemporal covariance functions, but with the temporal domain being replaced by the "height" domain; Gneiting et al. (2006) and gne (2010) provide excellent expositions of such functions.

A relevant concern in our current application is the lack of *separability*, i.e., the covariance function should not factorize into a purely spatial component and a purely height component. Separability would imply that the spatial association in the LiDAR signals remains invariant across heights and, similarly, the association among signals at different heights remains the same for each spatial location. This assumption is too stringent for our application; see, e.g., the disparity in empirical semivariogram parameter estimates presented in the Supplementary material. Furthermore, separable covariance functions violate the so-called "screening" effect (Stein, 2005) and the resulting associations can be sensitive to small perturbations in spatial locations.

Based upon the above, we use a slightly simpler version of a highly flexible class of covariance functions developed by Gneiting (2002),

$$C_u(\ell, \ell'; \boldsymbol{\theta}_u) := \frac{\sigma_u^2}{(a|x - x'|^2 + 1)^\gamma} \exp\left(-\frac{c \|\mathbf{s} - \mathbf{s}'\|}{(a|x - x'|^2 + 1)^{\gamma/2}}\right), \quad (2)$$

where $\ell = (\mathbf{s}, x)$, $\ell' = (\mathbf{s}', x')$, and $\boldsymbol{\theta}_u = \{\sigma_u^2, a, \gamma, c\}$, with σ_u^2 , a , and c all greater than 0 and $\gamma \in [0, 1]$. The parameter γ describes the space-height interaction. Values of γ close to 1 indicate strong space and height interaction. If γ is zero, then Eq. (2) is reduced to a separable covariance function, i.e., no space and height interaction. Observe that the above covariance function still assumes isotropy, i.e., the associations depend upon the distances between the spatial locations and the absolute difference between the heights. This, too, is unlikely in practice, but we are less concerned here because non-stationarity will be introduced in the covariance structures as a part of dimension reduction (Section 3.3).

3.2. Model for $y(\mathbf{s})$ and $z(\ell)$

The spatial process for AGB, $y(\mathbf{s})$, shares the same spatial domain as the LiDAR process and can be modeled using a Gaussian process over \mathcal{D} . Thus,

$$y(\mathbf{s}) = \mu_y(\mathbf{s}; \boldsymbol{\beta}_y) + w(\mathbf{s}) + \epsilon_y(\mathbf{s}), \quad (3)$$

where $\mu_y(\mathbf{s}; \boldsymbol{\beta}_y)$ captures large scale variation or trends in AGB, $w(\mathbf{s})$ is a zero-centered spatial process, and $\epsilon_y(\mathbf{s})$ is a white noise process with zero mean and variance τ_y^2 to capture measurement error in AGB.

We posit that the process for AGB is associated with the process for the LiDAR signals and desire to estimate this association. One possibility is to treat $w(\mathbf{s})$ as a shared process between AGB and LiDAR and introduce it as an additive component in Eq. (1). This, however, causes identifiability issues. First, an additional additive process in Eq. (1) may be difficult to identify from $u(\ell)$ using a single partial realization of the LiDAR process. Second, the AGB process is then governed by a single shared process, $w(\mathbf{s})$, and adding a second process, say $v(\mathbf{s})$, to capture departure from the shared component will, again, introduce identifiability problems. Both these problems can, in principle, be resolved in Bayesian settings if prior elicitation was possible on these different component processes. This, unfortunately, is difficult here and we do not pursue such approaches.

We prefer to treat $u(\ell)$ as a shared underlying process, common to both $z(\mathbf{s}; x)$ and $y(\mathbf{s})$. However, since the AGB has support over the spatial domain only, we assume that it is a continuous weighted average of $u(\ell)$ over \mathcal{X} . Therefore, we write $w(\mathbf{s})$ in Eq. (3) as

$$w(\mathbf{s}) = \int_{\mathcal{X}} \alpha(x) u(\mathbf{s}, x) dx + v(\mathbf{s}) \approx \sum_{j=1}^{n_x} \alpha(x_j) u(\mathbf{s}, x_j) + v(\mathbf{s}), \quad (4)$$

where $\alpha(x)$ is a weight function that maps height in \mathcal{X} to the real line and $v(\mathbf{s})$ is a zero-centered spatial process, independent of $u(\mathbf{s}, x)$, that captures features specific to AGB that are not shared with the LiDAR signal. Specifically, we assume $v(\mathbf{s})$ is a zero-mean Gaussian process with an exponential covariance function $C_v(\mathbf{s}, \mathbf{s}'; \boldsymbol{\theta}_v) = \sigma_v^2 \exp(-\phi_v \|\mathbf{s} - \mathbf{s}'\|)$, where $\boldsymbol{\theta}_v = \{\sigma_v^2, \phi_v\}$. More generally, a Matérn covariance function (Stein, 1999) with a prior on the smoothness parameter could have been used, but this does add to the computational burden without any discernible benefits in the substantive scientific inference we seek in the current application.

Rather than specify the weights $\alpha(x)$, we represent the integrated process in Eq. (4) as a linear combination of the $u(\ell)$'s over \mathcal{X} for any fixed \mathbf{s} and regard the $\alpha(x_j)$'s as unknown coefficients for the $u(\mathbf{s}, x_j)$'s. These coefficients capture the dependence of $w(\mathbf{s})$ on $u(\ell)$ and, hence, the association between the two processes. If they are all estimated to be effectively zero, then there is no association between the AGB and LiDAR processes, while significant departures of any of the coefficients from zero will indicate association between the processes.

Let \mathbf{u} be the $n \times 1$ vector with elements $u(\ell_i)$, $i = 1, 2, \dots, n$ stacked so that $\ell_i = (\mathbf{s}_j, x_k)$, where $i = (j-1)n_x + k$ with $j = 1, 2, \dots, n_s$ and $k \in \{1, 2, \dots, n_x\}$, and $\mathbf{C}_u(\boldsymbol{\theta}_u)$ is the corresponding $n \times n$ variance-covariance matrix with entries $\text{cov}(u(\ell_i), u(\ell_j))$. For the spatial process $v(\mathbf{s})$, we let \mathbf{v} be the $n_s \times 1$ vector with elements $v(\mathbf{s}_j)$ and $\mathbf{C}_v(\boldsymbol{\theta}_v)$ is the corresponding $n_s \times n_s$ spatial covariance matrix. Also, we assume linear fixed effects $\mu_z(\ell_i; \boldsymbol{\beta}_z) = \mathbf{q}_z(\ell_i)^\top \boldsymbol{\beta}_z$ and $\mu_y(\mathbf{s}_j; \boldsymbol{\beta}_y) = \mathbf{q}_y(\mathbf{s}_j)^\top \boldsymbol{\beta}_y$, where $\mathbf{q}_z(\ell_i)$ and $\mathbf{q}_y(\mathbf{s}_j)$ are $p_z \times 1$ and $p_y \times 1$ vectors of predictors or explanatory variables for $z(\ell_i)$ and $y(\mathbf{s}_j)$, respectively.

A joint Bayesian hierarchical model for $y(\mathbf{s}_j)$'s and $z(\ell_i)$'s, given measurements over \mathcal{S} and $\mathcal{S} \times \mathcal{X}$, respectively, is given by

$$\begin{aligned} p(\boldsymbol{\theta}) \times N(\boldsymbol{\beta}_y | \boldsymbol{\mu}_{\beta_y}, \mathbf{V}_{\beta_y}) \times N(\boldsymbol{\beta}_z | \boldsymbol{\mu}_{\beta_z}, \mathbf{V}_{\beta_z}) \times N(\boldsymbol{\alpha} | \boldsymbol{\mu}_{\alpha}, \mathbf{V}_{\alpha}) \\ \times N(\mathbf{v} | \mathbf{0}, \mathbf{C}_v(\boldsymbol{\theta}_v)) \times N(\mathbf{u} | \mathbf{0}, \mathbf{C}_u(\boldsymbol{\theta}_u)) \\ \times \prod_{j=1}^{n_s} N(y(\mathbf{s}_j) | \mathbf{q}_y^\top(\mathbf{s}_j) \boldsymbol{\beta}_y + \boldsymbol{\alpha}^\top \mathbf{u}(\mathbf{s}_j) + v(\mathbf{s}_j), \tau_y^2) \\ \times \prod_{i=1}^n N(z(\ell_i) | \mathbf{q}_z^\top(\ell_i) \boldsymbol{\beta}_z + u(\ell_i), \tau_z^2(x_k)), \end{aligned} \quad (5)$$

where $\boldsymbol{\theta} = \{\boldsymbol{\theta}_u, \boldsymbol{\theta}_v, \tau_y^2, \tau_z^2\}$ with $\tau_z^2 = (\tau_z^2(x_k))_{k=1}^{n_x}$, $\boldsymbol{\beta}_z$ and $\boldsymbol{\beta}_y$ are regression slopes for each $\mathbf{q}_z(\ell_i)$ and $\mathbf{q}_y(\mathbf{s}_j)$, respectively, $\mathbf{u}(\mathbf{s}_j)$ is the vector with elements $u(\mathbf{s}_j, x_k)$ for x_k 's in \mathcal{X} yielding LiDAR signals corresponding to \mathbf{s}_j , $\boldsymbol{\alpha}$ is an $n_x \times 1$ vector of unknown coefficients, viz. the $\alpha(x_j)$'s, for the elements in $\mathbf{u}(\mathbf{s}_j)$, and $p(\boldsymbol{\theta})$ are joint prior distributions on the process parameters for $u(\mathbf{s}, x)$ and $v(\mathbf{s})$. Further specifications customarily assume that

$$p(\boldsymbol{\theta}) \propto p(\boldsymbol{\theta}_u) \times p(\boldsymbol{\theta}_v) \times IG(\tau_y^2 | a_{\tau_y}, b_{\tau_y}) \times \prod_{k=1}^{n_x} IG(\tau_z^2(x_k) | a_{\tau_z}, b_{\tau_z}), \quad (6)$$

where $p(\boldsymbol{\theta}_u) = p(a, \gamma, c) \times IG(\sigma_u^2 | a_u, b_u)$ and $p(\boldsymbol{\theta}_v) = p(\phi_v) \times IG(\sigma_v^2 | a_v, b_v)$, with IG denoting the inverse-Gamma distribution. When the number of space-height coordinates n is large, estimating Eq. (5) is computationally expensive and, depending upon the available computational resources, possibly unfeasible.

3.3. Predictive process counterparts for dimension reduction

To implement the computations necessary for estimating Eq. (5) when n is large, we exploit reduced rank processes to achieve dimension reduction. Such processes usually arise as basis expansions of the original process with fewer number of basis functions than the number of data points. This yields “low-rank” processes. Every choice of basis functions yields a process and there are far too many choices to enumerate here; see, e.g., Wikle (2010) for an excellent overview of these methods. Here, we opt for a particularly convenient choice, the predictive process (Banerjee et al., 2008; Finley et al., 2009), which derives the basis functions from taking the conditional expectation of the original process, often called the “parent” process, given its realizations over a fixed set of points, often referred to as “knots”.

Let $\mathcal{S}_u^* = \{\mathbf{s}_{u,1}^*, \mathbf{s}_{u,2}^*, \dots, \mathbf{s}_{u,n_u}^*\}$ and $\mathcal{S}_v^* = \{\mathbf{s}_{v,1}^*, \mathbf{s}_{v,2}^*, \dots, \mathbf{s}_{v,n_v}^*\}$ be two sets of spatial knots to be used for constructing the predictive process counterparts of $u(\ell)$ and $v(\mathbf{s})$, labeled $\tilde{u}(\ell)$ and $\tilde{v}(\mathbf{s})$, respectively. Let $\mathcal{X}^* = \{x_1^*, x_2^*, \dots, x_{n_x}^*\}$ be a set of knots for heights in the LiDAR signal. Dimension reduction is achieved because the number of knots, i.e., n_u^* , n_v^* , and n_x^* , is much smaller than the original number of observations n_s and n_x . Implementation details for the predictive process version of Eq. (5) used in the subsequent analyses is detailed in the Supplemental material.

3.4. Bayesian prediction

As noted in Section 1, we seek predictive inference for $z(\ell_0)$ at any arbitrary space-height coordinate ℓ_0 and for $y(\mathbf{s}_0)$ at any arbitrary spatial location \mathbf{s}_0 . The posterior predictive distributions and corresponding sampling algorithms that yield this inference are defined in the Supplemental material. In Section 4, we use posterior predictive inferences at i) unobserved locations to create prediction maps of the LiDAR signals and AGB, and to assess models' predictive performance using holdout set validation, and ii) observed locations to

Table 1

Parameter credible intervals, 50% (2.5%,97.5%), goodness-of-fit, and predictive validation. Bold entries indicate where the true value is missed, bold goodness-of-fit metrics indicate best fit, and bold prediction metrics indicate best predictive performance.

Parameter	True	Models			
		$n_u^* = 300$ and $S_u^* = S$	$n_u^* = 200$	$n_u^* = 100$	$n_u^* = 50$
$\beta_{y,0}$	20	20.11 (19.3,21.06)	20.21 (19.48,21.05)	20.28 (19.46,21.32)	20.5 (19.95,21.12)
α_1	-2	-1.92 (-2.25,-1.61)	-2.09 (-2.49,-1.73)	-2.03 (-2.63,-1.45)	-1.41 (-2.19,-0.71)
α_2	0	0.06 (-0.3,0.46)	0.38 (-0.13,0.83)	0.25 (-0.62,1.03)	-0.1 (-1.09,0.87)
α_3	2	1.57 (1.06,2.01)	1.21 (0.71,1.77)	1.12 (0.26,1.99)	0.91 (-0.09,2.01)
α_4	1	0.91 (0.54,1.29)	0.77 (0.34,1.19)	0.93 (0.26,1.6)	0.68 (-0.19,1.59)
α_5	5	4.94 (4.64,5.23)	4.9 (4.52,5.27)	4.82 (4.23,5.43)	5.04 (4.38,5.68)
σ_u^2	0.2	0.2 (0.19,0.22)	0.21 (0.2,0.22)	0.22 (0.21,0.23)	0.21 (0.2,0.22)
a	12	12.91 (11.62,14.46)	11.53 (10.1,13.36)	10.58 (9.29,12.63)	10.62 (8.77,13.19)
γ	0.9	0.89 (0.84,0.93)	0.94 (0.85,0.97)	0.92 (0.83,0.95)	0.92 (0.83,0.97)
c	5	5.17 (4.62,5.93)	5.24 (4.7,5.87)	3.82 (3.35,4.29)	3.09 (2.7,3.5)
σ_v^2	0.5	0.64 (0.34,1.27)	0.68 (0.37,1.24)	0.51 (0.15,1.47)	0.6 (0.12,1.74)
ϕ_v	2	1.68 (0.41,3.87)	1.7 (0.49,3.8)	1.38 (0.21,5.26)	7.6 (0.25,9.99)
p_D for AGB & LiDAR		95.01	95.94	96.14	92.39
DIC for AGB & LiDAR		28,399.73	29,238.34	30,140.16	30,647.04
G for AGB & LiDAR		6883.54	7705.30	8566.07	8779.24
P for AGB & LiDAR		7507.49	8215.54	9007.16	9386.19
D = G + P for AGB & LiDAR		14,391.03	15,920.84	17,573.23	18,165.43
RMSPE for AGB & LiDAR		0.80	0.80	0.81	0.82
CRPS for AGB & LiDAR		2054.52	2056.32	2076.22	2090.47
GRS for AGB & LiDAR		- 712.04	-743.83	-864.92	-962.09
95% prediction interval coverage for AGB& LiDAR %		95.27	95.1	95.39	94.9
RMSPE for AGB observed LiDAR		2.79	2.74	2.78	2.81
CRPS for AGB observed LiDAR		159.32	156.28	157.71	158.69
GRS for AGB observed LiDAR		-311.06	- 305.18	-308.63	-308.17
95% prediction interval coverage for AGB observed LiDAR %		89	91	93	93
95% prediction interval width for AGB observed LiDAR		9.30	9.70	9.74	9.96
CPU time		101.7 h	60.0 h	27.5 h	13.8 h

provide replicated data (see, e.g., Gelman et al., 2013) used to assess candidate models' fit.

4. Data analysis

The proposed Markov chain Monte Carlo (MCMC) sampler and prediction algorithms, detailed in the Supplementary material, were implemented in C++. All code needed to fit the proposed models and reproduce the subsequent results are available in the Supplemental material. Posterior inference for subsequent analysis were based upon three chains of 50,000 MCMC iterations (with a burn-in of 5000 iterations). The computations were conducted on a Linux workstation using two Intel Nehalem quad-Xeon processors.

In the subsequent simulation experiment Section 4.1 and PEF data analysis Section 4.2 candidate models were compared based

on parameter estimates, fit to the observed data, out-of-sample prediction, and posterior predictive distribution coverage. Model choice was assessed using the deviance information criterion or DIC and model complexity p_D (Spiegelhalter et al., 2002) and a posterior predictive loss criterion $D = G + P$ (Gelfand and Ghosh, 1998), where smaller values of DIC and D indicate preferred models. For both analyses, a 25% holdout set, comprising locations selected at random, served to assess out-of-sample prediction. Prediction accuracy for the holdout locations was measured using root mean squared prediction error (RMSPE) (Yeniay and Goktas, 2002) as well as CRPS and GRS given in Eqs. (21) and (27), respectively, in Gneiting and Raftery (2007). Smaller values RMSPE and CRPS, and larger values of GRS, indicate improved predictive ability. The percent of holdout locations that fell within their respective posterior predictive distribution 95% CI was also computed along with the average interval width.

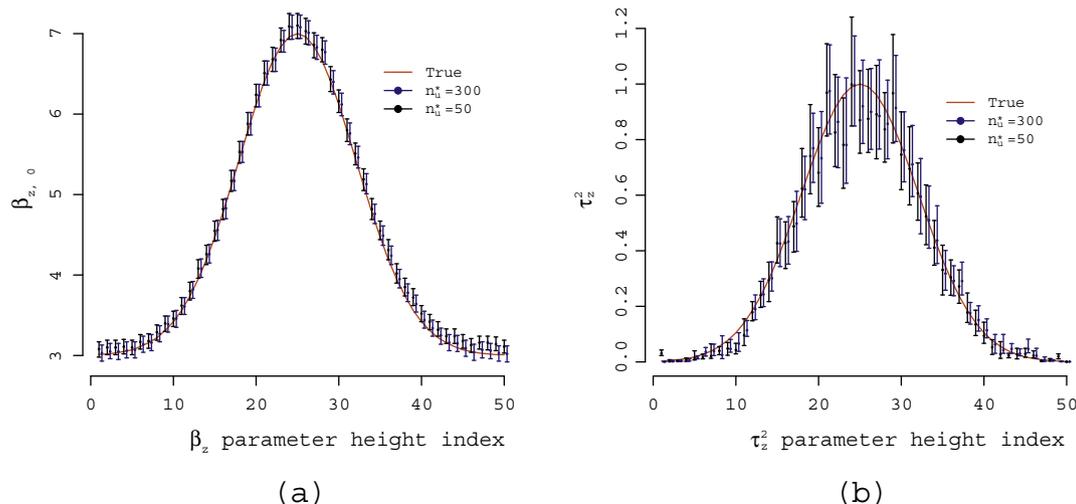


Fig. 2. Parameter posterior summaries, 50% point symbol and 95% credible interval bars. Posterior summaries are jittered slightly along the x-axis to facilitate comparison.

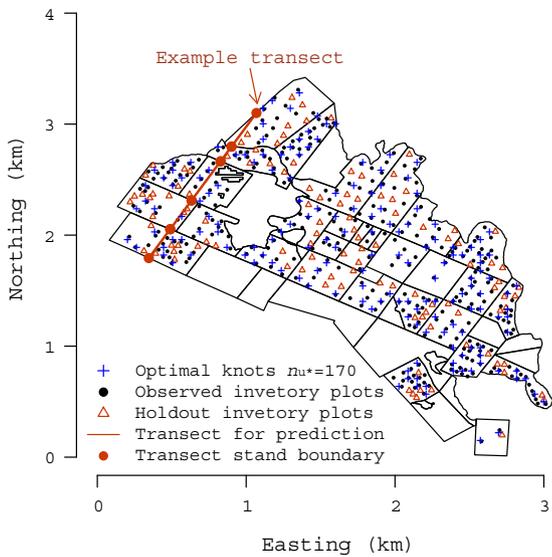


Fig. 3. Locations of observed and holdout PSPs, predictive process knots, and illustrative transect on the PEF.

4.1. Simulation experiment

Using the true parameter values given in the first column of Table 1 and Fig. 2, we simulated AGB and LiDAR signals from the full GP joint likelihood for AGB and LiDAR in Eq. (5) for $n_s = 400$ coinciding locations in S on a regular grid within a $[0, 4] \times [0, 4]$ domain and $n_x = 50$ heights within $[0, 5]$. The AGB signal was regressed on a global intercept (β_y) while the LiDAR signals were regressed on the 50 height-specific intercepts and non-spatial variances; thus τ_z and β_z^2 are both 50×1 . A subset of 100 locations from the 400 were withheld to assess out-of-sample predictions. Each of our candidate predictive process models used $n_x^* = 5$ equally spaced knots for height in the $[0, 5]$ interval and $n_u^* = n_s = 300$ with $S_y^* = S$. Candidate models differed on the number of knots n_u^* . We considered models with $n_u^* = 300$ and $S_u^* = S$ and with n_u^* equaling 200, 100, and 50 knots, respectively, selected on a regular grid within the domain.

Table 2
Parameter credible intervals, 50% (2.5%,97.5%), and goodness-of-fit for the $n_u^* = 339$ and $n_v^* = 339$ models. Bold parameter values indicated values that differ from zero where appropriate and bold goodness-of-fit metrics indicate best fit.

Parameter	Height knot models					
	$n_x^* = 2$	$n_x^* = 3$	$n_x^* = 4$	$n_x^* = 5$	$n_x^* = 6$	$n_x^* = 7$
$\beta_{y,0}$	0.21 (-0.43,0.86)	0.01 (-0.7,0.76)	-0.06 (-0.77,0.65)	0.39 (-0.51,1.35)	0.09 (-0.76,0.98)	-0.12 (-1.01,0.71)
$\beta_{y,Elev}$	0.02 (0,0.03)	0.02 (0.01,0.04)	0.03 (0.01,0.04)	0.02 (-0.01,0.04)	0.02 (0,0.04)	0.03 (0.01,0.05)
α_1	-0.08 (-0.15,-0.01)	-0.06 (-0.15,0.03)	-0.04 (-0.09,0.01)	0.02 (-0.04,0.07)	-0.05 (-0.1,0.01)	-0.04 (-0.09,0.02)
α_2	0.33 (0.27,0.39)	0.06 (-0.11,0.24)	-0.16 (-0.23,-0.1)	-0.05 (-0.11,0.03)	-0.11 (-0.18,-0.04)	-0.07 (-0.13,0)
α_3	-	0.43 (0.33,0.52)	0.1 (0.05,0.16)	0.05 (-0.02,0.12)	-0.1 (-0.17,-0.03)	-0.09 (-0.15,-0.03)
α_4	-	-	0.26 (0.13,0.39)	0.2 (0.14,0.27)	0.04 (-0.03,0.12)	-0.01 (-0.06,0.06)
α_5	-	-	-	0.28 (0.13,0.47)	0.11 (0.03,0.18)	0.06 (-0.02,0.12)
α_6	-	-	-	-	0.18 (0.06,0.35)	0.17 (0.07,0.28)
α_7	-	-	-	-	-	0.04 (-0.13,0.28)
σ_u^2	0.1 (0.1,0.11)	0.16 (0.14,0.17)	0.48 (0.44,0.52)	0.61 (0.57,0.65)	0.97 (0.88,1.07)	1.15 (1.07,1.26)
a	1.12 (0.99,1.28)	1.41 (1.24,1.56)	0.75 (0.7,0.8)	1.12 (1.05,1.18)	0.99 (0.93,1.05)	1.05 (0.99,1.11)
γ	0.99 (0.98,0.99)	1 (1,1)	1 (0.99,1)	1 (1,1)	0.99 (0.98,1)	1 (1,1)
c	16.39 (14.03,19.62)	10.94 (9.89,12.14)	8.89 (8.16,9.72)	8.55 (7.98,9.1)	8.07 (7.54,8.59)	8.44 (7.94,9.01)
σ_v^2	0.09 (0.07,0.15)	0.11 (0.07,0.18)	0.1 (0.06,0.13)	0.1 (0.07,0.17)	0.08 (0.07,0.11)	0.1 (0.08,0.12)
ϕ_v	3.6 (2.23,7.49)	3.27 (1.86,4.93)	3.59 (1.84,5.32)	2.14 (1.64,3.55)	4.51 (3.34,6.58)	3.8 (3.04,4.82)
τ_y^2	0.03 (0.01,0.04)	0.03 (0.02,0.04)	0.03 (0.02,0.03)	0.03 (0.02,0.04)	0.02 (0.01,0.03)	0.02 (0.02,0.03)
p_D	112.67	113.14	105.05	104.81	104.28	100.52
DIC	25,249.95	23,914.36	19,508.47	15,141.14	12,678.65	9627.62
G	10,201.51	8243.9	6932.84	4466.97	3636.59	3165.57
P	10,606.52	8770.24	7522.97	4918.89	4188.48	3595.77
D	20,808.03	17,014.13	14,455.8	9385.86	7825.07	6761.33

Parameter estimates and performance metrics for all candidate models are given in Table 1. With the exception of α_3 for n_u^* equal to 200 and 100, and a few of the covariance parameters for n_u^* equal to 100 and 50, the 95% CIs for all parameters included the true values. Importantly, the α estimates—used to relate information between LiDAR signals and AGB—remain consistent in sign and magnitude as the spatial process associated with the signals is modeled over a reduced number of knots. Fig. 2 provides the posterior summaries for the 50 height-specific intercepts and non-spatial variances associated with the LiDAR signals; results for only two candidate models are provided due to the large number of parameters and minimal difference in estimates among the models. These estimates also seem robust to a coarser representation of the underlying process (Fig. 2).

Not surprisingly, for the joint outcome vector, goodness-of-fit and out-of-sample prediction is best for the full model, i.e., $n_u^* = 300$ (rows labeled AGB & LiDAR in Table 1). Interestingly, in an interpolation setting when LiDAR is observed, RMSPE, CRPS and GRS all show that AGB prediction improves slightly when moving from the full model to the $n_u^* = 200$ knot model (rows labeled AGB | observed LiDAR in Table 1). In general, goodness-of-fit and predictive performance is not substantially degraded for the predictive process models when compared to the full model. The last row in Table 1 gives the CPU time for the candidate models. A 6-fold decrease in knots between the full model and $n_u^* = 50$ knot model results in a 7-fold decrease in computing time.

4.2. Forest LiDAR and biomass data analysis

4.2.1. Data preparation and exploratory data analysis

Pre-processing the raw G-LiHT LiDAR data followed methods detailed in Cook et al. (2013) and produced a complete 15×15 m grid across the PEF, where each pixel contained a LiDAR signal. Prior to analysis, these LiDAR signals were further processed to remove excess zeros and coarsened to remove small-scale noise. Specifically, the maximum tree height across the PEF was approximately 22.8 m and hence LiDAR signal values beyond this height were zero and subsequently removed. Small-scale anomalies that occurred across each signal's 113 values were smoothed by averaging every two consecutive measurements. Truncation above forest canopy extent and smoothing resulted in signals of length $n_x = 39$ within

Table 3

Parameter credible intervals, 50% (2.5%,97.5%), and goodness-of-fit for the $n_u^* = 170$ and $n_v^* = 339$ models. Bold parameter values indicated values that differ from zero where appropriate and bold goodness-of-fit metrics indicate best fit.

Parameter	Height knot models					
	$n_x^* = 2$	$n_x^* = 3$	$n_x^* = 4$	$n_x^* = 5$	$n_x^* = 6$	$n_x^* = 7$
$\beta_{y,0}$	-0.09 (-0.76,0.59)	-0.44 (-1.16,0.27)	-0.25 (-0.88,0.37)	-0.15 (-0.9,0.57)	-0.15 (-0.82,0.49)	-0.2 (-0.92,0.48)
$\beta_{y,Elev}$	0.03 (0.01,0.04)	0.03 (0.02,0.05)	0.03 (0.02,0.04)	0.03 (0.01,0.04)	0.03 (0.01,0.04)	0.03 (0.01,0.04)
α_1	-0.2 (-0.42,0.04)	-0.4 (-0.66,-0.14)	-0.03 (-0.17,0.1)	-0.07 (-0.22,0.07)	-0.1 (-0.26,0.03)	-0.13 (-0.32,0.08)
α_2	0.72 (0.53,0.89)	-0.45 (-0.79,0.01)	-0.28 (-0.5,-0.08)	-0.25 (-0.44,-0.06)	-0.13 (-0.28,0)	-0.09 (-0.26,0.09)
α_3	-	0.62 (0.37,0.86)	0.23 (0.05,0.41)	-0.13 (-0.31,0.05)	-0.28 (-0.44,-0.14)	-0.24 (-0.45,-0.04)
α_4	-	-	0.33 (0.08,0.55)	0.24 (0.11,0.42)	0.14 (-0.03,0.31)	-0.06 (-0.2,0.1)
α_5	-	-	-	0.1 (-0.19,0.41)	0.11 (-0.05,0.31)	0.12 (-0.06,0.32)
α_6	-	-	-	-	0.09 (-0.22,0.31)	0.12 (-0.12,0.37)
α_7	-	-	-	-	-	0.05 (-0.3,0.41)
σ_u^2	0.05 (0.05,0.06)	0.09 (0.07,0.1)	0.24 (0.22,0.28)	0.35 (0.32,0.39)	0.39 (0.35,0.44)	0.42 (0.38,0.48)
a	0.8 (0.64,1)	1.02 (0.82,1.23)	0.84 (0.74,0.94)	1.23 (1.12,1.35)	1.25 (1.14,1.37)	1.49 (1.36,1.64)
γ	0.98 (0.9,0.99)	0.99 (0.96,1)	0.99 (0.97,1)	0.99 (0.98,1)	1 (0.99,1)	1 (0.99,1)
c	6.02 (5.32,6.71)	4.75 (4.19,5.39)	3.23 (2.85,3.6)	3.05 (2.77,3.39)	2.81 (2.53,3.15)	2.89 (2.55,3.29)
σ_v^2	0.08 (0.05,0.16)	0.09 (0.05,0.14)	0.06 (0.04,0.09)	0.08 (0.04,0.18)	0.05 (0.04,0.08)	0.06 (0.04,0.1)
ϕ_v	2.63 (1.32,5.02)	2.51 (1.43,4.54)	3.92 (2.02,8.4)	1.42 (1.13,3.61)	8.17 (3.76,16.33)	4.1 (2.6,7.84)
τ_z^2	0.04 (0.02,0.05)	0.03 (0.02,0.04)	0.03 (0.02,0.04)	0.04 (0.02,0.05)	0.02 (0.01,0.05)	0.03 (0.02,0.04)
p_D	118.32	121.42	110.41	112.36	114.27	110.13
DIC	26,952.72	26,222.81	24,446.58	22,809.18	22,325.5	21,889.75
G	10,703.8	9529.07	8085.34	6678.27	6402.89	6131.33
P	10,868.98	9840.64	8480.8	7080.34	6885.51	6656.74
D	21,572.78	19,369.71	16,566.14	13,758.61	13,288.4	12,788.07

the [0, 22.8] m height interval. Fig. 1 (b) illustrates the processed signals over the PSPs within two MUs.

As described in Section 1, important current and future LiDAR acquisition missions sparsely sample the domain of interest. The sampling designs, e.g., transects or clusters, aim to collect LiDAR data across the domain and also at forest inventory plot locations. To mimic the sparseness of these anticipated datasets and associated inferential challenges, only LiDAR signals that spatially coincided with PSPs were used for candidate model parameter estimation.

Candidate models were assessed without and with predictor variables. The set of candidate models without predictor variables (i.e., $\mathbf{q}_z = \mathbf{1}$ and $\mathbf{q}_y = 1$, i.e., intercepts only) mimic a worst-case settings where we do not have complete coverage, wall-to-wall, predictor variables. The set of candidate models with predictor variables use ground surface topographic characteristics derived from G-LiHT's digital elevation model to help explain variability in AGB and LiDAR signals. Although we considered a host of aspect, slope, and roughness predictor variables in exploratory data analysis using the proposed models (following suggested topographic transformations in Stage, 1976), only elevation consistently explained a substantial portion of variability in observed AGB and LiDAR signals. Therefore, the set of

models with predictor variables was fit using $\mathbf{q}_z^T(\ell) = (1, \text{Elev}(\mathbf{s}))$ and $\mathbf{q}_y^T(\mathbf{s}) = (1, \text{Elev}(\mathbf{s}))$, where Elev is ground elevation (m).

To better assess the information contribution of latent LiDAR regressors for AGB prediction an additional set of intercept only models were fit with $v(\mathbf{s})$ set to zero. For this set of models, only information from the latent LiDAR regressors, i.e., via α , is available to explain variability in AGB.

4.2.2. Candidate model results and discussion

For brevity, in the main text we only present results for the set of candidate models what include elevation as a predictor variable. Results for the intercept only and $v(\mathbf{s})$ set to zero models are offered in the Supplemental material. Results were comparable among all candidate model sets; however, models with the elevation predictor and $v(\mathbf{s})$ showed consistent, albeit marginal, improvement in fit and predictive performance.

Candidate models were formed by varying n_x^* , n_u^* , and n_v^* along with knot location following the approximately optimal knot design criteria described in the Supplementary material. We present results for n_x^* from 2 to 7, n_u^* equal 339 and 170, and $n_v^* = 339$. All candidate

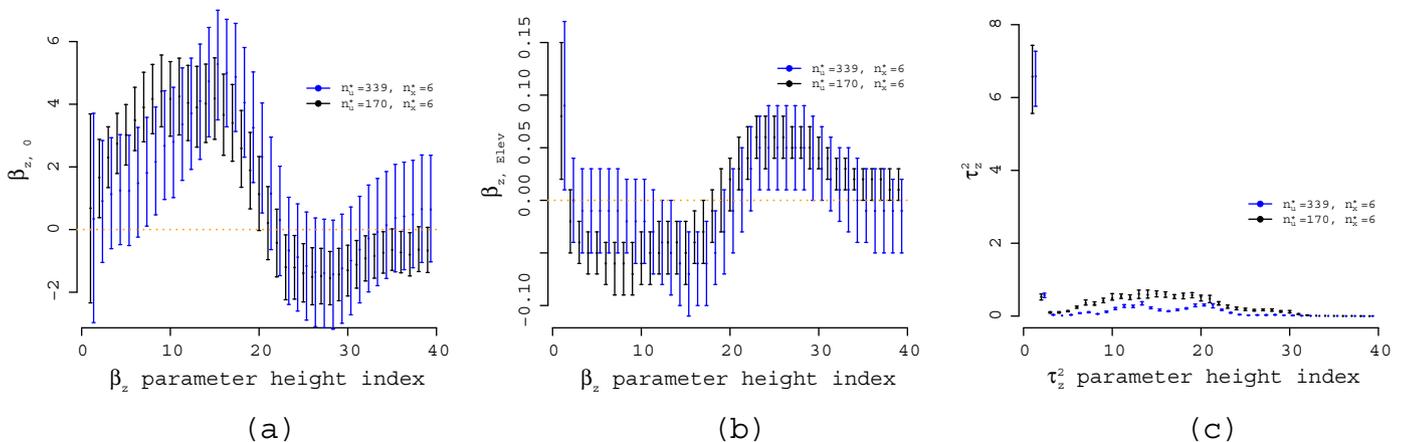


Fig. 4. Parameter posterior summaries, 50% point symbol and 95% credible interval bars. Posterior summaries are jittered slightly along the x-axis to facilitate comparison.

Table 4
Prediction metrics for the $n_u^* = 339$ and $n_v^* = 339$ models. Bold values indicate best predictive performance.

Parameter	Height knot models					
	$n_x^* = 2$	$n_x^* = 3$	$n_x^* = 4$	$n_x^* = 5$	$n_x^* = 6$	$n_x^* = 7$
RMSPE for AGB & LiDAR	0.883	0.842	0.83	0.781	0.777	0.77
CRPS for AGB & LiDAR	1944.36	1868.3	1846.69	1754.68	1773.21	1777.23
GRS for AGB & LiDAR	-1895.37	-1381.9	-1359.45	- 1084.02	-1230.43	-1376.24
95% prediction coverage for AGB & LiDAR	93.5	93.8	94.2	93.4	94.9	95.7
RMSPE for AGB observed LiDAR	0.313	0.308	0.304	0.311	0.306	0.305
CRPS for AGB observed LiDAR	19.84	19.36	19.16	19.49	19.26	19.22
GRS for AGB observed LiDAR	140	151.73	154.43	151.4	153.75	153.93
95% prediction interval coverage for AGB observed LiDAR	90.2	93.8	93.8	96.4	95.5	95.5
95% prediction interval width for AGB observed LiDAR	1	1.1	1.12	1.23	1.19	1.22

models were fit using a subset of $n_s = 339$ PSPs selected at random from the complete set of 451 PSPs. The remaining 112 PSP were used for out-of-sample prediction validation. Observed and holdout PSPs along with knot locations are illustrated in Fig. 3. Here too, an example transect is identified along with locations where the transect crosses MU boundaries. This example transect is used to help visualize and assess results.

Parameter estimates and goodness-of-fit metrics for the aforementioned choices of n_x^* are provided in Tables 2 and 3 for n_u^* equal 339 and 170, respectively. For both choices of n_u^* , increasing the number of height knots n_x^* resulted in improved fit (noted by lower values of DIC and D). This makes sense because a greater number of knots provides an improved representation of the LiDAR signal. This result also holds for the intercept only and $v(s) = 0$ candidate models (Supplementary material Tables S1, S2, S5, and S6).

The regression slope parameter estimates for $\beta_{y,Elev}$ in Tables 2 and 3 suggest elevation explains a significant amount of variation in AGB, with greater biomass occurring on higher elevation PSPs. Fig. 4 provides posterior summaries for the LiDAR signal's height-specific intercept $\beta_{z,0}$, elevation regression slope parameter $\beta_{z,Elev}$, and non-spatial variance parameter estimates for two candidate models. For heights of less than ~ 10 m (corresponding to β_z index ~ 20 in Fig. 4 (b)), higher elevation is associated with fewer energy returns, and between ~ 10 and 18 m (corresponding to β_z index ~ 20 –35 in Fig. 4 (b)) higher elevation is associated with greater energy returns. This is not surprising, given $\beta_{y,Elev}$ estimates suggest greater AGB is associated with higher elevation and tall dense canopies are indicative of forest with greater AGB.

Also, unlike the synthetic data analysis, panels (a) and (b) in Fig. 4, show differences in precision between parameter estimates at different levels of n_u^* . Specifically, we see more precise estimates of the intercept and elevation regression slope parameters at $n_u^* = 170$ versus $n_u^* = 339$. This is likely due to a phenomena called spatial confounding (see, e.g., Hanks et al., 2015), which is most pronounced when a predictor variable and random effect are correlated, in our case elevation seems to be correlated spatially with $\tilde{u}(\ell)$. The greater the resolution on the spatial process, i.e. moving from $n_u^* = 170$ to

339, the greater the influence of spatial confounding on the estimates of $\beta_{1,Elev}$. Spatial confounding can result in wider regression coefficient credible intervals but should not have deleterious effects on prediction, i.e., the inferential focus of our analysis.

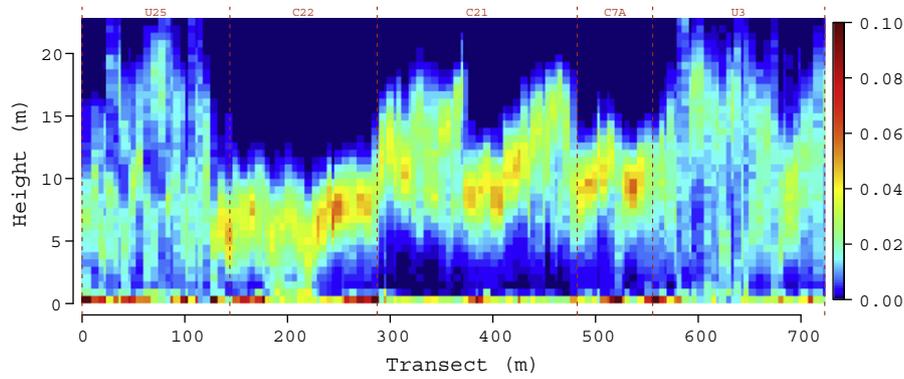
Fig. 4 (c) shows lower residual variances for $n_u^* = 339$ across heights. This is not surprising, given the additional information about the signal supplied by the higher resolution spatial process representation.

Inference on α , which act as weights for the LiDAR process, help us gauge the usefulness of the latent LiDAR regressors for explaining variability in AGB. The increasing subscript value on α 's in Tables 2 and 3 correspond to increasing knot heights in \mathcal{X}^* . Regardless of the choice for n_u^* or n_v^* , estimates of α and knot height are positively associated, for example, estimates of α for the $n_x^* = 6$ model in Table 2 increase from α_1 (corresponding to knot x_1^* which is near the ground) to α_6 (corresponding to x_6^* which is near the maximum forest canopy height). The intuition here is that the latent LiDAR process $\tilde{u}(\ell)$ tend to have larger values at heights where energy return is greater (i.e., where the signal encounters tree material such as leaves, branches, and boles) and small where energy return is low (i.e., where there is mostly empty space in the vertical profile of the forest, or dense over-story intercepts the majority of the signal). Typically, more mature forest with large diameter and tall trees have higher AGB compared with younger lower canopy or sparsely populated forest. Therefore, we expect greater AGB in regions returning much of the LiDAR signal from greater heights and, conversely, lower AGB in regions returning much of the signal at lower heights.

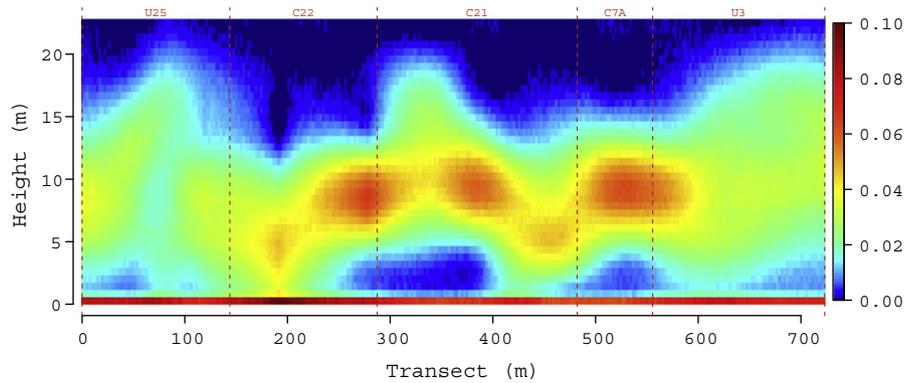
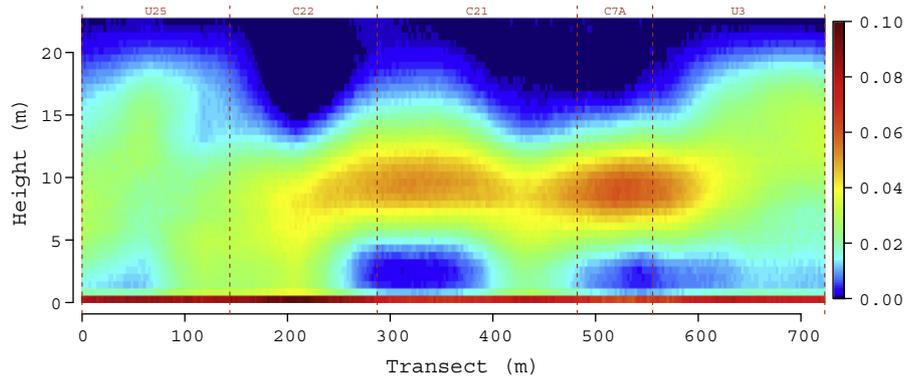
Tables 4 and 5 provide out-of-sample prediction validation results corresponding to the models presented in Tables 2 and 3, respectively. For joint prediction of AGB and LiDAR, using either level of n_u^* , RMSPE favors $n_x^* = 7$ whereas selection results based on CRPS and CRS are mixed. Importantly, however, holdout validation results suggest there is very little difference in predictive ability among the range of height knots beyond $n_x^* = 2$ or 3. If interest is in predicting AGB at a location given observed LiDAR, the majority of the prediction metrics favor n_x^* between 4 and 6 as indicated in the lower portion of Tables 4 and 5.

Table 5
Prediction metrics for the $n_u^* = 170$ and $n_v^* = 339$ models. Bold values indicate best predictive performance.

Parameter	Height knot models					
	$n_x^* = 2$	$n_x^* = 3$	$n_x^* = 4$	$n_x^* = 5$	$n_x^* = 6$	$n_x^* = 7$
RMSPE for AGB & LiDAR	0.876	0.846	0.824	0.798	0.793	0.792
CRPS for AGB & LiDAR	1916.97	1867.7	1814.08	1755.98	1748.6	1741.51
GRS for AGB & LiDAR	-1009.11	-1052.07	- 976.27	-1060.06	-1024.2	-1103.49
95% prediction coverage for AGB & LiDAR	94.6	94.2	93	91.8	91.6	91.5
RMSPE for AGB observed LiDAR	0.304	0.302	0.298	0.303	0.296	0.298
CRPS for AGB observed LiDAR	19.11	18.92	18.74	19.02	18.62	18.75
GRS for AGB observed LiDAR	153.35	156.75	158.25	154.26	160.24	158.42
95% prediction interval coverage for AGB observed LiDAR	92.9	93.8	93.8	92.9	93.8	94.6
95% prediction interval width for AGB observed LiDAR	1.05	1.09	1.05	1.06	1.07	1.08



(a) Observed LiDAR signal

(b) Predicted LiDAR signal with $n^*u = 339, n^*x = 6$ (c) Predicted LiDAR signal with $n^*u = 170, n^*x = 6$ **Fig. 5.** Posterior predictive median for LiDAR signals along the example transect denoted in Fig. 3.

Results for the intercept only and $v(\mathbf{s}) = 0$ candidate models are presented in the Supplementary material. For the $v(\mathbf{s}) = 0$ models, fit and prediction is only influenced by the choice of n_x^* and n_u^* . Comparing results between the intercept only model and $v(\mathbf{s}) = 0$ suggest that inclusion of $v(\mathbf{s})$ has little effect on the *best* model selected using goodness-of-fit and out-of-sample prediction validation metric within each model set, i.e., n_u^* equal 339 and 170 (Supplementary material Tables S1–S8). This suggests that the underlying process seems to be driven by features shared between AGB and LiDAR and there is negligible information on features specific to AGB that are not shared by LiDAR. Hence, the shared AGB–LiDAR process pursued in this work. However, inclusion of $v(\mathbf{s})$ does marginally improve fit to the data and prediction. This improvement suggests there is

some spatial structure in the residuals of AGB that is not captured by information from the LiDAR signals.

It is useful to consider a 2-dimensional slice through the data to further assess candidate model results. Fig. 5 (a) is the side-view of the observed LiDAR signals along the example transect denoted in Fig. 3. Analogous to the portrayal in Fig. 1 (b), larger values of the signal correspond to greater density of tree material; hence, one could imagine Fig. 5 (a) is like looking at the side of a forest (15 m in width and ~700 m in length) where greater values correspond to denser forest. Lower values in Fig. 5 (a) could occur because there is no forest (i.e., above the canopy extent), sparse forest, or overstorey acts to block the LiDAR signal from penetrating into the lower portions of the forest. MU boundaries are also superimposed on Fig. 5 (a) and

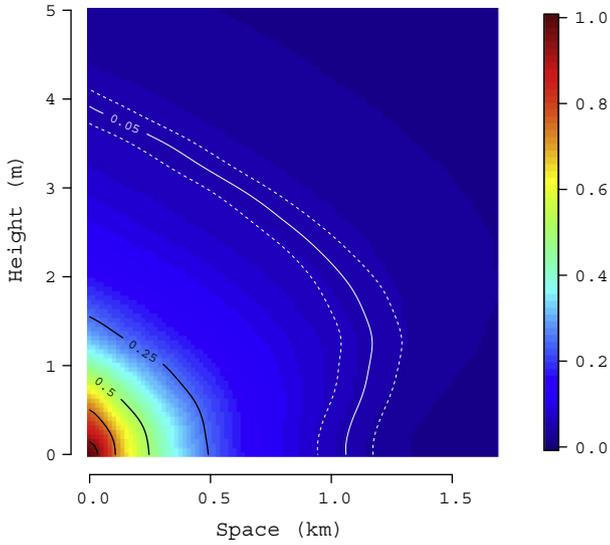


Fig. 6. Space-height correlation posterior median surface and contours. Median (solid white lines) and associated 95% credible interval (dotted white lines) for 0.05 correlation contour.

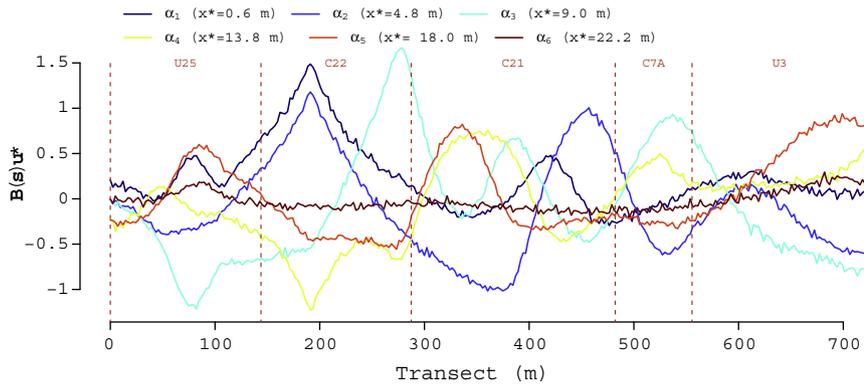
clearly show how different silvicultural treatments (i.e., tree harvesting) result in different vertical and horizontal distributions of tree material. For example, MU C22 is a young, short stature, forest versus the older, taller, and more vertically homogeneous forest in MU U3. As seen in Fig. 3, no PSPs or knots in S_u^* fall on the transect and

hence the LiDAR signals in Fig. 5 (a) were not used for parameter estimation.

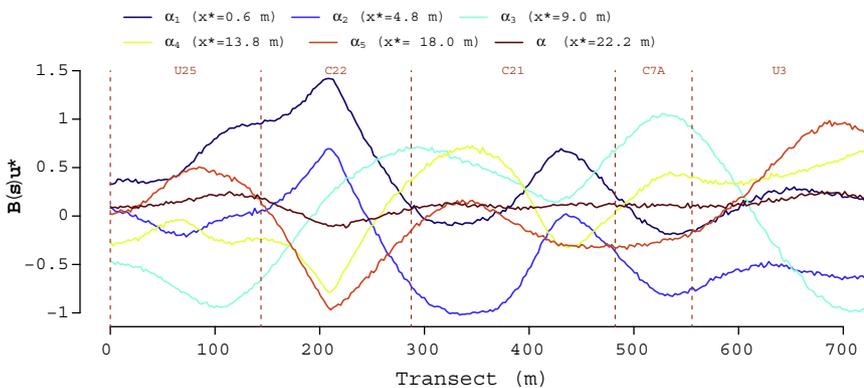
For brevity, we subsequently consider $n_x^* = 6$ candidate models but note that, in general, values of $n_x^* \geq 4$ yield comparable results. Signal prediction along the example transect using the $n_x^* = 6$ models are given in panels (b) and (c) in Fig. 5. Comparison between these predictions and observed data, Fig. 5 (a), shows the candidate models capture the dominant trends in LiDAR signals. Reducing u^* process knots by half, i.e., moving from the $n_u^* = 339$ to $n_u^* = 170$ model, does not greatly affect the vertical and horizontal distribution of predicted signal values.

The observed signal data, Fig. 5, suggest a strong space-height process interaction. The strength of this interaction is captured by γ in covariance function (Eq. (2)), where values close to one indicate strong interaction and values close to zero indicate weak interaction. Parameter estimates for γ in Tables 2 and 3 do indeed corroborate the presence of strong interaction between space and height. Fig. 6 summarizes estimated space-height correlation and shows the median posterior correlation surface and associated contours using posterior samples from the $n_u^* = 170$ and $n_x^* = 6$ model. Here, at a given height the spatial correlation is small (i.e., 0.25) at ~ 0.5 km and negligible (i.e., 0.05) at ~ 1 km. This makes sense because the average areal extent of the MUs is a bit less than a half kilometer. Within a given signal, i.e., at a given spatial location, the correlation drops to 0.05 at ~ 4 m. Again, looking at Figs. 1 (b) and 5(a), we see fairly weak correlation in any given signal beyond several meters for most MUs.

Panels (a) and (b) in Fig. 7 display the posterior median for each latent LiDAR regressor along the example transect that correspond to the $n_x^* = 6$ model α estimates in Tables 2 and 3, respectively. Figure legends also include the x^* knot height associated with each latent



(a) Predicted latent LiDAR process with $n_u^* = 339, n_x^* = 6$



(b) Predicted latent LiDAR process with $n_u^* = 170, n_x^* = 6$

Fig. 7. Posterior predictive median of latent LiDAR signal regressors along the example transect denoted in Fig. 3. The legend relates each regressor to the corresponding element in α with the predictive process knot height in meters is given in parentheses.

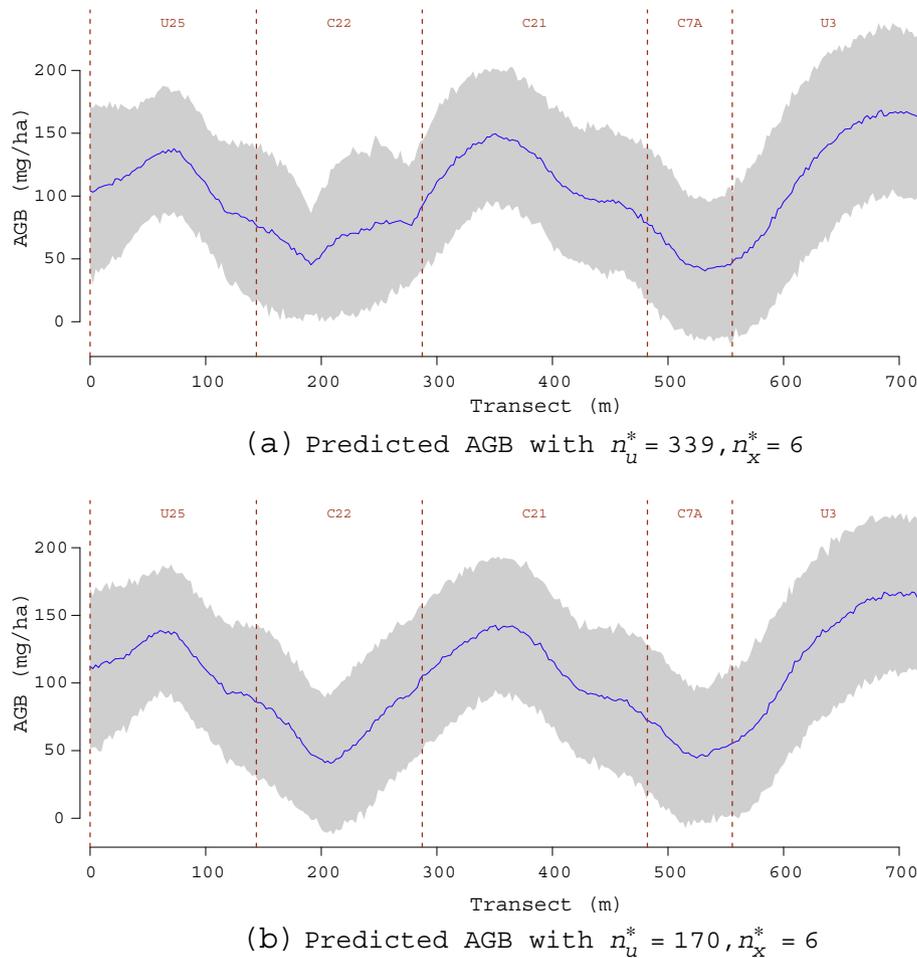


Fig. 8. Posterior predictive median for AGB along the example transect denoted in Fig. 3. MU identifiers are provided across the top of each panel.

regressor. The latent regressors interpretation becomes clear when Fig. 5 (a) is considered with panel (a) or (b) in Fig. 7. For example, Fig. 5 (a) shows most energy returns in MU C7A (between ~490 and 560 m along the example transect) are between ~5 and 12 m in height, hence we see large values of the latent regressor associated with $x^* = 9$ m and α_3 in panels (a) and (b) in Fig. 7. Similarly, paucity of energy returns in the ~1–5 m height at ~300–370 m along the example transect, results in small values of the latent regressors associated with x^* equal to 0.6, and 4.8 m. Deviations seen between the latent regressor lines in panels (a) and (b) in Fig. 7 and trends in Fig. 5 (a) are due to process smoothing that results from lack of PSPs and S_{ij}^* occurring on the example transect. This smoothing also accounts for difference between panels (a) and (b) in Fig. 7.

Significant α parameter estimates suggest LiDAR signal trends captured by the low-rank models are useful for explaining variability in AGB. The impact of latent LiDAR regressors is seen in Fig. 8 where both models capture AGB trends within MUs and along the example transect. Clearly spatial smoothing occurs—there should likely be more abrupt changes in median AGB across MUs—however there is nothing to inform AGB predictions except for the elevation predictor variable, representation of the LiDAR signals, and residual spatial random effects, all three of which are smoothly varying across the domain. Other candidate models, including those presented in the Supplementary material, produce similar AGB profiles. We could certainly add a MU indicator or additional location specific predictors to help inform AGB prediction. However, again, these data are rarely available in applied settings and a key objective of this analysis was

to assess the usefulness of the latent LiDAR regressors for modeling AGB. Indeed, even lacking additional location specific predictor variables the candidate models yield very useful AGB data products that are critical inputs to forest management and MRV systems. For example, Fig. 9 offers candidate models' AGB posterior predictive median and associated measure of uncertainty at a 15×15 meter resolution for the entire PEF. This figure shows the candidate models deliver nearly identical AGB prediction and uncertainty maps despite the large reduction in space-height process dimension. As expected, more precise AGB prediction occurs in proximity to observed PSP as shown by narrow 95% CI intervals in panels (b) and (d) in Fig. 9.

5. Summary

We have developed and implemented a class of Bayesian hierarchical models to jointly model LiDAR signals and AGB and effectively exploit the information from the high-dimensional LiDAR signals to explain AGB variability. We account for spatial dependence among and within the high-dimensional LiDAR signals and predict the LiDAR signals and AGB at arbitrary spatial locations and heights. We circumvent computational bottlenecks presented by the LiDAR signal dimensionality and number of spatial locations by applying reduced-rank predictive processes, a collapsed MCMC framework, and some efficient numerical linear algebra.

We opted for a fully process-based approach using covariance functions to exploit the easy constructibility and interpretability of

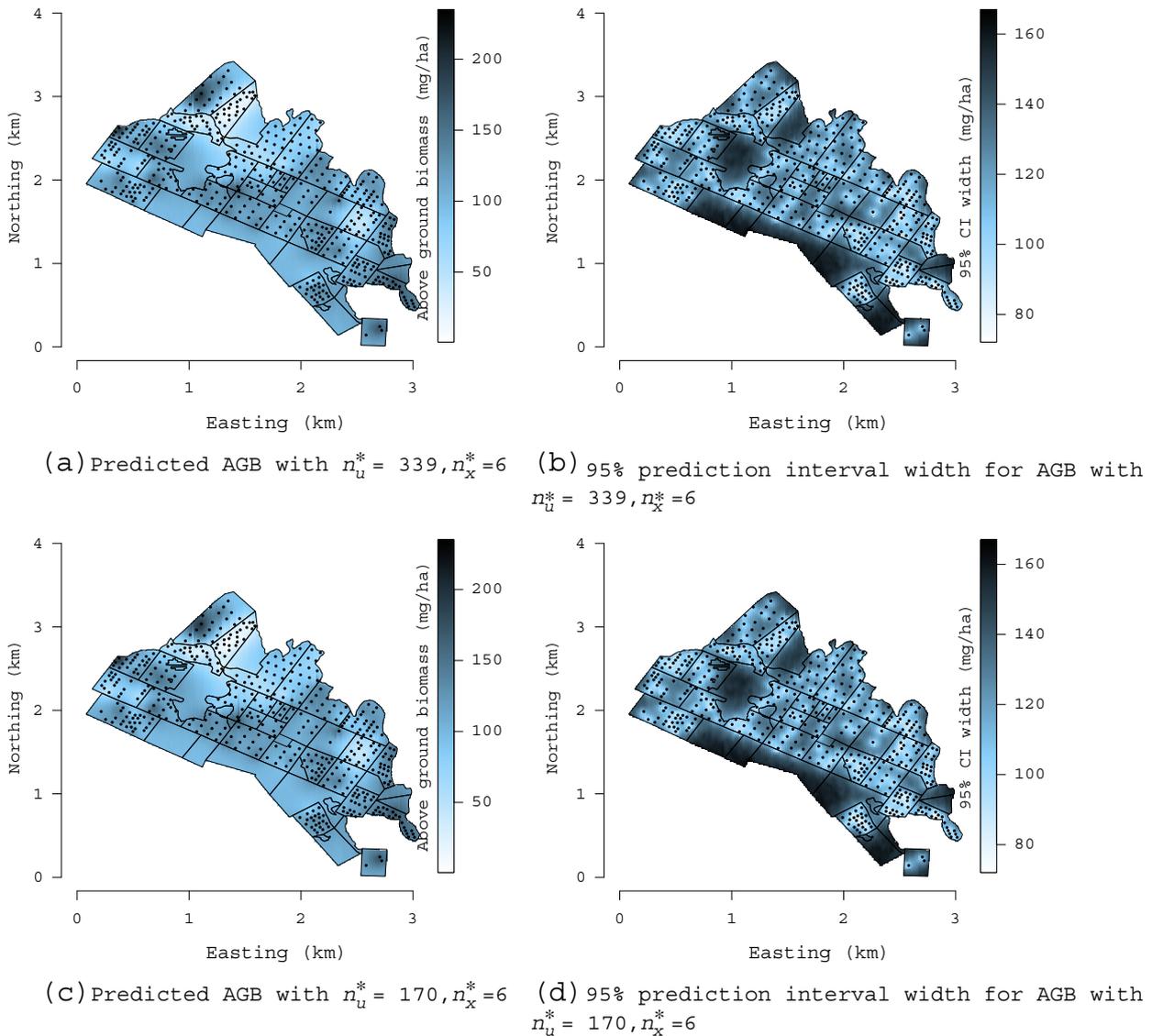


Fig. 9. Posterior predictive distribution's median and width of 95% prediction interval for AGB.

the joint models. Alternative approaches could build upon existing functional data models that treat the high-dimensional signals as a function of space and height. For instance, one could possibly adapt the approach of Yang et al. (2015), who mapped agricultural soil properties, to build joint AGB-LiDAR models. Properties of these models and, in particular, their scalability to massive datasets still need to be explored.

Substantive contributions from the current PEF analysis include LiDAR-based maps of AGB with associated uncertainty that can inform analyses of MU specific silvicultural experiments and also serve as baseline estimates, with uncertainty, for future management and experiments. More broadly, we believe this modeling framework will be employed for future explorations and analysis relating LiDAR and similar high-dimensional signal data—generated by the missions detailed in Section 1—with AGB and other forest variables of interest. Future methodological extensions include analyzing several forest variables (e.g., AGB by tree species or structural variables such as density and basal area) perhaps correlated among themselves, as well as accounting for spatiotemporal associations. There is also considerable interest in adapting the proposed framework to model non-Gaussian forest variables such as forest/non-forest, fire

risk categories, and species or functional types. We plan to extend this joint modeling framework to accommodate additional sparsely sampled high-dimensional signal data such as hyper-spectral data that is similar to LiDAR but records information across the electromagnetic spectrum and can provide information on forest species or tree health status.

Our focus was on modeling the space-height structure of LiDAR signals to improve the prediction of *i*) signals at non-sampled locations and *ii*) AGB at locations where the signal may or may not have been observed. If one was interested in modeling vegetation structure, such as leaf area density, then a MacArthur-Horn transformation (MacArthur and Horn, 1969) could be applied in either a pre-processing step prior to model fitting, or in a posterior predictive fashion (one-for-one using samples from $z(\ell)$'s posterior distribution) to generate posterior distributions of the transformed signals. Using the proposed joint model, future work could test if such signal transformation increases the explained variability in AGB (or similar forest variables of interest), via the α coefficients.

Future analysis of LiDAR and forestry data will need to cope with massive amounts of data and increasing demands on model scalability. Here, we could considerably enhance the scalability

of the predictive process using the multi-resolution extensions in Katzfuss (2016), where we construct a sequence of nested predictive processes over a nested partition of the spatial domain. Alternatively, recent developments in massively scalable sparsity-inducing Nearest-Neighbor Gaussian Processes or NNGPs (Datta et al., 2016) can be exploited. Our framework seamlessly accommodates such processes—we replace $u(\ell)$ and $v(s)$ in Eq. (5) with their NNGP counterparts instead of predictive processes. Rather than dimension reduction, scalability will be achieved exploiting sparsity structures in the resulting precision matrices.

Acknowledgments

Andrew Finley was supported by the National Science Foundation (NSF) DMS-1513481, EF-1137309, EF-1241874, and EF-1253225, as well as NASACarbon Monitoring System grants. Sudipto Banerjee was supported by NSF DMS-1513654 and IIS-1562303.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.rse.2016.12.004>.

References

- Abdalati, W., Zwally, H., Bindschadler, R., Csatho, B., Farrell, S., Fricker, H., Harding, D., 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. *Proc. IEEE* 98, 735–751.
- 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 Sens. Environ.* 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.
- 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 J. Sel. Top. Appl. Earth Obs. Remote Sens.* 6, 6–14.
- Banerjee, S., Gelfand, A.E., Finley, A.O., Sang, H., 2008. Gaussian predictive process models for large spatial data sets. *70*, 825–848.
- Blair, J., Hofton, M., 1999. Modeling laser altimeter return waveforms over complex vegetation using high-resolution elevation data. *Geophys. Res. Lett.* 26, 2509–2512.
- CMS, 2010. NASA carbon monitoring system. Accessed: 5-22-2014. <http://carbon.nasa.gov>.
- 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 Sens.* 5, 4045–4066.
- Datta, A., Banerjee, S., Finley, A.O., Gelfand, A.E., 2016. Hierarchical nearest-neighbor gaussian process models for large geostatistical datasets. *J. Am. Stat. Assoc.* In press.
- Finley, A., Banerjee, S., MacFarlane, D., 2011. A hierarchical model for quantifying forest variables over large heterogeneous landscapes with uncertain forest areas. *J. Am. Stat. Assoc.* 106, 31–48.
- Finley, A.O., Banerjee, S., Weiskittel, A.R., Babcock, C., Cook, B.D., 2014. Dynamic spatial regression models for space-varying forest stand tables. *Environmetrics* 25, 596–609.
- Finley, A.O., Sang, H., Banerjee, S., Gelfand, A.E., 2009. Improving the performance of predictive process modeling for large datasets. *Comput. Stat. Data Anal.* 53, 2873–2884.
- GEDI, 2014. Global ecosystem dynamics investigation LiDAR. Accessed: 1-5-2015. <http://science.nasa.gov/missions/gedi/>.
- Gelfand, A.E., Ghosh, S.K., 1998. Model choice: a minimum posterior predictive loss approach. *Biometrika* 85, 1–11.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. *Bayesian Data Analysis*. third, Chapman & Hall/CRC Texts in Statistical Science, Chapman & Hall/CRC.
- Gneiting, T., 2002. Nonseparable, stationary covariance functions for space-time data. *J. Am. Stat. Assoc.* 97 (458), 590–600.
- Gneiting, T., Genton, M.G., Guttorp, P., 2006. Geostatistical space-time models, stationarity, separability and full symmetry. In: Finkenstaedt, B., Held, L., Isham, V. (Eds.), *Statistical Methods for Spatio-temporal Systems*. Chapman and Hall CRC Press, pp. 151–175.
- Gneiting, T., Guttorp, P., 2010. Continuous-parameter spatio-temporal processes. In: Gelfand, A.E., Diggle, P., Fuentes, M. (Eds.), *Handbook of Spatial Statistics*. Chapman and Hall/CRC, pp. 427–436.
- Gneiting, T., Raftery, A.E., 2007. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.* 102, 359–378.
- Gonzalez, P., Asner, G., Battles, J., Lefsky, M., Waring, K., Palace, M., 2010. Forest carbon densities and uncertainties from LiDAR, quickbird, and field measurements in California. *Remote Sens. Environ.* 114, 1561–1575.
- Hanks, E.M., Schliep, E.M., Hooten, M.B., Hoeting, J.A., 2015. Restricted spatial regression in practice: geostatistical models, confounding, and robustness under model misspecification. *Environmetrics* 26, 243–254. *Env.2331*.
- 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.
- ICESat-2, 2015. Ice, cloud, and land elevation satellite-2. Accessed: 1-5-2015. <http://icesat.gsfc.nasa.gov/>.
- 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. *Int. J. Appl. Earth Obs. Geoinf.* 23, 109–118.
- Katzfuss, M., 2016. A multi-resolution approximation for massive spatial datasets. *J. Am. Stat. Assoc.* In press.
- 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 Sens. Environ.* 115, 2850–2860.
- MacArthur, R., Horn, H., 1969. Foliage profile by vertical measurements. *Ecology* 50, 802–804.
- Mateu, J., Müller, W., 2012. *Spatio-temporal Design: Advances in Efficient Data Acquisition*. John Wiley & Sons, Ltd.
- Muss, J., Mladenoff, D., Townsend, P., 2011. A pseudo-waveform technique to assess forest structure using discrete lidar data. *Remote Sens. Environ.* 115, 824–835.
- Næsset, E., 2011. Estimating above-ground biomass in young forests with airborne laser scanning. *Int. J. Remote Sens.* 32, 473–501.
- 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 Sens. Environ.*, 274–287.
- 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. *Clim. Chang.* 1–16.
- Popescu, S., Zhao, K., 2008. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote Sens. Environ.* 112, 767–781.
- 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. *Can. J. For. Res.* 33, 2116–2128.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. *J. R. Statistical Soc. B* 64, 583–639.
- Stage, A., 1976. An expression for the effect of aspect, slope, and habitat type on tree growth. *For. Sci.* 22, 457–460.
- Stark, S.C., Enquist, B.J., Saleska, S.R., Leitold, V., Schiatti, J., Longo, M., Alves, L.F., Camargo, P.B., Oliveira, R.C., 2015. Linking canopy leaf area and light environments with tree size distributions to explain amazon forest demography. *Ecol. Lett.* 18, 636–645.
- Stein, M.L., 1999. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer.
- Stein, M.L., 2005. Space-time covariance functions. *J. Am. Stat. Assoc.* 100, 310–321.
- Stockton, N., 2014. How a flying laser built a 3-d map of a massive Alaskan forest. Accessed 2015-06-01. <http://www.wired.com/2014/12/alaska-laser-survey-3d-map>.
- Tonolli, S., Dalponte, M., Neteler, M., Rodeghiero, M., Vescovo, L., Gianelle, D., 2011. Fusion of airborne LiDAR and satellite multispectral data for the estimation of timber volume in the Southern Alps. *Remote Sens. Environ.* 115, 2486–2498.
- UN-REDD, 2009. The UN-REDD programme. Accessed: 5-22-2014. <http://www.un-redd.org/>.
- UNFCCC, 2015. United Nations framework convention on climate change. Accessed: 6-6-15. <http://unfccc.int/2860.php>.
- West, P., 2004. *Tree and Forest Measurement*. Springer. Tree and Forest Measurement.
- Wikle, C.K., 2010. Low-rank representations for spatial processes. In: Gelfand, A.E., Diggle, P., Fuentes, M., Guttorp, P. (Eds.), *Handbook of Spatial Statistics*. Chapman and Hall/CRC, pp. 107–118.
- Xia, G., Miranda, M.L., Gelfand, A.E., 2006. Approximately optimal spatial design approaches for environmental health data. *Environmetrics* 17, 363–385.
- Yang, W.H., Wikle, C.K., Holan, S.H., Myers, D.B., Sudduth, K.A., 2015. Bayesian analysis of spatially-dependent multi-dimensional functional predictors. *Stat. Sin.* 25, 205–223.
- Yeniay, O., Goktas, A., 2002. A comparison of partial least squares regression with other prediction methods. *Hacet. J. Math. Stat.* 31, 99–111.