2019

Resolution dependence in an area-based approach to forest inventory with airborne laser scanning

Packalen, Petteri

Elsevier BV

http://dx.doi.org/10.1016/j.rse.2019.01.022

Title Resolution Dependence in an Area-based Approach to Forest inventory with Airborne Laser Scanning

Authors Petteri Packalen[a], Jacob Strunk[b], Tuula Packalen[c], Matti Maltamo[d] and Lauri Mehtätalo[e]

[a] petteri.packalen@uef.fi, School of Forest Sciences, Faculty of Science and Forestry, University of Eastern Finland, P.O. Box 111, 80101 Joensuu, Finland. Phone +358-029-445-1111. Corresponding author.

[b] jstrunk@fs.fed.us, USDA Forest Service, 3625 93rd Ave SW, Olympia, WA 98512.

[c] tuula.packalen@luke.fi, Natural Resources Institute Finland (Luke), Yliopistokatu 6, 80100 Joensuu, Finland.

[d] matti.maltamo@uef.fi, School of Forest Sciences, Faculty of Science and Forestry, University of Eastern Finland, P.O. Box 111, 80101 Joensuu, Finland.

[e] lauri.mehtatalo@uef.fi, School of Computing, University of Eastern Finland, P.O. Box 111, 80101 Joensuu, Finland.

Abstract

In an Area Based Approach (ABA) to forest inventories using Airborne Laser Scanning (ALS) data, the sample plot size may vary or the cell size may differ from the plot size. Although this resolution mismatch may cause bias and increase in prediction error, it has not been thoroughly studied. The aim of this study was to clarify the meaning of resolution dependence in ABA, and to further identify its causal factors and quantify their effects. In general, a number of factors contribute to resolution dependence in ABA forest inventories, including the varying point density of the ALS data, the type of response variable, how the predictor variables are computed, and the properties of the prediction model. For quantification, we used field plots with mapped tree locations, which enabled the generation of different sized sample plots inside a larger plot. Plot level above ground biomass (AGB) was the response variable employed in all the models. The error rate seemed to increase when the prediction plots were larger than

the fitting plots, and vice versa. The maximum BIAS was 1.50 % and the maximum change of RMSE compared to its value in native resolution was 0.97 % when there was a 4-fold difference in resolution. This indicates that the resolution effect is small in most real-world use cases, however, resolution effect should be carefully considered in ALS-assisted large area inventories that target unbiased estimates of forest parameters.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Keywords

Scale dependence, resolution invariance, Airborne laser scanning, Forest inventory, Lidar

Introduction

Resolution dependence

The construction of a spatial model often involves the specification of input parameters at some chosen spatial scale, i.e. resolution. The resolution may vary with model predictions and may be different from that used in the model construction. This is quite typical in remote sensing and a large number of studies have addressed this issue from different point of views (e.g. Strahler et al., 1986; Turner et al., 1989; Raffy, 1992; Marceau and Hay, 1999; Simic et al., 2004; Chasmer et al., 2009). In general, the conversion from a high resolution to a lower resolution is called *upscaling* and the conversion from low resolution to a higher resolution is analogously called *downscaling* (Liang, 2004). Several authors have proposed frameworks in order to solve scaling problems and compensate for scaling effects (Wu and Li, 2009).

The terms *scale* and *resolution* refer to different concepts that depend on the context, discipline, and author. In ecology, for example, the concept of *scale* often refers to the extent and the concept of *resolution* to grain size, but there is not necessarily a connection between the grain size and spatial

resolution. In some studies, however, spatial scale refers to the grain size or the grain size is assumed to be equal to the spatial resolution, i.e. pixel size. Correspondingly, *scale invariance* and *resolution invariance* are interpreted differently depending on the circumstances. In this study, we will use the term *resolution* to refer to grain size, although in some previous remote sensing studies the term *scale* was used to refer to more or less the same concept (e.g. Zhao et al., 2009). We will explain later in greater detail what resolution means in the context of our study.

Area-based approach to forest inventory

Forest inventories employing Airborne Laser Scanning (ALS) data have become common in many countries (Nilsson et al., 2017). The ALS-based forest inventory methods (Hyyppä et al., 2008) can be divided into two groups: the area-based approach (ABA) (Means et al., 2000; Næsset, 2002; Magnussen et al., 2013) and single-tree detection (Hyyppä et al. 2001; Koch et al., 2006; Lähivaara et al., 2014). Operational forest inventories employing ALS data are most often implemented with the ABA (Maltamo et al., 2014, chapters 11-13). In ABA, metrics used as predictor variables are calculated from the ALS returns within a plot or grid cell. Using training plots with field-measured stand attributes, a model is formulated between the stand attributes and ABA metrics. This model is then used to predict the stand attributes for each grid cell. A model is typically based on linear (Næsset, 2002) or non-linear regression (Packalén et al., 2011), but non-parametric regression techniques, such as k nearest neighbor (Hudak et al., 2008) and artificial neural networks (Niska et al., 2010), have also been used. A large proportion of studies have focused on forest management inventories (e.g. Næsset, 2004; Packalén and Maltamo, 2007), but a large number of studies have also focused on statistical inference and sampling designs (e.g. Chirici et al., 2016; Gregoire et al., 2016). In forest management inventories, cell level predictions are typically aggregated to the stand level (Næsset, 2004), whereas in large area strategic inventories, cell level predictions may be used as auxiliary information either in sampling design or in the estimation phases (Grafström et al., 2017; Strunk et al., 2012).

Several factors that affect the performance of ABA have been studied, such as the number of sample plots (Junttila et al., 2013), plot positioning error (Gobakken and Næsset, 2009), ALS echo density (Thomas et al., 2006), ALS sensor and flying configurations (Keränen et al., 2016), and seasonal effects (Næsset, 2005). The effect of sample plot size has also been addressed in many studies (Gobakken and Næsset, 2008; Frazer et al., 2011; Ruiz et al., 2014). In general, a reduction in prediction error with larger plot sizes is expected due to spatial averaging of errors (Goetz and Dubayah, 2011). Chen et al. (2016) showed within a coherent framework that AGB prediction error decreases as plot size increases, and determined the relative contributions of each error source.

Resolution issues in ABA originate from the use of non-constant size plots and/or cells. The most convenient option is use the same sized plots and cells, but this is not always possible. There are several reasons why an analyst may have a need to deal with data at different resolutions:

1) *Sample plot sizes vary*. It is common to measure sample plots of different sizes in a range of forest types to maximize the efficiency of field work – small plots in young or homogeneous forests, and large plots in sparse, mature or heterogeneous forests. A similar approach is used in some ABA inventories.
2) *Cell sizes vary*. It may be desirable to split cells (corresponding plot size) into pieces and merge them with neighboring cells. This approach is useful if we want to avoid "mixed" cells in stand borders in order to reduce prediction errors on borders. For example, grid cells located on the border between a sapling and a sawlog stand may greatly overestimate timber volume or mean diameter in the saplings.
3) *Sample plot size is different to cell size*, although plots are of similar size and cells are of similar size. For instance, an end user may require a certain grid cell size, but an existing set of sample plots or measurement protocols must be used e.g. for economic reasons.
4) *A combination of cases 1, 2 and 3*.

The effect of varying resolution in the ABA arena is an almost unexplored field of study. Zhao et al. (2009) developed methods for the scale-invariant estimation of forest biomass using functional regression models. The models are called functional regression models because the predictors, namely canopy height distribution (CHD) and canopy height quantile functions (CHQ), are themselves functions or functional data. The CHD and CHQ were computed from a ALS-derived canopy height model (CHM) rather than from the ALS point cloud. There are also ABA studies that address resolution or scale but from an entirely different viewpoint than this study. For example, Magnussen et al. (2016) demonstrated a statistical method for upscaling ABA predictions in cases where the response variable is not additive, whereas Chen et al. (2016) proposed a new uncertainty analysis method built on model-based inference that can characterize biomass uncertainty across multiple spatial resolutions. Our study is distinct from the previous research because we clarify the factors that cause resolution issues in a typical ABA, and with an empirical setting, we evaluate how serious the consequences might be if the requirements of resolution invariance are violated. There is a need to understand more thoroughly the resolution dependence in the ABA arena. For instance, we did not find any studies that discuss the resolution issues that originate from the irregular point pattern of ALS data.

Regular or irregular point spacing?

We examine resolution effects in the context of ABA wherein metrics are computed from an ALS point cloud, in contrast to the majority of remote sensing studies where resolution issues are addressed in the context of raster images. This distinction is important: an image is a regular tessellation whereas ALS data are irregularly spaced in a horizontal domain. When metrics are computed from images, the number of observations (i.e. pixels) is constant per unit area. In an irregular point pattern, such as ALS data, the point density per unit area varies. The implication of areas with higher and lower point densities is that areas with higher point density are more heavily weighted than areas with lower point density. For example, let us assume that the left side of a rectangular grid cell has a double point density compared to the right side and a computed metric is e.g. the mean height of points. The mean value of the metric over the cell area is then equal to:

$$\frac{2 \times mean\ height\ of\ the\ left\ side + 1 \times mean\ height\ of\ the\ right\ side}{2+1}. \quad (1)$$

However, there is no reason to assume that areas with a higher point density are more important than areas with a low point density. The un-equal weighting is simply an undesirable consequence of irregularities in the data acquisition process that can result, for example, from scan pattern, varying scanning angle, or overlapping flight lines. The effects of variable point density on ABA inference are also resolution dependent because the weighting scheme for points within cells also changes as the resolution changes. For example, when the resolution becomes finer, areas with low point density are less likely to be combined with high point density areas.

Objectives

The overall aim of this study is to clarify the meaning of resolution dependence in the context of ABA, to identify the underlying causal factors, and to quantify their effects. Our study design enables us to examine the effects of using up to nine-fold differences in plot and cell sizes. Specific objectives are:

- To identify requirements for resolution invariance.
- How do the bias and error rates behave when the requirements of resolution invariance are violated?
- What is the direction and magnitude of bias with the commonly used ALS height percentiles when the resolution is changed?
- To present a simple approach to achieve resolution invariance in ABA by considering varying point density.

Theoretical background

Definition of resolution invariance

Consider a situation where a predictive model based on field plots is used to make prediction for a large area. In principle, we examine two alternatives to compute the prediction for the area: (1) Divide the area

to cells of a certain size (e.g. 15 x 15m), compute the predictors for each cell and use them for predictions, and average or sum the cell-level predictions over the cells. (2) Compute the predictors directly for the whole area and apply them for the whole area as if it were one cell. We define *resolution invariance* in ABA to mean that both approaches will yield an identical prediction. In this study, we focus on resolution discrepancy common in practice due to a resolution mismatch between the plot area used in model fitting and the area covered by a cell from an ALS-derived predictor grid. We do not assume that the least squares model coefficients are identical when fitted at the different resolutions, but instead look at the behavior of predictions when fitted at one resolution and used for predictions at a different resolution. In this study, we only consider continuous response variables, although the theory and findings partly apply to categorical response variables as well.

Resolution invariant ABA model

Consider a situation where spatial resolution affects the values of ALS-derived metrics only through the number of ALS echoes, which is fixed and proportional to cell size. This is true in a grid where the number of ALS echoes per grid cell is the same for all cells of a given size. Obviously, this is not the case in most real world applications, because ALS data consist of irregularly spaced echoes in horizontal space.

Consider a case where the cells of area a_i (e.g. 625 m^2 in Figure 1, left panel) are formed by k_i non-overlapping subcells of areas a_{ij} (e.g. 156.25 m^2 in Figure 1, mid panel or 69.44 in the right panel), so that $a_i = \sum_{j=1}^{k_i} a_{ij}$. The forest attribute for cell *i* is y_i and for its subcell *j* it is $y_{ij.}$ Variable *y* is said to be additive if $a_i y_i = \sum_{j=1}^{k} a_{ij} y_{ij}$. The measures of mean or total quantity per area unit: total volume, above ground biomass, basal area and number of stems, are additive. Measures of mean or total quantity per tree: dominant height, mean height or mean diameter, for example, are not additive. There are, however, special cases when the mean diameter or height of trees is additive, this happens whenever the number of trees per area unit is constant.

Consider a metric x_{ij}, which is based on the ALS measurements of the canopy on subcell *j* within the

larger cell *i*. It is typically the mean of echo heights, a certain quantile of the echo heights, or proportion of echoes above, below, or between fixed heights. A metric is said to have a *resolution invariant mean* if the (optionally) weighted mean of subcell metrics is equal to the metric computed for the entire cell *i*:

$$\sum_{j=1}^{k} w_{ij} x_{ij} = x_i, \quad (2)$$

where weights w_{ij} is rescaled subcell area a_{ij}, scaled such that subcell-level weights sum to one at cell-level:

$$w_{ij} = \frac{a_{ij}}{\sum_{j=1}^{k} a_{ij}}. \quad (3)$$

Assuming a grid where the number of ALS echoes per cell and subcell is constant, a_{ij} can be alternatively the number of echoes per subcell. Some ALS metrics satisfy the condition of a resolution invariant mean, such as the mean height of echoes without any height threshold and many metrics of proportion above, below, or between certain fixed heights. Height quantiles are an example of ALS metrics that are not resolution invariant as they are not unbiased for the corresponding population parameters: upper sample quantiles produce an underestimation of the corresponding population quantiles and lower quantiles produce an overestimation. The biases are greatest for the extreme quantiles (Reiss, 1989), and are affected by the shape of the underlying distribution of ALS echoes, as well as by the number of ALS echoes. In the ALS arena, quantiles are typically called percentiles and hereafter we will follow this convention.

Consider an ABA forest inventory, where the prediction is based on the model:

$$y_i = f(x_i; \boldsymbol{\beta}) + e_i, \quad (4)$$

where x_i is the vector of ALS metrics in the larger cell *i* and the residual e_i has a mean of zero. Let us assume that the parameter vector $\boldsymbol{\beta}$ is fixed and known so that:

$$E(y_i) = f(x_i; \boldsymbol{\beta}), \quad (5)$$

192 which is the expected value of y_i. The model is called resolution invariant if the weighted sum of subcell-
193 level predictions equals the cell-level prediction:

194 $$E(y_i) = \sum_{j=1}^{k} w_{ij} f(x_{ij}; \boldsymbol{\beta}). \quad (6)$$

195 Under the linear model, this yields:

196 $$E(y_i) = \sum_{j=1}^{k} w_{ij} \left[\beta^{(1)} + \beta^{(2)} x_{ij}\right] = \beta^{(1)} \sum_{j=1}^{k} w_{ij} + \beta^{(2)} \sum_{j=1}^{k} w_{ij} x_{ij}. \quad (7)$$

197 Assuming that x_{ij} have a resolution invariant mean (equation 2), this yields

198 $$E(y_i) = \beta^{(1)} + \beta^{(2)} x_i, \quad (8)$$

199 which is exactly the same as the direct prediction at the cell level. Equations 7 and 8 generalize to
200 multiple linear models as well.

201 The additivity of the response variable is usually not achieved if nonlinear transformations are used in the
202 predictors. For example, $a_i \ln(y_i)$ is not equal to $\sum_{j=1}^{k_i} a_{ij} \ln(y_{ij})$. One might consider overcoming this
203 problem by using back-transformed predictions from a linear model of transformed y, but the right-hand
204 side of the prediction model would not be a linear function of predictors. In addition, the treatment of the
205 back-transformation bias of such models would be an additional issue in such an approach.

206 To summarize, resolution invariance in ABA is obtained at least when all of the following four conditions
207 are met:

208 Condition 1: ALS echoes are regularly spaced in the horizontal dimension
209 Condition 2: the forest variable (y) of interest is additive
210 Condition 3: the ALS-based metrics (x) have a resolution invariant mean
211 Condition 4: a linear model is used without nonlinear transformations in y

212 Condition 1 has not received much attention in the literature thus far because the literature on this topic
213 tends to deal with regularly spaced data. Condition 2 is not always satisfied because some inventory

attributes are not additive in nature. Commonly used metrics in ABA do not satisfy condition 3, but in principle, some ABA metrics are satisfactory provided condition 1 is also satisfied. Condition 4 can be satisfied, for example, with a classical linear regression model (Searle, 1971).

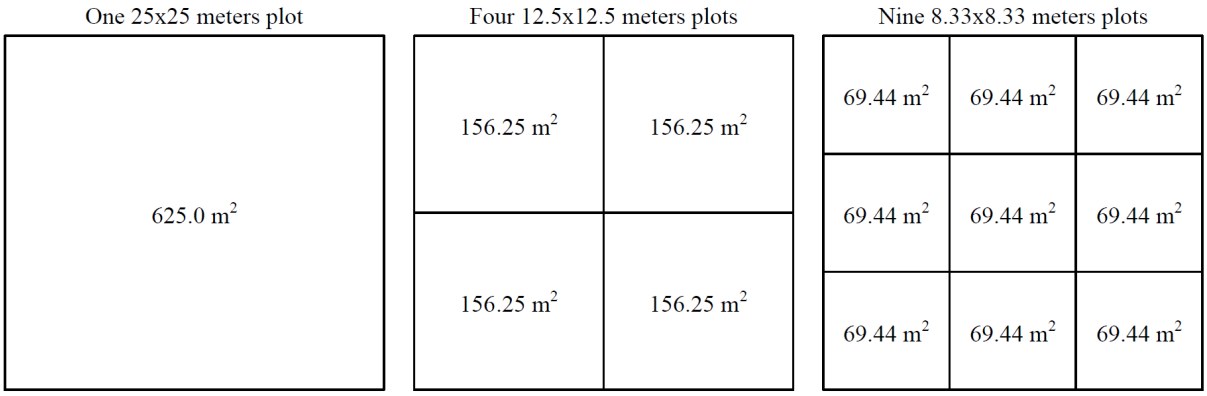
Material

Field data

The study area is a boreal managed forest area in eastern Finland (62°31'N, 30°10'E). Field measurements were carried out in the summer of 2010 and sample plots were deliberately placed to represent the tree size and species variation for the study area. The allocation of field plots was based on the development stage of the forest and the dominant tree species. Scots pine (*Pinus sylvestris* L.) is the dominant tree species representing about 75 % of the volume and the remainder consists of Norway spruce (*Picea abies* [L.] Karst.) and a mixture of native deciduous species. We used field measurements from 58 sample plots (25 x 25 m).

For trees with either a diameter at breast height (DBH) exceeding 4 cm or height exceeding 4 m, the DBH, height, and tree species were recorded. The DBH of each tree was estimated as the average of the maximum diameter and the diameter perpendicular to the maximum diameter. The above ground biomass of the individual trees was calculated as a function of DBH and tree height using the species-specific models developed by Repola (2008) and Repola (2009).

Each 25 x 25 m plot was divided into four 12.5 x 12.5 m plots and nine 8.33 x 8.33 m plots (Figure 1). Delineation of sub-plot boundaries was feasible because the tree locations were known. For a more detailed description of tree location determination see Packalen et al. (2015). In total, there were 58 25 x 25 m plots, 232 12.5 x 12.5 m plots and 522 8.33 x 8.33 m plots. Hereafter, we refer to these as 25.0 m, 12.5 m and 8.33 m resolutions. Plot level biomass per hectare (AGB $Mg \cdot ha^{-1}$) for a plot (or a sub-plot) was computed as the sum of tree biomass on a plot (or a sub-plot), expanded by the ratio of one hectare over the plot area. A summary of AGB by plot size is provided in Table 1.



238

239 Figure 1. Plot setting.

240 Table1. Mean, standard deviation, minimum and maximum above ground biomass (AGB) by plot size.

	Plot size		
	8.33 m	12.5 m	25.0 m
Mean AGB ($Mg \cdot ha^{-1}$)	109.2	109.2	109.2
Std AGB ($Mg \cdot ha^{-1}$)	59.2	45.8	37.1
Min AGB ($Mg \cdot ha^{-1}$)	0.0	25.0	50.6
Max AGB ($Mg \cdot ha^{-1}$)	541.2	310.5	212.0

241 *ALS data*

242 ALS data were collected on June 26, 2009 using an Optech ALTM Gemini laser scanning system. The
243 nominal pulse density was approximately 12 pulses per square meter. The test site was scanned from an
244 altitude of approximately 600 m above ground level, with a field of view of 26 degrees and side overlap
245 between transects of 55 %. Pulse repetition frequency was set to 125 kHz. A Digital Terrain Model
246 (DTM) was constructed by first classifying points as ground and non-ground hits according to the
247 approach described by Axelsson (2000). A raster DTM of 0.5 m spatial resolution was then obtained by
248 interpolation using Delaunay triangulation. Heights above ground (dZ) for ALS echoes were calculated
249 by differencing their elevations above the ellipsoid from corresponding DTM elevations. Hereafter, we
250 refer to dZ as echo height.

251

Methods

Predictor variables

The ALS metrics used as predictor variables were computed in the same way for each plot at different resolutions. The metrics were computed with either first or last echoes. First echoes contain original echo categories "first of many" and "only" and last echoes contain "last of many" and "only". Subscript $_f$ denotes that a predictor variable was computed using a set of first echoes and correspondingly subscript $_l$ denotes that last echoes were used. Regression models contain the following predictor variables:

- Height percentiles [$h5_f$, $h10_f$, $h20_f$, $h30_f$, $h40_f$, $h50_f$, $h60_f$, $h70_f$, $h80_f$, $h90_f$, $h95_f$ and $h95_l$] using echoes with dZ at least 2 m. These variables do not have *resolution invariant means*.
- Mean height [$havg_f$] of echoes. This variable has a *resolution invariant mean* assuming the number of echoes per plot is constant.
- Mean height [$havg_{f>2m}$] of echoes with dZ at least 2 m. This variable does not have a *resolution invariant mean* except if the echo density above 2 m is constant, or the varying echo density is taken into account in some other way (see LLS-RXN).
- Proportions of echoes below the fixed height thresholds of 2 m [$p2m_f$] and 10 m [$p10m_l$] in relation to all echoes of the same echo category. For instance, if the total number of echoes of the same echo category (e.g. first echoes) within a plot is 100 and 40 of these are below 2 m, then $p2m_f$ is assigned a value of 0.4. These variables have *resolution invariant means* assuming the number of echoes per plot is constant.
- Proportions of echoes below the varying height threshold of $havg_l$ [$pavg_l$] in relation to all echoes of the same echo category. For instance, if the total number of echoes of the same echo category (here last echoes) within a plot is 100 and 60 of these are below the height defined by $havg_l$ (e.g. 15 m), then $pavg_l$ is assigned a value of 0.6. This variable does not have a *resolution invariant mean*.

Broadly similar predictor variables have been used in previous ABA studies and operational inventory projects (see e.g. chapters 1, 11 and 12 in Maltamo et al. 2014).

Model forms

We formulated a set of regression models to reveal the effect of changing resolution. Plot level AGB was the response variable in all the models. AGB is an additive response variable and no transformations were carried out. Predictor variables and model forms were selected such that they represent a realistically wide variety of model options in ABA. We examined two types of models; linear (LLS) and nonlinear (NLS). Regression models contain two types of predictor variables; those that have resolution invariant mean (XR) and those that have not (XNR). Note that the predictor variables categorized here as XR have the property of a resolution invariant mean in the statistical sense but are not resolution invariant in the ABA arena because of the non-regular pattern of ALS echoes and/or threshold value used to compute the metric. The name of the equation depicts the type of model and predictors, for instance, LLS-XR denotes that a model is linear and that the predictor variables have a resolution invariant mean.

$$AGB = \beta_0 + \beta_1 \mathrm{havg_f} + \beta_2 \mathrm{p10m_l} + \varepsilon\,, \qquad \text{(LLS-XR, 9)}$$

$$AGB = \beta_0 + \beta_1 \mathrm{h95_f} + \beta_2 \mathrm{pavg_l} + \varepsilon\,, \qquad \text{(LLS-XNR, 10)}$$

$$AGB = \lambda e^{(\beta_0 + \beta_1 \mathrm{havg_f} + \beta_2 \mathrm{p2m_l})} + \varepsilon\,, \qquad \text{(NLS-XR, 11)}$$

$$AGB = \lambda e^{(\beta_0 + \beta_1 \mathrm{h95_l} + \beta_2 \mathrm{pavg_l})} + \varepsilon\,, \qquad \text{(NLS-XNR, 12)}$$

where $\beta_0, \ldots, \beta_2$ are coefficients to be estimated from data, ε the residual vector and λ that forces a zero mean of residuals. When λ is omitted, the NLS-XR and NLS-XNR models provide a slightly non-zero mean of residuals because there is no intercept in the selected model form. The calibration factor is used for convenience because we require the NLS models to provide a zero mean of residuals when model fitting and prediction are done at the same resolution.

In addition to the four models described above, we constructed a model that is entirely resolution invariant. This model makes an implicit assumption that areas with a higher point density are more important than areas with a low point density. This cannot be avoided if resolution invariance is desired and metrics are computed directly from ALS echoes. Resolution invariance is obtained by taking into

account the varying number of echoes per unit of area while computing predictor variables. The model resembles the LLS-XR except that the mean height of ALS echoes was computed from above 2 m echoes. The predictor variable $\mathrm{havg_{f>2m}}$ was rescaled as follows:

$$\Omega(\mathrm{havg_f}) = n_{f>2m}\,\mathrm{havg_{f>2m}}\,, \tag{13}$$

where $n_{f>2m}$ is the number of first echoes having dZ at least 2 m per area unit (e.g. per square meter) in a plot or subplot. The predictor variable $\mathrm{p10m_l}$ was rescaled as follows:

$$\Omega(\mathrm{p10m_l}) = n_l\,\mathrm{p10m_l}\,, \tag{14}$$

where n_l is the number of last echoes per unit area in a plot or subplot. Note that rescaling is always conducted using the number of echoes that are used to compute the metric in question, i.e. height threshold must be taken into account if applied before computing a metric. Thus, the resolution invariant model is as follows:

$$AGB = \beta_0 + \beta_1\Omega(\mathrm{havg_{f>2m}}) + \beta_2\Omega(\mathrm{p10m_l}) + \varepsilon\,, \tag{LLS-RXN, 15}$$

where RXN denotes that XN type predictor variables were rescaled. We also fitted a set of HPERC models to demonstrate the bias properties for the commonly used predictor variable $\mathrm{h[5\ldots95]_f}$ (height percentile computed by excluding echoes near to the ground) when combined with another common predictor $\mathrm{p2m_f}$ (proportion of echoes below 2 m). HPERC models have the form:

$$AGB = \beta_0 + \beta_1\mathrm{p2m_f} + \beta_2\mathrm{h[5\ldots95]_f} + \varepsilon\,, \tag{HPERC, 16}$$

where height percentile varies between 5 and 95. The model form was selected such that it fits well with all the percentiles.

Model fitting, prediction and validation

We fitted the models at the resolutions of 8.33 m, 12.5 m and 25.0 m with the method of least squares using *lm* and *nls* functions available in the R environment (R Development Core Team 2011). Then we

predicted at all resolutions with every model. The performance of models was assessed by cross-validation. First, one 25 m plot was excluded and a model was fitted with other plots either at 8.33 m, 12.5 m or 25.0 m resolution. Then the prediction was made to the excluded 25 m plot, either at 8.33 m, 12.5 m or 25.0 m resolution. Thus, each model was fitted three times (× 58 considering cross-validation) and was then used to predict at its native resolution and two other resolutions. This was repeated with models that violate the resolution invariance conditions (LLS-XR, LLS-XNR, NLS-XR, NLS-XNR) and, as a *proof-of-concept*, with the proposed resolution invariant model (LLS-RXN). All the predictions were aggregated to the 25.0 m level by computing the mean value of sub-plots. The example in Figure 2 describes a model that is first fitted at the 12.5 m resolution, then used to predict at the 8.33 m resolution, and finally predictions are aggregated to the 25.0 m resolution. This setting enables an analysis of resolution effect when a plot size is either smaller or bigger in model fitting than in prediction.

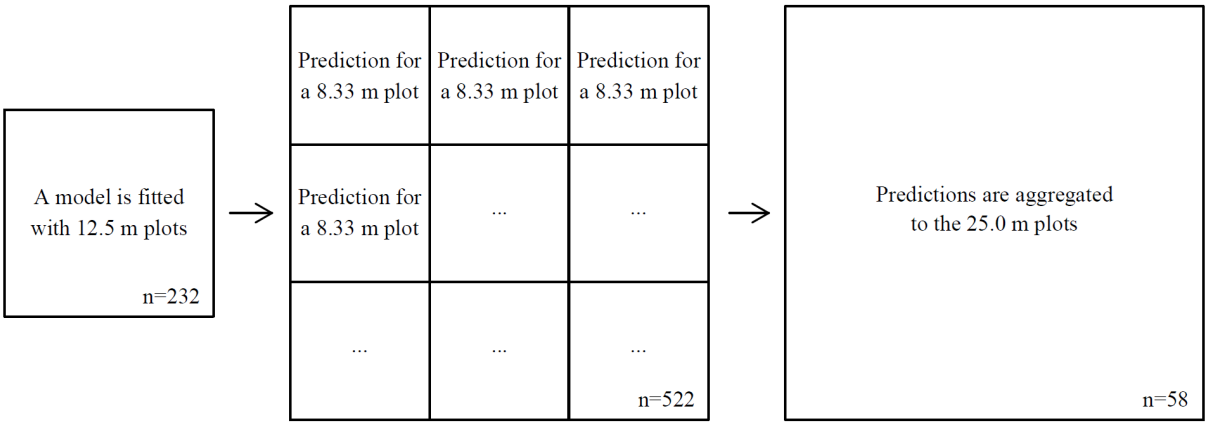


Figure 2. An example of model fitting, prediction and aggregation workflow. For the simplicity, cross-validation is not shown in the figure.

The purpose of HPERC models is to demonstrate the bias properties for the commonly used predictor variable $h[5\ldots95]_f$ when combined with another common predictor $p2m_f$. This was done by fitting HPERC models at resolutions of 8.33 m and 25.0 m. For each height percentile, the 8.33 m model was used to directly predict at a resolution of 25.0 m, whereas the 25.0 m model was used to predict to a resolution of 8.33 m and the predictions were then aggregated to the 25.0 m resolution.

We assess the performance at the 25.0 m resolution using the estimated relative BIAS and RMSE.

$$\text{RMSE-\%} = 100 \times \frac{\sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}}{\bar{y}}, \quad (17)$$

$$\text{BIAS-\%} = 100 \times \frac{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)}{n}}{\bar{y}}, \quad (18)$$

where $\hat{y}_i$ is the predicted value in plot i, y_i is the observed value in plot i, n is the number of plots, and $\bar{y}$ is the mean of observed values.

Results

Error rate

The RMSE values and the change of RMSE compared to its value in native resolution are presented in Table 2. The RMSE decreased if the prediction resolution was higher than the fitting resolution, and conversely, the RMSE increased if the prediction resolution was lower than the fitting resolution. This trend was obvious but not consistent in every case. The effect of varying resolution was slightly stronger when fitting at high resolution and predicting at low resolution than vice versa. In most cases, the lowest error rate was obtained when a model was fitted at the largest resolution and prediction was carried out at the finest resolution. The resolution effect tended to be much smaller with predictor variables that had a resolution invariant mean (LLS-XR and NLS-XR) than with variables that did not have this property (LLS-XNR and NLS-XNR). The model type (LLS or NLS) did not have any apparent effect. The resolution invariant LSS-RXN model had a constant RMSE value under different resolution predictions when the regression coefficients were fixed (i.e. they are solved at a certain resolution). This was due to the exact same predictions at different resolutions. In this instance too, the error rate decreased as the fitting resolution became lower. The RMSE values ranged from 16.74 to 20.73 %.

Table 2. RMSE values (%) for the different models and resolutions. Fitting resolution is by rows and prediction resolution by columns. The change of RMSE compared to its value in native resolution (diagonal) is shown in parenthesis. Absolute RMSE values ($Mg \cdot ha^{-1}$) are in appendix A1.

	Predict

Model		Plot size	8.33 m	12.5 m	25.0 m
LLS-XR	Fit	8.33 m	17.38 (0.00)	17.54 (0.16)	17.53 (0.15)
		12.5 m	16.81 (-0.10)	16.91 (0.00)	16.91 (0.00)
		25.0 m	16.74 (-0.05)	16.80 (0.01)	16.79 (0.00)
LLS-XNR	Fit	8.33 m	19.90 (0.00)	20.46 (0.56)	20.73 (0.82)
		12.5 m	19.01 (-0.34)	19.35 (0.00)	19.50 (0.15)
		25.0 m	19.18 (-0.21)	19.35 (-0.04)	19.39 (0.00)
NLS-XR	Fit	8.33 m	18.37 (0.00)	18.32 (-0.05)	18.40 (0.03)
		12.5 m	17.49 (-0.03)	17.52 (0.00)	17.58 (0.06)
		25.0 m	17.24 (-0.10)	17.28 (-0.06)	17.34 (0.00)
NLS-XNR	Fit	8.33 m	19.01 (0.00)	19.98 (0.97)	20.25 (1.24)
		12.5 m	18.07 (-0.57)	18.64 (0.00)	18.67 (0.03)
		25.0 m	18.27 (-0.33)	18.63 (0.02)	18.61 (0.00)
LLS-RXN	Fit	8.33 m	17.77 (0.00)	17.77 -(0.00)	17.77 (0.00)
		12.5 m	17.75 (0.00)	17.75 -(0.00)	17.75 (0.00)
		25.0 m	17.68 (0.00)	17.68 (0.00)	17.68 (0.00)

Bias

The BIAS values for the different models and resolutions are presented in Table 3. The resolution effect with respect to BIAS was smaller with predictor variables that had a resolution invariant mean (LLS-XR and NLS-XR) than with variables that did not have this property (LLS-XNR and NLS-XNR). If the prediction resolution was lower than the fitting resolution it resulted in an overestimation, and conversely, if the prediction resolution was higher than the fitting resolution it produced an underestimation. However, the direction of BIAS was different in NLS-XR than in the other cases. There was no clear difference in terms of BIAS between the linear (LLS) and non-linear (LLS) models. The LLS-RXN models showed small BIAS, which was consistent with the same model at different resolutions. This small BIAS is due to cross-validation; without cross-validation the LLS-RXN models show zero BIAS.

376 Table 3. BIAS values (%) for the different models and resolutions. Fitting resolution is by rows and
377 prediction resolution by columns. Negative values denote an underestimation and positive values an
378 overestimation (eq. 18). Absolute RMSE values ($Mg \cdot ha^{-1}$) are in appendix A2.

			Predict		
Model		Plot size	8.33 m	12.5 m	25.0 m
LLS-XR	Fit	8.33 m	-0.03	0.12	0.21
		12.5 m	-0.16	-0.02	0.07
		25.0 m	-0.21	-0.09	-0.01
LLS-XNR	Fit	8.33 m	-0.03	1.46	3.03
		12.5 m	-1.36	0.00	1.45
		25.0 m	-2.60	-1.35	0.01
NLS-XR	Fit	8.33 m	-0.02	-0.71	-1.48
		12.5 m	0.56	0.05	-0.55
		25.0 m	0.97	0.55	0.05
NLS-XNR	Fit	8.33 m	-0.02	1.50	2.61
		12.5 m	-1.32	0.06	1.08
		25.0 m	-2.26	-0.94	0.07
LLS-RXN	Fit	8.33 m	-0.13	-0.13	-0.13
		12.5 m	-0.11	-0.11	-0.11
		25.0 m	-0.12	-0.12	-0.12

379 Figure 3 shows the effect of using varying height percentile on prediction BIAS (HPERC models). A
380 clear trend in the opposing directions was observed. Fitting with the 8.33 m resolution and predicting with
381 the 25.0 m resolution led to a positive BIAS with low percentiles and a negative BIAS with high
382 percentiles, while fitting with the 25.0 m resolution and predicting with the 8.33 m resolution led to a
383 negative BIAS with low percentiles and positive BIAS with high percentiles. The more extreme the
384 percentile, the more serious the BIAS; the highest BIAS was 2.8%.We can interpolate between the nodes
385 in the graph and estimate that (in this case) unbiased predictions could be obtained with the 43^{th}
386 percentile.

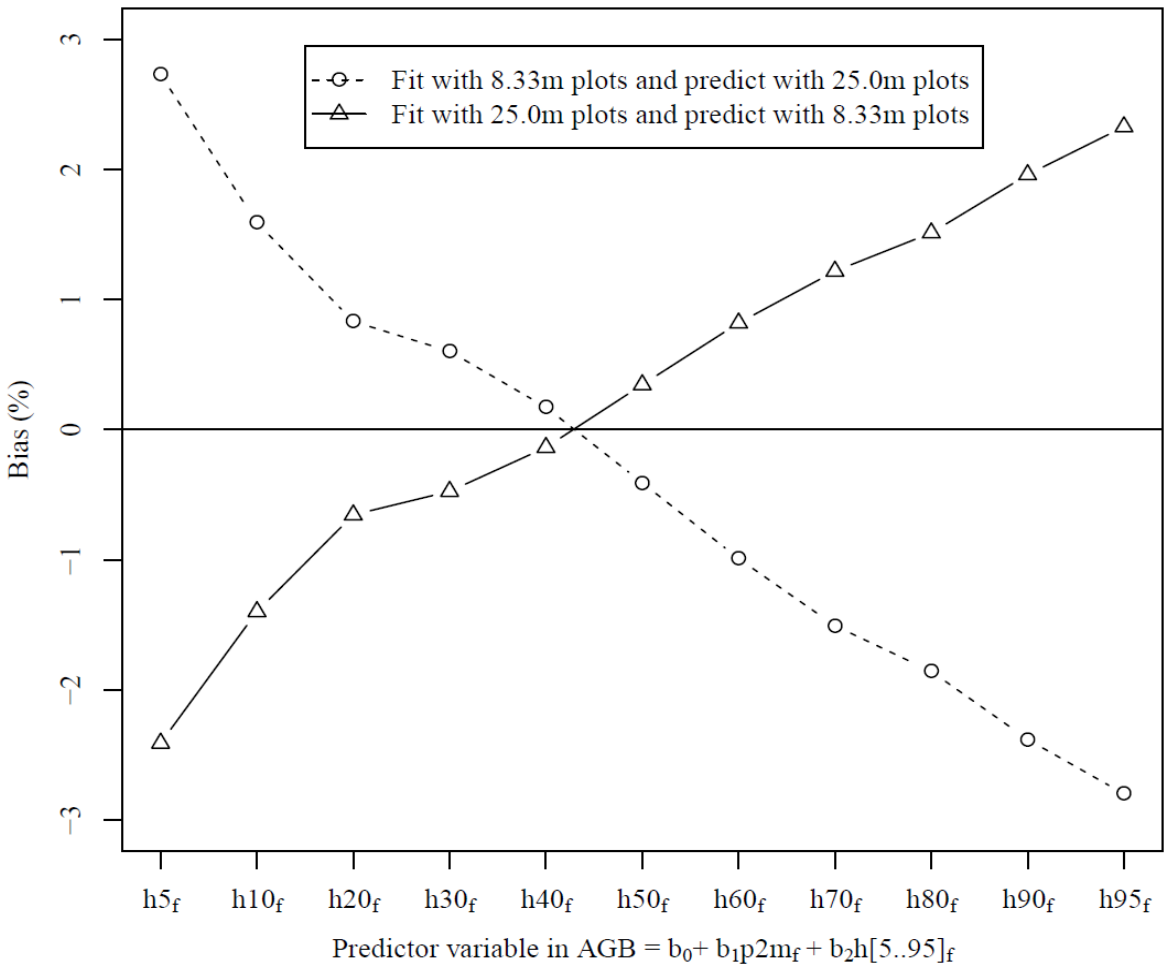


Figure 3. Behavior of BIAS using fixed $p2m_f$ variable and varying height percentile $h[5…95]_f$ (HPERC models). Line -○- depicts the case where a model is fitted at the resolution of 8.33 m and the prediction is done at the resolution of 25.0 m. Line -△- depicts the opposite; a model is fitted at the 25.0 m resolution and the prediction is done at the 8.33 m resolution.

Discussion

The main findings of the behavior of RMSE and BIAS with respect to differing fitting and prediction resolutions are summarized in Figure 4. The RMSE value increases when the prediction resolution is lower than the fitting resolution and decreases when the fitting resolution is lower than the prediction resolution. The direction of BIAS was not consistent across all models. In most cases, however, the result was an overestimation when the prediction resolution was lower than the fitting resolution, and vice versa. This indicates that the direction of BIAS also depends on the model form. Predicting at resolutions

lower than the fitting resolution leads to an overestimation with low height percentiles (h[low]) and an underestimation with high height percentiles (h[high]), while predicting at resolutions higher than the fitting resolution leads to an underestimation with low percentiles and an overestimation with high percentiles. This is due to the bias of sample percentiles as estimators of population percentiles. The sample percentiles overestimate low population percentiles and underestimate the high percentiles (Hyndman and Fan, 1996).

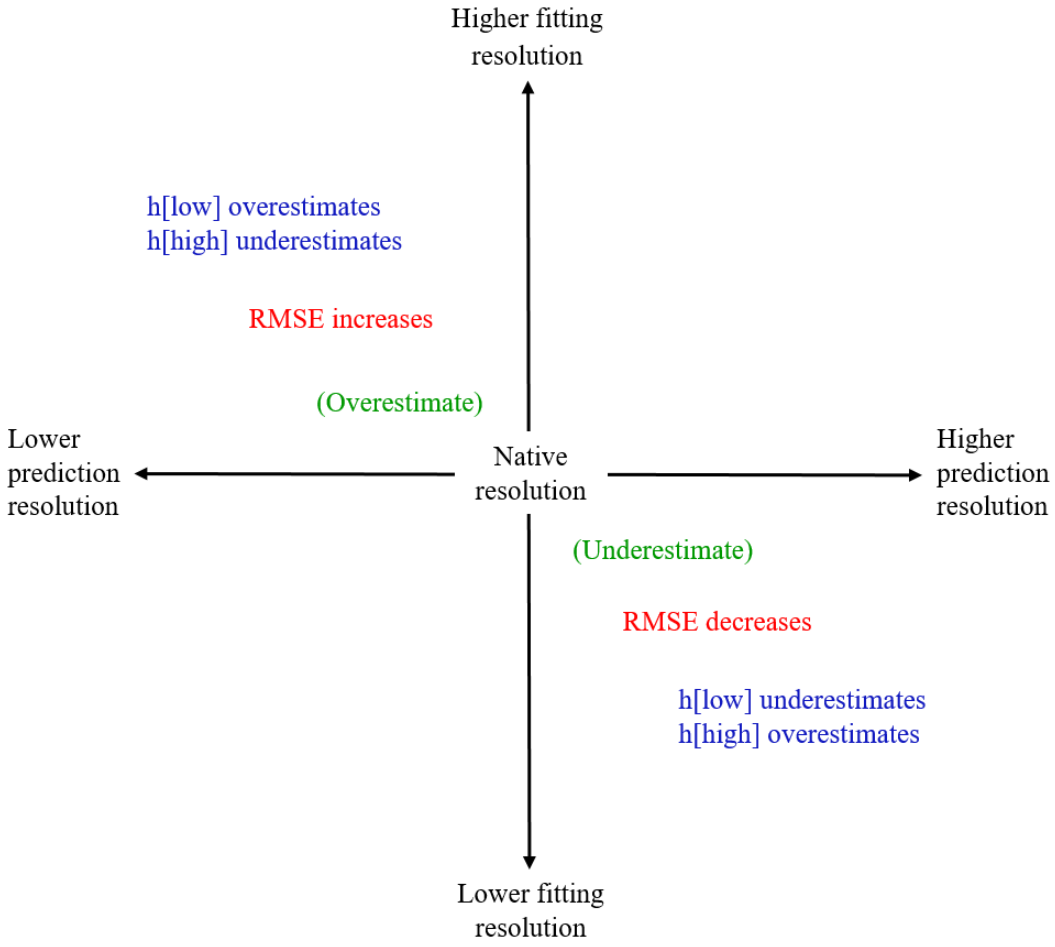


Figure 4. Main findings of the study with respect to lower vs. higher fitting and prediction resolutions. Overestimate and underestimate in brackets denotes that the direction of BIAS was not consistent across all models, but in most cases it was as depicted. h[low] and h[high] refer to low and high height percentiles, respectively.

Our results indicate that the lowest error rate is obtained when a model is fitted at the lowest resolution and when prediction is done at the highest resolution. A reason for this may be that the metrics computed from larger plots lead to regression coefficients that are closer to the true value than those obtained with small plot sizes. This may be due to larger measurement errors in predictions that use smaller plots (Gobakken and Næsset, 1999), which reduce regression coefficient estimates towards zero (Lappi, 1993; Carroll et al., 2006), or because the edge tree effect, which is similar to a measurement error, is greater in the smaller plots (Packalen et al., 2015). However, as there is a tradeoff between plot size and the number of plots it is not obvious which plot size is optimal given the need for sufficient numbers of plots as well as sufficiently large plots. In the real world, resources available for field measurements are typically limited, and increasing the plot areas would result in a fewer numbers of plots.

The level of BIAS in this study seems to be quite low even if several conditions of resolution invariance are violated at the same time. The maximum bias of 3.03 % is observed when the resolution is changed 9-fold. When the resolution is changed 4-fold, the maximum bias is 1.50 %. In a real world scenario, the change of resolution is hardly ever as high as 4-fold. The effect of resolution on BIAS is smaller with predictor variables that have a resolution invariant mean than with variables that do not have this property. In the case of the linear model and predictor variables with a resolution invariant mean (LLS-XR), the maximum bias is only 0.21 % when the resolution is changed 9-fold. However, it should be noted that it is not exactly a resolution invariant because the number of ALS points per cell varies.

The Pearson correlation coefficient between the BIAS (Table 3) and change of RMSE (Table 2, figures in parenthesis) is 0.78 when all 54 cases are considered. This is expected behavior because $RMSE^2 = BIAS^2$ + variance. There is a trend that BIAS is positive when RMSE increases with respect to the native resolution, and vice versa when it decreases. The NLS-XR model was an exception: underestimation increased and overestimation decreased RMSE with respect to the native resolution. We can conclude that the decrease or increase of RMSE values with respect to native resolution does not depend on the sign of BIAS but depends on model form and predictor variables.

435 Resolution invariant prediction is feasible with LLS-RXN. The main idea is to rescale predictor variables
436 in order to take into account the varying number of echoes per unit of area (eq. 13-14). Otherwise, LLS-
437 RXN is the same as LLS-XR. “Resolution invariance” means that prediction BIAS is always exactly zero,
438 although there is a tradeoff in terms of RMSE. LLS-RXN has consistently higher RMSE values than
439 LLS-XR, because the weighting does not contribute to the relationship between response and predictor
440 values – they simply enforce a desired condition of resolution invariance. Therefore, a greater RMSE is
441 the price to be paid for exact resolution invariance. It is also debatable whether weighting with the
442 number of echoes per unit area makes sense. LLS-RXN corrects for the effects of irregular point density
443 resulting from variability in the ALS data acquisition process. Consequently, areas with a higher point
444 density receive higher weighting than areas with lower point density – where in practice, we are equally
445 interested in every location. An alternative approach to revise irregular point density would be to
446 standardize ALS point density per area unit by interpolating a continuous surface and computing metrics
447 from this surface, instead of computing them directly from the ALS echoes. A canopy height model
448 (CHM) is an example of this kind of surface: condition 1 is fulfilled if plots and cells are arranged such
449 that the number of pixels per unit area is the same in every plot and cell. Zhao et al. (2009) used this
450 approach to derive canopy height distributions. Chirici et al. (2016) studied the decrease in accuracy if
451 CHM metrics are used instead of echo metrics in the estimation of forest aboveground biomass. They
452 found that the R^2 decreased from 0.58 to 0.56 with a linear model, and from 0.54 to 0.48 with the k-NN
453 technique. Undoubtedly, the use CHM metrics increases the prediction error but this trade-off may be
454 acceptable. On the other hand, if unbiasedness is not a strict requirement, the effect of a slight imbalance
455 in the number of echoes per unit area may not have very much practical significance.

456 Note that the LLS-RXN model accounts for the effect of varying point density both within and between
457 plots (and cells). In classical ABA, between plots (or cells) variation in point density is not an issue
458 because the value of metrics do not depend on the number of points per plot. However, an intrinsic
459 property of the classical ABA is that higher point density areas within a plot get more weight than lower

point density areas within a plot. The use of a continuous surface, such as computing metrics from the CHM, does not have an issue with either within a plot or between plots variation in point density.

It is possible to use only ALS metrics that fulfill the condition of a resolution invariant mean, but there are several commonly used metrics that do not obey this condition. Height percentile is the most common category of this type and is frequently used because there is a strong relationship between upper percentiles and the main attributes of interest in forest inventories (Næsset, 2002). It is also common practice to include only vegetation echoes when computing height metrics. Typically, this is implemented by excluding echoes below 1-2 m (Næsset, 2004; Vauhkonen et al., 2012; Chen et al., 2016). The application of a height threshold makes the situation even more difficult from a resolution viewpoint, because it causes the number of echoes considered by a metric to vary considerable cell by cell. Figure 3 demonstrates the behavior of BIAS with a linear model mimicking a real world use case when the resolution is changed 9-fold. The model contains one cover metric and varying height percentiles 5…95 computed with a height threshold of 2 m. This shows that bias increases more or less symmetrically towards the extreme percentiles. If resolution bias is of concern, then it can be mitigated to some degree by simply avoiding extreme percentiles.

Drawing absolute conclusions with regard to the level of BIAS or RMSE is difficult when resolution invariance conditions are not met and the resolution is changed. For example, one may use a non-linear model form that is exceedingly sensitive to resolution, or a predictor variable that may be extremely sensitive to resolution, such as maximum echo height. In cases when forest inventory attributes that are not additive in nature, such as Lorey's Height (mean height weighted by basal area) or DGM (diameter of the basal area median tree), resolution invariance cannot be achieved with the methods examined here, and the observations made here with regard to BIAS and RMSE do not apply.

The use of 3D remote sensing data is becoming increasingly widespread and different point cloud generation approaches are now available. Point cloud reconstruction from image pairs is currently the most common alternative (St-Onge et al., 2008) to ALS and also new Lidar techniques have emerged

(Swatantran et al., 2016). The observations made in this study apply partly to image point clouds but there are differences that must be taken into account. For example, some image point cloud software produces a regularly spaced surface model instead of a true point cloud, in which case varying point density is not an issue.

Conclusions

A number of factors contribute to resolution dependence in ABA forest inventories. These include the varying point density of the ALS data, the type of response variable used, how the predictor variables are computed, and the properties of the model. Complete resolution invariance is feasible using the LLS-RXN approach, or by computing metrics from a continuous surface interpolated with ALS data (i.e. CHM) and by meeting other conditions too.

The maximum BIAS was 1.50 % and the maximum change of RMSE compared to its value in native resolution was 0.97 % when there was a 4-fold difference in resolution. This indicates that the resolution effect is small in most real-world use cases, especially as the difference in plot or cell size is usually considerably smaller than the 4-fold. The effect of resolution on BIAS and change of RMSE compared to its value in native resolution was much smaller with XR than XNR type metrics. Therefore, if resolution effect is of concern, XR type metrics are recommended, although they are not resolution invariant in an ABA context because the number of ALS echoes per unit area is not constant. In this study, the model type (LLS or NLS) did not have any clear effect on BIAS or change of RMSE.

Unbiasedness is not a strict requirement in most stand level forest management inventories. In that case, resolution invariance may be of limited practical importance, particularly because at the stand level the error rate will typically greatly exceed the level of bias caused by the resolution invariance. For strategic inventories that cover large areas, the importance of resolution invariance is greater. In a model-based or model-assisted approach to estimation over large areas, it is assumed that model residuals are an unbiased sample of the deviations that would be observed between predictions and observations. In the case of resolution differences between the fitting and prediction datasets, this property does not hold, and the bias

may be large relative to the sampling variation. Attention to resolution dependence will thus be most relevant when large area inferences are made from ALS-assisted inventories.

References

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. In: International Archives of Photogrammetry and Remote Sensing, Amsterdam, The Netherlands, 16–22 July 2000, Vol. XXXIII, part B4, 110–117.

Carroll, R.J., Ruppert, D., Stefanski, L.A., Crainiceanu, C.M., 2006. Measurement Error in Nonlinear Models: A Modern Perspective (Second ed.). Chapman and Hall/CRC (484 p).

Chasmer, L., Barr, A., Hopkinson, C., McCaughey, H., Treitz, P., Black, A., Shashkov, A., 2009. Scaling and assessment of GPP from MODIS using a combination of airborne lidar and eddy covariance measurements over jack pine forests. Remote Sensing of Environment113, 82–93.

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

Chirici, G., McRoberts, R.E., Fattorini, L., Mura, M., Marchetti, M., 2016. Comparing echo-based and canopy height model-based metrics for enhancing estimation of forest aboveground biomass in a model-assisted framework. Remote Sensing of Environment 174, 1–9.

Frazer, G.W., Magnussen, S., Wulder, M.A., Niemann, K.O., 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. Remote Sensing of Environment 115, 636-649.

Gobakken, T., Næsset, E. 2009. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. Canadian Journal of Forest Research 39, 1036-1052.

Goetz, S., Dubayah, R., 2011. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Management 2:3, 231–244.

Grafström, A., Zhao, X., Nylander, M., Petersson, H., 2017. A new sampling strategy for forest inventories applied to the temporary clusters of the Swedish national forest inventory. Canadian Journal of Forest Research 47, 1161–1167.

Gregoire, T.G., Næsset, E., McRoberts, R.E., Ståhl, G., Andersen, H.-E., Gobakken, T., Ene, L., Nelson, R., 2016. Statistical rigor in LiDAR-assisted estimation of aboveground forest biomass. Remote Sensing of Environment 173, 98–108.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., Falkowski, M.J., 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment 112, 2232–2245.

Hyndman, R. J., Fan, Y., 1996. Sample quantiles in statistical packages. American Statistician 50, 361–365.

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

Hyyppä, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. IEEE Transactions on Geoscience and Remote Sensing 39, 969-975.

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

Keränen, J., Maltamo, M., Packalen, P., 2016. Effect of flying altitude, scanning angle and scanning mode on the accuracy of ALS based forest inventory. International Journal of Applied Earth Observation and Geoinformation 52, 349–360.

Koch, B. Heyder, U., Weinacker, H., 2006. Detection of individual tree crowns in airborne lidar data. Photogrammetric Engineering & Remote Sensing 72, 357–363.

559 Lähivaara, T., Seppänen, A., Kaipio, J.P., Vauhkonen, J., Korhonen, L., Tokola, T., Maltamo, M., 2014.
560 Bayesian approach to tree detection based on airborne laser scanning data. IEEE Transactions on
561 Geoscience and Remote Sensing 52, 2690–2699.

562 Lappi, J., 1993. Metsäbiometrian menetelmiä. Silva Carelica 24. University of Joensuu (182 p).

563 Liang, S., 2004. Quantitative Remote Sensing of Land Surfaces. John Wiley & sons, Inc. Hoboken, New
564 Jersey (560 p).

565 Magnussen, S., Næsset, E., Gobakken, T., 2013. Prediction of tree-size distributions and inventory
566 variables from cumulants of canopy height distributions. Forestry 86, 583–595.

567 Maltamo, M., Næsset, E., Vauhkonen, J. 2014. Forestry applications of airborne laser scanning -concepts
568 and case studies. Managing Forest Ecosystems 27. Springer (464 p).

569 Marceau, D.J., Hay, G.J., 1999. Remote Sensing Contributions to the Scale Issue. Canadian Journal of
570 Remote Sensing 25, 357–366.

571 Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., Hendrix, C.J., 2000. Predicting forest
572 stand characteristics with airborne scanning LiDAR. Photogrammetric Engineering & Remote Sensing
573 66, 1367–1372.

574 Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical
575 two-stage procedure and field data. Remote Sensing of Environment 80, 88–99.

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

578 Næsset, E., 2005. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on
579 biophysical stand properties derived from small-footprint airborne laser data. Remote Sensing of
580 Environment 98, 356–370.

581 Nilsson, M., Nordkvist, K., Jonzén, J., Lindgren, N., Axensten, P., Wallerman, J., Egberth, M., Larsson,
582 S., Nilsson, L., Eriksson, J., Olsson, H., 2017. A nationwide forest attribute map of Sweden predicted
583 using airborne laser scanning data and field data from the National Forest Inventory. Remote Sensing of
584 Environment 194, 447–454.

585 Niska, H., Skön, J-P., Packalén, P., Tokola, T., Maltamo, M., Kolehmainen, M., 2010. Neural Networks
586 for the Prediction of Species-Specific Plot Volumes Using Airborne Laser Scanning and Aerial
587 Photographs. IEEE Transactions on Geoscience and Remote Sensing 48, 1076–1085.

588 Packalén, P., Maltamo, M., 2007. The k-MSN method for the prediction of species-specific stand
589 attributes using airborne laser scanning and aerial photographs. Remote Sensing of Environment 109,
590 328–341.

591 Packalén, P., Mehtätalo, L., Maltamo, M., 2011. ALS-based estimation of plot volume and site index in a
592 eucalyptus plantation with a nonlinear mixed-effect model that accounts for the clone effect. Annals of
593 Forest Science 68, 1085–1092.

594 Packalen, P., Strunk, J., Pitkänen, J., Temesgen, H., Maltamo, M., 2015. Edge-tree Correction for
595 Predicting Forest Inventory Attributes Using Area-based Approach With Airborne Laser Scanning. IEEE
596 Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8, 1274–1280.

597 R Development Core Team, 2011. R: A language and environment for statistical computing. R
598 Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-
599 project.org/.

600 Raffy, M., 1992. Change of scale in models of remote sensing: A general method for spatialization of
601 models. Remote Sensing of Environment 40, 101–112.

602 Reiss, R.-D., 1989. Approximate distributions of order statistics with applications to nonparametric
603 statistics. Springer-Verlag (355 p).

604 Repola, J., 2008. Biomass equations for birch in Finland. Silva Fennica 42, 605–624.

605 Repola, J., 2009. Biomass equations for Scots pine and Norway spruce in Finland. Silva Fennica 43, 625–
606 647.

607 Ruiz, A.L., Hermosilla, T., Mauro, F., Godino, M., 2014. Analysis of the Influence of Plot Size and
608 LiDAR Density on Forest Structure Attribute Estimates. Forests 5(5), 936–951.

609 Searle, S.R., 1971. Linear models. John Wiley & Sons, New York (532 p).

610 Simic, A., Chen, J.M., Liu, J., Csillag, F., 2004. Spatial scaling of net primary productivity using subpixel
611 information. Remote Sensing of Environment 93, 246–258.

612 St-Onge, B., Vega, C., Fournier, R.A., Hu, Y., 2008. Mapping canopy height using a combination of
613 digital stereo-photogrammetry and lidar. International Journal of Remote Sensing 29, 3343–3364.

614 Strahler, A.H., Woodcock, C.E., Smith, J.A., 1986. On the nature of models in remote sensing. Remote
615 Sensing of Environment 20, 121–139.

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

618 Swatantran, A., Tang, H., Barrett, T., DeCola, P., Dubayah, R., 2016. Rapid, High-Resolution Forest
619 Structure and Terrain Mapping over Large Areas using Single Photon Lidar. Scientific Reports 6, Article
620 number: 28277.

621 Thomas, V., Treitz, P., McCaughey, J.H., Morrison I., 2006. Mapping stand-level forest biophysical
622 variables for a mixedwood boreal forest using lidar: an examination of scanning density, Canadian
623 Journal of Forest Research 36, 34–47.

624 Turner, M. G., Dale, V. H., Gardner, R. H., 1989. Predicting across scales: Theory development and
625 testing. Landscape Ecology 3, 245–252.

626 Vauhkonen, J., Seppänen, A., Packalén, P., Tokola, T., 2012. Improving species-specific plot volume
627 estimates based on airborne laser scanning and image data using alpha shape metrics and balanced field
628 data. Remote Sensing of Environment 124, 534–541.

629 Wu, H., Li, Z-L., 2009. Scale Issues in Remote Sensing: A Review on Analysis, Processing and
630 Modeling. Sensors 9, 1768–1793.

631 Zhao, K., Popescu, S.C., Nelson, R.F., 2009. Lidar remote sensing of forest biomass: A scale-invariant
632 estimation approach using airborne lasers. Remote Sensing of Environment 113, 182–196.

633

634

635

636

637

638

639

640

641

642

643

644

List of Figure Captions

Figure 1. Plot setting.

Figure 2. An example of model fitting, prediction and aggregation workflow.

Figure 3. Behavior of BIAS using fixed $p2m_f$ variable and varying height percentile $h[5…95]_f$ (HPERC models).

Figure 4. Main findings of the study with respect to lower vs. higher fitting and prediction resolutions. Overestimate and underestimate in brackets denotes that the direction of BIAS was not consistent across all models, but in most cases it was as depicted. h[low] and h[high] refer to low and high height percentiles, respectively.

Appendix A.

Table A1

Absolute RMSE values ($Mg \cdot ha^{-1}$) for the different models and resolutions. Fitting resolution is by rows and prediction resolution by columns. The change of RMSE compared to its value in native resolution (diagonal) is shown in parenthesis.

			Predict		
Model		Plot size	8.33 m	12.5 m	25.0 m
LLS-XR	Fit	8.33 m	18.98 (0.00)	19.16 (0.17)	19.15 (0.17)
		12.5 m	18.35 (-0.11)	18.47 (0.00)	18.46 (0.02)
		25.0 m	18.28 (-0.06)	18.35 (0.01)	18.34 (0.00)
LLS-XNR	Fit	8.33 m	21.74 (0.00)	22.34 (0.61)	22.63 (0.90)
		12.5 m	20.76 (-0.37)	21.13 (0.00)	21.30 (0.17)
		25.0 m	20.94 (-0.23)	21.13 (-0.04)	21.17 (0.00)
NLS-XR	Fit	8.33 m	20.06 (0.00)	20.01 (-0.06)	20.10 (0.03)
		12.5 m	19.10 (-0.03)	19.13 (0.00)	19.19 (0.06)
		25.0 m	18.83 (-0.11)	18.87 (-0.07)	18.94 (0.00)
NLS-XNR	Fit	8.33 m	20.76 (0.00)	21.82 (1.06)	22.11 (1.35)
		12.5 m	19.73 (-0.62)	20.35 (0.00)	20.39 (0.04)
		25.0 m	19.95 (-0.36)	20.34 (0.02)	20.32 (0.00)
LLS-RXN	Fit	8.33 m	19.40 (0.00)	19.40 (0.00)	19.40 (0.00)
		12.5 m	19.38 (0.00)	19.38 (0.00)	19.38 (0.00)
		25.0 m	19.31 (0.00)	19.31 (0.00)	19.31 (0.00)

675 Table A2

676 Absolute BIAS values ($Mg \cdot ha^{-1}$) for the different models and resolutions. Fitting resolution is by rows
677 and prediction resolution by columns. Negative values denote an underestimation and positive values an
678 overestimation (eq. 18).

			Predict		
Model		Plot size	8.33 m	12.5 m	25.0 m
LLS-XR	Fit	8.33 m	-0.03	0.13	0.23
		12.5 m	-0.18	-0.02	0.07
		25.0 m	-0.23	-0.09	-0.01
LLS-XNR	Fit	8.33 m	-0.04	1.59	3.31
		12.5 m	-1.48	0.00	1.59
		25.0 m	-2.84	-1.47	0.02
NLS-XR	Fit	8.33 m	-0.02	-0.77	-1.61
		12.5 m	0.61	0.06	-0.60
		25.0 m	1.06	0.60	0.05
NLS-XNR	Fit	8.33 m	-0.03	1.63	2.85
		12.5 m	-1.44	0.06	1.18
		25.0 m	-2.47	-1.03	0.08
LLS-RXN	Fit	8.33 m	-0.14	-0.14	-0.14
		12.5 m	-0.12	-0.12	-0.12
		25.0 m	-0.13	-0.13	-0.13

679